

Información Importante

La Universidad Santo Tomás, informa que el(los) autor(es) ha(n) autorizado a usuarios internos y externos de la institución a consultar el contenido de este documento a través del Catálogo en línea del CRAI-Biblioteca y el Repositorio Institucional en la página Web de la CRAI-Biblioteca, así como en las redes de información del país y del exterior con las cuales tenga convenio la Universidad.

Se permite la consulta a los usuarios interesados en el contenido de este documento, para todos los usos que tengan **finalidad académica**, nunca para usos comerciales, siempre y cuando mediante la correspondiente cita bibliográfica se le dé crédito al trabajo de grado y a su autor.

De conformidad con lo establecido en el Artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, la Universidad Santo Tomás informa que “los derechos morales sobre documento son propiedad de los autores, los cuales son irrenunciables, imprescriptibles, inembargables e inalienables.”

**Centro de Recursos para el Aprendizaje y la Investigación, CRAI-Biblioteca
Universidad Santo Tomás, Bucaramanga**

Tendencias y características de los viajeros que visitan la ciudad de Bucaramanga. Una solución de inteligencia de negocios basada en minería de datos.

Liliana Pabuena Villarreal

Trabajo de Grado para optar al título de Ingeniería Industrial

Director

Javier Hernández Cáceres

Ingeniero Industrial

Universidad Santo Tomás, Bucaramanga

División de Ingenierías y Arquitectura

Facultad de Ingeniería Industrial

2017

Dedicatoria

Dedico este trabajo primeramente a Dios, por haberme dado la vida y permitirme haber llegado hasta este importante momento de mi formación profesional. También agradecer a Dios por estar conmigo en cada paso que doy, por brindarme la fortaleza que mi corazón necesita e iluminar mi mente, y lo más importante, por haber puesto en mi camino aquellas personas que han sido mi soporte y compañía durante toda esta etapa educativa.

Agradecer a mis padres por todos y cada uno de los sacrificios que han hecho por mí por hacer una excelente persona, capaz de lograr lo que quiere y amar lo que anhela alcanzar. A mi madre Susana Villarreal por siempre creer en mí y por su apoyo incondicional. A mi padre Leonel Pabuena porque gracias a él aprendí que los sueños se alcanzan a pesar de los obstáculos que la vida nos quiera colocar, para hacer de cada uno, la versión original y mejorada de lo que realmente somos. A ustedes les estaré siempre inmensamente agradecida por ser su hija, y a Dios por regalarme los mejores padres.

A mi esposo Jonathan García, el amor de mi vida, agradecerle por hacer de mí la persona que hoy en día soy, porque sé que sin su amor y su apoyo no sé dónde estaría. Gracias por los consejos, las palabras de apoyo, el amor incondicional, por querer siempre que sea la mejor aunque en ocasiones no lo logre. A él agradezco el gran sacrificio de apoyarme con nuestro hijo para que sacará mi profesión adelante, y por ese hermoso sacrificio sé que estoy logrando esta gran meta. Por ti amor lo logré.

A mi hijo hermoso Angelus Gavri'el García por quien luchó día a día, por quien anhelo ser mejor para ser su orgullo y porque quiero que en un futuro sienta que soy para él, una mujer de admirar y con mucha fuerza para luchar por lo que ama y desea. A todos ustedes los amo.

Tabla de contenido

Introducción.	12
1. Definición del Problema.	14
1.1. Descripción del Problema.	14
1.2. Formulación del Problema.	26
1.3. Justificación del Problema.	27
1.4. Objetivos.	29
1.4.1. Objetivo General.	29
1.4.2. Objetivos Específicos.	29
1.5. Alcance.	30
2. Marco Referencial.	33
2.1. Marco Teórico.	33
2.1.1. Minería de Datos y Extracción de conocimiento a partir de los datos.	33
2.1.2. Enchancing Tourism Intermediaries with the Data Mining Process.	41
2.1.3. A Knowledge Destination Framework for Tourism Sustainability: A Business Intelligence Application from Sweden.	43
2.1.4. A Supplier-Oriented Knowledge Application – The Destination Management Information System (DMIS).	45
2.1.5. Atributos relevantes en el perfil del turista del interior de la ciudad de Mar de la Plata. 46	46
2.2. Marco Conceptual.	48
2.3. Marco Legal.	51
2.4. Marco Histórico.	57
2.5. Estado del Arte.	64
3. Metodología.	73
3.1. Fundamentos Epistemológicos.	73
3.2. Diseño de la Investigación.	77
3.2.1. Tipo de Investigación.	77
3.2.2. Fases, etapas o momentos del proyecto.	77
3.2.3. Técnicas e instrumentos de Recolección en Investigación Cuantitativa.	79
4. Análisis e Interpretación De Resultados.	80
4.1. Desarrollo del Proceso de Minería de Datos con el modelo CRISP-DM.	80

Conclusiones.....	202
Recomendaciones.....	204
Bibliografía.....	207
Apéndice.....	213

Lista de Figuras

Figura 1. Indicadores año 2015 Turismo en Colombia	14
Figura 2. Flujo de Vehículos por departamento. Enero-Julio 2015-2016	15
Figura 3. Total de Agencias por Departamento – Julio 2016	18
Figura 4. Proceso Obtenido desde el software WEKA	104
Figura 5. Asignaciones de cada clúster utilizando el software WEKA	105
Figura 6. Asignaciones del clúster 0 utilizando el software WEKA	106
Figura 7. Asignaciones del clúster 1 utilizando el software WEKA	107
Figura 8. Asignaciones del clúster 2 utilizando el software WEKA	108
Figura 9. Proceso para generar reglas de asociación numRules=20	110
Figura 10. Proceso para generar reglas de asociación numRules=30	111
Figura 11. Atributo Edad de los encuestados	112
Figura 12. Atributo Género de los encuestados	113
Figura 13. Atributo CiudadOrigen de los encuestados	114
Figura 14. Atributo PaísOrigen de los encuestados	115
Figura 15. Atributo EstadoCívil de los encuestados	116
Figura 16. Atributo NivelEscolaridad de los encuestados	117
Figura 17. Atributo Ocupación de los encuestados	118
Figura 18. Atributo N°VecesViajado de los encuestados	119
Figura 19. Atributo MotivoViaje de los encuestados	120
Figura 20. Atributo N°DíasPermanencia de los encuestados	121
Figura 21. Atributo N°CompañeroViaje de los encuestados	122
Figura 22. Atributo LugarAlojamiento de los encuestados	123
Figura 23. Atributo ActividadPlaneada de los encuestados	124
Figura 24. Atributo MedioTransportedeLlegada de los encuestados	125
Figura 25. Atributo MedioTransporteenlaCiudad de los encuestados	126
Figura 26. Atributo LugarAlimentación de los encuestados	127
Figura 27. Atributo GastoPromedio de los encuestados	128
Figura 28. Atributo CalificaciónGastoPromedio de los encuestados	129
Figura 29. Atributo AspectoqueImpactodelaCiudad de los encuestados	130
Figura 30. Atributo AspectoMotivaaRegresar de los encuestados	131
Figura 31. Atributo GradoSatisfacciónEntrenamiento de los encuestados	132
Figura 32. Atributo GradoSatisfacciónComercial de los encuestados	133
Figura 33. Atributo GradoSatisfacciónCultural de los encuestados	134
Figura 34. Atributo GradoSatisfacciónSeñalizaciónTurística de los encuestados	135
Figura 35. Atributo GradoSatisfacciónCulturaCiudadana de los encuestados	136
Figura 36. Atributo GradoSatisfacciónAlojamiento de los encuestados	137
Figura 37. Atributo GradoSatisfacciónCalidadAtenciónRecibida de los encuestados	138
Figura 38. Atributo GradoSatisfacciónCalidadAlojamiento de los encuestados	139
Figura 39. Atributo GradoSatisfacciónDisponibilidadAlojamiento de los encuestados	140
Figura 40. Atributo GradoSatisfacciónPrecioAlojamiento de los encuestados	141
Figura 41. Atributo GradoSatisfacciónOficinaTurismooTransportadoras de los encuestados	142
Figura 42. Atributo GradoSatisfacciónHoteluHospedaje de los encuestados	143
Figura 43. Atributo GradoSatisfacciónPersonaenlaCalle de los encuestados	144
Figura 44. Atributo GradoSatisfacciónPolicia de los encuestados	145
Figura 45. Atributo GradoSatisfacciónPolicia de Tránsito de los encuestados	146

Figura 46. Atributo GradoSatisfacciónPoliciaCarreteras de los encuestados	147
Figura 47. Atributo GradoSatisfacciónInformaciónTurística de los encuestados	148
Figura 48. Análisis de Clúster totales estadísticos de los viajeros	150
Figura 49. Edad viajeros que visitan la ciudad de Bucaramanga	151
Figura 50. Género viajeros que visitan la ciudad de Bucaramanga	152
Figura 51. CiudadOrigen viajeros que visitan la ciudad de Bucaramanga	154
Figura 52. Ocupación viajeros que visitan la ciudad de Bucaramanga	155
Figura 53. GastoPromedio viajeros que visitan la ciudad de Bucaramanga	157
Figura 54. AspectoMotivaaRegresar viajeros que visitan la ciudad de Bucaramanga	158
Figura 55. AspectoqueImpactodelaCiudad viajeros que visitan la ciudad de Bucaramanga	160
Figura 56. Histograma Género-Edad	161
Figura 57. Histograma Género-EstadoCívil	162
Figura 58. Histograma Género-NivelEscolaridad	163
Figura 59. Histograma Género-Ocupación	164
Figura 60. Histograma Género-FuenteIngreso	165
Figura 61. Histograma Género-N°VecesViajado	166
Figura 62. Histograma Género-Ocupación	167
Figura 63. Histograma Edad-Ocupación	168
Figura 64. Histograma Edad-NivelEscolaridad	169
Figura 65. Histograma Edad-EstadoCivil	170
Figura 66. Histograma Edad-FuenteIngreso	171
Figura 67. Asignaciones de cada clúster utilizando el software WEKA	175
Figura 68. Asignaciones del clúster 0 utilizando el software WEKA	176
Figura 69. Asignaciones del clúster 1 utilizando el software WEKA	176
Figura 70. Asignaciones del clúster 2 utilizando el software WEKA	177
Figura 71. Asignaciones de cada clúster utilizando el software WEKA	179
Figura 72. Asignaciones del clúster 0 utilizando el software WEKA	180
Figura 73. Asignaciones del clúster 1 utilizando el software WEKA	180
Figura 74. Asignaciones del clúster 2 utilizando el software WEKA	181
Figura 75. Asignaciones de cada clúster utilizando el software WEKA	183
Figura 76. Asignaciones del clúster 0 utilizando el software WEKA	183
Figura 77. Asignaciones del clúster 1 utilizando el software WEKA	184
Figura 78. Asignaciones del clúster 2 utilizando el software WEKA	184
Figura 79. Asignaciones de cada clúster utilizando el software WEKA	186
Figura 80. Asignaciones del clúster 0 utilizando el software WEKA	187
Figura 81. Asignaciones del clúster 1 utilizando el software WEKA	187
Figura 82. Asignaciones del clúster 2 utilizando el software WEKA	188
Figura 83. Regla de Asociación	189

Lista de Tablas

<i>Tabla 1. Estadísticas de Turismo a Mayo de 2014</i>	16
<i>Tabla 2. Comportamiento de vehiculos y pasajeros en terminales de transporte terrestre</i>	18
<i>Tabla 3. Fases, etapas o momentos del proyecto</i>	78
<i>Tabla 4. Técnicas e instrumentos de Recolección en Investigación Cuantitativa</i>	79
<i>Tabla 5. Campo de datos personales de los encuestados</i>	83
<i>Tabla 6. Campo de motivos de viaje y clasificación</i>	84
<i>Tabla 7. Ficha Técnica sobre la encuesta</i>	85
<i>Tabla 8. Campos discretizados de la Edad</i>	87
<i>Tabla 9. Campos discretizados de Género</i>	87
<i>Tabla 10. Campos discretizados de EstadoCívil</i>	87
<i>Tabla 11. Campos discretizados de NivelEscolaridad</i>	88
<i>Tabla 12. Campos discretizados de Ocupación</i>	88
<i>Tabla 13. Campos discretizados de Fuente de Ingreso</i>	89
<i>Tabla 14. Campos discretizados de N°Veces Viajado</i>	89
<i>Tabla 15. Campos discretizados de Motivo Viaje</i>	89
<i>Tabla 16. Campos discretizados de N°Días Permanencia</i>	90
<i>Tabla 17. Campos discretizados de N°Compañero Viaje</i>	90
<i>Tabla 18. Campos discretizados de Lugar Alojamiento</i>	91
<i>Tabla 19. Campos discretizados de Actividad Planeada</i>	91
<i>Tabla 20. Campos discretizados de Medio Transporte de Llegada</i>	91
<i>Tabla 21. Campos discretizados de Medio Transporte en la Ciudad</i>	92
<i>Tabla 22. Campos discretizados de Lugar Alimentación</i>	92
<i>Tabla 23. Campos discretizados de Gasto Promedio</i>	92
<i>Tabla 24. Campos discretizados de Calificación Gasto Promedio</i>	93
<i>Tabla 25. Campos discretizados de Aspecto que Impacta de la Ciudad</i>	93
<i>Tabla 26. Campos discretizados de Aspecto Motivado a Regresar</i>	94
<i>Tabla 27. Campos discretizados de Grado Satisfacción Entretenimiento</i>	94
<i>Tabla 28. Campos discretizados de Grado Satisfacción Comercial</i>	95
<i>Tabla 29. Campos discretizados de Grado Satisfacción Señalización Turística</i>	95
<i>Tabla 30. Campos discretizados de Grado Satisfacción Cultura Ciudadana</i>	95
<i>Tabla 31. Campos discretizados de Grado Satisfacción Alojamiento</i>	96
<i>Tabla 32. Campos discretizados de Grado Satisfacción Calidad de Atención Recibida</i>	96
<i>Tabla 33. Campos discretizados de Grado Satisfacción Calidad Alojamiento</i>	97
<i>Tabla 34. Campos discretizados de Grado Satisfacción Disponibilidad Alojamiento</i>	97
<i>Tabla 35. Campos discretizados de Grado Satisfacción Precio Alojamiento</i>	97
<i>Tabla 36. Campos discretizados de Grado Satisfacción Oficina Turismo o Transportadoras</i>	98
<i>Tabla 37. Campos discretizados de Grado Satisfacción Hotel u Hospedaje</i>	98
<i>Tabla 38. Campos discretizados de Grado Satisfacción Personas en la Calle</i>	99
<i>Tabla 39. Campos discretizados de Grado Satisfacción Policía</i>	99
<i>Tabla 40. Campos discretizados de Grado Satisfacción Policía de Tránsito</i>	99
<i>Tabla 41. Campos discretizados de Grado Satisfacción Policía Carreteras</i>	100
<i>Tabla 42. Campos discretizados de Grado Satisfacción Información Turística</i>	100
<i>Tabla 43. Edad viajeros que visitan la ciudad de Bucaramanga</i>	150
<i>Tabla 44. Género viajeros que visitan la ciudad de Bucaramanga</i>	151
<i>Tabla 45. Ciudad Origen viajeros que visitan la ciudad de Bucaramanga</i>	152

<i>Tabla 46. Ocupación viajeros que visitan la ciudad de Bucaramanga</i>	155
<i>Tabla 47. Gasto Promedio viajeros que visitan la ciudad de Bucaramanga</i>	156
<i>Tabla 48. Aspecto Motivaa Regresar viajeros que visitan la ciudad de Bucaramanga</i>	157
<i>Tabla 49. Aspecto que impacta de la Ciudad viajeros que visitan la ciudad de Bucaramanga</i>	159
<i>Tabla 50. Clustering Validación Proceso inicial</i>	174
<i>Tabla 51. Clustering Validación Segundo Proceso</i>	178
<i>Tabla 52. Clustering Validación Tercer Proceso</i>	182
<i>Tabla 53. Clustering Validación Cuarto Proceso</i>	185
<i>Tabla 54. Proceso para Análisis de la percepción.</i>	193

Resumen

El sector del turismo es uno de los más grandes e importantes a nivel mundial, generando aproximadamente 212 millones de empleos, contribuyendo así al PIB de muchos países.

La ciudad de Bucaramanga no cuenta actualmente con un observatorio de turismo ni estudios sobre los viajeros, en comparación a otras ciudades importantes donde en estos observatorios se tienen en cuenta actores de turismo como los prestadores de servicios hoteleros, restaurantes, comercio, cultura, etc. Las entidades encargadas del Turismo de Bucaramanga no cuentan con dicho análisis de las variables pues no existe información recopilada sobre esto.

Por esto se realizó un análisis de la ciudad que permita determinar datos estadísticos como: Procedencia de los visitantes, motivo del viaje, servicios utilizados, lugar de alimentación, tipo de transporte utilizado, nivel de satisfacción en la ciudad, costo de la ciudad, entre otras variables.

Con los datos recopilados se realizó una segmentación y descripción de hábitos de los viajeros realizando un análisis de clúster, posteriormente se clasificaron de acuerdo a las tendencias y características aplicando reglas de asociación, y finalmente se agruparon por el nivel de gastos, percepción y satisfacción luego de visitar la ciudad por medio de un análisis de clúster.

Este proceso de extracción del conocimiento se desarrolló utilizando la metodología CRISP-DM y fue implementado sobre el uso libre de la aplicación SPSS y WEKA, y comprende la recolección de la información y extracción de esta con ayuda de la base de datos diseñada para tal fin, su transformación, validación, cálculo de clúster y reglas de asociación.

Abstract

The tourism sector is one of the largest and most important in the world, generating approximately 212 million jobs, thus contributing to the GDP of many countries.

The city of Bucaramanga does not currently have an observatory of tourism or studies on travelers, compared to other important cities where these observatories take into account tourism actors such as providers of hotel services, restaurants, commerce, culture, etc.

The entities in charge of the Tourism of Bucaramanga do not have this analysis of the variables since there is no information collected about this.

For this reason, a city analysis was carried out to determine statistical data such as: Origin of visitors, reason for travel, services used, place of feeding, type of transportation used, level of satisfaction in the city, cost of the city, between other variables.

With the data collected, a segmentation and description of habits of the travelers was carried out by means of a cluster analysis, later they were classified according to the tendencies and characteristics applying rules of association, and finally they were grouped by the level of expenses, perception and satisfaction later to visit the city by means of a cluster analysis.

This knowledge extraction process was developed using the CRISP-DM methodology and was implemented on the free use of the SPSS and WEKA application, and includes the collection of information and extraction of this information with the help of the database designed for this purpose , Its transformation, validation, cluster calculation and association rules.

Introducción.

Actualmente las organizaciones manejan gran cantidad de datos, alguno de estos organizados de manera adecuada, y otros que utilizan diversos sistemas para el almacenamiento de los datos repetidos, desorganizados y erróneos. Las nuevas organizaciones almacenan poca cantidad de datos en comparación con aquellas que cuentan con una trayectoria más amplia, que por lo general cuentan con gran cantidad de datos históricos. Es por esto que podemos decir que toda empresa, institución o negocio requiere del almacenamiento de datos adecuado para poder realizar informes y tener la información clara y precisa que permita tener acceso y conocimiento a la información que se requiera para realizar cualquier tipo de análisis.

Cuando los reportes que se llegasen a realizar se quedan solo en la parte experimental, es necesario implementar un método diferente, una metodología que convierta dichos datos en una información más útil.

Estas metodologías presentan el conocimiento implícito, del cual se puede derivar un análisis para realizar proyecciones, encontrar patrones en los datos, realizar asociaciones con estos datos, agrupaciones de clúster, encontrar tendencias, entre otras operaciones. Dichas metodologías se encuentran bajo el nombre de minería de datos, el cual tiene bases muy fuertes de la inteligencia artificial y el análisis estadístico. Con estas bases se crea una ciencia encargada de recoger, modificar y limpiar los datos que permiten extraer conocimiento significativo para cualquier investigación.

En la ciudad de Bucaramanga, luego de realizar un estudio detallado al sector del turismo, se pudo observar que las entidades públicas y privadas de la región dichos datos, algunos completos, otros incompletos, algunos por el sector hotelero, otros por medios de transporte de llegada a la ciudad,

entre otros. En este sector se puede realizar un estudio de minería de datos, donde se puedan encontrar patrones que permitan plantear predicciones muy útiles para la ciudad y la región, como una manera importante para fomentar y mejorar el turismo.

El estudio por parte de la información recopilada mediante encuestas realizadas a personas que visitan la ciudad, contó con un registro recopilado de 400 encuestados, ya que en la ciudad de Bucaramanga no existe información estadística sobre los aspectos requeridos para el análisis.

El desarrollo del trabajo siguió la metodología CRISP-DM, una metodología estándar internacional en proyectos de minería de datos, presentando para cada fase las parametrizaciones y resultados obtenidos.

Se utilizó software de uso libre, entre ellos SPSS como herramienta base para el almacenamiento de la información obtenida de los turistas y WEKA para implementar un análisis a la información almacenada e implementar el proceso de minería de datos.

Finalmente se presenta un análisis de resultados y sus debidas conclusiones.

1. Definición del Problema.

1.1.Descripción del Problema.

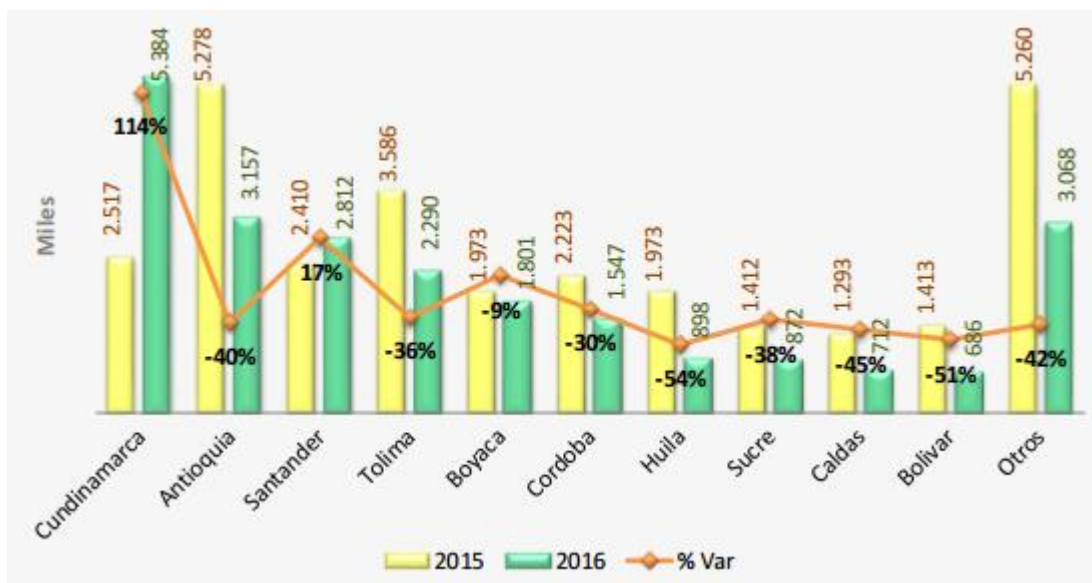
La industria turística es una de las más importantes a nivel mundial, ya que generan un aproximado de 212 millones de empleos y representan una importante contribución al PIB de cada país en los últimos años. En una industria tan grande como esta, ha jugado un papel fundamental las Tecnologías de Información, más conocidas como las TIC, que mejoran la eficiencia de las organizaciones que hacen uso de esta, y brindan una mejor entrega del servicio.

Según las estadísticas realizadas por el Ministerio de Comercio en el mes de Julio de 2016, se puede observar, teniendo en cuenta indicadores del año 2015, que



Figura 1. Indicadores año 2015 Turismo en Colombia. .. Adaptado de: [19].

En el mes de Julio transitaron en total 2.987.254 vehiculos por los peajes del país. Para el periodo de Enero-Julio transitaron 23.228.947, siendo Cundinamarca (23,2%), Antioquia (13,6%) y Santander (12,1%) los departamentos más representativos.



Φιγυρα 2. Φλυφο δε ζεηή'χυλοσ πορ δεπαρταμεντο. Ενερο-θυλιο 2015-2016 . Αδαπαδο δε: INV IAS, χ(λχυλοσ OEE. Χιφρασ οργαניζαδασ πορ ήλιμο μεσ ρεγιστραδο δε 2015

Según estadísticas del Ministerio de comercio del año 2011, se observa que el total de llegadas de turistas internacionales a la ciudad de Bucaramanga es mayor que el de Pereira, Armenia y Manizales, conocidos como el Triángulo del Café, sin tener en cuenta los turistas que ingresaron al país por una frontera.[24]

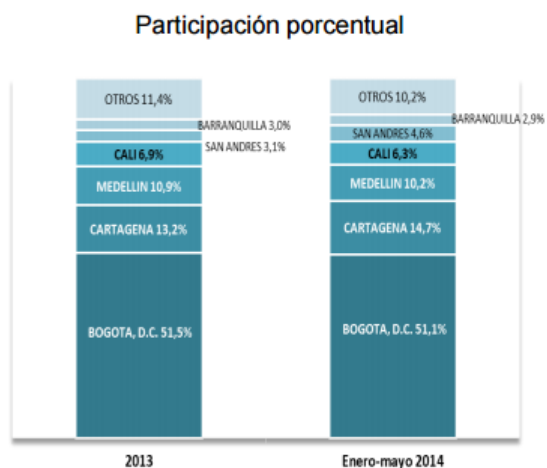
En el año 2014 se mantiene esta tendencia, por lo cual es clara la necesidad de contar con los datos estadísticos de Bucaramanga y sus visitantes. [20]

Tabla 1.

Estadísticas de Turismo a Mayo de 2014

Numero de llegadas y variación anual

Destino reportado	Total llegadas			Variación anual	
	2013	Enero-mayo 2014	Mayo 2014	2013	Enero-mayo 2014
	BOGOTA, D.C.	888.852	380.587	75.639	7,5%
CARTAGENA	227.034	109.677	19.454	11,8%	14,0%
MEDELLIN	187.552	75.757	15.096	17,7%	16,0%
CALI	119.327	47.186	9.412	11,0%	14,2%
SAN ANDRES	53.835	34.404	6.438	5,3%	32,1%
BARRANQUILLA	52.387	21.234	3.676	4,9%	8,4%
CUCUTA	30.592	13.107	2.252	15,9%	23,3%
SANTA MARTA	21.688	8.773	1.594	-1,2%	-4,3%
BUCARAMANGA	19.183	6.871	1.103	-7,7%	-0,5%
PEREIRA	17.974	5.226	1.092	2,9%	-21,0%
OTROS	107.876	42.040	7.798	0,9%	8,3%
TOTAL	1.726.300	744.862	143.554	8,5%	12,0%



Nota: Principal ciudad de destino. Extranjeros no residentes, Mayo 2014. Adaptado de: [20]

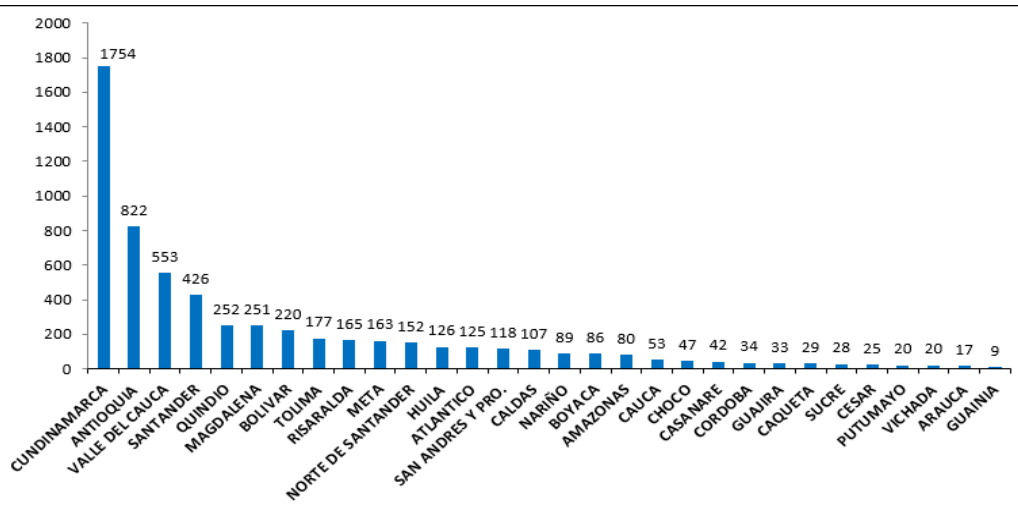
Bucaramanga cuenta con un observatorio de turismo, donde es analizado el flujo de turistas, basados en cuatro subsectores: Servicios de alojamiento, restaurantes y bares, agencias de viajes y turismo, y transporte turístico especial. Esta información es recopilada desde 2011 por medio de PIT (Puestos de Información Turística), que se encuentran ubicados en 5 puntos estratégicos de la ciudad: en el Aeropuerto Internacional Palonegro, en la Terminal de Transporte, en Hoteles de 3 y 4 estrellas, en el Ecoparque Cerro del Santísimo y centros comerciales como El Cacique, Parque Caracolí y Cañaveral. Esta información no se publica ni es conocida, ya que el propósito de los PIT es brindar información más no recopilarla. El instituto municipal de Cultura y Turismo es el encargado de la recopilación de la información de los PIT. [16]

Para estudios estadísticos debemos tener presente las cuentas satélites de turismo (CST), que son una extensión de cuentas naciones que busca ampliar el conocimiento sobre actividades turísticas según recomendaciones de la Organización Mundial del Turismo (OMT), para construir una

herramienta básica de análisis que permita la formulación de políticas de promoción y comercialización del sector. [7]

Basado en estadísticas realizadas por ANATO, se considera a la ciudad de Bucaramanga entre las 10 ciudades más turistas de Colombia, junto a Cartagena, Bogotá y Medellín. Durante el mes de Julio del presente año han visitado a esta ciudad cerca de 2.399 extranjeros. Además el comportamiento mensual de movimiento de pasajeros en el aeropuerto (salidas y llegadas) ha sido 140.784; y comportamiento mensual de movimiento de pasajeros internacionales (salidas y llegadas) de 9.005. [4]

Santander cuenta con un total de 426 agencias de turismo, de las cuales 16 son agencias mayoristas, 131 agencias operadoras y 279 agencias de viajes y turismo, con una participación porcentual del 7,06. De esta totalidad el 40% corresponde a la ciudad de Bucaramanga, representado por 170 oficinas.



Φιγυρα 3. Total δε Αγενχιασ πορ Δεπαρταμεντο □ θυλιο 2016. . Αδαπαδο δε: ANATO θυλιο 2016

Además presenta una ocupación hotelera en lo corrido de este año de 46,08%, aumentando un 2,50% en relación con años anteriores. Adicional a esto se conoce que la tarifa promedio por Región para realizar visitas y conocer lugares turísticos en Bucaramanga, es de \$159.207, con variación del 13,09% con indicadores anteriores.

El comportamiento de vehiculos y pasajeros en terminales de transporte terrestre muestra el flujo constante de turistas y visitantes que mensualmente arriban a la ciudad de Bucaramanga.

Tabla 2.

Comportamiento de vehiculos y pasajeros en terminales de transporte terrestre

		BUCARAMANGA	
Mes	Año	Vehículos	Pasajeros
2015	Enero	26.953	257.902
	Febrero	22.487	179.285
	Marzo	24.472	199.703
	Abril	24.458	213.864
	Mayo	22.728	196.243
	Junio	22.858	234.604
	Julio	23.356	230.200
	Agosto	22.322	208.584
	Septiembre	21.386	190.869
	Octubre	22.678	221.990
	Noviembre	21.622	209.125
	Diciembre	24.857	282.149
TOTAL		280.177	2.624.518
2016	Enero	19.142	224.391
	Febrero	18.443	183.149
	Marzo	21.429	215.006
	Abril	20.539	230.982
	Mayo		
	Junio		
	Julio		
	Agosto		
	Septiembre		
	Octubre		
	Noviembre		
	Diciembre		
TOTAL		79.553	853.528

Nota: Datos referentes a la afluencia de vehículos y pasajeros en terminales de transporte. Adaptado de: ANATO

Implementando la minería de datos, será necesario que todas las empresas de la ciudad de Bucaramanga, cámara de comercio, alcaldía y demás entidades caractericen los visitantes y así mejorar aspectos para el turismo de la ciudad.

Además existen estudios teóricos como el realizado por Tzu-Ching Lin denominado **Enhancing Tourism Intermediaries with the Data Mining Process**, llevado a cabo en la isla de Taiwán, donde se desarrolló un estudio para la industria del turismo, donde dicho proyecto permitía la administración y mejora de las relaciones con los clientes del turismo a largo plazo, sabiendo que los beneficiarios directos relacionados con el turismo serían restaurantes, hoteles, vendedores, comercio, artesanos, sitios de recreación, museos, discotecas, bares, guías turísticos, etc.

En Taiwán se enfocaron en promocionar la isla como un destino turístico, como estrategia para que las agencias de viajes tuvieran un rol importante en la operación del turismo. Ellos mismos fueron quienes actuaron como intermediarios entre proveedores de hoteles, transporte, viajes y servicios que se les prestarían a los consumidores. Por esta razón las agencias de viaje fueron y son actualmente reconocidas por el gobierno como parte fundamental en la cadena de distribución en el incremento de ingreso y salida de turistas. [10]

Adicionalmente estas agencias fijaron sus ojos en los viajeros y en mejorar las plataformas de información y comunicación, implementar un CRM. Gracias a este estudio, se logró un común acuerdo entre administradores, donde se coincidía que el mayor beneficio de la minería de datos (como herramienta de análisis) era la retención de cada uno de los clientes de las diferentes agencias. Finalmente, quien efectuó este análisis notó que las agencias de viajes no hacían un uso

adecuado a la minería de datos, y se debió agregar un cuarto proceso, igual de valioso, para estas agencias de turismo intermediaras.

De igual manera existe otro estudio teórico denominado **A knowledge destination framework for tourism sustainability: A business intelligence application from Sweden**, llevado a cabo por Mid Sweden University. Este proyecto se realizó mediante diversas fuentes de información de datos de clientes viajeros dispersos en los distintos destinos turísticos como los servidores web de turistas, comportamiento de navegación web, base de datos con transacciones, etc... . Estos datos que fueron recolectados permanecieron por cierto tiempo sin usar. Por este motivo, se fomenta la creación de conocimiento y adquisición de procesos en los diferentes destinos turísticos aplicando minería de datos, considerándola como una nueva alternativa para descubrir las preferencias de los consumidores del sector turístico, sobre todo en temas como hotelería y restaurantes, teniendo en cuenta la opinión de los sitios web a analizar como TripAdvisor. [15]

Por otro parte encontramos la teoría propuesta por Robin J. B. Ritchie titulada **A supplier-oriented knowledge application – The destination management information system (DMIS)**, donde enfatiza explícitamente el uso del diseño DMIS. Este diseño, se basa en reconocer que el cliente se crea mediante técnicas de segmentación de clientes y servicios de evaluación de desempeño. [33]

Los datos que son recolectados, se almacenan, analizan y evalúan en DMIS, incluyendo datos demográficos de los turistas, características de compra, motivos y percepción de marcas, como también la información de clientes y patrones de consumo de productos. [12]

Y finalizando con estos estudios que aportan a la investigación que se realizará, encontramos un artículo llevado a cabo por la Red de Revistas Científicas de América Latina y el Caribe, España

y Portugal, sobre los **Atributos Relevantes en el Perfil del Turista del Interior en la Ciudad de Mar de Plata.**

Este artículo habla sobre el estudio que se realizó sobre los “Atributos Relevantes en el Perfil del Turista del Interior en la Ciudad de Mar de Plata”, con el propósito de presentar los grupos de turistas del interior con características diferentes que se identificaron en el verano de 2004-2005, utilizando técnicas de aprendizaje no supervisadas. [31]

Debido a la poca información y estudios realizados en Colombia sobre la afluencia de viajeros y las tendencias se ha convertido en un problema de desinformación para la población. Al tema específico del turismo en la ciudad de Bucaramanga se presenta pocos estudios sobre este, del mismo modo que en la ciudad de Pereira, donde es evidente percibir el poco y hasta vacío conocimiento de esta información, que es necesaria para dar a conocer a quienes les interesa. Entre los estudios que han llevado a cabo, se encuentran los siguientes:

ESTUDIO DE COMPETITIVIDAD EN EL SECTOR TURISMO EN COLOMBIA, elaborado pro Angela María Bravo y David Rincón, estudiantes de la Universidad del Rosario, cuya investigación fue realizada con el objetivo de conocer la evolución que ha tenido el sector turismo en el mundo y en Colombia para analizar el impacto que este ha generado en la economía nacional. La información obtenida de fuentes secundarias permitió conocer el desarrollo y el avance del sector a través de cifras y estadísticas con las cuales se identificaron las oportunidades y ventajas que tiene Colombia para lograr ser un país turísticamente competitivo y a su vez se evidenciaron los factores de mejora. Las Políticas Sectoriales ayudaron a identificar la normativa del sector y la manera en que cada uno de los procedimientos se debe llevar. Se realizó un análisis de competitividad en el país en el cual se resalta que a pesar de las falencias que aún se presentan en

las actividades y servicios que se prestan para este sector, Colombia ha logrado un avance significativo con la correcta aplicación de la Políticas que el Gobierno Nacional creó y las buenas prácticas que día a día se convierten en un estilo de vida para los empresarios y actores participantes del sector. [32]

De igual manera existe otro estudio denominado MEDICIÓN DEL TURISMO INTERNO EN COLOMBIA, EXPERIENCIAS Y RETOS, realizado por personal calificado del DANE, en el margen de la 3° Conferencia Internacional sobre Medición y Análisis económico del Turismo regional. Este proyecto liderado por la CAN (Comunidad Andina de Naciones), la Dirección de Síntesis y Cuentas Nacionales (DSCN) del Departamento Administrativo Nacional de Estadísticas (DANE) de Colombia realizó en los años 2010-2011 un ejercicio piloto de estimación de una Cuenta Satélite de Turismo de Colombia referida al período 2000 – 2005. Este ejercicio permitió apreciar la importancia del turismo interno dentro del turismo interior, al arrojar que aproximadamente el 75% del gasto del turismo interior corresponde al turismo interno, es decir que supera con creces al turismo receptor. Además de recalcar su importancia en términos de participación dentro del total, el ejercicio evidenció la importancia de conocerlo mejor ya que los datos disponibles eran insuficientes para caracterizarlo con la profundidad requerida. Con este fin, el Vice Ministerio de Turismo acordó mediante convenio 146 de 2011 con la Dirección de Metodología y Producción Estadística (DIMPE) del DANE la realización de la Encuesta de Gasto en Turismo Interno (EGIT) con el objetivo de “consolidar el sistema de estadísticas de turismo con el fin de disponer de información que permita el diseño eficiente de políticas, estrategias, planes y programas turísticos y fortalecer el proyecto de cuenta satélite de turismo de Colombia”. [27]

También existe un estudio realizado por la Superintendencia de Industria y Comercio en el año 2011, en la ciudad de Bogotá denominado ESTUDIO MERCADO, ESTUDIO DE AGENCIAS DE VIAJES EN COLOMBIA: El presente estudio pretende evaluarlos diferentes aspectos de la dinámica del sector de las agencias de viajes mediante una aproximación al mismo, a partir del establecimiento de la estructura de mercado, y la realización del correspondiente análisis de los principales participantes en este y las medidas regulatorias implementadas en dicho sector. Lo anterior, con el propósito de lograr un enriquecimiento académico respecto de este sector, de manera que sea una herramienta útil para posteriores estudios económicos. [39]

Sin embargo se presentan ideas y postulaciones que no están muy de acuerdo con el crecimiento y expansión del turismo a nivel nacional e internacional, pues se piensa que aún no somos competentes como país para lograr que la industria turística alcance logros mayores comparados con países donde este tema es una meta alcanzada. Por esto surgen estudios que intentan dar a conocer sus ideas de cómo se debería lograr el auge turístico basado en ciertos parámetros como en el proyecto COMPETITIVIDAD: EL DESAFÍO PARA ALCANZAR UN TURISMO DE CLASE MUNDIAL, documento político elaborado por el Ministerio de Comercio, Industria y Turismo, La necesidad de fortalecer las ventajas competitivas, a partir de la dotación de factores comparativos que Colombia tiene para el turismo, surge como estrategia y técnica especializada de la gestión pública y privada del sector desde mediados de los años 1990. Dentro de estos esfuerzos de fortalecer la competitividad turística de Colombia como destino turístico, se debe resaltar referentes que permiten dar continuidad a la gestión adelantada en el tema como el estudio de Michael Portes, Estudio de competitividad del sector turismo y la articulación de la gestión de planificación turística con la suscripción de convenios de competitividad turística departamentales.

[23]

Aunque otro problema o idea que surge es la relacionada con la violencia como causante del poco turismo en Colombia, considerando la causa principal para que muchas personas decidan viajar. Esto es un problema que afecta directamente la economía de nuestro país debido a que disminuyen notablemente ingresos que solventarían ciertas falencias económicas actuales que presenta Colombia. Basado en este ideal el Ministerio de Comercio, Industria y Turismo, en compañía de la Policía de Turismo realizan un estudio sobre **SEGURIDAD TURÍSTICA: RETO COMPETITIVO DE COLOMBIA (PLAN ESTRÁTEGICO)**, Considerando que la seguridad es una de las condiciones fundamentales para el fortalecimiento de la competitividad del turismo en Colombia, es necesario que la prioridad de la gestión del Estado sea la de buscar los mecanismos de coordinación y de gestión que permitan articular medidas prácticas en el tema, para que la actividad turística consolide su reactivación, se valore su importancia socioeconómica, se garantice su calidad y se le reconozca como factor de paz. El Gobierno Nacional, desde el Ministerio de Comercio Industria y Turismo, pretende ser la fuerza motora para continuar con la creación de mejores condiciones para los destinos y los turistas, promoviendo la organización y coordinación de acciones de cooperación para la seguridad que requiere el turismo. [\[22\]](#)

Por lo anterior el gobierno nacional colombiano ha reglamentado normas y leyes que regulan el sector turístico, con el fin de prestar un buen servicio a los turistas y visitantes, de igual manera proteger y evitar riesgos que impidan en buen desarrollo de la actividad y generen a su vez inconformidad en el sector. Basado en esto se establecen leyes como:

DECRETO 0348 de 2015, donde se reglamenta el Servicio Público de Transporte Terrestre Automotor Especial.

DECRETO 1075 DE 1997, sanciones a los prestadores turísticos.

DECRETO 1095 DE MAYO 30 DE 1994, por el cual se adopta el Código de Ética Profesional del Agente de Viajes y Turismo.

LEY 1558 10 DE JULIO DE 2012, cuya ley tiene por objeto el fomento, el desarrollo, la promoción, la competitividad del sector y la regulación de la actividad turística, a través de los mecanismos necesarios para la creación, conservación, protección y aprovechamiento de los recursos y atractivos turísticos nacionales, resguardando el desarrollo sostenible y sustentable y la optimización de la calidad, estableciendo los mecanismos de participación y concertación de los sectores público y privado en la actividad.

LEY 300 DE 1996, protección al consumidor (Viajero y/o Turista)

DECRETO 972 DE 1997, por el cual se reglamentan los planes de servicios y descuentos especiales en materia de turismo para la tercera edad.

DECRETO 2438 DE 2010, por el cual se dictan unas disposiciones relacionadas con la responsabilidad de las Agencias de Viajes en la prestación de servicios turísticos.

LEY 336 DE 1996, cuya ley tiene por objeto unificar los principios y los criterios que servirán de fundamento para la regulación y reglamentación del Transporte Público Aéreo, Marítimo, Fluvial, Férreo, Masivo, Terrestre y su operación en el Territorio Nacional.

LEY 105 DE 1993, por la cual se dictan disposiciones básicas sobre el transporte, se redistribuyen competencias y recursos entre la Nación y las Entidades Territoriales, se reglamenta la planeación en el sector transporte y se dictan otras disposiciones.

1.2. Formulación del Problema.

¿Cómo podemos agrupar y definir reglas de asociación de los viajeros y turistas de la ciudad de Bucaramanga, para así poderlos caracterizar?

1.3. Justificación del Problema.

Actualmente es normal que las empresas del sector de Turismo cuenten con diversas herramientas de registro de turistas y visitantes, donde se puede contar con información como ingreso y salida de dichos usuarios. De igual manera se puede realizar estadísticas de consumo, conociendo que actividades desean realizar, que atractivos desearían conocer, que sitios turísticos visitan con más frecuencia, entre otras.

El problema radica en que en ocasiones esta información es inservible ya que los datos que son recolectados y a su vez son importantes para las empresas, no es muy claro que hacer con ella. Es por este motivo que mediante este proyecto, utilizando minería de datos a la información recolectada, se llevará a cabo un estudio donde se analizaran dichos datos, que a su vez se obtendrá por medio de encuestas efectuadas a un grupo determinado de personas, teniendo en cuentas ciertas características como hotelería, agencias de viaje, preferencia al viajar, sitios turísticos, etc., y a partir de esto establecer patrones de consumo y comportamiento, caracterizando a cada uno de los visitantes, para generar beneficios al sector turístico de la ciudad de Bucaramanga.

El desarrollo de este proyecto tiene como finalidad establecer las causas, motivos y características, además de conocer con qué frecuencia realiza estos viajes, que influyen para que las personas viajen a la ciudad de Bucaramanga, aplicando un instrumento estadístico, en este caso la minería de datos, para evaluar el porcentaje de la población que viaja, valga la redundancia.

Los resultados que se generen mediante esta investigación serán la base para tener el conocimiento de la tendencia y característica de la población que viaja y los factores determinantes de esta práctica. Dichos resultados de esta investigación podrán ser de suma importancia para medianas y grandes empresas del sector turístico en la ciudad. Si estos resultados son exitosos, será de gran

apoyo para el sector a la hora de tomar decisiones. Es importante resaltar los beneficios finales que tendrán las empresas turísticas de Bucaramanga, como el Departamento, pues permitirá que se logre potenciar el turismo, a su vez generar empleo para la mayoría de los sectores, sino que aportará información concreta y exacta para estudios que a futuro desee realizar el Ministerio de Comercio, Industria y Turismo en la ciudad de Bucaramanga.

Además toda la información recolectada y los conocimientos que se adquirieran, permitirán identificar la necesidad de seguir aplicando investigaciones en los estratos socioeconómicos de la ciudad, con el fin de crear estrategias que promuevan e incentiven el deseo de viajar y conocer más de nuestro país, y otras culturas, quitando de su mentalidad cualquier tipo de indisposición y poca fuerza de voluntad de atreverse a conocer y experimentar nuevas y únicas experiencias que solo viajando y disfrutando, se puede lograr.

1.4.Objetivos.**1.4.1. Objetivo General.**

Establecer las principales tendencias y características de los viajeros que visitan la ciudad de Bucaramanga por medio de técnicas de minería de datos.

1.4.2. Objetivos Específicos.

- Segmentar y describir hábitos de los viajeros que visitan la ciudad de Bucaramanga por medio de un Análisis de Clúster para la industria turística de Bucaramanga.
- Establecer relaciones y tendencias existentes entre las características de los viajeros y los aspectos que rodean su visita a la ciudad por medio de reglas de asociación.
- Determinar en los viajeros el nivel de gastos, percepción y satisfacción luego de visitar la ciudad de Bucaramanga por medio de un Análisis de Clúster.

1.5. Alcance.

La ciudad de Bucaramanga es un municipio colombiano, capital del departamento de Santander. Declarada según el Banco Mundial en su estudio de "ciudades competitivas para estudios y crecimientos" como la ciudad más próspera de América Latina y ejemplo para el mundo.¹² Está ubicada al nororiente del país sobre la Cordillera Oriental, rama de la Cordillera de los Andes, a orillas del Río de Oro. Bucaramanga junto con Floridablanca, Girón y Piedecuesta conforman el área metropolitana con un total de 1 132 339 habitantes, siendo la quinta aglomeración urbana más poblada del país. Se encuentra a 384 km de Bogotá, capital del país, además está comunicada con las demás ciudades del país por carretera. Para el transporte aéreo, cuenta con el Aeropuerto Internacional Palonegro.

Basados en el último censo realizado en la ciudad, existe una población de 528.269 habitantes, distribuidos en estratos socioeconómicos tres (25,8%) y cuatro (32,1%), evidenciando un poder de adquisición de mediano poder. Las principales actividades económicas que desarrolla el municipio, están relacionadas con el comercio y la prestación de servicios, donde se comercializan productos agrícolas, de ganadería y avicultura. [28]

El turismo se ha convertido en uno de los sectores más importantes para el desarrollo del país, principalmente en el aspecto social, puesto a que esta industria permite la recreación y el aprovechamiento del tiempo libre, generando ingresos que permitan mejorar la calidad de los servicios que esta puede ofrecer.

Es por esto que en Santander, especialmente en la ciudad de Bucaramanga, exista una diferencia competitiva frente a las demás ciudades y regiones del país, aprovechando con los recursos que se cuenta, y de este modo utilizar tecnologías que ofrezcan al mercado, por medio de ideas

innovadoras, proyectos que marquen la diferencia y eviten el estancamiento de una actividad, que tiende a generar, haciendo uso adecuado, de grandes beneficios para la ciudad y el departamento, logrando que el sector del turismo se posicione como una industria que favorezca al desarrollo social, económico, cultural y ambiental.

Teniendo en cuenta que la información existente sobre los turistas de la ciudad de Bucaramanga, se encuentra recopilada en los PIT, ubicados en 5 puntos estratégicos de la ciudad, se debe establecer los indicadores de turismo que permitan determinar el tipo de satisfacción.

Debido a que por medio de la estadística estamos limitados a muestreo e interpretación a pequeña escala, se hace necesario un estudio de minería de datos, utilizando información, que permita agrupar y definir nuevas relaciones con los datos, determinando reglas y asociando datos de acuerdo a determinadas características.

Para llevar a cabo este proyecto de investigación, se realizará una prueba piloto con tamaño de la muestra de 400 personas encuestadas entre hombres y mujeres mayores de edad que arriben a la ciudad de Bucaramanga o en su defecto, se dirijan a otros lugares del país. Mediante este proyecto podremos consultar información determinante e importante en hoteles, restaurantes, parques, sitios turísticos, terminales de transporte y aeropuerto en la ciudad de Bucaramanga, que permitirán el almacenamiento y recopilación de datos, que serán de suma importancia para tomar decisiones para el mejoramiento del sector turístico de la ciudad. De igual forma se enfocará en la caracterización de los visitantes y cuáles son sus preferencias al momento de elegir esta ciudad como centro turístico. Todo esto se llevara a cabo por medio de un análisis de datos a una muestra de la población de estrato 3 y 4 que viajan a la ciudad de Bucaramanga, y bajo qué condiciones y características lo prefieren.

El resultado permitirá una adecuada toma de decisiones para el mejoramiento continuo y optimo del turismo, por medio de la identificación de variables, que permitirán el diseño y aplicación de la encuesta, que nos permitirá obtener la recopilación de información y datos importantes, que nos ayuden a analizar y evaluar las tendencias y características de los viajeros que visitan la ciudad de Bucaramanga.

2. Marco Referencial.

2.1. Marco Teórico.

Dentro del estudio de Tendencias y características de los viajeros que visitan la ciudad de Bucaramanga podemos visualizar técnicas que se han aplicado para la investigación, desarrollo y logro de los resultados deseados en el análisis. Por lo cual se cuenta con los siguientes estudios teóricos.

2.1.1. Minería de Datos y Extracción de conocimiento a partir de los datos.

Este estudio fue realizado por Arun K. Pujari en el año 2011, el cual expone que la minería de datos (Data Mining) es el proceso de detectar la información procesable de los conjuntos grandes de datos. Utiliza el análisis matemático para deducir los patrones y tendencias que existen en los datos. Normalmente, estos patrones no se pueden detectar mediante la exploración tradicional de los datos porque las relaciones son demasiado complejas o porque hay demasiado datos. Este proceso utiliza nuevas técnicas para buscar la información, muchas basadas en procesos estadísticos o de inteligencia artificial, que permite identificar patrones y relaciones en la información para crear modelos que permitan caracterizar y hasta predecir el comportamiento de esta; la extracción de conocimiento (KDD, por sus siglas en inglés) abarca la minería de datos, pero además incluye la preparación de la información y la interpretación de los resultados con el fin de dar significado al modelo planteado.[30].

Para llevar a cabo este proceso, se deben tener en cuenta las etapas de KDD (Proceso de Extracción del Conocimiento). Dichas etapas inician con un flujo de datos y finalizan con un conocimiento extraído a partir de esos datos.

- Selección: seleccionar o segmentar los datos de acuerdo a su relevancia con base en los objetivos planteados para la extracción de conocimiento.
- Preprocesado: se refiere a la limpieza de la información, eliminando información innecesaria, consolidando información y unificando formatos que provenga de múltiples fuentes.
- Transformación: consiste en transformar la información recolectada con el objeto de que pueda ser utilizable y navegable (consultar la información desde diferentes puntos de vista) preparándola para la etapa de minería de datos.
- Minería de datos: extracción de patrones en los datos.
- Interpretación y evaluación: convertir los patrones encontrados en conocimiento, que sirvan como fundamento para la toma de decisiones.
- Visualización de datos: busca ofrecer profundidad en el análisis y un entendimiento más intuitivo de los datos y sus patrones definidos en la minería de datos.

La minería de datos es la principal etapa de KDD ya que se enfoca en la búsqueda de las relaciones y patrones ocultos, por lo que se profundiza en el presente documento.

Se encuentra que el alcance que tuvo este estudio basado en la Minería de Datos deriva de las similitudes entre buscar valiosa información de negocios en grandes bases de datos. Dadas bases de datos de suficiente tamaño y calidad, la tecnología de Data Mining puede generar nuevas oportunidades de negocios al proveer estas capacidades:

- Predicción automatizada de tendencias y comportamientos. Data Mining automatiza el proceso de encontrar información predecible en grandes bases de datos. Preguntas que

tradicionalmente requerían un intenso análisis manual, ahora pueden ser contestadas directa y rápidamente desde los datos.

- Descubrimiento automatizado de modelos previamente desconocidos. Las herramientas de Data Mining barren las bases de datos e identifican modelos previamente escondidos en un sólo paso. Otros problemas de descubrimiento de modelos incluye detectar transacciones fraudulentas de tarjetas de créditos e identificar datos anormales que pueden representar errores de tipeado en la carga de datos. [41]

La minería de datos genera cinco elementos principales que serán de suma importancia al momento de tomar decisiones de acuerdo al análisis que se desee realizar, pues permite:

- Extraer, transformar y cargar datos de transacciones en el sistema de almacenamiento de datos.
- Almacenar y manejar los datos en un sistema de base de datos multidimensional.
- Proporcionar acceso a los datos para los analistas de negocios y profesionales de tecnologías de la información.
- Analizar datos por un software de aplicación.
- Presentar datos en un formato útil, como un gráfico o tabla.

Adicionalmente este estudio permite conocer las diferentes áreas de investigación que existen dentro de la minería de datos al igual que las estrategias para el desarrollo de las mismas, entre las cuales encontramos:

- Estadísticas: Uno de los fundamentos de la minería de datos es la estadística, su amplia difusión y fundamentación teórica la valida, sin embargo, sus resultados deben ser

interpretados por expertos debido a su complejidad en la interpretación. Se suelen usar modelos estadísticos como los lineales para llevar a cabo la minería de datos.

- Máquinas de aprendizaje: Es un proceso automatizado de aprendizaje, tomando el aprendizaje como el equivalente a reglas basadas en observaciones, la generalización de comportamientos basada en ejemplos.
- Aprendizaje supervisado: Es el aprendizaje basado en ejemplos, el cual produce una función que establece la correspondencia entre la entrada y la salida deseada (ejemplo).
- Aprendizaje sin supervisión: Es el aprendizaje basado en la observación y el descubrimiento, por lo que no existe un conocimiento a priori, no existe una categorización de la información.
- Programación matemática.[13]

Implementando adecuadamente las técnicas de almacenamiento de datos permitirá que la mayoría de las decisiones que se tomen en una empresa u organización, se realicen con base a la información obtenida de los datos que se tienen almacenados sobre la actividad de dicha organización; generalmente éste almacenamiento está consolidado como una base de datos en primera instancia, la cual puede ser transformada en una bodega de datos para facilitar la toma de decisiones, pero inclusive y mejor aún se puede hacer una búsqueda inteligente de patrones o tendencias con las herramientas de Minería de Datos como:

- **Bases de datos:** Una base de datos, es el almacenamiento organizado de datos que tienen una dependencia y que han sido recolectados y explotados por una organización o empresa en particular, dicha explotación se hace con programas creados para la manipulación de la misma.

- **Bodegas de datos:** También conocidas como Almacenes de Datos, es un concepto relativamente nuevo, orientado al manejo de grandes volúmenes de datos, provenientes de diversas fuentes, de muy diversos tipos. Estos datos cubren largos períodos de tiempo, lo que trae consigo que se tengan diferentes esquemas de las bases de datos fuentes. La concentración de esta información está orientada a su análisis para apoyar la toma de decisiones oportunas y fundamentadas. Una Bodega de Datos, es un conjunto de datos integrados orientados a una materia, que varían con el tiempo y que no son transitorios, los cuales sirven de soporte en el proceso de toma de decisiones de la administración.
- **Base de datos NOSQL:** Base de datos no relacional, distribuida, de alto rendimiento y altamente escalable. Estas bases de datos son una categoría de sistemas de bases de datos que no utilizan SQL como lenguaje de consulta principal. Existen cuatro grandes categorías: Almacenes clave-valor, almacenes de documentos, almacenes de grafos y familias de bases de datos columnares.
- **Bases de datos clave-valor:** El modelo clave-valor procede del modo de acceso a memoria en la programación en ensamblador. La dirección de la posición de memoria actúa como el valor que almacena en esa dirección de memoria. Otro ejemplo típico del modelo valor-clave es el concepto de tabla hash que tiene una función que transforma la clave de un índice utilizado para encontrar su valor asociado.
- **Bases de datos orientadas a grafos:** Estas bases de datos organizan la información en grafos dirigidos. Son óptimas para hacer operaciones de consulta sobre las relaciones entre miembros y son extremadamente rápidas. La información se almacena dividiéndola en trozos más básicos, nodos (chunks) y estableciendo relaciones. Este tipo de bases de datos es muy eficiente para el caso de múltiples contactos.

- **Bases de datos orientadas a documentos:** Estos abarcan una gran colección de formatos y codificación binaria. Los formatos de marcación estándar tales como XML y JSON se combinan con formatos propietarios como PDF de Adobe o Word de Microsoft. El lenguaje de consultas UnQL se desarrolló principalmente para consulta de documentos y objetos de datos marcados en JSON.

Se debe tener muy claro los pasos a seguir para la realización de un proyecto de minería de datos, pues son siempre los mismos, independientemente de la técnica específica de extracción de conocimiento usada. El proceso de minería de datos pasa por las siguientes fases:

- **Filtrado de datos:** El formato de los datos contenidos en la fuente de datos (base de datos, Data Warehouse...) nunca es el idóneo, y la mayoría de las veces no es posible ni siquiera utilizar ningún algoritmo de minería sobre los datos en bruto. Mediante el preprocesado, se filtran los datos (de forma que se eliminan valores incorrectos, no válidos, desconocidos... según las necesidades y el algoritmo a usar), se obtienen muestras de los mismos (en busca de una mayor velocidad de respuesta del proceso), o se reducen el número de valores posibles (mediante redondeo, clustering,...).
- **Selección de Variables:** Aún después de haber sido preprocesados, en la mayoría de los casos se tiene una cantidad ingente de datos. La selección de características reduce el tamaño de los datos eligiendo las variables más influyentes en el problema, sin apenas sacrificar la calidad del modelo de conocimiento obtenido del proceso de minería. Los métodos para la selección de características son básicamente dos:
 - Aquellos basados en la elección de los mejores atributos del problema.

- Aquellos que buscan variables independientes mediante test de sensibilidad, algoritmos de distancia o heurísticos.
- Extracción de Conocimiento: Mediante una técnica de minería de datos, se obtiene un modelo de conocimiento, que representa patrones de comportamiento observados en los valores de las variables del problema o relaciones de asociación entre dichas variables. También pueden usarse varias técnicas a la vez para generar distintos modelos, aunque generalmente cada técnica obliga a un preprocesado diferente de los datos.
- Interpretación y Evaluación: Una vez obtenido el modelo, se debe proceder a su validación, comprobando que las conclusiones que arroja son válidas y suficientemente satisfactorias. En el caso de haber obtenido varios modelos mediante el uso de distintas técnicas, se deben comparar los modelos en busca de aquel que se ajuste mejor al problema. Si ninguno de los modelos alcanza los resultados esperados, debe alterarse alguno de los pasos anteriores para generar nuevos modelos. [\[41\]](#)

Sumando a los pasos que se deben seguir en la Minería de Datos, hay que tener muy claro como funciona esta. Si bien la tecnología de información a gran escala ha ido evolucionando por separado las transacciones y sistemas de análisis, la minería de datos proporciona un enlace entre los dos. El software de minería de datos analiza las relaciones y patrones en los datos de transacción almacenados sobre la base de consultas de los usuarios de composición abierta. Existen varios tipos de software de análisis que están disponibles: estadísticos, de aprendizaje automático, y redes neuronales. En general, se buscan cualquiera de estos cuatro tipos de relaciones:

- Clases: Los datos almacenados se utilizan para localizar datos en grupos predeterminados. Por ejemplo, una cadena de restaurantes podría minar datos de compra del cliente para

determinar cuando los clientes los visitan y lo que normalmente ordenan. Esta información podría utilizarse para aumentar el tráfico al tener especiales del día.

- Grupos: Los elementos de datos se agrupan de acuerdo a las relaciones lógicas o preferencias de los consumidores. Por ejemplo, los datos pueden ser extraídos para identificar segmentos de mercado o afinidades de los consumidores.
- Asociaciones: Los datos pueden ser minados para identificar asociaciones. El ejemplo de cerveza-pañal es un ejemplo de minería asociativa.
- Patrones secuenciales: Los datos se minan para anticipar patrones de comportamiento y tendencias. Por ejemplo, un distribuidor especializado en sistemas al aire libre puede predecir la probabilidad de compra de una mochila basado en la compra de un consumidor de sacos de dormir y calzado de senderismo. [40]

Diferentes niveles de análisis disponibles:

- Redes neuronales artificiales: modelos predictivos no lineales que aprenden a través de la formación y se asemejan a redes neuronales biológicas en su estructura.
- Algoritmos genéticos: Técnicas de optimización que usan procesos tales como combinación genética, mutación y selección natural en un diseño basado en los conceptos de evolución natural.
- Árboles de decisión: estructuras en forma de árbol que representan conjuntos de decisiones. Estas decisiones generan reglas para la clasificación de un conjunto de datos. Métodos específicos de árboles de decisión incluyen Árboles de Clasificación y Regresión (CART) y Detección Automática de Interacción (Chi Cuadrado CHAID). CART y CHAID son técnicas de árboles de decisión para la clasificación de un conjunto de datos. Constituyen

un conjunto de reglas que se pueden aplicar a un nuevo (sin clasificar) conjunto de datos para predecir cuáles registros tendrán un resultado determinado. CART segmenta un conjunto de datos mediante la creación de dos vías divididas, mientras que CHAID segmenta utilizando pruebas de chi cuadrado para crear divisiones en múltiples direcciones. CART normalmente requiere menos preparación de datos que CHAID.

- Método del vecino más próximo: Una técnica que clasifica cada registro en un conjunto de datos basado en una combinación de las clases de registro(s) k más similares a él en un conjunto de datos históricos (donde $k \geq 1$). A veces se llama técnica del vecino k -más cercano.
- Inducción de reglas: La extracción de reglas if-then útiles basados en la significación estadística.
- Visualización de datos: La interpretación visual de las complejas relaciones de datos multidimensionales. Las herramientas de gráficos se utilizan para ilustrar las relaciones de datos.

2.1.2. Enhancing Tourism Intermediaries with the Data Mining Process.

Este tercer estudio teórico encontrado para la investigación, está el realizado por Tzu-Ching Lin en el año 2012, denominado **Enhancing Tourism Intermediaries with the Data Mining Process**, llevado a cabo en la isla de Taiwán, donde se desarrolló un estudio para la industria del turismo, donde dicho proyecto permitía la administración y mejora de las relaciones con los clientes del turismo a largo plazo, sabiendo que los beneficiarios directos relacionados con el turismo serían restaurantes, hoteles, vendedores, comercio, artesanos, sitios de recreación, museos, discotecas, bares, guías turísticos, etc.

Su objetivo implicó la elaboración de una estrategia para las Agencias de Viajes, donde fueran ellas quienes cumplieran un rol importante en la operación del turismo. Por tal motivo ellos actuaron como Intermediarios entre los proveedores de hoteles, transporte, viajes y servicios ofrecidos a los consumidores. De esta manera se logró que las agencias de viajes fuesen reconocidas por el gobierno como parte fundamental de la cadena de distribución en el aumento de ingreso y salida de turistas, hasta el día de hoy.

De esta manera las agencias se interesaron por los viajeros y mejoraron sus plataformas teniendo como base un CRM. Con el estudio realizado, todos los administradores estuvieron de acuerdo que el mayor beneficio de la minería de datos era la de retener los clientes en cada una de las agencias. Luego del estudio realizado, el autor del estudio se dio cuenta que las agencias de viajes no aplicaron adecuadamente la minería de datos, de acuerdo con esto añadieron un 12 cuarto proceso valioso para las agencias de turismo intermediarias de la siguiente manera:

Paso 1. Obtener información de los clientes. En este paso, es posible para las agencias de viajes obtener información valiosa de los tiquetes, reservaciones y procesos de registro. Esta información incluye solicitudes especiales de las acomodaciones y la duración de la estadía, esto ayuda a los agentes de viaje a modificar sus campañas de marketing e incrementar las reservas.

Paso 2 Clasificar los datos recolectados de los clientes. Utilizando la minería de datos para compilar datos recolectados, se clasifica a los clientes en segmentos predefinidos que permiten un tamaño y estructura adecuados para ser monitoreados. Al final, las agencias de viajes pueden analizar los datos demográficos para sectorizar los clientes los cuales relacionan las compras y demandas de los clientes.

Paso 3 Identificando segmentos objetivos. La minería de datos es utilizada para identificar grupos de clientes con comportamiento similar y medir las reacciones de los clientes a las diferentes ofertas. La estrategia del CRM puede visualizarse con más claridad luego de este paso y es utilizado para establecer relaciones de larga duración con los clientes.

Paso 4 Mejorando los servicios. Los clientes de alta rentabilidad se pueden mantener por el uso de la minería de datos desde que dichas técnicas ayuden a mejorar el servicio de acuerdo con las preferencias de cada cliente. Por ejemplo, una agencia de viajes crea una campaña directa al email y se lo envía únicamente a los clientes objetivos.

De acuerdo con el proceso realizado de forma empírica por la entidad, las agencias de viaje mejoran sus estrategias de marketing y maximizan sus ganancias. [10]

2.1.3. A Knowledge Destination Framework for Tourism Sustainability: A Business Intelligence Application from Sweden.

Este proyecto llevado a cabo por estudiantes de Mid Sweden University en el 2013, quienes mediante diversas fuentes de información de datos de clientes viajeros dispersos en los distintos destinos turísticos como los servidores web de turistas, comportamiento de navegación web, base de datos con transacciones, etc... , recolectaron datos que permanecieron por cierto tiempo sin usar. Por este motivo, se fomentó la creación de conocimiento y adquisición de procesos en los diferentes destinos turísticos aplicando minería de datos, considerándola como una nueva alternativa para descubrir las preferencias de los consumidores del sector turístico, sobre todo en temas como hotelería y restaurantes, teniendo en cuenta la opinión de los sitios web a analizar como TripAdvisor.

Basados en el Framework (Marco de referencia) se analizó que debían realizarse diferentes etapas que permitieran conocer dichas preferencias que tenían los consumidores en el sector turístico.

Dichas etapas eran:

Capa de creación del conocimiento. Aplica métodos de extracción de información, convierte los orígenes de datos a una forma accesible para los subproveedores. Por ejemplo, en el lado del cliente/turista, el conocimiento se genera por mecanismos de feedback como encuestas y evaluaciones en línea. Después el conocimiento implícito puede ser explícito visualizando información de seguimiento de turistas. (Ejemplo: comportamiento de las búsquedas en la web) a través de minería. En el lado del proveedor, el conocimiento sobre productos, perfiles, información disponible, información sobre recursos y calidad de los recursos, así como calidad de vida de los residentes y satisfacción en el trabajo.

La capa de aplicación del conocimiento. Ofrece servicios a usuarios finales que inteligentemente informan acerca de los recursos de los destinos, elementos de los proveedores y actividades de los clientes. Así las aplicaciones para clientes / turistas particularmente comprenderá los servicios de localización que apoyan la construcción de comunidad y el consumo inteligente (por ejemplo, a través de servicios de recomendación por ser sensible al contexto, así como de adaptación para el usuario.

En cambio, los servicios inteligentes para los proveedores de destino interesados, se centran en aplicaciones de inteligencia de negocios relacionados con el turismo. [33]

2.1.4. A Supplier-Oriented Knowledge Application – The Destination Management Information System (DMIS).

Mediante esta teoría propuesta en el año 2002 por Robin J. B. Ritchie, se enfatiza explícitamente el uso del diseño DMIS. Este diseño, se basa en reconocer que el cliente se crea mediante técnicas de segmentación de clientes y servicios de evaluación de desempeño. [33] Los datos que son recolectados, se almacenan, analizan y evalúan en DMIS, incluyendo datos demográficos de los turistas, características de compra, motivos y percepción de marcas, como también la información de clientes y patrones de consumo de productos.

Dado que el uso eficaz de una DMIS requiere no sólo las aplicaciones de tecnología sofisticada, pero particularmente demanda establecer un aprendizaje organizacional, es crucial para integrar las partes interesadas públicas y privadas en para definir los requisitos de conocimientos. En el estudio de establecer el " Destino de Turismo Neutro ", los estudiosos del turismo obtuvieron suficiente evidencia empírica necesaria para caracterizar el turismo sostenible con el medio ambiente. En consecuencia, los turistas con la " huella ecológica más pequeña " relativamente se caracterizan por los siguientes atributos:

1. Socio- demográficos:
 - Mediana edad hasta tercera edad
 - Educación superior
2. El comportamiento en el viaje y estilos de las vacaciones:
 - Sitios de campamento
 - Apartamentos privados
3. Las actitudes positivas hacia actividades basadas en la naturaleza:

- Apreciación de la naturaleza y el disfrute de la belleza natural y el paisaje
- Interacción con la naturaleza y la preferencia por actividades basadas en la naturaleza (por ejemplo, senderismo, ciclismo, pesca, escalada, observación de la naturaleza, etc.)

En el documento y en la investigación realizada hay fundamentos teóricos y empíricos para enfocar la encuesta hacia los puntos más relevantes para tener en cuenta por el sector público y privado en cuanto a las características de los turistas. Los clientes dejan evidencias de sus actividades en la web como, búsquedas, planeación de viajes, reservaciones, consumos, feedback en sus sitios webs o blogs.

2.1.5. Atributos relevantes en el perfil del turista del interior de la ciudad de Mar de la Plata.

Finalmente para culminar los estudios y técnicas teóricas utilizadas para la investigación y análisis del proyecto, se encuentra el estudio realizado por la Red de Revistas Científicas de América Latina, el Caribe, España y Portugal en los años 2004-2005, publicado en un artículo, se hablaba dicho estudio tendría “El propósito del trabajo consistía en presentar los grupos de turistas del interior con características distintivas para su identificación durante la temporada de verano 2004-2005 en la ciudad de Mar del Plata a través de la utilización de técnicas de aprendizaje no supervisado.”

Las variables que se tuvieron en cuenta señalan los aspectos más relevantes en el trazado de los rasgos distintivos de aquellas personas que eligen dicha ciudad como destino de sus vacaciones, utilizando para este análisis a los turistas que residen en las provincias del interior de Argentina, dicho trabajo es empírico, seleccionando a priori las personas por su lugar de procedencia. En su metodología de desarrollo se puede observar que la información necesaria para el proceso de

minería de datos era inexistente, por lo tanto utilizaron varias técnicas de recopilación de información [15] para su posterior procesamiento:

- Encuesta Coincidental:

Lugar de toma de datos: Puntos fijos en estaciones de servicio, Estación de ómnibus, Estación de Trenes y Aeropuerto, se encuesta a los turistas en el momento de la partida de la ciudad.

Tamaño de la muestra: 300 casos semanales

Muestreo: probabilístico según ajuste al peso relativo de los medios de transporte utilizados por los turistas para ingresos a la ciudad.

- Encuesta de Ocupación Hotelera

Muestra: 117 hoteles

Estratificación: por categorías y por zonas

Medio de Recolección: relevamiento telefónico

- Encuesta de Perfil y Comportamiento

Lugar de toma de datos: balnearios de la ciudad

Muestreo: probabilístico por conglomerados.

Tamaño de la muestra: 1.100 casos proyectados, 1.072 casos realizados.

2.2. Marco Conceptual.

TURISMO: El turismo es un fenómeno social, cultural y económico relacionado con el movimiento de las personas a lugares que se encuentran fuera de su lugar de residencia habitual por motivos personales o de negocios/profesionales. Estas personas se denominan *visitantes* (que pueden ser *turistas* o *excursionistas*; residentes o no residentes) y el turismo tiene que ver con sus actividades, de las cuales algunas implican un *gasto turístico*.

Como tal, el turismo tiene efectos en la economía, en el entorno natural y en las zonas edificadas, en la población local de los lugares visitados y en los visitantes propiamente dichos. Debido a estos diversos impactos, la amplia gama y variedad de factores de producción requeridos para producir los bienes y servicios adquiridos por los visitantes y al amplio espectro de agentes involucrados o afectados por el turismo, es necesario adoptar un enfoque global del desarrollo, la gestión y la supervisión del turismo. Este enfoque es muy recomendable con miras a la formulación y puesta en práctica de políticas de turismo nacionales y locales, así como de acuerdos internacionales u otros procesos en materia de turismo. [26]

VIAJES: Un viaje es el traslado que realizan las personas desde el lugar donde habitan hacia por ejemplo otras ciudades dentro del mismo país o bien hacia otro país, y que puede tener un fin turístico como ser por descanso o vacaciones, para visitar a algún pariente o amigo que se encuentra viviendo en otra ciudad o país, o bien estar motivado por una necesidad de tipo profesional. El viaje aunque puede realizarse a pie, generalmente y más que nada por una razón de distancias se lleva a cabo en un medio transporte que puede ser público o privado. Los autos, los autobuses, los trenes, los barcos o los aviones son los medios de transporte más utilizados para emprender un viaje largo. [8]

CARACTERÍSTICA: Una característica es un rasgo o una singularidad que identifica a alguien o a algo. Normalmente se emplea el término en plural, pues son varios los elementos que sirven para describir las distintas realidades. Es valioso destacarse que esas cuestiones que singularizan y diferencian pueden estar asociadas a la personalidad, el carácter, el físico de una persona o al aspecto simbólico. [9]

TENDENCIA: Se considera que una tendencia es una corriente o una preferencia generalizada por la cual las personas se inclinan a preferir determinados fines o determinados medios por sobre otros. Las tendencias pueden consistir en un sentido general a la predisposición que tienen las cosas, los animales o las personas hacia una situación específica. También se considera que una tendencia es una fuerza física que permite que un cuerpo se incline o deslice hacia otro. Otro significado del término refiere a una manera específica de pensar, entender o razonar un aspecto particular del pensamiento (una forma de pensar de pensar la política, una forma de pensar el arte, una forma de pensar la economía, etc.). [37]

INTELIGENCIA DE NEGOCIOS: La Inteligencia de Negocios es una solución tecnológica cuyo objetivo primordial es el apoyar la toma de decisiones, esta solución tecnológica está conformada por metodologías y herramientas que permiten cumplir dicho objetivo. Sin embargo, la Inteligencia de negocios puede adaptarse a diversos alcances dependiendo de las necesidades del negocio. La inteligencia de Negocios busca transformar los datos en información para finalmente transformar la información en conocimientos. [38]

MINERÍA DE DATOS: La minería de datos es el proceso de detectar la información procesable de los conjuntos grandes de datos. Utiliza el análisis matemático para deducir los patrones y tendencias que existen en los datos. Normalmente, estos patrones no se pueden detectar mediante

la exploración tradicional de los datos porque las relaciones son demasiado complejas o porque hay demasiado datos. [\[18\]](#)

BASES DE DATOS: Una base de datos es una colección de información organizada de forma que un programa de ordenador pueda seleccionar rápidamente los fragmentos de datos que necesite. Una base de datos es un sistema de archivos electrónico. Las bases de datos tradicionales se organizan por campos, registros y archivos. [\[17\]](#)

2.3. Marco Legal.

El gobierno Nacional Colombiano ha reglamentado la Normatividad del Turismo, encargado de revisar las leyes, decretos y resoluciones en materia de turismo durante 1990 – 2016. Dicha normatividad permite que se conozcan las regulaciones, políticas, obligaciones y servicio que deben prestar las diversas agencias de viaje, transporte y demás entidades que hacen parte del sector turístico del país.

- **DECRETO 0348 DE 2015:** Decreto que tiene como objeto reglamentar la prestación del Servicio Público de Transporte Terrestre Automotor Especial y establecer los requisitos que deben cumplir las empresas interesadas en obtener y mantener la habilitación de ésta modalidad, las cuales deberán operar de forma eficiente, segura, oportuna y económica, cumpliendo con los principios rectores de transporte como el de libre competencia y el de la iniciativa privada, a las cuales solamente se les aplicarán las restricciones establecidas por la ley y los Convenios Internacionales. [\[25\]](#)
- **DECRETO 1075 DE 1997:** Por el cual se señala el procedimiento para la imposición de sanciones a los prestadores de servicios turísticos. El Ministerio de Desarrollo Económico impondrá sanciones, de oficio o a petición de parte, a los prestadores de servicios turísticos, cuando incurran en cualquiera de las siguientes conductas:[\[2\]](#)
 - a) Presentar documentación falsa o adulterada al Ministerio de Desarrollo Económico o a las entidades oficiales que la soliciten;
 - b) Utilizar publicidad engañosa o que induzca a error al público sobre precios, calidad o cobertura del servicio turístico ofrecido;

c) Ofrecer información engañosa o dar lugar a error en el público respecto a la modalidad del contrato, la naturaleza jurídica de los derechos surgidos del mismo y sus condiciones o sobre las características de los servicios turísticos ofrecidos y los derechos y obligaciones de los turistas;

d) Incumplir los servicios ofrecidos a los turistas;

e) Incumplir las obligaciones frente a las autoridades de turismo;

f) Infringir las normas que regulan la actividad turística;

g) Operar sin el previo registro de que trata el artículo 61 de la Ley 300 de 1996.

- **DECRETO 1095 30 DE MAYO DE 1994:** Por el cual se adopta el Código de Ética Profesional del Agente de Viajes y Turismo. En el ejercicio profesional el Agente de Viajes actuará respetando las normas establecidas en la Constitución y conforme a las leyes de la República. Su actividad se entiende como de interés público, por lo tanto deberá obrar de acuerdo con los postulados de la moral, la buena fe, y con diligencia y eficacia en la prestación del servicio, procurando la defensa de los intereses del usuario. [21]
- **+LEY 1558 10 DE JULIO DE 2012:** La presente ley tiene por objeto el fomento, el desarrollo, la promoción, la competitividad del sector y la regulación de la actividad turística, a través de los mecanismos necesarios para la creación, conservación, protección y aprovechamiento de los recursos y atractivos turísticos nacionales, resguardando el desarrollo sostenible y sustentable y la optimización de la ' calidad, estableciendo los

mecanismos de participación y concertación de los sectores público y privado en la actividad. [29]

- **LEY 300 DE 1996:** Por la cual se expide la ley general de turismo y se dictan otras disposiciones. Importancia de la industria turística. El turismo es una industria esencial para el desarrollo del país y en especial de las diferentes entidades territoriales, regiones, provincias y que cumple una función social. El Estado le dará especial protección en razón de su importancia para el desarrollo nacional. La industria turística se regirá con base en los siguientes principios generales:

1. Concertación. En virtud del cual las decisiones y actividades del sector se fundamentarán en acuerdos para asumir responsabilidades, esfuerzos y recursos entre los diferentes agentes comprometidos, tanto del sector estatal como del sector privado nacional e internacional para el logro de los objetivos comunes que beneficien el turismo.

2. Coordinación. En virtud del cual las entidades públicas que integran el sector turismo actuarán en forma coordinada en el ejercicio de sus funciones.

3. Descentralización. En virtud del cual la actividad turística es responsabilidad de los diferentes niveles del Estado en sus áreas de competencia y se desarrolla por las empresas privadas y estatales, según sus respectivos ámbitos de acción.

4. Planeación. En virtud del cual las actividades turísticas serán desarrolladas de acuerdo con el plan sectorial de turismo, el cual formará parte del plan nacional de desarrollo.

5. Protección al ambiente. En virtud del cual el turismo se desarrollará en armonía con el desarrollo sustentable del medio ambiente. [1]

- **DECRETO 972 DE 1997:** Por el cual se reglamentan los planes de servicios y descuentos especiales en materia de turismo para la tercera edad. Entidades que deberán prestar servicios a la tercera edad. Las entidades del orden nacional, regional y local que reciban recursos del Estado para desarrollar actividades de turismo, incluirán en sus planes los referentes a servicios y descuentos especiales para la tercera edad, para lo cual elaborarán las fichas de inversión correspondientes que serán presentadas a las oficinas de planeación del nivel estatal que corresponda. [11]
- **DECRETO 2438 DE 2010:** Por el cual se dictan unas disposiciones relacionadas con la responsabilidad de las Agencias de Viajes en la prestación de servicios turísticos. Las Agencias de Viajes en la prestación de sus servicios, deberán observar las siguientes reglas:
 1. Extender a los usuarios un comprobante que especifique los servicios contratados.
 2. Suministrar en forma completa la información sobre los servicios solicitados por los usuarios, indicando al viajero con precisión la hora estimada de llegada y de salida del destino y la duración de la estadía.
 3. Informar al usuario la facultad del organizador del viaje de efectuar modificaciones al plan o servicio turístico contratado en eventos de fuerza mayor o caso fortuito, sin que se requiera aceptación del usuario.

4. Llevar un archivo con todos los soportes, eventualidades y circunstancias en las que se desarrolló el plan o servicio turístico. [3]

- **LEY 105 DE 1993:** Por la cual se dictan disposiciones básicas sobre el transporte, se redistribuyen competencias y recursos entre la Nación y las Entidades Territoriales, se reglamenta la planeación en el sector transporte y se dictan disposiciones. o. Sector y Sistema Nacional del Transporte. Integra el sector Transporte, el Ministerio de Transporte, sus organismos adscritos o vinculados y la Dirección General Marítima del Ministerio de Defensa Nacional, en cuanto estará sujeta a una relación de coordinación con el Ministerio de Transporte. Conforman el Sistema Nacional de Transporte, para el desarrollo de las políticas de transporte, además de los organismos indicados en el inciso anterior, los organismos de tránsito y transporte, tanto terrestre, aéreo y marítimo e infraestructura de transporte de las entidades territoriales y demás dependencias de los sectores central o descentralizado de cualquier orden, que tengan funciones relacionadas con esta actividad.

[5]

- **LEY 336 DE 1996:** La presente ley tiene por objeto unificar los principios y los criterios que servirán de fundamento para la regulación y reglamentación del Transporte Público Aéreo, Marítimo, Fluvial, Férreo, Masivo, Terrestre y su operación en el Territorio Nacional, de conformidad con la Ley 105 de 1993, y con las normas que la modifiquen o sustituyan. La seguridad, especialmente la relacionada con la protección de los usuarios, constituye prioridad esencial en la actividad del Sector y del Sistema de Transporte. Para los efectos pertinentes, en la regulación del transporte público las autoridades competentes exigirán y verificarán las condiciones de seguridad, comodidad y accesibilidad requeridas

para garantizarle a los habitantes la eficiente prestación del servicio básico y de los demás niveles que se establezcan al interior de cada Modo, dándoles prioridad a la utilización de medios de transporte masivo. [6]

2.4. Marco Histórico.

El turismo como tal, nace en el siglo XIX, como una consecuencia de la Revolución industrial, con desplazamientos cuya intención principal es el ocio, descanso, cultura, salud, negocios o relaciones familiares. Estos movimientos se caracterizan por su finalidad de otros tipos de viajes motivados por guerras, movimientos migratorios, conquista, comercio, entre otros. No obstante el turismo tiene antecedentes históricos claros.

En la edad antigua, en la Grecia clásica se daba gran importancia al ocio, y el tiempo libre lo dedicaban a la cultura, diversiones, religión y deporte. Los desplazamientos más destacados eran los que realizaban con motivo de asistir a los Juegos Olímpicos de la Edad Antigua (las olimpiadas son los 4 años que transcurren entre dos juegos consecutivos en la ciudad de Olimpia), a las que acudían miles de personas y donde se mezclaban religión y deporte. También existían peregrinaciones religiosas, como las que se dirigían a los oráculos de Delfos y de Dodoma.

Durante el Imperio romano los romanos frecuentaban aguas termales (termas de Caracalla), eran asiduos de grandes espectáculos, como los teatros, y realizaban desplazamientos habituales hacia la costa (muy conocido es el caso de una villa de vacaciones a orillas del mar). Estos viajes de placer fueron posibles debido a tres factores fundamentales: la Paz romana, el desarrollo de importantes vías de comunicación y la prosperidad económica que posibilitó a algunos ciudadanos medios económicos y tiempo libre.

Durante la Edad Media hay en un primer momento un retroceso debido a la mayor conflictividad y recesión económica consiguiente. En esta época surge un tipo de viaje nuevo, las peregrinaciones religiosas. Éstas ya habían existido en la época antigua y clásica pero tanto el Cristianismo como

el Islam las extenderían a mayor número de creyentes y los desplazamientos serían mayores. Son famosas las expediciones desde Venecia a Tierra Santa y las peregrinaciones por el Camino de Santiago (desde el 814 en que se descubrió la tumba del santo), fueron continuas las peregrinaciones de toda Europa, creándose así mapas, mesones y todo tipo de servicios para los caminantes). En el mundo Islámico el Hajjad o peregrinación a La Meca es uno de los cinco Pilares del Islam obligando a todos los creyentes a esta peregrinación al menos una vez en la vida.

Las peregrinaciones continúan durante la Edad Moderna. En Roma mueren 1500 peregrinos a causa de una plaga de peste bubónica. Es en este momento cuando aparecen los primeros alojamientos con el nombre de hotel (palabra francesa que designaba los palacios urbanos). Como las grandes personalidades viajaban acompañadas de su séquito (cada vez más numeroso) se hacía imposible alojar a todos en palacio, por lo que se crearon estas construcciones.

Ésta es también la época de las grandes expediciones marítimas de españoles, británicos y portugueses que despiertan la curiosidad y el interés por viajar.

A finales del siglo XVI surge la costumbre de mandar a los jóvenes aristócratas ingleses a hacer el Grand Tour al finalizar sus estudios con el fin de complementar su formación y adquirir ciertas experiencias. Era un viaje de larga duración (entre 3 y 5 años) que se hacía por distintos países europeos, y de ahí proceden las palabras: turismo, turista, etc.

El Grand Tour es un viaje motivado por la necesidad de instrucción de estos jóvenes aristócratas que en un futuro habrán de gobernar su país. Del conocimiento "in situ" de la grandeza de Roma, París o Atenas así como de los debates en los cafés de los grandes centros termales, los viajeros

deberían aprender cómo llevar las riendas de un Imperio como el británico. Para algunos autores éste es el auténtico fenómeno fundacional del turismo moderno ya que surge como un fenómeno revolucionario en paralelo al resto de transformaciones que se dan en la Ilustración.

También en esta época hay un resurgir de las termas, que habían decaído durante la Edad Media. No sólo se asiste a ellas por consejo médico, sino que también se pone de moda la diversión y el entretenimiento en los centros termales como por ejemplo en Bath (Inglaterra). También de esta época data el descubrimiento de los baños de barro como remedio terapéutico, playas frías (Niza, Costa Azul) a donde iban a tomar los baños por prescripción médica.

Con la Revolución industrial se consolida la burguesía que volverá a disponer de recursos económicos y tiempo libre para viajar. En la Edad Contemporánea el invento de la máquina de vapor supone una reducción espectacular en los transportes, que hasta el momento eran tirados por animales. Las líneas férreas se extienden con gran rapidez por toda Europa y Norteamérica. También el uso del vapor en la navegación reduce el tiempo de los desplazamientos.

Comienza a surgir el turismo de montaña o salud: Se construyen famosos sanatorios y clínicas privadas europeas, muchos de ellos llegan a nuestros días como pequeños hoteles con encanto. Es también la época de las playas frías (Costa azul, Canal de la Mancha,...).

En 1841 Thomas Cook organiza el primer viaje organizado de la historia. Aunque fue un fracaso económico se considera un rotundo éxito en cuanto a precedente del paquete turístico, pues se percató de las enormes posibilidades económicas que podría llegar a tener esta actividad, creando así en 1851 la primera Agencia de Viajes del mundo "Thomas Cook & son".

En 1867 inventa el bono o Boucher, documento que permite la utilización en hoteles de ciertos servicios contratados y pre pagados a través de una agencia de viajes.

Al estallar la Primera Guerra Mundial en el verano de 1914 se considera que había aproximadamente 150.000 turistas americanos en Europa. Tras finalizar la guerra comenzó la fabricación en masa de autocares y automóviles. En esta época las playas y los ríos se convierten en el centro del turismo en Europa comenzando a adquirir gran importancia el turismo de costa.

El avión, utilizado por minorías en largas distancias, se va desarrollando tímidamente para acabar imponiéndose sobre las compañías navieras. La crisis del 1929 repercute en el negativamente en el sector turístico limitando su desarrollo hasta bien entrado en 1932. La Segunda Guerra Mundial paraliza absolutamente el turismo en el mundo y sus efectos se extienden hasta el año 1949.

Entre 1950 y 1973 se comienza a hablar del boom turístico. El turismo internacional crece a un ritmo superior de lo que lo había hecho en toda la historia. Este desarrollo es consecuencia del nuevo orden internacional, la estabilidad social y el desarrollo de la cultura del ocio en el mundo occidental. En esta época se comienza a legislar sobre el sector. También éstos son los años en los que se desarrollan los grandes núcleos urbanos y se hace evidente la masificación, surge también el deseo de evasión, escapar del estrés de las ciudades y despejar las mentes de presión.

En estos años se desarrolla la producción de automóviles en cadena que los hace cada vez más asequibles, así como la construcción de carreteras y autopistas, permite un mayor flujo de viajeros. De hecho, la nueva carretera de los Alpes que atraviesa Suiza de Norte a Sur supuso la pérdida de

la hegemonía de este país como núcleo receptor, ya que ahora los turistas cruzan Suiza para dirigirse a otros países con mejor clima.

El avión de hélice es sustituido por el de reacción, lo que supone un golpe definitivo para las compañías navieras, que se ven obligadas a destinar sus barcos a los cruceros o al desguace.

Todos estos factores nos llevan a la era de la estandarización del producto turístico. Los grandes tour operadores lanzan al mercado millones de paquetes turísticos idénticos. En la mayoría de los casos se utiliza el vuelo chárter, que abarata el producto y lo populariza. Al principio de este período (1950) había 25 millones de turistas, y al finalizar (1973) había 190 millones.

No obstante esta etapa también se caracteriza por la falta de experiencia, lo que implica las siguientes consecuencias como la falta de planificación (se construye sin hacer ninguna previsión ni de la demanda ni de los impactos medioambientales y sociales que se pueden sufrir con la llegada masiva de turistas) y el colonialismo turístico (hay una gran dependencia de los tour operadores extranjeros estadounidenses, británicos y alemanes fundamentalmente).

En los años 80`s se produce una internacionalización muy marcada de las grandes empresas hosteleras y de los tour operadores, que buscan nuevas formas de utilización del tiempo libre (parques temáticos, deporte, riesgo, salud,...) y aplican técnicas de marketing, pues el turista cada vez tiene mayor experiencia y busca nuevos productos y destinos turísticos, lo que crea una fuerte competencia entre ellos. La multimedia y las comunicaciones transforman el sector, modificando el diseño de los productos, la prestación del servicio, la comercialización del mismo de una manera más fluida.

La década de los 90 incluye grandes acontecimientos como la caída de los regímenes comunistas europeos, la Guerra del Golfo, la reunificación alemana, las Guerras yugoslavas, etc., que inciden de forma directa en la historia del turismo.

Se trata de una etapa de madurez del sector que sigue creciendo aunque de una manera más moderada y controlada. Se limita la capacidad receptiva (adecuación de la oferta a la demanda, se empieza a controlar la capacidad de aforo de monumentos, etc.), se diversifica la oferta (nuevos productos y destinos), se diversifica la demanda (aparecen nuevos tipos diferentes de turistas) y se mejora la calidad (al turista no le importa gastar más si la calidad es mejor).

El turismo entra como parte fundamental de la agenda política de numerosos países desarrollando políticas públicas que afectan a la promoción, planificación y comercialización como una pieza clave del desarrollo económico. Se mejora la formación desarrollando planes educativos especializados. El objetivo de alcanzar un desarrollo turístico sostenible mediante la captación de nuevos mercados y la regulación de la estacionalidad.

También las políticas a nivel supranacional consideran el desarrollo turístico con elementos tan importantes como el Tratado de Maastricht en 1992 (libre tráfico de personas y mercancías, ciudadanía europea,...), y en el 1995 la entrada en vigor Schengen y se eliminan los controles fronterizos en los países de la UE.

Existe de nuevo un abaratamiento de los viajes por vía aérea por medio de las compañías de bajo costo y la liberación de las compañías en muchos países y la feroz competencia de las mismas.

Esta liberalización afecta a otros aspectos de los servicios turísticos como la gestión de aeropuertos. [\[34\]](#)

2.5. Estado del Arte.

Para este proyecto se han tenido en cuenta experiencias y estudios realizados con anterioridad, que han permitido el aporte de información y conocimientos al tema de investigación sobre Tendencias y características de los viajeros que visitan la ciudad de Bucaramanga. Entre estos estudios se encuentra el Estudio de la Competitividad en el Sector Turismo en Colombia, realizado por Angela María Bravo Mantilla y David Ernesto Rincón Ambrosio, estudiantes de la Facultad Administración de Negocios Internacionales de la Universidad del Rosario en el año 2013. Esta investigación se realizó con el objetivo de conocer la evolución que ha tenido el sector turismo en el mundo y en Colombia para analizar el impacto que este ha generado en la economía nacional. La información obtenida de fuentes secundarias permitió conocer el desarrollo y el avance del sector a través de cifras y estadísticas con las cuales se identificaron las oportunidades y ventajas que tiene Colombia para lograr ser un país turísticamente competitivo y a su vez se evidenciaron los factores de mejora. Las Políticas Sectoriales ayudaron a identificar la normativa del sector y la manera en que cada uno de los procedimientos se debe llevar. Se realizó un análisis de competitividad en el país en el cual se resalta que a pesar de las falencias que aún se presentan en las actividades y servicios que se prestan para este sector, Colombia ha logrado un avance significativo con la correcta aplicación de las Políticas que el Gobierno Nacional creó y las buenas prácticas que día a día se convierten en un estilo de vida para los empresarios y actores participantes del sector. Se destacaron factores como la Política de Mercadeo y Promoción turística, los programas de educación y certificación que han sido apoyados por el Ministerio de Comercio, Industria y Turismo y la planificación para las diferentes obras que se han llevado a cabo por medio de herramientas como los Inventarios turísticos y los fondos que están disponibles para estas provenientes de entidades públicas y/o privadas, Fondo de promoción turística y

donaciones, entre otros. De igual manera, se destacó la importancia de los Convenios de Competitividad Turística como aspecto fundamental en el proceso de posicionar a Colombia como un destino de talla mundial ya que en estos se define la problemática por 12 departamentos, las estrategias para solucionarla y el seguimiento que cada región le ha hecho a sus compromisos. Se mostró el cumplimiento de los objetivos específicos propuestos mediante una tabla en la que se describe la acción que se llevó a cabo para la respuesta satisfactoria de cada uno de ellos. Finalmente se dieron recomendaciones para el desarrollo del sector con base al análisis realizado en toda la investigación, resaltando las ciudades que han tenido un mayor crecimiento en los dos últimos años, las acciones que hacen falta implementar en las principales ciudades del país, la necesidad de certificar el talento humano mediante cursos ofrecidos por entidades como el SENA, el mejoramiento de la infraestructura colombiana y la continuidad que se le debe dar a programas como las “marcas país”.

El segundo estudio se denomina Competitividad= El desafío para alcanzar un turismo de clase mundial, llevado a cabo por Ministerio de Comercio, Industria y Turismo en el año 2009, cuyo objetivo se centraba en analizar la necesidad de fortalecer las ventajas competitivas, a partir de la dotación de factores comparativos que Colombia tiene para el turismo, surge como estrategia y técnica especializada de la gestión pública y privada del sector desde mediados de los años 1990. Dentro de este esfuerzo de fortalecer la competitividad turística de Colombia como destino turístico, es importante resaltar algunos referentes que permiten dar continuidad a la gestión adelantada en este tema, como los siguientes:

- El estudio de Michael Porter, importante teórico de la competitividad, que en su documento “Creación de la ventaja competitiva de Colombia – Informe MONITOR” (1994), identificó al turismo como sector de gran potencialidad nacional.
- La orientación filosófica del marco legal del sector, que busca la creación de ventajas competitivas, reflejada en las disposiciones, competencias institucionales e instrumentos creados en la Ley 300 de 1996, Ley General de Turismo.
- El Estudio de Competitividad del Sector Turismo (1996) que identificó los clúster o regiones que caracterizan la oferta turística colombiana. El documento analizó acertadamente las limitantes y potencialidades competitivas en los temas de entorno base del turismo, potencial de atractivos, infraestructura regional, oferta de servicios turísticos, conectividad, calidad de componentes turísticos, entorno empresarial, papel del sector público, nacional y local, países y clúster competidores y análisis de microclusters.
- La articulación de la gestión de planificación turística con la suscripción de Convenios de competitividad turística departamentales, que a partir de matrices priorizadas y consensuadas con las regiones, identificaron las principales barreras que afectan la capacidad de competencia del sector y de sus respectivos actores en 30 regiones del país (2001). Sus insumos fueron fundamentales para los Planes Sectoriales de Turismo 2003-2006 y 2007-2010.
- La declaración del turismo como sector de clase mundial y eje para hacer de Colombia un destino de clase mundial, definido en la Política Nacional de Productividad y Competitividad, la cual diseñó unos programas específicos que buscan fortalecer la capacidad productiva en el entorno empresarial, desarrollando sectores de manera estratégica, entre los cuales se encuentra el turismo.

Uno de sus principales mecanismos para aplicar esta política es la creación de los convenios de competitividad.

- El reciente impulso a la evaluación del rol en materia de competitividad que generan de las autoridades regionales de turismo, medición propuesta en los avances obtenidos sobre indicadores de competitividad turística.

- La definición de agendas conjuntas regionales y nacionales para fortalecer los convenios de competitividad turística, a través de espacios de concertación propiciados por el Ministerio de Comercio, Industria y turismo, tales como “Colombia Crece”, “Colombia Compite” y especialmente “Turismo Compite”.

- La Política Nacional de Turismo – Plan Sectorial de Turismo 2008-2010: “Colombia Destino Turístico de Clase Mundial”, cuyo objetivo general busca “consolidar los procesos regionales de turismo de tal manera que se disponga de una oferta de productos y destinos altamente competitivos (de clase mundial) para los mercados nacionales e internacionales”. De igual forma, el Plan establece entre sus estrategias principales el mejoramiento y consolidación de la competitividad del sector, así como la calidad de los servicios, los productos, las empresas y destinos turísticos.

También se encontró en el estudio de Seguridad Turística= Reto competitivo de Colombia (Plan Estratégico) dirigido por Ministerio de Comercio, Industria y Turismo – Policía de Turismo en el año 2003. Esta investigación buscaba poder caracterizar la problemática específica de la seguridad en el turismo, teniendo en cuenta la percepción que sobre el tema tienen los turistas y cómo incide en su satisfacción. Aunque Colombia aún tiene problemas de seguridad, el trabajo imperante por garantizar seguridad a los viajeros ha arrojado buenos resultados, ya que la percepción del turista

tanto nacional como internacional ha sido transformada, pues el número de viajeros que visitan el país ha venido aumentando de manera importante, así como también se ha notado incremento en el turismo doméstico. Las situaciones de inseguridad afectan la tranquilidad del turista y condiciona las acciones de la Policía de Turismo. Estas limitantes requieren soluciones integradas con otros actores del desarrollo turístico, las cuales vienen siendo estructuradas como prioridad de acción nacional y regional a través de los diferentes acuerdos regionales de competitividad que ha venido suscribiendo el Viceministerio de Turismo con los principales destinos turísticos potenciales y reales del país. Para el análisis de la problemática de destino se recopilaron los planteamientos propuestos por el Estudio de Competitividad del Sector Turismo. De igual forma, se complementó esta información con los datos suministrados por la Policía de Turismo a nivel regional, por algunas oficinas departamentales de turismo y por los Consejos de Seguridad realizados a la fecha.

El cuarto estudio se titula Medición del Turismo interno en Colombia, Experiencias y Retos, realizado por Marlon Pinot de Libreros, Javier Andrés Rojas Pulido y Carlos Arturo Mora Martínez, personal capacitado del DANE en el año 2013. Dicho proyecto fue liderado por la CAN (Comunidad Andina de Naciones), la Dirección de Síntesis y Cuentas Nacionales (DSCN) del Departamento Administrativo Nacional de Estadísticas (DANE) de Colombia realizó en los años 2010-2011 un ejercicio piloto de estimación de una Cuenta Satélite de Turismo de Colombia referida al período 2000 – 2005. Este ejercicio permitió apreciar la importancia del turismo interno dentro del turismo interior, al arrojar que aproximadamente el 75% del gasto del turismo interior corresponde al turismo interno (véase cuadro No 1), es decir que supera con creces al turismo receptor. Además de recalcar su importancia en términos de participación dentro del total, el ejercicio evidenció la importancia de conocerlo mejor ya que los datos disponibles eran

insuficientes para caracterizarlo con la profundidad requerida. Con este fin, el Vice Ministerio de Turismo acordó mediante convenio 146 de 2011 con la Dirección de Metodología y Producción Estadística (DIMPE) del DANE la realización de la Encuesta de Gasto en Turismo Interno (EGIT) con el objetivo de “consolidar el sistema de estadísticas de turismo con el fin de disponer de información que permita el diseño eficiente de políticas, estrategias, planes y programas turísticos y fortalecer el proyecto de cuenta satélite de turismo de Colombia”.

El Departamento Administrativo Nacional de Estadística (DANE) en agosto de 2003 realizó la Encuesta de Turismo a Hogares – ETUR que tenía como objetivos recopilar información sobre el turismo realizado por los residentes en Colombia, (interno y emisor), disponer de información confiable para caracterizar al sector para así mismo promoverlo, y alimentar el Centro de Información Turística (CITUR) en cabeza del Ministerio de Comercio, Industria y Turismo. La cobertura geográfica de la encuesta correspondía a la zona urbana del país, y se obtuvieron resultados desagregados para Barranquilla, Bogotá, Bucaramanga, Cali, Cartagena, Medellín, Pereira, Santa Marta y resto urbano. El período de referencia fue diferente, de acuerdo con los temas indagados: para viajes con pernoctación, este período fue de un año calendario (agosto 2002 a julio 2003) y para viajes a segunda vivienda y excursionismo, el mes anterior al periodo de recolección (julio 2003). En el caso de los viajes con pernoctación, por ser el período de referencia tan largo, se presentó un efecto notorio de recordación por parte de las personas encuestadas, por ello se obtuvieron resultados inconsistentes, con una propensión a viajar demasiado baja para el periodo de estudio (1,8 viajes por persona-año) inconsistente con las demás informaciones disponibles tales como los flujos del turismo emisor (número de viajes), ocupación hotelera y datos de otros países de la región. Las variables investigadas fueron: disponibilidad de segunda vivienda, características socioeconómicas de los miembros del hogar encuestado, caracterización de los

viajes realizados (destino, número de viajes, número y caracterización de las personas que participaron en los viajes, pernoctaciones, motivo principal de los viajes y fecha de realización), y gastos asociados a los viajes. Uno de los inconvenientes de esta investigación fue la definición del viaje como variable asignada a un grupo de personas viajando juntas y no a los individuos, lo que obligó a un nuevo procesamiento de la información con otra unidad estadística para su uso en la compilación de la Cuenta Satélite, con el fin de obtener información coherente, tanto con el sistema de estadísticas de turismo definido internacionalmente como con los datos de oferta (en particular provenientes de la encuesta a hoteles). Luego entre octubre de 2006 y marzo de 2007 se incluyó un módulo de turismo en la Gran Encuesta Integrada de Hogares (GEIH) con el fin de generar la información necesaria para cuantificar la demanda de servicios turísticos y establecer indicadores para determinar su incidencia en la economía nacional. Desafortunadamente fue difícil aprovechar esta información porque el diseño del cuestionario no hacía posible asociar el gasto realizado por cada persona a los viajes realizados; en particular no se pudo establecer el vínculo de las estructuras de gasto al motivo de viaje y a la forma de alojamiento, variables que necesariamente deben tomarse en cuenta para caracterizar el comportamiento turístico, en particular el gasto. Finalmente, en el año 2011, el ejercicio de Cuenta Satélite de Turismo antes mencionado dimensionó la importancia del turismo interno dentro de la actividad turística nacional y puso en evidencia la falta de información oportuna, suficiente y confiable sobre el mismo y la urgencia de contar con ella.

Al finalizar dicho estudio se concluyó que se deben seguir tomando las medidas para que esta experiencia sea la primera de una serie de mediciones: se ha aprendido mucho no solamente sobre el sector turismo, sino sobre como mejor observarlo, en particular la adaptación del diseño muestral a la observación de un fenómeno escaso (sólo 15% de ocurrencia) lo cual requiere un diseño especial para volverlo eficiente

(y por consiguiente, menos costoso en términos de trabajo de campo) y mejorar la distribución geográfica de la muestra y ampliar la cobertura a otras ciudades capitales, en particular algunas con atractivos turísticos tales como Santa Marta, Armenia, San Andrés o Neiva. La duración de los viajes debe también medirse de manera diferente a lo hecho ahora, asignando todo viaje a su período de terminación. Adicionalmente, el turismo es un fenómeno con un alto componente estacional. El diseño actual no permite apreciarlo en toda su magnitud. Es necesario plantear muy precisamente como se quiere medir esta estacionalidad, y por ejemplo el efecto de la ley Emiliani¹¹ o de los diferentes festivales regionales sobre los viajes personales. También se intuye que a mayor ingreso, mayor propensión a viajar, una particularidad que no se pudo captar con el diseño actual.

Por último, se tuvo en cuenta el estudio más relevante para la investigación del proyecto. Este estudio se denominó Estudio de Mercado, Estudio de Agencias de Viajes en Colombia, realizado en el año 2011 por la Superintendencia de Industria y Comercio, cuyo objetivo pretende evaluar los diferentes aspectos de la dinámica del sector de las agencias de viajes mediante una aproximación al mismo, a partir del establecimiento de la estructura de mercado, y la realización del correspondiente análisis de los principales participantes en este y las medidas regulatorias implementadas en dicho sector. Lo anterior, con el propósito de lograr un enriquecimiento académico respecto de este sector, de manera que sea una herramienta útil para posteriores estudios económicos.

Asimismo, se busca conocer el mercado de las agencias de viajes a nivel nacional, dadas sus condiciones como un factor importante en la economía y en el sector del turismo. En este sentido, a nivel mundial, se estima que la participación del sector turismo es del 5% del PIB mundial.

Para la elaboración de este estudio se contó con información proporcionada por las 50 agencias de viajes que mayores ventas reportaron en el año 2011. Asimismo, se recopiló información de

autoridades administrativas como el DANE, el Registro Nacional del Turismo, la Aeronáutica Civil, el Ministerio de Comercio, Industria y Turismo, y participantes del mercado que incluyen aerolíneas nacionales y extranjeras con operaciones en Colombia, la Asociación Colombiana de Agencias de Viajes y Turismo-ANATO, Fondo de Promoción Turística, entre otras fuentes.

Se busca, por tanto, lograr la consolidación de la información recopilada, para que sea una herramienta útil para recomendar estrategias en el papel de las agencias de viajes para mantener la dinámica del mercado, con el fin de buscar el beneficio de los consumidores y la protección de la competencia.

Los propósitos principales de este estudio son:

- Analizar el comportamiento económico del sector turístico con relación a las agencias de viajes a nivel nacional.
- Determinar los índices de concentración en el mercado de agencias de viajes a nivel nacional.
- Analizar la regulación del sector de las agencias de viajes y turismo.
- Detectar las barreras de entrada para el sector de las agencias de viajes en Colombia.
- Brindar recomendaciones con relación al mercado de agencias de viajes y turismo a nivel nacional para el fortalecimiento de las agencias de viajes.

3. Metodología.

Para realizar esta investigación se requiere de una metodología independiente de las herramientas de minería de datos, la cual contenga procesos de revisión de negocio, que permita organizar, estructurar y definir parámetros iniciales de negocio. Dicha metodología será cuantitativa pues se analizarán y lograr obtener mediciones de las variables que influyen en este para la obtención de datos y resultados. [35] Es por esto que la metodología que se necesita para la investigación es la CRISP-DM, por su uso de herramientas libres y herramientas de pago.

3.1.Fundamentos Epistemológicos.

El enfoque que se llevará a cabo para la metodología antes nombrada, será un enfoque cuantitativo. Por este motivo se realizarán encuestas a 400 personas de estrato 3 y 4 que viajan a Bucaramanga, ya sea visitantes o turistas, conociendo su preferencia de esta ciudad al viajar, por qué motivos viaja, que desean encontrar durante su estadía, y demás características fundamentales para llevar a cabo el análisis, y de este modo generar ideas para la toma de decisiones enfocadas en aumento del sector turístico de la ciudad. Al final de la investigación estos resultados servirán como punto de partida para las empresas del sector turísticos que deseen implementar estrategias para obtener beneficios referentes a la satisfacción del servicio al cliente.

Esta investigación es de tipo exploratorio puesto que el objetivo principal será dar una percepción sobre las tendencias y características de los viajeros que visitan la ciudad de Bucaramanga. De igual forma es de tipo descriptivo, ya que se enfoca en preferencias e intereses que posee una población específica, y que basado en estos se realizará un análisis de los resultados que se obtengan. También es un investigación de tipo exploratorio, donde se identificará una situación para conocer y analizar las percepciones de las variables a medir. Y por último, identificamos que

la investigación es de tipo correlacional, pues mide de relación entre las variables que antes fueron analizadas, identificando su grado de aceptación o rechazo con ayuda de la metodología CRISP-DM.

1. COMPRESION DEL DOMINIO: Esta fase inicial se enfoca en la comprensión de los objetivos de proyecto y exigencias desde una perspectiva institucional, luego convirtiendo este conocimiento de los datos en la definición de un problema de minería de datos y en un plan preliminar diseñado para alcanzar los objetivos:

- Establecimiento de los objetivos del proyecto.
- Evaluación de la situación.
- Establecimiento de los objetivos de la minería de datos

2. COMPRESIÓN DE LOS DATOS: La fase de entendimiento de datos comienza con la colección de datos inicial y continúa con las actividades que le permiten familiarizarse primero con los datos, identificar los problemas de calidad de datos, descubrir los primeros conocimientos en los datos, y descubrir subconjuntos interesantes para formar hipótesis en cuanto a la información oculta:

- Recopilación inicial de datos.
- Descripción de los datos.
- Exploración de los datos.
- Verificación de calidad de datos.

3. PREPARACIÓN DE DATOS: La fase de preparación de datos cubre todas las actividades necesarias para construir el conjunto de datos final (los datos que serán provistos en las herramientas de modelado) de los datos en brutos iniciales. Las tareas de

preparación de datos probablemente van a ser realizadas muchas veces y no en cualquier orden prescripto. Las tareas incluyen la selección de tablas, registros, y atributos, así como la transformación y la limpieza de datos para las herramientas que modelan:

- Selección de los datos.
- Limpieza de datos.
- Construcción de datos.
- Integración de datos.

4. MODELADO: En esta fase, varias técnicas de modelado son seleccionadas y aplicadas, y sus parámetros son calibrados a valores óptimos. Típicamente hay varias técnicas para el mismo tipo de problema de minería de datos. Algunas técnicas tienen requerimientos específicos sobre la forma de datos. Por lo tanto, volver a la fase de preparación de datos es a menudo necesario:

- Selección de la técnica de modelado.
- Diseño de la evaluación.
- Construcción del modelo.
- Evaluación del modelo.

5. EVALUACIÓN: .En esta etapa del proyecto, se ha construido un modelo (o modelos) que parecen tener la alta calidad desde la perspectiva de análisis de datos. Antes del proceder al despliegue final del modelo, es importante evaluarlo a fondo y la revisión de los pasos ejecutados para crearlo, para comparar el modelo correctamente obtenido con los objetivos de negocio. Un objetivo clave es determinar si hay alguna cuestión importante de negocio

que no ha sido suficientemente considerada. En el final de esta fase, una decisión en el uso de los resultados de minería de datos debería ser obtenida:

- Evaluación de resultados
- Revisar el proceso
- Validación de resultados obtenidos

6. DESPLIEGUE: La creación del modelo no es generalmente el final del proyecto. Incluso si el objetivo del modelo es de aumentar el conocimiento de los datos, el conocimiento ganado tendrá que ser organizado y presentado en el modo en el que el cliente pueda usarlo. Ello a menudo implica la aplicación de modelos "vivos" dentro de un proceso de toma de decisiones de una organización, por ejemplo, en tiempo real la personalización de página Web o la repetida obtención de bases de datos de mercadeo. Dependiendo de los requerimientos, la fase de desarrollo puede ser tan simple como la generación de un informe o tan compleja como la realización repetida de un proceso cruzado de minería de datos a través de la empresa. En muchos casos, es el cliente, no el analista de datos, quien lleva el paso de desarrollo. Sin embargo, incluso si el analista realizara el esfuerzo de despliegue, esto es importante para el cliente para entender de frente que acciones necesita para ser ejecutadas en orden para hacer uso de los modelos creados actualmente:

- Generación de informe final
- Documento con experiencias

3.2. Diseño de la Investigación.

3.2.1. Tipo de Investigación.

La investigación de este estudio es cuantitativa de tipo exploratorio ya que al no existir investigaciones previas sobre el tema pero sin culminar, se efectuará este tipo de investigación para la recolección de datos, mediante encuestas donde se hará énfasis en los viajeros que llegan a la ciudad de Bucaramanga. De igual forma será también de tipo descriptivo, donde se seleccionarán variables que se medirán y analizarán por medio de las 400 encuestas que se realizarán a personas de estrato 3 y 4. De este modo se tendrá una especificación y caracterización sobre las personas encuestadas conociendo edad, sexo, nivel de escolaridad, preferencias, etc. Esta encuesta será de tipo explicativa, pues se conocerán los motivos por los cuales las personas entrevistadas viajan a la ciudad, y con los resultados obtenidos, analizar e interpretarlos, con el fin de cumplir con la hipótesis o planteamiento inicial del proyecto, la tendencia y característica de los viajeros que visitan la ciudad de Bucaramanga. Por último se identifica una investigación correlacional, para medir el grado de relación e interacción entre las variables del análisis, identificando aspectos negativos y positivos al momento de realizar el tratamiento de la información obtenida.

Además se tiene en cuenta la intervención del investigador en el escenario será no experimental, debido a que se busca sólo la medición de las variables y el comportamiento de estas. Y finalmente según la ubicación temporal del proyecto, será vertical o transversal, ya que el estudio se efectuará en un periodo determinado de un mes.

3.2.2. Fases, etapas o momentos del proyecto.

Tabla 3.

Fases, etapas o momentos del proyecto.

FASES	ACTIVIDADES A REALIZAR
<p>Comprensión del Dominio: se enfoca en la comprensión de los objetivos de proyecto, luego convirtiendo este conocimiento de los datos en la definición de un problema de minería de datos y en un plan preliminar diseñado para alcanzar los objetivos.</p>	<ul style="list-style-type: none"> ● Establecimiento de los objetivos del proyecto. ● Evaluación de la situación. ● Establecimiento de los objetivos de la minería de datos. ● Realizar el plan de proyecto.
<p>Comprensión de los Datos: fase de entendimiento de datos comienza con la colección de datos inicial y continúa con las actividades que le permiten familiarizarse primero con los datos, identificar los problemas de calidad de datos, descubrir los primeros conocimientos en los datos, y descubrir subconjuntos interesantes para formar hipótesis en cuanto a la información oculta.</p>	<ul style="list-style-type: none"> ● Recopilación inicial de datos. ● Descripción de los datos. ● Exploración de los datos. ● Verificación de los datos.
<p>Preparación de Datos: cubre todas las actividades necesarias para construir el conjunto de datos final (los datos que serán provistos en las herramientas de modelado) de los datos en brutos iniciales. Las tareas incluyen la selección de tablas, registros, y atributos, así como la transformación y la limpieza de datos para las herramientas que modelan</p>	<ul style="list-style-type: none"> ● Selección de datos. ● Limpieza de datos. ● Construcción de datos. ● Integración de datos. ● Formateo de datos.
<p>Modelado: varias técnicas de modelado son seleccionadas y aplicadas, y sus parámetros son calibrados a valores óptimos. Típicamente hay varias técnicas para el mismo tipo de problema de minería de datos. Algunas técnicas tienen requerimientos específicos sobre la forma de datos. Por lo tanto, volver a la fase de preparación de datos es a menudo necesario.</p>	<ul style="list-style-type: none"> ● Selección de la técnica de modelado. ● Diseño de la evaluación. ● Construcción del modelado. ● Evaluación del modelado.
<p>Evaluación: En esta etapa del proyecto, se ha construido un modelo (o modelos) que parecen tener la alta calidad desde la perspectiva de análisis de datos. Antes del proceder al despliegue final del modelo, es importante evaluarlo a fondo y la revisión de los pasos ejecutados para crearlo, para comparar el modelo correctamente obtenido con los objetivos de negocio. En el final de esta fase, una decisión en el uso de los resultados de minería de datos debería ser obtenida.</p>	<ul style="list-style-type: none"> ● Evaluación de resultados. ● Revisar el proceso. ● Validación de resultados obtenidos.

<p>Despliegue: Dependiendo de los requerimientos, la fase de desarrollo puede ser tan simple como la generación de un informe o tan compleja como la realización repetida de un proceso cruzado de minería de datos a través de la empresa.</p>	<ul style="list-style-type: none"> • Planificación de despliegue • Planificación de la monitorización y del mantenimiento. • Generación del informe final. • Documento con experiencias. • Revisión del proyecto.
--	--

Nota: Etapas a desarrollar durante el proyecto.

3.2.3. Técnicas e instrumentos de Recolección en Investigación Cuantitativa.

Tabla 4.

Técnicas e instrumentos de Recolección en Investigación Cuantitativa.

TECNICA	REGISTRO
Encuesta	Cuestionario: Se realizarán las mismas preguntas a todos los encuestados. Se llevará a cabo de manera personal y autodirigida para obtener datos correctos y exactos.
Entrevista	Guía de Entrevista: Preguntas abiertas y cerradas, donde se tendrá en cuenta la opinión del entrevistado, enfocadas en el tema, y se tendrá como evidencia fotos al momento de realizarlas.
Análisis de contenido cuantitativo	Análisis de datos por medio de Minería de datos
Observación	Realizar una pequeña charla a los entrevistados donde se le explicara de manera breve en que consiste la metodología de la encuesta y qué se quiere lograr con esto.
Pruebas estandarizadas	Protocolo al momento de realizar la encuesta para lograr el objetivo de la totalidad de encuestas a realizar.
Datos secundarios	Información recolectada y estudios realizados por otros investigadores.

Nota: Técnicas utilizadas para la recolección de datos de la investigación cuantitativa.

4. Análisis e Interpretación De Resultados.

4.1. Desarrollo del Proceso de Minería de Datos con el modelo CRISP-DM.

Comprensión del Dominio

Objetivos institucionales de las entidades encargadas del Turismo. Los objetivos de las personas encargadas del turismo son los mismos planteados en la investigación, es decir, los objetivos de establecimiento de tendencias y características de los viajeros que visitan la ciudad de Bucaramanga.

Dentro de estos objetivos, especificamos algunos como:

- Recolectar los datos de los viajeros que visitan la ciudad.
- Clasificar los datos recolectados de los viajeros.
- Identificar los segmentos objetivos.
- Segmentar y describir hábitos de los viajeros que visitan la ciudad de Bucaramanga mediante un Análisis de Clúster para la industria turística de Bucaramanga.
- Establecer relaciones y tendencias existentes entre las características de los viajeros y los aspectos que rodean su visita a la ciudad de Bucaramanga utilizando reglas de asociación.
- Determinar en los viajeros que visitan la ciudad de Bucaramanga en nivel de gastos, percepción y satisfacción luego de visitar la ciudad por medio de un análisis de clúster.

Para el turismo es evidente el cambio en el entorno y los requerimientos por parte de proveedores y consumidores, ya que constantemente los destinos cambian debido a situaciones, temporadas o celebraciones de ciertos eventos culturales.

Evaluación de la Situación

- Recursos disponibles: Se cuenta como único recursos disponible con el autor de proyecto, encargado de diseñar las encuestas basado en los modelos de encuestas, facilidad de uso y herramientas ofimáticas y software para automatizar las encuestas.
- Datos: Los datos con los que se cuenta son los obtenidos al realizar la encuesta.

La información con la que cuenta esta base de datos son:

- Datos personales del viajero: Edad, Género, Ciudad de Origen, País de Origen, Estado Civil, Fuente de Ingreso.
- Datos sobre el viaje: N° de Veces que ha Viajado, N° de Días de Permanencia, N° de Compañeros de Viaje, Lugar de Alojamiento, Medio de Transporte de Llegada a la ciudad, Medio de Transporte utilizado en la ciudad, Motivo de Viaje, Actividad Planeada.
- Datos de satisfacción: Lugar de Alimentación, Gasto Promedio, Calificación del Gasto Promedio, Aspecto que lo Impacto de la ciudad, Aspecto que lo Motiva a Regresar.

Recursos Computacionales

Hardware: Estos recursos computacionales para el procesamiento y análisis de la información son proporcionados por el autor del proyecto.

- Computador portátil Dell Inspiron N4050 Intel® Core™ i5-2450M CPU @2.50GHz con 4 GB de RAM.

Software: Se utilizan Herramientas de Office y el uso del software es de manera gratuita, como SPSS y WEKA, para la limpieza de datos y brindando lenguajes y entornos de programación de aplicaciones estadísticas y minería de datos requeridos en cualquier etapa del proceso CRISP-DM.

La publicación de los resultados no contiene información que afecte a los viajeros.

Determinación de los objetivos de la minería de datos

- Segmentar y describir hábitos de los viajeros que visitan la ciudad de Bucaramanga mediante un Análisis de Clúster para la industria turística de Bucaramanga.
- Establecer relaciones y tendencias existentes entre las características de los viajeros y los aspectos que rodean su visita a la ciudad de Bucaramanga utilizando reglas de asociación.
- Determinar en los viajeros que visitan la ciudad de Bucaramanga en nivel de gastos, percepción y satisfacción luego de visitar la ciudad por medio de un análisis de clúster.

Comprensión de los datos

Recolección de datos iniciales. El conjunto de datos se recolectó en su totalidad, implementando encuestas, donde dicha información se alojó posteriormente en una base de datos para el análisis de esta. La información se encuentra almacenada en una tabla estructurada con una serie de preguntas con todos los campos de posibles respuestas, con el fin de lograr control de los datos seleccionados, evitando que los encuestados dejaran datos incompletos o datos erróneos. Esto no evita que se encuentren datos poco comunes como un viajero menor de 18 años con una fuente de ingreso ALTA. Dichos datos serán considerados dentro de este proyecto como “Outliers” o datos que alteran la muestra, pero que pertenecen a un grupo a analizar.

Toda la información obtenida se maneja en la etapa de preparación de datos.

Cabe resaltar que cada viajero encuestado solo puede participar una sola vez en este proceso.

El proceso de recolección de datos se llevó a cabo de la siguiente manera:

- Se diseñó la encuesta restringiendo cada uno de los campos de respuestas para que el encuestado solo pudiera seleccionar aquella que cumpliera con lo preguntado.

- La encuesta fue desarrollada en idioma español, pero si existía la posibilidad de encuestar a un extranjero cuyo idioma fuese diferente, se optaba por traducirla y brindar asesoría para su fácil diligenciamiento.
- La información recolectada fue computada en el software SPSS.
- Se validó que toda la información estuviera correctamente diligenciada y no presentara errores al traspasar los datos obtenidos en la encuesta.

Tabla 5.

Campo de datos personales de los encuestados.

CAMPO	TIPO	DESCRIPCIÓN
Edad.	Selección múltiple.	Menosde18Años, 18-25Años, 26-35Años, 36-45Años, 46-55Años, 56-65Años, 66-75Años, Másde75Años.
Género.	Selección múltiple.	Masculino, Femenino.
Ciudad de Origen.	Selección múltiple.	Texto.
País de Origen.	Selección múltiple.	Texto.
Estado Civil.	Selección múltiple.	Soltero, Casado, UniónLibre, Viudo, Divorciado.
Fuentes de Ingreso.	Selección múltiple.	Estudiante, Empleado, Independiente, Pensionado/Jubilado, Desempleado, Familia, Otros.

Nota: Caracterización de los encuestados teniendo en cuenta sus datos personales.

Tabla 6.

Campo de motivos de viaje y clasificación.

CAMPO	TIPO	DESCRIPCIÓN
N°VecesViajado.	Selección múltiple	0-1Veces, 1-2Veces, 3-10Veces, 11-20Veces, Másde20Veces.
N°DíasPermanencia.	Selección múltiple.	0-1Día, 1-2Días, 3-10Días, 11-20Días, Másde20Días.
N°CompañeroViaje.	Selección múltiple.	0-1Personas, 1-2Personas, 3-10Personas, 11-20Personas, Másde20Personas.
LugarAlojamiento.	Selección múltiple.	Hotel, Familia, Amigos, Hostal, Otro.
MedioTransportedeLlegada.	Selección múltiple.	Avión, TransportePúblico, Particular, Otro.
MedioTransporteenlaCiudad.	Selección múltiple.	TransportePúblico, Particular, Caminar, Otro.
MotivoViaje.	Selección múltiple.	Turismo, Comercio, Aventura, Educación, Salud, Cultura, Deportes, Otro.
ActividadPlaneada.	Selección múltiple.	Turismo, Comercio, Aventura, Educación, Salud, Cultura, Deportes, Otro.
LugarAlimentación.	Selección múltiple.	Restaurante, ComifdaRápidaenRestaurantes, ComidaRápidaenlaCalle, Familia, Amigos, Otro.
GastoPromedio.	Selección múltiple.	Menosde\$100.000, \$100.000-\$150.000, \$150.000-\$300.000,

		\$300.000-\$500.000, \$500.000-\$700.000, \$700.000-\$1.000.000, Másde\$1.000.000.
CalificaciónGastoPromedio.	Selección múltiple.	MuyBajo, Bajo, Medio, Alto, MuyAlto.
Aspectoque ImpactodelaCiudad.	Selección múltiple.	AmbienteAgradable, ClimaAgradable, CiudadRuidosa, AltaContaminación, GenteAmable, CiudadSegura, MuchaIndigencia, Otro.

Nota: Clasificación e información de los encuestados sobre el motivo de viaje.

Ficha Técnica: Viajeros que visitan la ciudad de Bucaramanga

Tabla 7.

Ficha Técnica sobre la encuesta.

Ficha Técnica de los viajeros que Visitan la ciudad de Bucaramanga	
Números Total de encuestados	400 viajeros
Total de Atributos	38 Atributos
Países tenidos en cuenta para el análisis	Todos los Países
Fecha de realización de la encuesta	20 de Mayo al 03 de Junio de 2017
Origen del Viajero	Nacional e Internacional
Edades	Entre Menos de 18 años hasta Más de 75 Años
Tipo de Viajero	Turista o Visitante
Nivel socioeconómico	Todos son tenidos en cuenta
Nivel de Escolaridad	Todos son tenidos en cuenta

Religión	No son tenidos en cuenta
----------	--------------------------

Nota: Ficha técnica de los viajeros que visitan la ciudad de Bucaramanga.

Exploración y Validación de los datos. Para este proceso se utilizó la herramienta SPSS, permitiendo explorar de manera ágil los datos y realizar el análisis con las múltiples opciones disponibles.

La encuesta se validó con anterioridad para evitar redundancia de datos, datos incompletos o erróneos al ingresar la información. Las medidas tomadas fueron:

- Los datos se validaron para evitar información incompleta, es decir, se puede seleccionar la respuesta adecuada dependiendo de los datos personales e intereses individuales de cada encuestado.
- En la encuesta todos los campos eran obligatorios.
- En la variable Edad va desde Menores de 18 Años hasta Más de 75 Años, como una instancia para quienes eran encuestados y se encontraban en esos rangos de edad.
- En el campo N°VecesViajado a la ciudad se consideró la instancia 0-1Veces para determinar si es su primer viaje realizado a la ciudad, aclarando este punto en el formato a diligenciar.

Preparación de los datos

Selección y Limpieza de datos. Para este proceso se utilizó la herramienta SPSS para la creación de una base de datos, donde se pudieran aplicar ajustes a los datos obtenidos mediante la creación de nuevas tablas para la interpretación de cada objetivo planteado en el proyecto.

Las tablas 11 a la 45 se muestran de manera detallada como fueron discretizados con su respectiva columna de “Campo” y “Valor” para mejorar la interpretación de los datos.

Los datos resultantes se pueden observar en la descripción de las siguientes tablas:

Tabla 8.

Campos discretizados de la Edad.

Edad	Valor
Menosde18Años	1
18-25Años	2
26-35Años	3
36-45Años	4
46-55Años	5
56-65Años	6
66-75Años	7
Másde75Años	8

Nota: Valor asignado para cada atributo de la Edad.

Tabla 9.

Campos discretizados de Género.

Género	Valor
Masculino	1
Femenino	2

Nota: Valor asignado para cada atributo de Género..

Tabla 10.

Campos discretizados de EstadoCívil.

EstadoCívil	Valor
--------------------	--------------

Soltero	1
Casado	2
UniónLibre	3
Viudo	4
Divorciado	5

Nota: Valor asignado para cada atributo de EstadoCívil.

Tabla 11.

Campos discretizados de NivelEscolaridad.

NivelEscolaridad	Valor
Primaria	1
Secundaria	2
Técnico	3
Tecnólogo	4
Pregrado	5
Especialización	6
Maestría	7
Doctorado	8
Ninguno	9

Nota: Valor asignado para cada atributo de NivelEscolaridad.

Tabla 12.

Campos discretizados de Ocupación.

Ocupación	Valor
Estudiante	1
Empleado	2
Independiente	3

Pensionado/Jubilado	4
Desempleado	5

Nota: Valor asignado para cada atributo de Ocupación.

Tabla 13.

Campos discretizados de FuentedeIngreso.

FuentedeIngreso	Valor
Estudiante	1
Empleado	2
Independiente	3
Pensionado/Jubilado	4
Desempleado	5
Familia	6
Otros	7

Nota: Valor asignado para cada atributo de FuentedeIngreso..

Tabla 14.

Campos discretizados de N°VecesViajado.

N°VecesViajado	Valor
0-1Veces	1
1-2Veces	2
3-10Veces	3
11-20Veces	4
Másde20Veces	5

Nota: Valor asignado para cada atributo de FuentedeIngreso.

Tabla 15.

Campos discretizados de MotivoViaje.

MotivoViaje	Valor
--------------------	--------------

Turismo	1
Comercio	2
Aventura	3
Educación	4
Salud	5
Cultura	6
Deportes	7
Otro	8

Nota: Valor asignado para cada atributo de MotivoViaje.

Tabla 16.

Campos discretizados de N°DíasPermanencia.

N°DíasPermanencia	Valor
0-1Día	1
1-2Días	2
3-10Días	3
11-20Días	4
Másde20Días	5

Nota: Valor asignado para cada atributo de N°DíasPermanencia.

Tabla 17.

Campos discretizados de N°CompañeroViaje.

N°CompañeroViaje	Valor
0-1Persona	1
1-2Personas	2
3-10Personas	3
11-2Personas	4

Másde20Personas	5
-----------------	---

Nota: Valor asignado para cada atributo de N°CompañeroViaje..

Tabla 18.

Campos discretizados de LugarAlojamiento.

LugarAlojamiento	Valor
Hotel	1
Familia	2
Amigos	3
Hostal	4
Otro	5

Nota: Valor asignado para cada atributo de LugarAlojamiento.

Tabla 19.

Campos discretizados de ActividadPlaneada.

ActividadPlaneada	Valor
Turismo	1
Comercio	2
Aventura	3
Educación	4
Salud	5
Cultura	6
Deportes	7
Otro	8

Nota: Valor asignado para cada atributo de ActividadPlaneada..

Tabla 20.

Campos discretizados de MedioTransportedeLlegada.

MedioTransportedeLlegada	Valor
Avión	1
TransportePúblico	2
Particular	3
Otro	4

Nota: Valor asignado para cada atributo de MedioTransportedeLlegada..

Tabla 21.

Campos discretizados de MedioTransporteenlaCiudad.

MedioTransporteenlaCiudad	Valor
TransportePúblico	1
Particular	2
Caminar	3
Otro	4

Nota: Valor asignado para cada atributo de MedioTransporteenlaCiudad..

Tabla 22.

Campos discretizados de LugarAlimentación.

LugarAlimentación	Valor
Restaurante	1
ComidaRápidaenRestaurantes	2
ComidaRápidaenlaCalle	3
Familia	4
Amigos	5
Otro	6

Nota: Valor asignado para cada atributo de LugarAlimentación.

Tabla 23.

Campos discretizados de GastoPromedio.

GastoPromedio	Valor
Menosde\$100.000	1
\$100.000-\$150.000	2
\$150.000-\$300.000	3
\$300.000-\$500.000	4
\$500.000-\$700.000	5
\$700.000-\$1.000.000	6
Másde\$1.000.000	7

Nota: Valor asignado para cada atributo de GastoPromedio.

Tabla 24.

Campos discretizados de CalificaciónGastoPromedio.

CalificaciónGastoPromedio	Valor
MuyBajo	1
Bajo	2
Medio	3
Alto	4
MuyAlto	5

Nota: Valor asignado para cada atributo de CalificaciónGastoPromedio.

Tabla 25.

Campos discretizados de AspectoqueImpactodelaCiudad.

AspectoqueImpactodelaCiudad	Valor
AmbienteAgradable	1
ClimaAgradable	2

CiudadRuidosa	3
AltaContaminación	4
GenteAmable	5
CiudadSegura	6
MuchaIndigencia	7
Otro	8

Nota: Valor asignado para cada atributo de AspectoqueImpactodelaCiudad.

Tabla 26.

Campos discretizados de AspectoMotivaaRegresar

AspectoMotivaaRegresar	Valor
AmbienteAgradable	1
GenteAmable	2
Bajocostodevida	3
Parques/SitiosTurísticos	4
Cultura	5
ComidaTípica	6
Todaslasanteriores	7
Ninguno	8
Novolveríanunca	9

Nota: Valor asignado para cada atributo de AspectoMotivaaRegresar.

Tabla 27.

Campos discretizados de GradoSatisfacciónEntretenimiento.

GradoSatisfacciónEntretenimiento	Valor
MásBaja	1
Baja	2

Media	3
Alta	4
MuyAlta	5

Nota: Valor asignado para cada atributo de GradoSatisfacciónEntretenimiento.

Tabla 28.

Campos discretizados de GradoSatisfacciónComercial.

GradoSatisfacciónComercial	Valor
MásBaja	1
Baja	2
Media	3
Alta	4
MuyAlta	5

Nota: Valor asignado para cada atributo de GradoSatisfacciónComercial.

Tabla 29.

Campos discretizados de GradoSatisfacciónSeñalizaciónTurística.

GradoSatisfacciónSeñalizaciónTurística	Valor
MásBaja	1
Baja	2
Media	3
Alta	4
MuyAlta	5

Nota: Valor asignado para cada atributo de GradoSatisfacciónSeñalizaciónTurística.

Tabla 30.

Campos discretizados de GradoSatisfacciónCulturaCiudadana.

GradoSatisfacciónCulturaCiudadana	Valor
MásBaja	1
Baja	2
Media	3
Alta	4
MuyAlta	5

Nota: Valor asignado para cada atributo de GradoSatisfacciónCulturaCiudadana.

Tabla 31.

Campos discretizados de GradoSatisfacciónAlojamiento.

GradoSatisfacciónAlojamiento	Valor
MásBaja	1
Baja	2
Media	3
Alta	4
MuyAlta	5

Nota: Valor asignado para cada atributo de GradoSatisfacciónAlojamiento.

Tabla 32.

Campos discretizados de GradoSatisfacciónCalidaddeAtenciónRecibida.

GradoSatisfacciónCalidaddeAtenciónRecibida	Valor
MásBaja	1
Baja	2
Media	3
Alta	4

MuyAlta	5
---------	---

Nota: Valor asignado para cada atributo de GradoSatisfacciónCalidaddeAtenciónRecibida.

Tabla 33.

Campos discretizados de GradoSatisfacciónCalidadAlojamiento.

GradoSatisfacciónCalidadAlojamiento	Valor
MásBaja	1
Baja	2
Media	3
Alta	4
MuyAlta	5

Nota: Valor asignado para cada atributo de GradoSatisfacciónCalidadAlojamiento.

Tabla 34.

Campos discretizados de GradoSatisfacciónDisponibilidadAlojamiento.

GradoSatisfacciónDisponibilidadAlojamiento	Valor
MásBaja	1
Baja	2
Media	3
Alta	4
MuyAlta	5

Nota: Valor asignado para cada atributo de GradoSatisfacciónDisponibilidadAlojamiento.

Tabla 35.

Campos discretizados de GradoSatisfacciónPrecioAlojamiento.

GradoSatisfacciónPrecioAlojamiento	Valor
MásBaja	1
Baja	2

Media	3
Alta	4
MuyAlta	5

Nota: Valor asignado para cada atributo de GradoSatisfacciónPrecioAlojamiento.

Tabla 36.

Campos discretizados de GradoSatisfacciónOficinaTurismooTransportadoras.

GradoSatisfacciónOficinaTurismooTransportadoras	Valor
MásBaja	1
Baja	2
Media	3
Alta	4
MuyAlta	5

Nota: Valor asignado para cada atributo de GradoSatisfacciónOficinaTurismooTransportadoras.

Tabla 37.

Campos discretizados de GradoSatisfacciónHoteluHospedaje.

GradoSatisfacciónHoteluHospedaje	Valor
MásBaja	1
Baja	2
Media	3
Alta	4
MuyAlta	5

Nota: Valor asignado para cada atributo de GradoSatisfacciónHoteluHospedaje.

Tabla 38.

Campos discretizados de GradoSatisfacciónPersonasenlaCalle.

GradoSatisfacciónPersonasenlaCalle	Valor
MásBaja	1
Baja	2
Media	3
Alta	4
MuyAlta	5

Nota: Valor asignado para cada atributo de GradoSatisfacciónPersonasenlaCalle.

Tabla 39.

Campos discretizados de GradoSatisfacciónPolicia.

GradoSatisfacciónPolicia	Valor
MásBaja	1
Baja	2
Media	3
Alta	4
MuyAlta	5

Nota: Valor asignado para cada atributo de GradoSatisfacciónPolicia.

Tabla 40.

Campos discretizados de GradoSatisfacciónPolicia de Tránsito.

GradoSatisfacciónPolicia de Tránsito	Valor
MásBaja	1
Baja	2
Media	3
Alta	4
MuyAlta	5

Nota: Valor asignado para cada atributo de GradoSatisfacciónPolicia de Tránsito.

Tabla 41.

Campos discretizados de GradoSatisfacciónPolicia Carreteras.

GradoSatisfacciónPoliciaCarreteras	Valor
MásBaja	1
Baja	2
Media	3
Alta	4
MuyAlta	5

Nota: Valor asignado para cada atributo de GradoSatisfacciónPolicia Carreteras.

Tabla 42.

Campos discretizados de GradoSatisfacciónInformaciónTurística,

GradoSatisfacciónInformaciónTurística	Valor
MásBaja	1
Baja	2
Media	3
Alta	4
MuyAlta	5

Nota: Valor asignado para cada atributo de GradoSatisfacciónInformaciónTurística..

Otra tarea muy importante para la limpieza de los datos es la detección de Outliers, los cuales son valores poco frecuentes o valores que se encuentran por fuera de los cuartiles, ya que se encuentran en los extremos. Para este proceso de minería de datos, la detección de Outliers se realiza en la fase de **Modelado**, debido a las características de la investigación se requiere analizar los datos por partes.

Construcción de datos e Integración de datos. En esta etapa se tuvo como objetivo construir un conjunto de datos o vista minable de cada viajero, estructurándolo en un solo registro. Dichos datos se normalizaron para evitar redundancias, y así realizar un análisis más efectivo.

La estructura de los datos minables y él porque de algunas normalizaciones se realizó utilizando todos los datos obtenidos por cada uno de los viajeros en un solo registro, los cuales se normalizaron en la etapa de “*Selección de los datos y Limpieza de datos*”. Estos datos fueron relacionados de la siguiente manera:

- Información de identificación del viajero: Edad, Género, CiudadOrigen, PaísOrigen, EstadoCivil, NivelEscolaridad.
- Información financiera del viajero: Ocupación, FuenteIngreso.
- Información sobre el viaje: N°VecesViajado, MotivoViaje, N°DíasPermanencia, N°CompañeroViaje, LugarAlojamiento, ActividadPlaneada.
- Información sobre los medios de transporte utilizados en el viaje: MedioTransportedeLlegada, MedioTransporteenlaCiudad.
- Información sobre la alimentación de la ciudad: LugarAlimentación.
- Información sobre el gasto: GastoPromedio, CalificaciónGastoPromedio.
- Información sobre el impacto al llegar a la ciudad: AspectoqueImpactodelaCiudad, AspectoMotivaaRegresar.
- Información sobre el grado de satisfacción: GradoSatisfacciónEntretenimiento, GradoSatisfacciónComercial, GradoSatisfacciónCultural, GradoSatisfacciónSeñalizaciónTurística, GradoSatisfacciónCulturaCiudadana, GradoSatisfacciónAlojamiento, GradoSatisfacciónCalidaddeAtenciónRecibida,

GradoSatisfacciónCalidadAlojamiento, GradoSatisfacciónDisponibilidadAlojamiento,
GradoSatisfacciónPrecioAlojamiento,
GradoSatisfacciónOficinaTurismoTransportadoras, GradoSatisfacciónHoteluHospedaje,
GradoSatisfacciónPersonasenlaCalle, GradoSatisfacciónPolicia,
GradoSatisfacciónPolicia deTránsito, GradoSatisfacciónPoliciaCarreteras,
GradoSatisfacciónInformaciónTurística.

Modelado

Seleccionar técnicas de modelado. Es la primera técnica seleccionada para dar respuesta a nuestro primer y tercer objetivo, el cual es segmentar y describir los hábitos de los viajeros que visitan la ciudad de Bucaramanga y segmentar de acuerdo con el nivel de satisfacción, teniendo en cuenta las características de los datos. Por este motivo es necesario realizar un análisis de Clustering.

La técnica de Clustering K-Means se seleccionó sobre otras ya que es un método no supervisado, lo cual indica que no es necesario utilizar DataSet de entrenamiento sino que el algoritmo identifica grupos de acuerdo a las características deseadas y posterior a esto los agrupa.

Con este proceso de modelado se realizaron ciertas actividades tales como:

- Proceso inicial de análisis exploratorio de datos, donde se realizan modelos matemáticos y esta información se analiza de forma gráfica, determinando distribución de variables y datos, datos estadísticos, detección de Outliers.
- Se ejecutó el modelo de Clustering para obtener grupos de datos de los viajeros.

- Aplicamos reglas de asociación para cada Clúster con el fin de encontrar y describir hechos que ocurren con mayor frecuencia dentro del conjunto de datos, y establecer relaciones entre las variables.

La segunda técnica utilizada para determinar nuestro segundo objetivo, el cual es describir las relaciones y tendencias entre las características de los viajeros y los aspectos de su visita. Para esto es necesario fue necesario el uso de las reglas de asociación.

Las reglas de asociación predicen cualquier atributo, y las posibles combinaciones entre ellos. Esta técnica permite determinar reglas y manejar datos de validación como la cobertura, la cual es el número de instancias por los cuales se predice correctamente la regla y aportan un nivel de confiabilidad alto y asertivo.

Se aplican las técnicas de Clustering y reglas de asociación a los grupos de datos que se fueron filtrando durante el análisis, y de ese modo se buscaron las combinaciones que nos proporcionaron información valiosa de los viajeros.

Generar el diseño de prueba. Para validar la calidad del modelo y de los datos analizados, la calidad de los clúster generados, la medida adecuada de clúster de acuerdo con su distribución, a calidad de los datos normalizados, la calidad de las reglas de asociación y su nivel de confianza, utilizamos varias medidas.

El diseño inicial tomó los datos del archivo de WEKA luego de la limpieza de la etapa de “Preparación de los datos” del modelo CRISP-DM. Se seleccionó el algoritmo SimpleKMeans de Clustering, donde se configuró tomando un valor de numClusters de 3, siendo este valor numClusters=3 un número óptimo para validar el modelado y permitir su aplicación a los 400 viajeros encuestados.

```

Number of iterations: 8
Within cluster sum of squared errors: 7165.0
Missing values globally replaced with mean/mode

Cluster centroids:

```

Attribute	Full Data (400)	Cluster#		
		0 (188)	1 (125)	2 (87)
Edad	2	3	2	6
Género	2	2	2	2
CiudadOrigen	Barrancabermeja	Bogotá	Barrancabermeja	Barrancabermeja
PaísOrigen	Colombia	Colombia	Colombia	Colombia
EstadoCivil	1	3	1	2
NivelEscolaridad	2	5	2	2
Ocupación	1	2	1	4
Fuente de Ingreso	2	2	6	4
N°VecesViajado	3	3	3	3
MotivoViaje	4	2	4	5
N°DíasPermanencia	2	2	5	2
N°CompañeroViaje	1	1	1	2
LugarAlojamiento	1	1	5	1
ActividadPlaneada	4	2	4	5
MedioTransporte de Llegada	2	2	2	2
MedioTransporte en la Ciudad	1	1	1	1
LugarAlimentación	1	1	6	1
GastoPromedio	5	5	4	5
CalificaciónGastoPromedio	3	3	3	3
Aspecto que Impacta de la Ciudad	2	2	2	3
Aspecto Motivado a Regresar	7	7	3	7
Grado Satisfacción Entrenimiento	4	4	3	4
Grado Satisfacción Comercial	4	4	3	3
Grado Satisfacción Cultural	3	3	3	3
Grado Satisfacción Señalización Turística	3	3	3	3
Grado Satisfacción Cultura Ciudadana	3	3	2	2
Grado Satisfacción Alojamiento	4	4	4	4
Grado Satisfacción Calidad de Atención Recibida	4	4	4	4
Grado Satisfacción Calidad Alojamiento	4	5	4	4
Grado Satisfacción Disponibilidad Alojamiento	4	4	4	4
Grado Satisfacción Precio Alojamiento	4	4	4	4
Grado Satisfacción Oficina Turismo o Transportadoras	3	4	4	3
Grado Satisfacción Hotel u Hospedaje	3	3	3	2
Grado Satisfacción Personas en la Calle	3	3	3	2
Grado Satisfacción Policía	2	2	2	3
Grado Satisfacción Policía de Tránsito	2	2	1	2
Grado Satisfacción Policía Carreteras	2	2	2	2
Grado Satisfacción Información Turística	3	3	3	2

Time taken to build model (full training data) : 0.36 seconds

=== Model and evaluation on training set ===

Clustered Instances

```

0      188 ( 47%)
1      125 ( 31%)
2       87 ( 22%)

```

Φιγυρα 4. Προχεςο Οβτενιδο δεσδε ελ σοφτωαρε ΩΕΚΑ. Αδαπααδο δε: Προχεςο Οβτενιδο δεσδε ελ σοφτωαρε ΩΕΚΑ.

Al seleccionar como dato óptimo nmuClusters=3, se obtuvieron tres clúster, donde 188 de los encuestados, que representan el 47%, su ubican en el Clúster 0; 125 de los encuestados (31%) pertenecen al Clúster 1 y 87 de los viajeros encuestados (22%) al Clúster 2.

Para la obtención de este resultado, se utilizó el método “Simple K-Means”, y posterior a esto se visualizó las asignaciones de cada clúster de la siguiente manera:

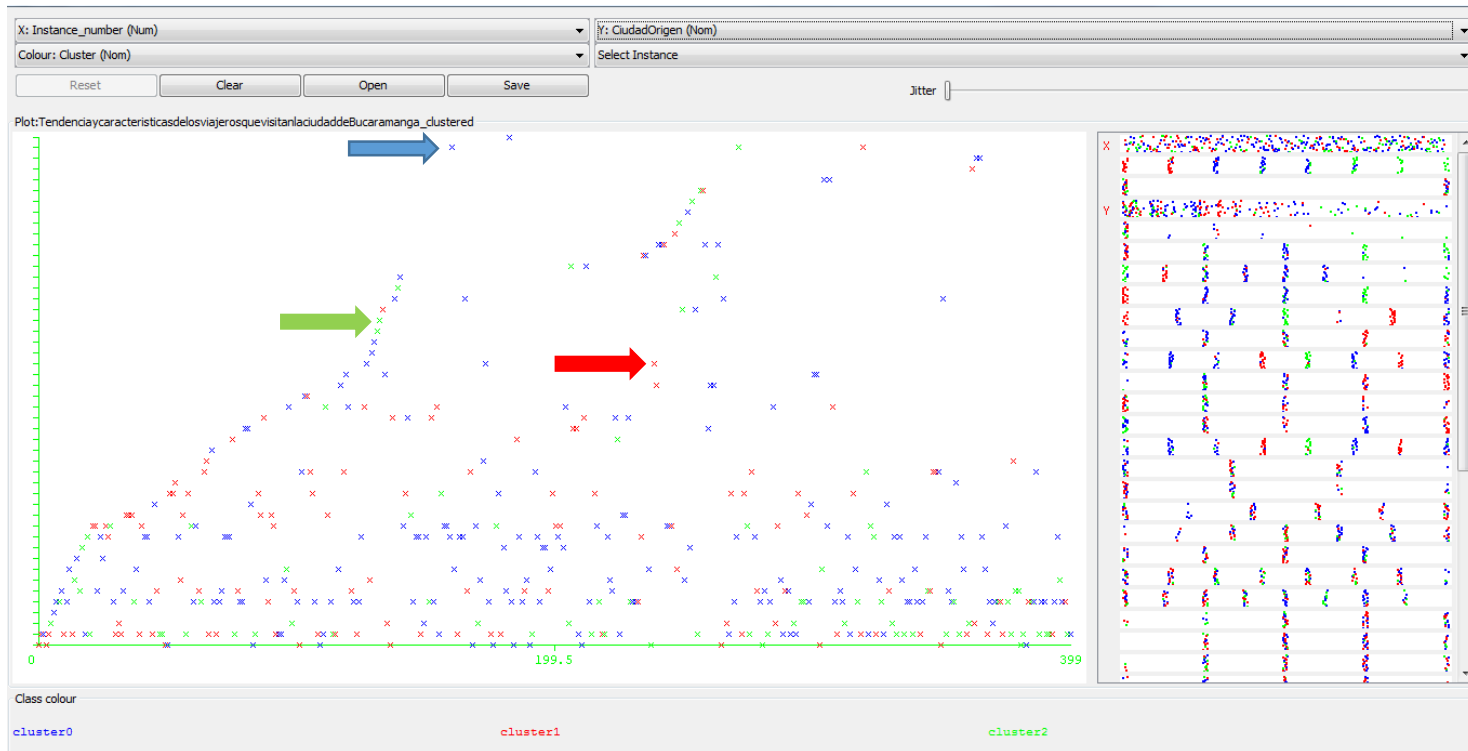
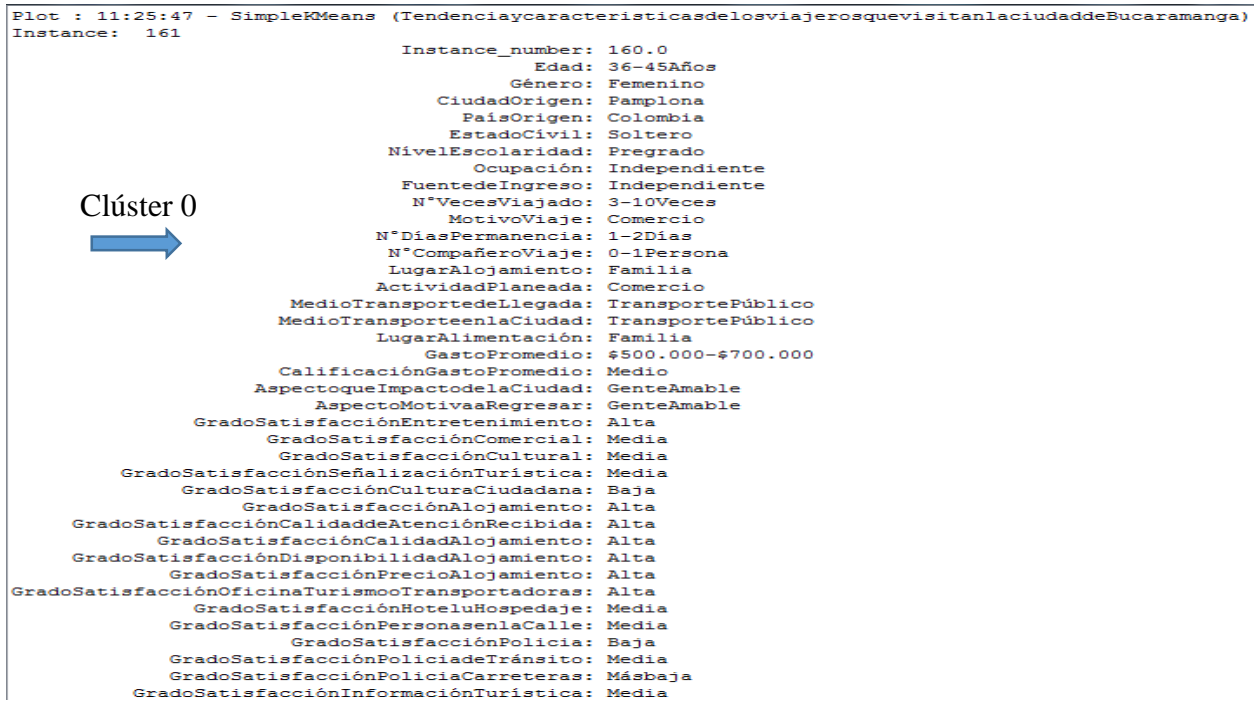


Figura 5. Asignaciones de cada clúster utilizando el software WEKA. Adaptado de: Proxenos Oβτενιδο δεσδε el software WEKA .

Mediante el siguiente cuadro se puede visualizar de manera más detallada cada uno de los datos analizados, es decir, los utilizados para este ejercicio (los que están señalados por las flechas):



Φιγυρα 6. Ασιγναχιονεσ δελ χλ /στερ 0 υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ.. Αδαπταδο δε: Προχεσο

Οβτενιδο δεσδε ελ σοφτωαρε ΩΕΚΑ .

Plot : 11:25:47 - SimpleKMeans (Tendencia y características de los viajeros que visitan la ciudad de Bucaramanga)
 Instance: 239

Clúster 1



Instance_number: 238.0
 Edad: 18-25Años
 Género: Masculino
 CiudadOrigen: Villavicencio
 PaísOrigen: Colombia
 EstadoCivil: Soltero
 NivelEscolaridad: Tecnólogo
 Ocupación: Desempleado
 Fuente de Ingreso: Familia
 N°VecesViajado: 1-2Veces
 MotivoViaje: Turismo
 N°DíasPermanencia: 11-20Días
 N°CompañeroViaje: 1-2Personas
 LugarAlojamiento: Hotel
 ActividadPlaneada: Turismo
 MedioTransporte de Llegada: Avión
 MedioTransporte en la Ciudad: Transporte Público
 LugarAlimentación: Restaurante
 Gasto Promedio: Más de \$1.000.000
 Calificación Gasto Promedio: Muy Bajo
 Aspecto que Impacta de la Ciudad: Ambiente Agradable
 Aspecto Motivado a Regresar: Parques/Sitios Turísticos
 Grado Satisfacción Entretenimiento: Muy Alta
 Grado Satisfacción Comercial: Baja
 Grado Satisfacción Cultural: Alta
 Grado Satisfacción Señalización Turística: Media
 Grado Satisfacción Cultura Ciudadana: Baja
 Grado Satisfacción Alojamiento: Más baja
 Grado Satisfacción Calidad de Atención Recibida: Baja
 Grado Satisfacción Calidad Alojamiento: Alta
 Grado Satisfacción Disponibilidad Alojamiento: Muy Alta
 Grado Satisfacción Precio Alojamiento: Muy Alta
 Grado Satisfacción Oficina Turismo o Transportadoras: Alta
 Grado Satisfacción Hotel u Hospedaje: Media
 Grado Satisfacción Personas en la Calle: Media
 Grado Satisfacción Policía: Baja
 Grado Satisfacción Policía de Tránsito: Más baja
 Grado Satisfacción Policía Carreteras: Muy Alta
 Grado Satisfacción Información Turística: Alta

Φιγυρα 7. Ασιγναχιονες δελ χλ /στερ 1 υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. Αδαπταδο δε: Π ροχεσο Οβτενιδο δεσδε ελ σοφτωαρε ΩΕΚΑ.

Plot : 11:25:47 - SimpleKMeans (TendenciaycaracteristicasdelosviajerosquevisitanlaciudaddeBucaramanga)
Instance: 133

Clúster 2



```

Instance_number: 132.0
  Edad: 66-75Años
  Género: Femenino
  CiudadOrigen: Quito
  PaísOrigen: Ecuador
  EstadoCívil: Viudo
  NivelEscolaridad: Especialización
  Ocupación: Pensionado/Jubilado
  Fuente de Ingreso: Pensionado/Jubilado
  N°VecesViajado: 0-1Veces
  MotivoViaje: Turismo
  N°DíasPermanencia: 3-10Días
  N°CompañeroViaje: 1-2Personas
  LugarAlojamiento: Hotel
  ActividadPlaneada: Turismo
  MedioTransportedeLlegada: Avión
  MedioTransporteenlaCiudad: TransportePúblico
  LugarAlimentación: Restaurante
  GastoPromedio: $700.000-$1.000.000
  CalificaciónGastoPromedio: Medio
  AspectoqueImpactodelaCiudad: CiudadRuidosa
  AspectoMotivaaRegresar: Todaslasanteriores
  GradoSatisfacciónEntretenimiento: MuyAlta
  GradoSatisfacciónComercial: MuyAlta
  GradoSatisfacciónCultural: Alta
  GradoSatisfacciónSeñalizaciónTurística: Alta
  GradoSatisfacciónCulturaCiudadana: Alta
  GradoSatisfacciónAlojamiento: MuyAlta
  GradoSatisfacciónCalidaddeAtenciónRecibida: Alta
  GradoSatisfacciónCalidadAlojamiento: Alta
  GradoSatisfacciónDisponibilidadAlojamiento: MuyAlta
  GradoSatisfacciónPrecioAlojamiento: Alta
  GradoSatisfacciónOficinaTurismooTransportadoras: Media
  GradoSatisfacciónHoteluHospedaje: Media
  GradoSatisfacciónPersonasenlaCalle: Media
  GradoSatisfacciónPolicia: Baja
  GradoSatisfacciónPolicia de Tránsito: Baja
  GradoSatisfacciónPoliciaCarreteras: Másbaja
  GradoSatisfacciónInformaciónTurística: Media
    
```

Figura 8. Ασιγναχιονες δελ χλ /στερ 2 υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. Αδαπταδο δε: Προχεσο Οβτενιδο δεσδε ελ σοφτωαρε ΩΕΚΑ .

Los aspectos a evaluar se realizaron con la medidas de validación internas específicas para los clúster, teniendo en cuenta que un buen clúster es donde todos los puntos están cercaos los unos de los otros y cada clúster debe estar separados entre sí.

Las medidas internas de validación de clustering que se utilizaron:

- AverageWithi centroid distance: Considera cada clúster y calcula la distancia promedio de cada punto en el clúster al centroide.
- Distribution Gini: Mide la diferencia entre valores de una frecuencia de distribución.
- Performance Silhouette: Este es el índice que nos da el valor de numClusters con el mejor rendimiento para nuestros clústers, pues es un método de interpretación y validación de datos en los clústers.

Para la validación de las reglas de asociación se debe tener en cuenta los siguientes aspectos:

- Las reglas de asociación descubren y describen hechos de los datos como las relaciones entre un conjunto de datos y de esta manera encontrar las reglas que predicen la ocurrencia de cada dato.
- Las medidas reales para determinar una buena regla de asociación son las de Soporte y Confianza, resultado de aplicar el algoritmo A priori.
- Aplicamos el algoritmo A priori a nuestra base de datos para obtener las relaciones y validar el soporte y la confianza, como se muestra a continuación

Apriori

=====

Minimum support: 0.45 (180 instances)

Minimum metric <confidence>: 0.9

Number of cycles performed: 11

Generated sets of large itemsets:

Size of set of large itemsets L(1): 20

Size of set of large itemsets L(2): 20

Size of set of large itemsets L(3): 3

Best rules found:

1. MedioTransportedeLlegada=2 CalificaciónGastoPromedio=3 196 ==> PaísOrigen=Colombia 196 conf:(1)
2. N°CompañeroViaje=1 232 ==> PaísOrigen=Colombia 231 conf:(1)
3. MedioTransportedeLlegada=2 MedioTransporteenlaCiudad=1 217 ==> PaísOrigen=Colombia 214 conf:(0.99)
4. MedioTransportedeLlegada=2 285 ==> PaísOrigen=Colombia 281 conf:(0.99)
5. N°VecesViajado=3 202 ==> PaísOrigen=Colombia 199 conf:(0.99)
6. CalificaciónGastoPromedio=3 263 ==> PaísOrigen=Colombia 258 conf:(0.98)
7. GradoSatisfacciónSeñalizaciónTurística=3 235 ==> PaísOrigen=Colombia 230 conf:(0.98)
8. GradoSatisfacciónCultural=3 227 ==> PaísOrigen=Colombia 222 conf:(0.98)
9. GradoSatisfacciónAlojamiento=4 219 ==> PaísOrigen=Colombia 214 conf:(0.98)
10. GradoSatisfacciónDisponibilidadAlojamiento=4 195 ==> PaísOrigen=Colombia 190 conf:(0.97)
11. GastoPromedio=5 193 ==> PaísOrigen=Colombia 188 conf:(0.97)
12. MedioTransporteenlaCiudad=1 CalificaciónGastoPromedio=3 193 ==> PaísOrigen=Colombia 188 conf:(0.97)
13. GradoSatisfacciónPrecioAlojamiento=4 210 ==> PaísOrigen=Colombia 204 conf:(0.97)
14. GradoSatisfacciónEntretenimiento=4 207 ==> PaísOrigen=Colombia 201 conf:(0.97)
15. GradoSatisfacciónCalidaddeAtenciónRecibida=4 217 ==> PaísOrigen=Colombia 210 conf:(0.97)
16. Género=1 188 ==> PaísOrigen=Colombia 181 conf:(0.96)
17. Género=2 212 ==> PaísOrigen=Colombia 204 conf:(0.96)
18. GradoSatisfacciónComercial=4 199 ==> PaísOrigen=Colombia 190 conf:(0.95)
19. GradoSatisfacciónPersonasenlaCalle=3 194 ==> PaísOrigen=Colombia 185 conf:(0.95)
20. MedioTransporteenlaCiudad=1 292 ==> PaísOrigen=Colombia 278 conf:(0.95)

Φιγυρα 9. Προχεςο παρα γενεραρ ρεγλασ δε ασοχιαχι $\int \nu$ νυμΡυλεσ=20. . Αδαπταδο δε: Προχεςο ο Οβτενιδο δεσδε ελ σοφτωαρε ΩΕΚΑ .

1. MedioTransportedeLlegada=2 CalificaciónGastoPromedio=3 196 ==> PaísOrigen=Colombia 196 conf:(1)
2. N°CompañeroViaje=1 MedioTransportedeLlegada=2 179 ==> PaísOrigen=Colombia 179 conf:(1)
3. CalificaciónGastoPromedio=3 GradoSatisfacciónCultural=3 168 ==> PaísOrigen=Colombia 168 conf:(1)
4. MedioTransportedeLlegada=2 GradoSatisfacciónCultural=3 165 ==> PaísOrigen=Colombia 165 conf:(1)
5. MedioTransportedeLlegada=2 GradoSatisfacciónSeñalizaciónTurística=3 165 ==> PaísOrigen=Colombia 165 conf:(1)
6. N°CompañeroViaje=1 CalificaciónGastoPromedio=3 163 ==> PaísOrigen=Colombia 163 conf:(1)
7. N°CompañeroViaje=1 232 ==> PaísOrigen=Colombia 231 conf:(1)
8. N°CompañeroViaje=1 MedioTransporteenlaCiudad=1 169 ==> PaísOrigen=Colombia 168 conf:(0.99)
9. N°DíasPermanencia=2 168 ==> PaísOrigen=Colombia 167 conf:(0.99)
10. CalificaciónGastoPromedio=3 GradoSatisfacciónSeñalizaciónTurística=3 167 ==> PaísOrigen=Colombia 165 conf:(0.99)
11. MedioTransportedeLlegada=2 MedioTransporteenlaCiudad=1 217 ==> PaísOrigen=Colombia 214 conf:(0.99)
12. MedioTransportedeLlegada=2 285 ==> PaísOrigen=Colombia 281 conf:(0.99)
13. N°VecesViajado=3 202 ==> PaísOrigen=Colombia 199 conf:(0.99)
14. CalificaciónGastoPromedio=3 263 ==> PaísOrigen=Colombia 258 conf:(0.98)
15. GradoSatisfacciónSeñalizaciónTurística=3 235 ==> PaísOrigen=Colombia 230 conf:(0.98)
16. GradoSatisfacciónCultural=3 227 ==> PaísOrigen=Colombia 222 conf:(0.98)
17. GradoSatisfacciónAlojamiento=4 219 ==> PaísOrigen=Colombia 214 conf:(0.98)
18. GradoSatisfacciónDisponibilidadAlojamiento=4 195 ==> PaísOrigen=Colombia 190 conf:(0.97)
19. GastoPromedio=5 193 ==> PaísOrigen=Colombia 188 conf:(0.97)
20. MedioTransporteenlaCiudad=1 CalificaciónGastoPromedio=3 193 ==> PaísOrigen=Colombia 188 conf:(0.97)
21. MedioTransporteenlaCiudad=1 GradoSatisfacciónSeñalizaciónTurística=3 179 ==> PaísOrigen=Colombia 174 conf:(0.97)
22. GradoSatisfacciónPrecioAlojamiento=4 210 ==> PaísOrigen=Colombia 204 conf:(0.97)
23. GradoSatisfacciónEntretenimiento=4 207 ==> PaísOrigen=Colombia 201 conf:(0.97)
24. GradoSatisfacciónHoteluHospedaje=3 171 ==> PaísOrigen=Colombia 166 conf:(0.97)
25. MedioTransporteenlaCiudad=1 GradoSatisfacciónCultural=3 167 ==> PaísOrigen=Colombia 162 conf:(0.97)
26. GradoSatisfacciónCulturaCiudadana=3 165 ==> PaísOrigen=Colombia 160 conf:(0.97)
27. MedioTransporteenlaCiudad=1 GradoSatisfacciónAlojamiento=4 165 ==> PaísOrigen=Colombia 160 conf:(0.97)
28. GradoSatisfacciónCalidaddeAtenciónRecibida=4 217 ==> PaísOrigen=Colombia 210 conf:(0.97)
29. GradoSatisfacciónPolicia=2 170 ==> PaísOrigen=Colombia 164 conf:(0.96)
30. Género=1 188 ==> PaísOrigen=Colombia 181 conf:(0.96)

Φιγυρα 10. Προχεσο παρα γενεραρ ρεγλασ δε ασοχιαχι $\int \nu \mu \rho \upsilon \lambda \epsilon \sigma = 30$. Αδαπταδο δε: Προχεσο Οβτενιδο δεσδε ελ σοφτωαρε ΩΕΚΑ .

Construcción del modelo

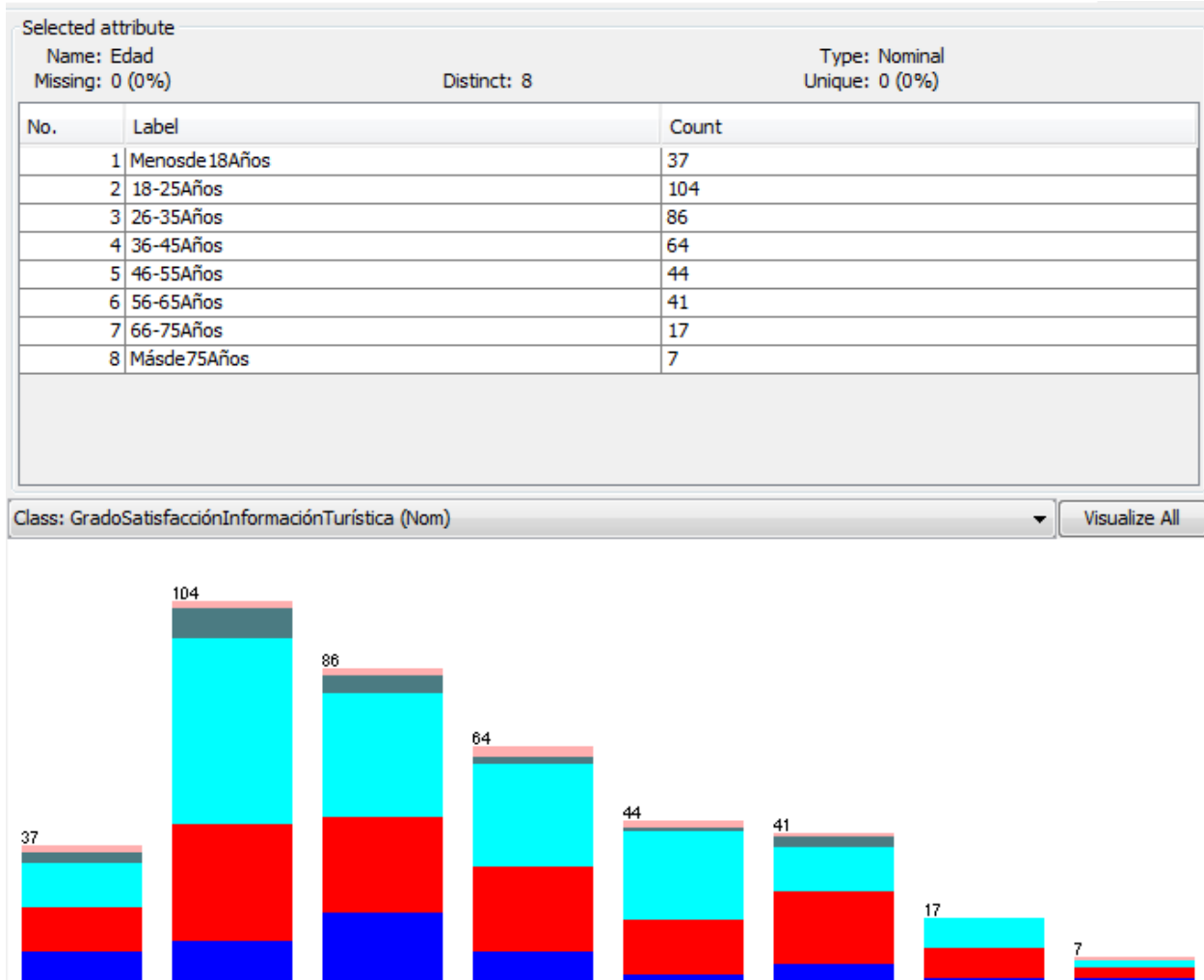
Desarrollo del Objetivo 1: Segmentar y describir hábitos de los viajeros que visitan la ciudad de Bucaramanga por medio de un Análisis de Clúster para la industria turística de Bucaramanga.

Para la aplicación de este modelo se utilizó:

- Herramienta WEKA con la información de los viajeros.

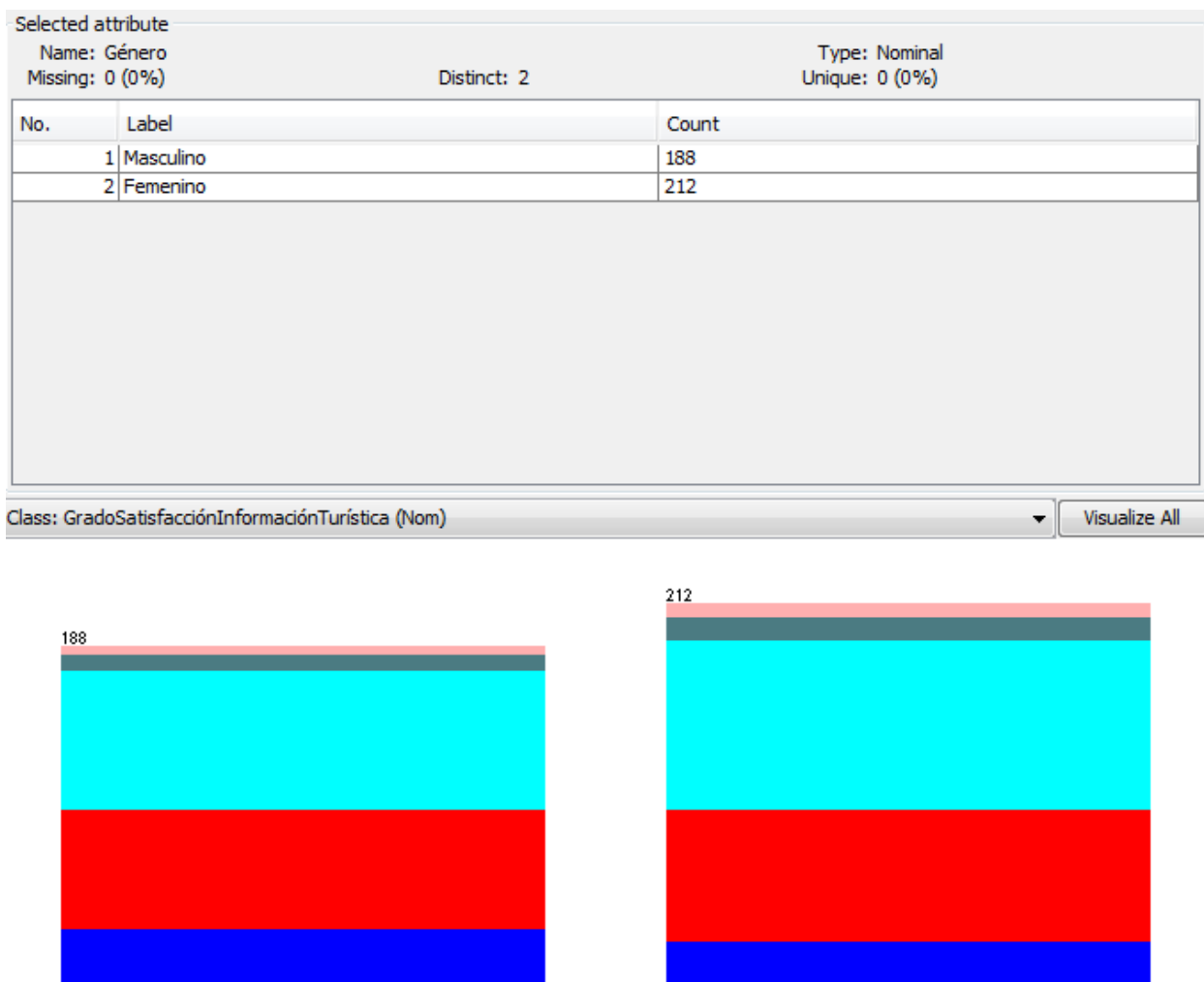
Proceso inicial de análisis exploratorio de datos (EDA): El primer proceso que realizamos fue determinar gráficamente como están distribuidos los datos, datos totales por categoría, variables representativas. Posteriormente determinamos como se agrupan los datos, es decir, para realizar el primer análisis de Clustering.

De los 400 viajeros encuestados, 104 de estos se encuentran entre los 18-25 años, representado con un 26% de la población total, pues entre este rango de edades encontramos gran parte de estudiantes que encuentran en la ciudad una oferta educativa accesible.



Φιγυρα 11. Ατριβυτο Εδαδ δε λοσ ενχυεσταδο. Αδαπταδο δε: Προχεσο Οβτενιδο πορ ελ σοφτωα ρε ΩΕΚΑ-

Se observa que el porcentaje más alto corresponde a la población femenina representada con un 53% de los encuestados; seguida de la población masculina representada con un 47% del total.



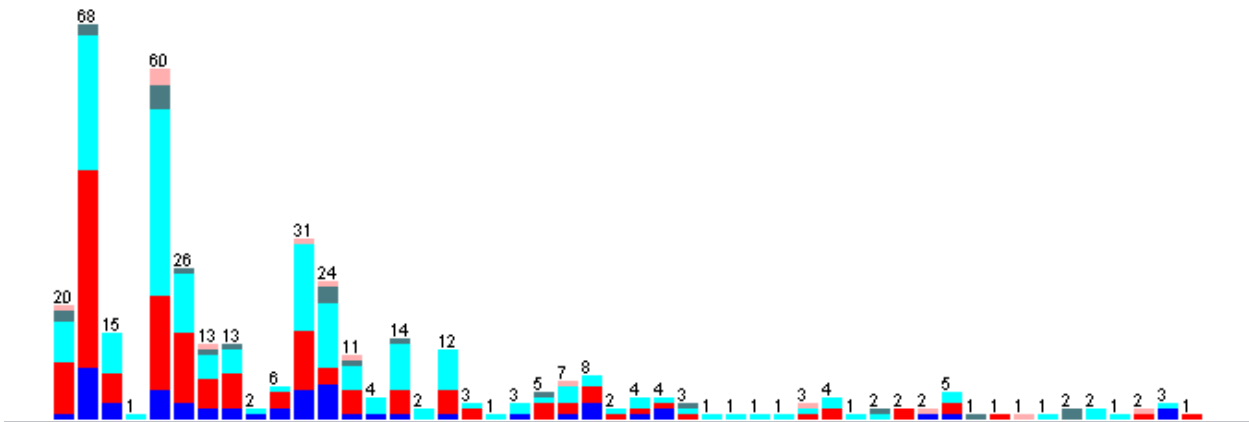
Φιγυρα 12. Ατριβυτο ΓΓνερο δε λοσ ενχνεσταδοσ. . Αδαπταδο δε: Προχεσο Οβτενιδο πορ ελ σοφ τωαρε ΩΕΚΑ.

Barrancabermeja es la Ciudad de Origen más representativa dentro del análisis realizado a los viajeros son Barrancabermeja con un 17%, frente a ciudades como Bogotá con un 15% y Medellín con un 7,75%.

Selected attribute
 Name: CiudadOrigen
 Missing: 0 (0%)
 Distinct: 48
 Type: Nominal
 Unique: 13 (3%)

No.	Label	Count
1	Cartagena	20
2	Barrancabermeja	68
3	Socorro	15
4	Arenal	1
5	Bogotá	60
6	Sangil	26
7	Calí	13
8	Barranquilla	13
9	NuevaYork	2
10	Pereira	6
11	Medellín	31
12	Cúcuta	24

Class: GradoSatisfacciónInformaciónTurística (Nom) Visualize All



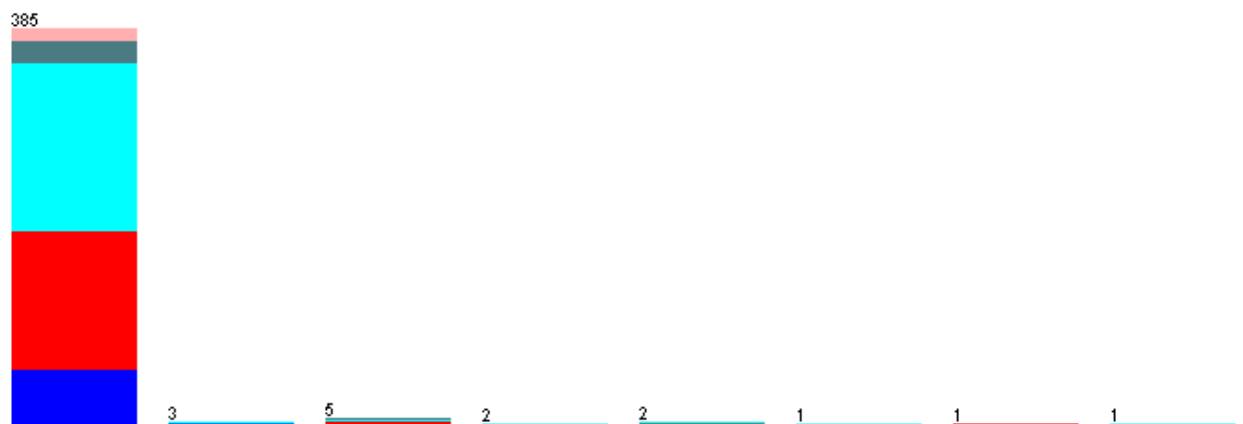
Φιγυρα 13. Ατριθυτο ΧιυδαδΟριγεν δε λοσ ενχυεσταδοσ. . Αδαπταδο δε: Προχυεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

El 96,25% de los viajeros encuestados son colombianos, siendo la población más representativa para el análisis del proyecto.

Selected attribute
 Name: PaísOrigen
 Missing: 0 (0%)
 Distinct: 8
 Type: Nominal
 Unique: 3 (1%)

No.	Label	Count
1	Colombia	385
2	EstadosUnidos	3
3	México	5
4	Argentina	2
5	Francia	2
6	Ecuador	1
7	Venezuela	1
8	España	1

Class: GradoSatisfacciónInformaciónTurística (Nom) Visualize All



Φιγυρα 14. Ατριθυτο ΠαίσΟριγεν δε λοσ ενχυεσταδοσ . Αδαπαδο δε: Προχυεσο Οβτενιδο πο ρ ελ σοφτωαρε ΩΕΚΑ.

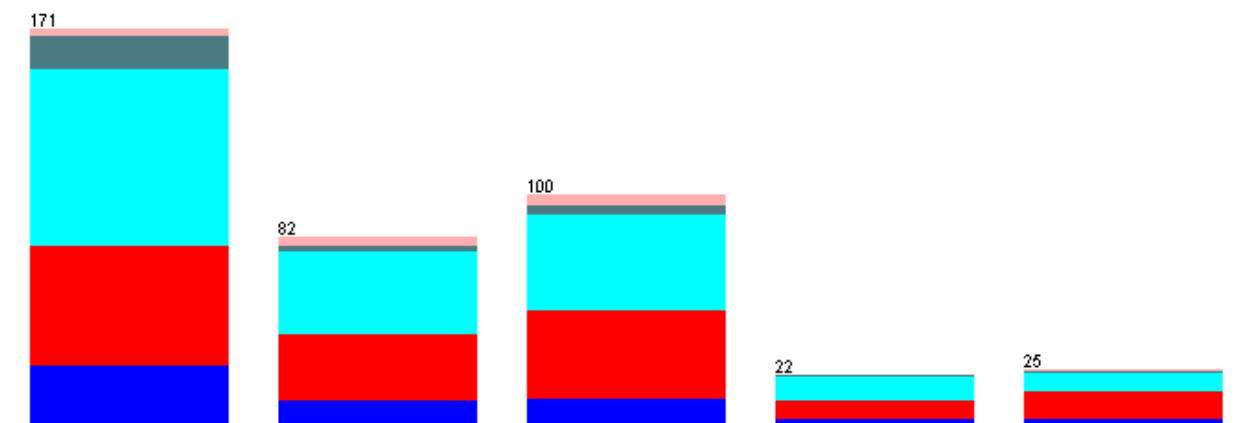
La población más representativa es la de viajeros cuyo estado civil es soltero con un 42,75%, pues parte de quienes se encuentran en este grupo son personas que oscilan entre Menos de 18 Años hasta los 25 Años, y sus actividades y planes no permiten consolidar su situación sentimental.

Selected attribute

Name: EstadoCívil Type: Nominal
 Missing: 0 (0%) Distinct: 5 Unique: 0 (0%)

No.	Label	Count
1	Soltero	171
2	Casado	82
3	UniónLibre	100
4	Viudo	22
5	Divorciado	25

Class: GradoSatisfacciónInformaciónTurística (Nom) Visualize All



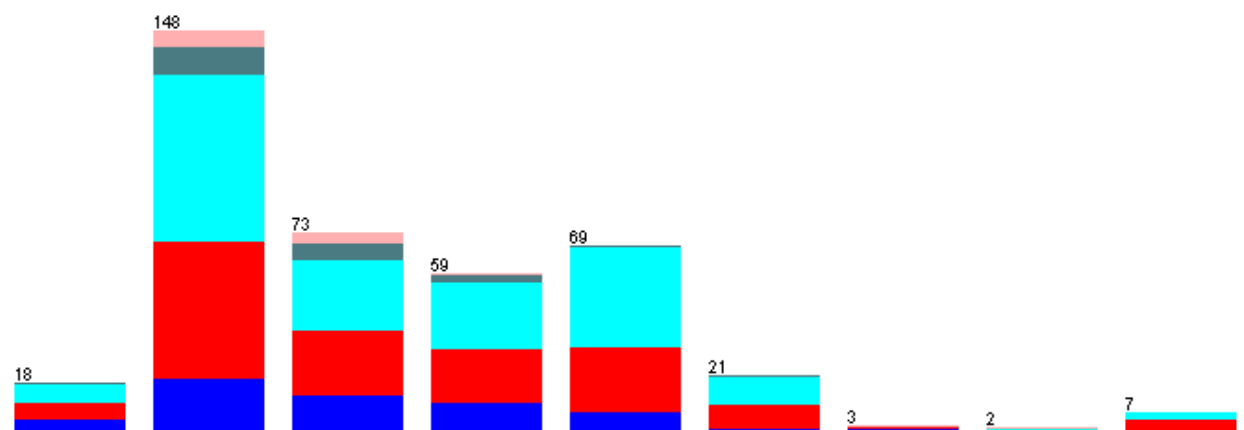
Φιγυρα 15. Ατρίβυτο ΕσταδοΧ΄πιλ δε λοσ ενχνεσταδοσ. – Αδαπταδο δε: Προχεςο Οβτενιδο πορ ελ σ οφτωαρε ΩΕΚΑ.

De los 400 viajeros encuestados, el porcentaje más representativo de la población analizada en la variable de Nivel de Escolaridad, se encuentran quienes aún están en la Secundaria con un 37% y visitan la ciudad en busca de ofertas educativas para inicio de su etapa universitaria; seguida de la población Técnica representada con un 18,25%.

Selected attribute
 Name: NivelEscolaridad
 Missing: 0 (0%)
 Distinct: 9
 Type: Nominal
 Unique: 0 (0%)

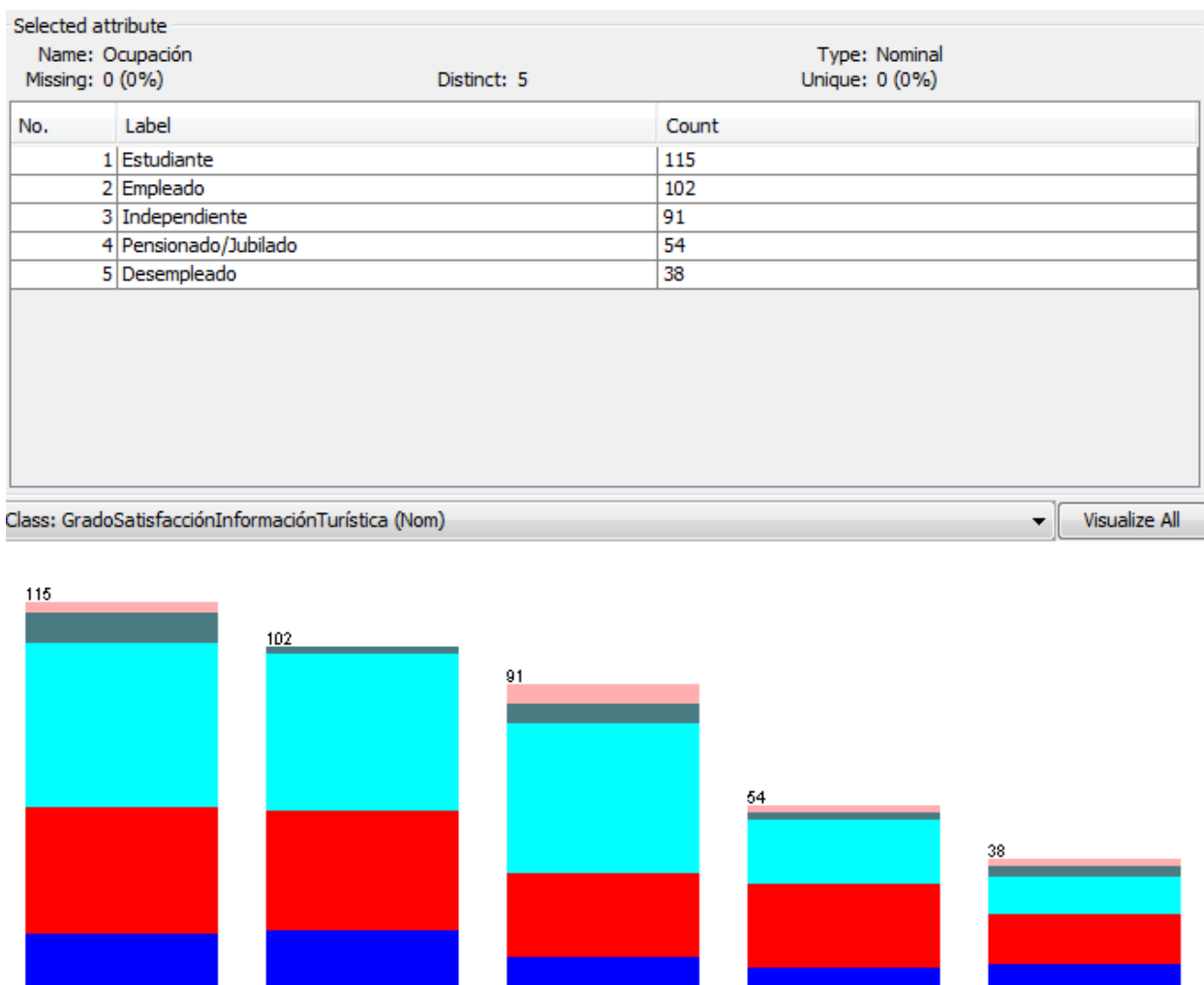
No.	Label	Count
1	Primaria	18
2	Secundaria	148
3	Técnico	73
4	Tecnólogo	59
5	Pregrado	69
6	Especialización	21
7	Maestría	3
8	Doctorado	2
9	Ninguno	7

Class: GradoSatisfacciónInformaciónTurística (Nom) Visualize All



Φιγυρα 16. Ατριβυτο Νιπελεσχολαριδαδ δε λοσ ενχυεσταδοσ.. Αδαπταδο δε: Προχεσο Οβτενιδ ο πορ ελ σοφτωαρε ΩΕΚΑ.

Analizando la variable Ocupación encontramos que la población más ruidosa fue la de Estudiantes representada con un 28,75% al ser parte importante para realizar este análisis puesto que visitan la ciudad en busca de intereses educativos ofertados por esta.



Φιγυρα 17. Ατριβυτο Οχυπαχι ίν δε λοσ ενχυεσταδοσ. . Αδαπαδο δε: Προχεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

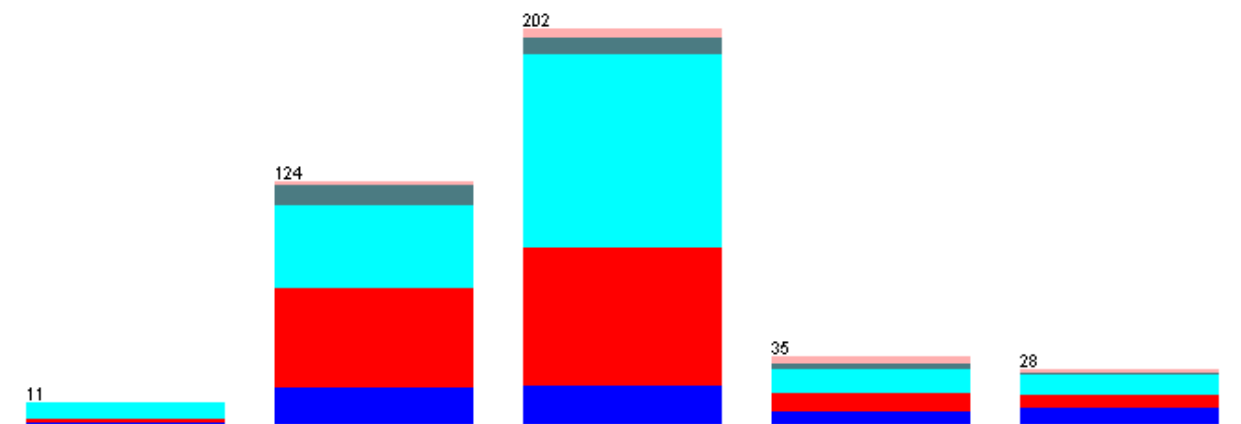
La mayoría de los viajeros visitan la ciudad entre unas 3-10 Veces, representando un 50.5%, al igual de quienes viajan entre 1-2 Veces, siendo un 31% del total analizado.

Selected attribute

Name: N°VecesViajado Type: Nominal
 Missing: 0 (0%) Distinct: 5 Unique: 0 (0%)

No.	Label	Count
1	0-1Veces	11
2	1-2Veces	124
3	3-10Veces	202
4	11-20Veces	35
5	Másde20Veces	28

Class: GradoSatisfacciónInformaciónTurística (Nom) Visualize All



Φιγυρα 18.Ατριβυτο Ν°ζεχεσσιαφαδολ δε λοσ ενχυεσταδοσ . Αδαπαδο δε: Προχεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

Durante el análisis de las respuestas de los viajeros, se identifica que el Motivo de Viaje por el cual arriban a la ciudad se debe a la Educación (25.5%), siendo Bucaramanga una de las ciudades con más oferta educativa en el país.

Selected attribute
 Name: MotivoViaje
 Missing: 0 (0%)
 Distinct: 8
 Type: Nominal
 Unique: 0 (0%)

No.	Label	Count
1	Turismo	16
2	Comercio	87
3	Aventura	21
4	Educación	102
5	Salud	62
6	Cultura	40
7	Deportes	20
8	Otro	52

Class: GradoSatisfacciónInformaciónTurística (Nom) Visualize All

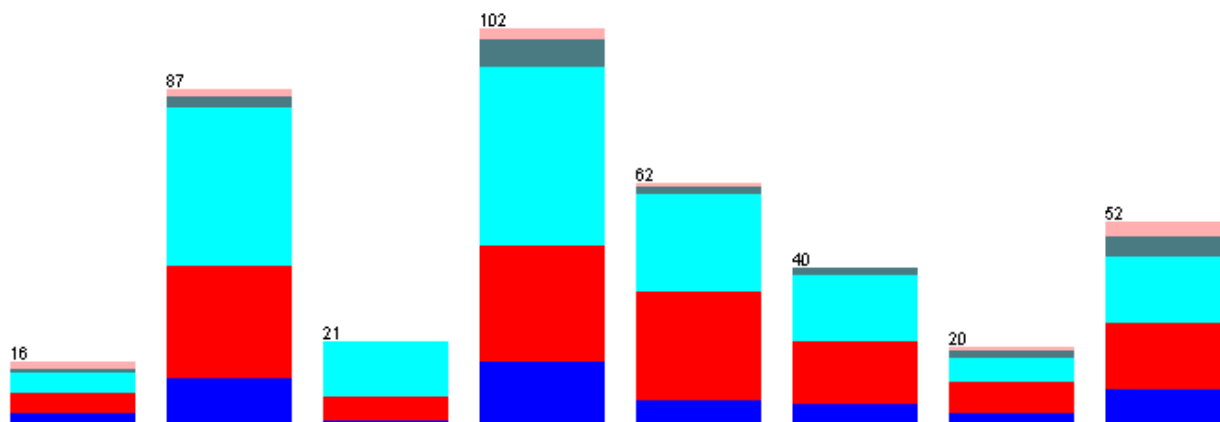
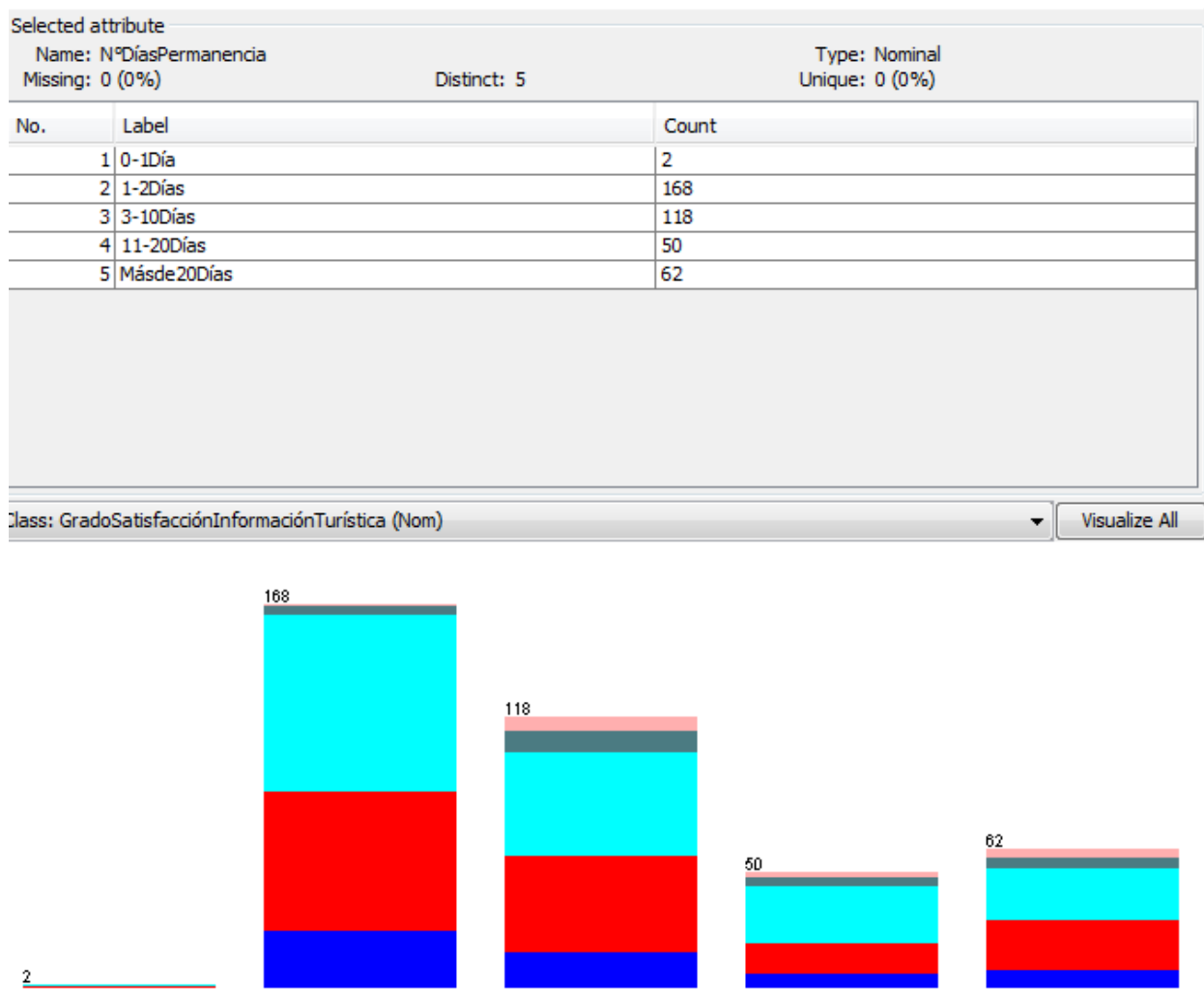


Figura 19. Atributo MotivoViaje de los encuestados. . Adaptado de: Proceso Obtenido por el software ΩΕΚΑ.

Dependiendo de las actividades que se planeen realizar en la ciudad, es notable observar que los Días de Permanencia más representativos de los viajeros son entre 1-2 Días con un 42%, ya que las actividades planeadas a ejecutar durante el viaje son actividades que pueden cumplirse dentro de este rango de días.



Φιγυρα 20. Ατριθυτο Ν°Δ΄ασΠερμανενχια δε λοσ ενχυεσταδοσ . Αδαπταδο δε: Προχεσο Οβτεν ιδο πορ ελ σοφτωαρε ΩΕΚΑ.

Gran parte de los viajeros deciden visitar la ciudad sin compañía alguna y disfrutar de las actividades que esta puede ofrecer. Esto se ve reflejado en el análisis con un 58% de los encuestados, comparado con aquellos que son acompañados por 1-2 Personas durante su estadía en Bucaramanga, representado con un 35.25%.

Selected attribute

Name: N°CompañeroViaje Type: Nominal
 Missing: 0 (0%) Distinct: 5 Unique: 0 (0%)

No.	Label	Count
1	0-1Persona	232
2	1-2Personas	141
3	3-10Personas	15
4	11-20Personas	8
5	Másde20Personas	4

Class: GradoSatisfacciónInformaciónTurística (Nom) Visualize All

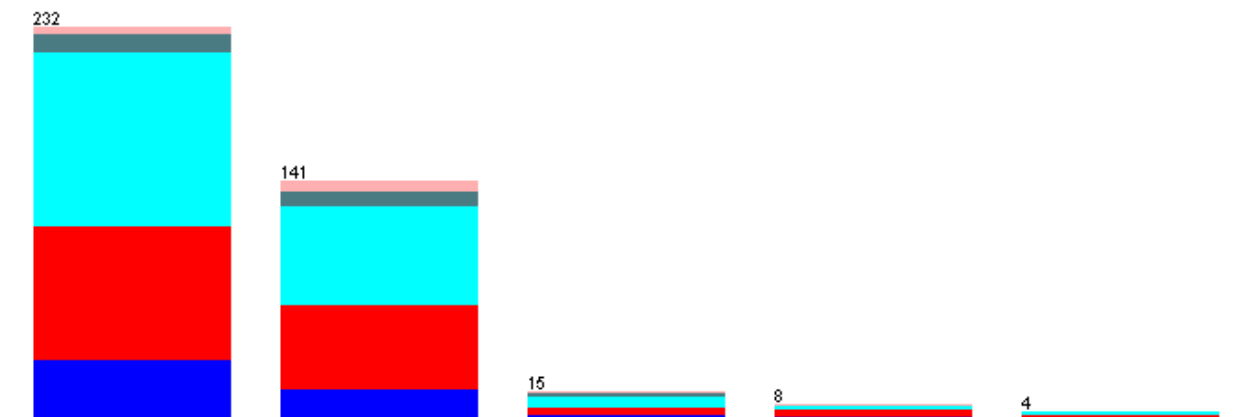


Figura 21. Atributo N°CompañeroViaje de los encuestados. . Adaptado de: Προχεςο Οβτενιδ ο πορ ελ σοφτωαρε ΩΕΚΑ.

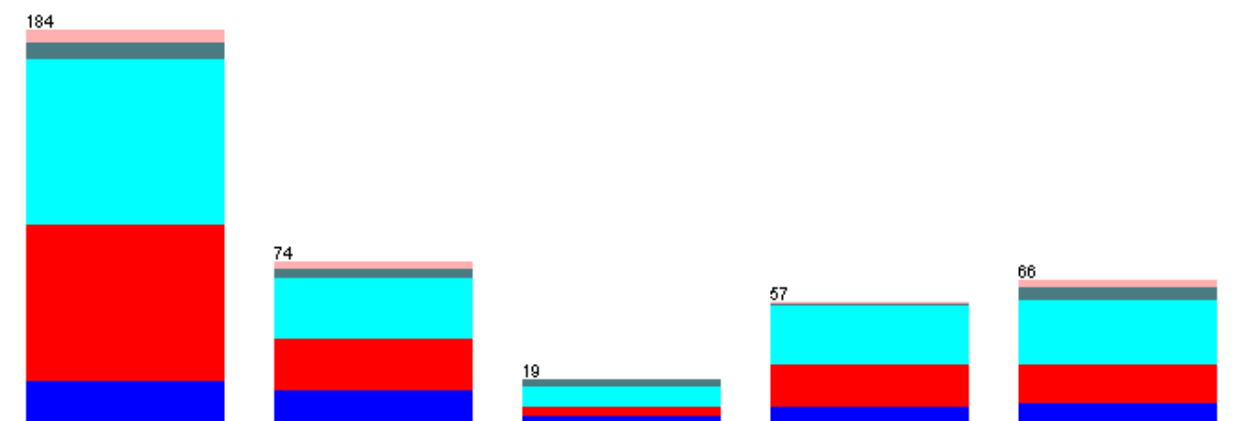
El Lugar de Alojamiento más utilizado por Los viajeros en Bucaramanga son los Hoteles representado con un 46% del total de los encuestados, pues brindan gran variedad de opciones que permiten que se acomoden a las necesidades de cada viajero.

Selected attribute

Name: LugarAlojamiento Type: Nominal
 Missing: 0 (0%) Distinct: 5 Unique: 0 (0%)

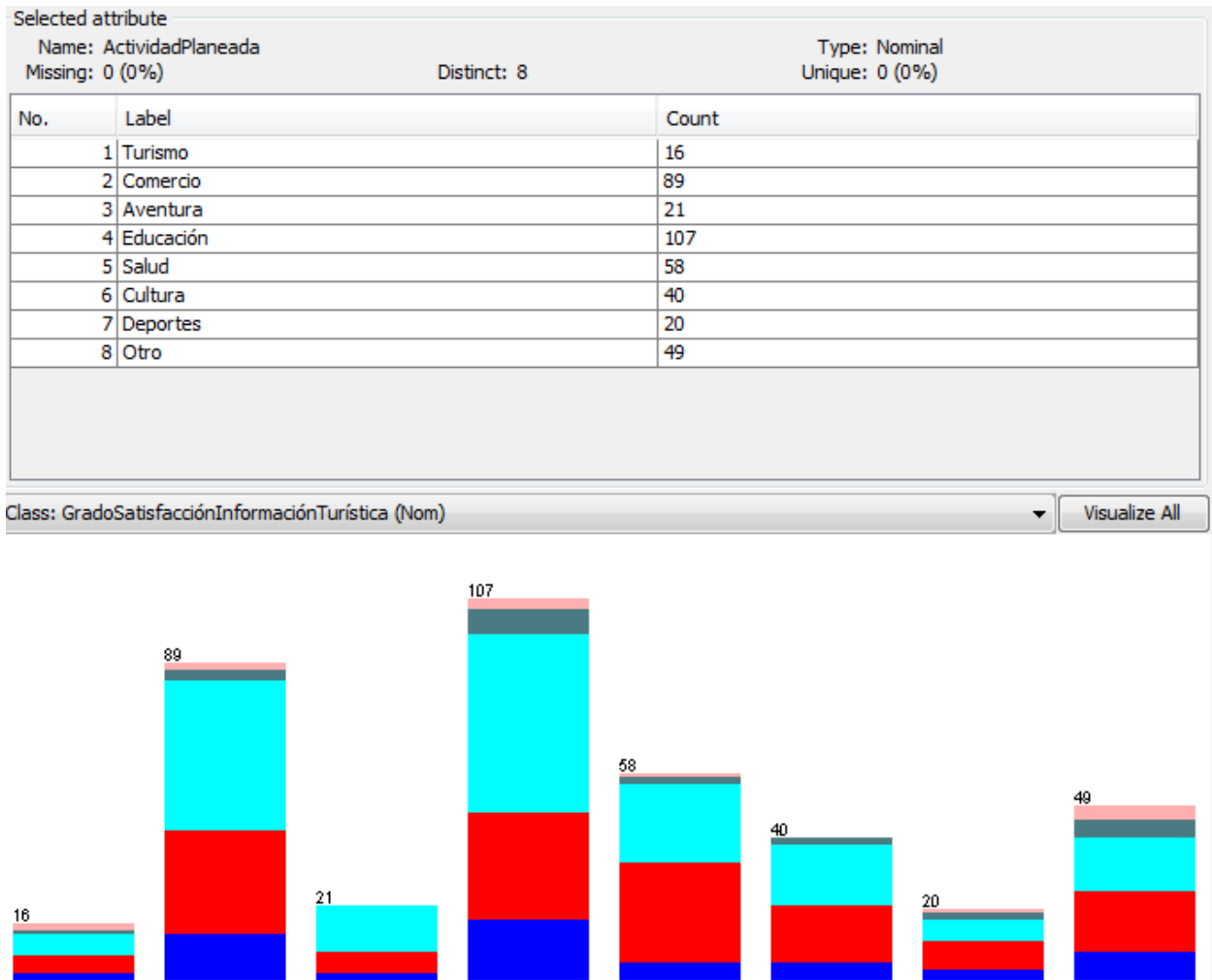
No.	Label	Count
1	Hotel	184
2	Familia	74
3	Amigos	19
4	Hostal	57
5	Otro	66

Class: GradoSatisfacciónInformaciónTurística (Nom) Visualize All



Φιγυρα 22. Ατριβυτο ΛυγαρΑλοφαμεντο δε λοσ ενχυεσταδοσ. .Αδαπαδο δε: Προχεσο Οβτενι δο πορ ελ σοφτωαρε ΩΕΚΑ.

Las Actividades Planeadas más realizadas en la ciudad son la Educación (26.75%) por su amplia oferta educativa, y el Comercio (22.25%) por la variedad e intercambio de productos y servicios ofertados para compra y venta.



Φιγυρα 23. Ατριβυτο ΑχτιπιδαδΠλανεαδα δε λοσ ενχυεσταδοσ. .Αδαπταδο δε: Προχεςο Οβτενι δο πορ ελ σοφτωαρε ΩΕΚΑ.

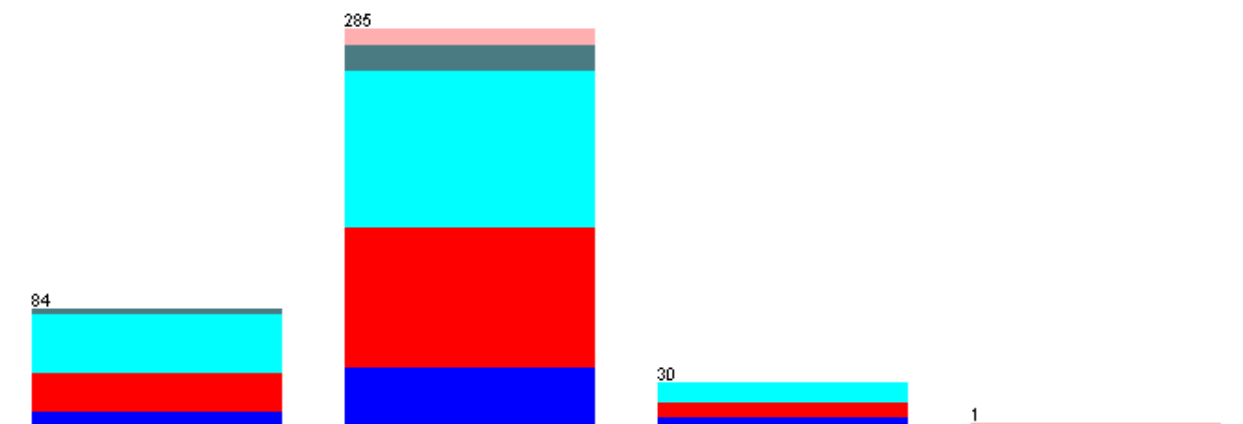
El Medio de Transporte más utilizado por los viajeros para llegar a la ciudad son el Transporte Público, bien sea por Buses, busetas y/o taxis que presten el servicio intermunicipal o departamental; representado con un 71.25% del total de la población encuestada.

Selected attribute

Name: MedioTransportedeLegada Type: Nominal
 Missing: 0 (0%) Distinct: 4 Unique: 1 (0%)

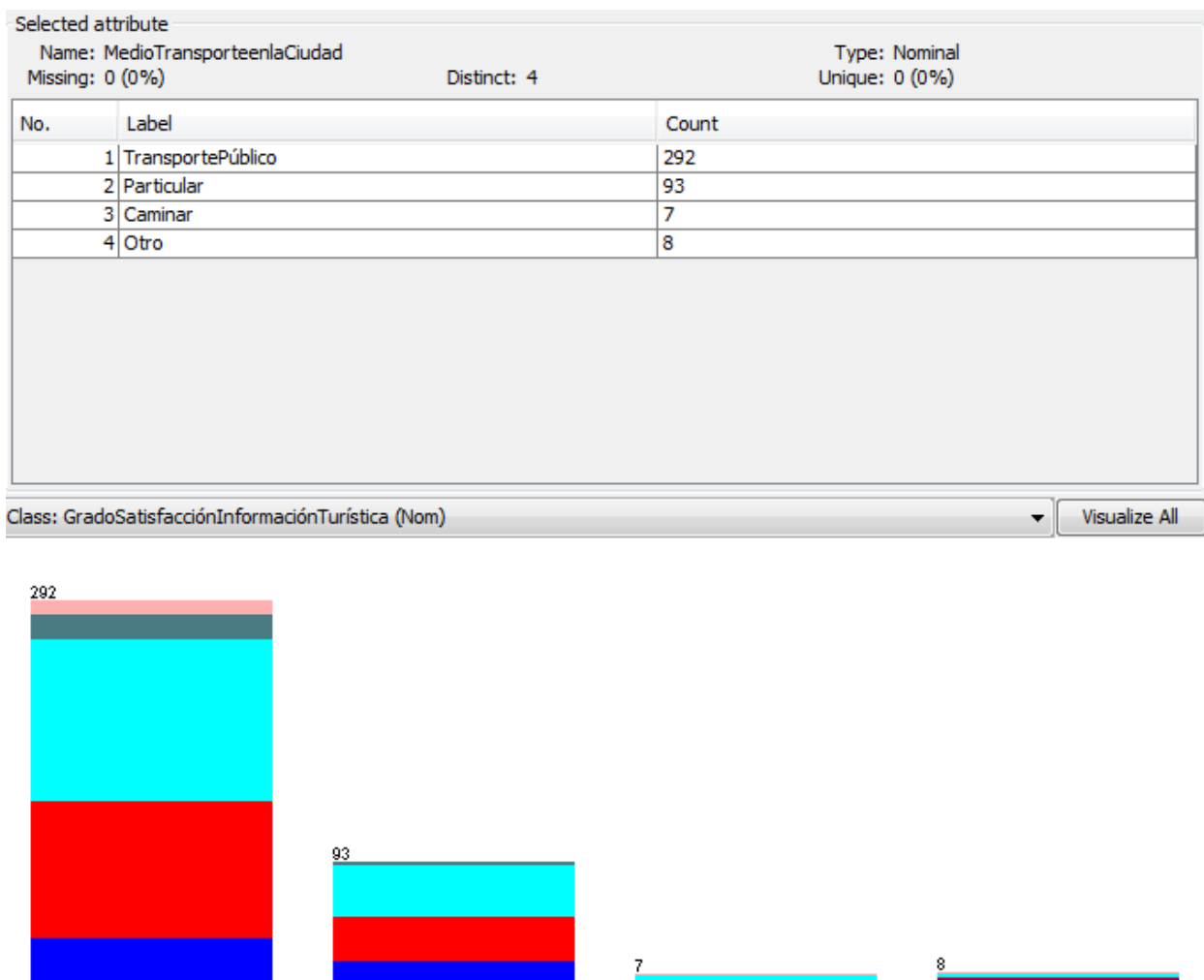
No.	Label	Count
1	Avión	84
2	TransportePúblico	285
3	Particular	30
4	Otro	1

Class: GradoSatisfacciónInformaciónTurística (Nom) Visualize All



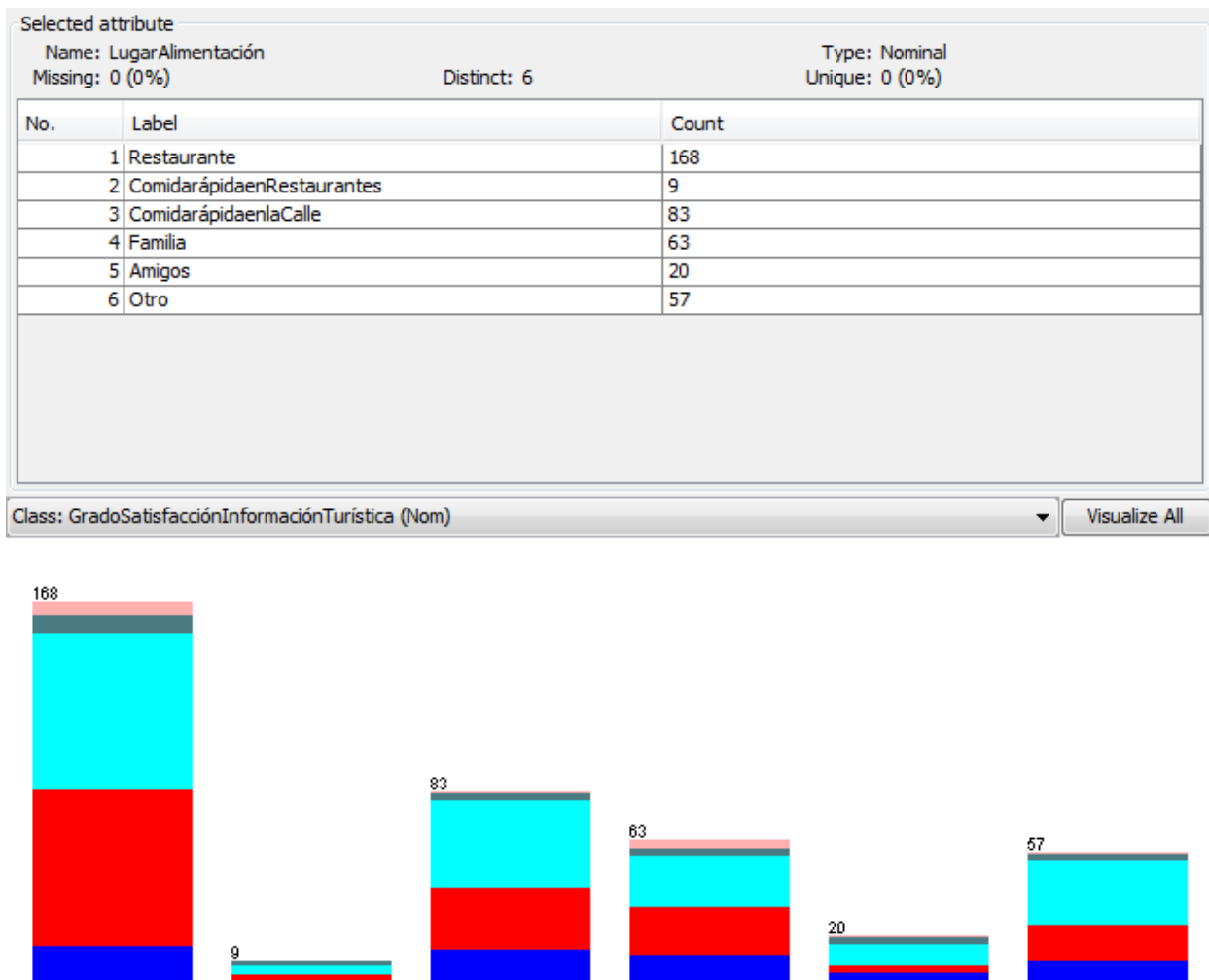
Φιγυρα 24. Ατριθυτο ΜεδιοΤρανσπορτεδεΑλεγαδα δε λοσ ενχυεσταδοσ. . Αδαπταδο δε: Προχ εσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

Dentro de la ciudad, los viajeros deciden movilizarse utilizando Medios de Transporte Público (73%), pues les genera un poco más de confianza y efectividad al momento de desplazarse por toda Bucaramanga. Esto permite un crecimiento económico positivo a la ciudad en temas de movilidad y transporte para el mejoramiento de esta frente al desarrollo constante del turismo.



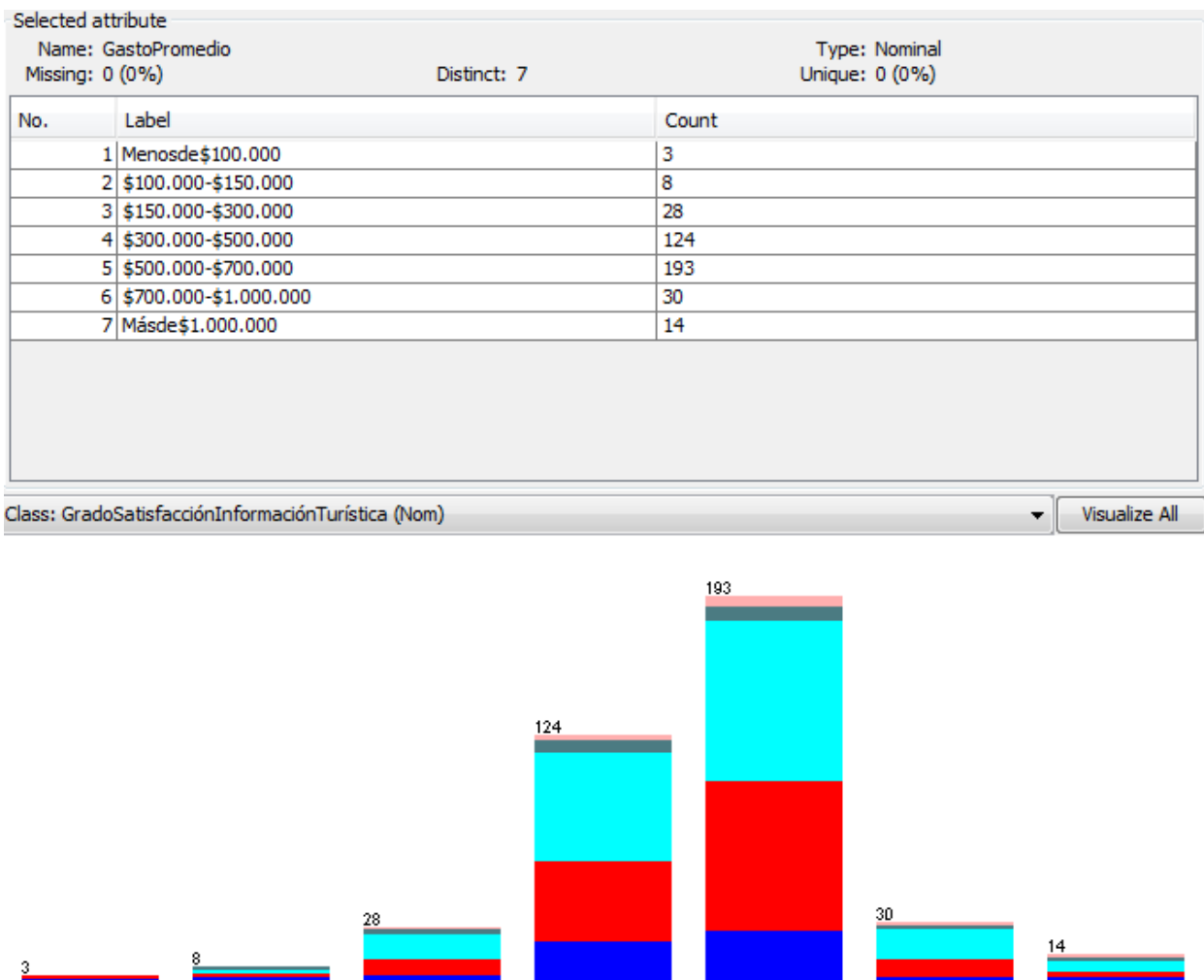
Φιγυρα 25. Ατριβυτο ΜεδιοΤρανσπορτεενλαΧιυδαδ δε λοσ ενχυεσταδοσ. . Αδαπταδο δε: Προχ εσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ .

Los Restaurantes son los Lugares de Alimentación más populares entre los viajeros (42%), pues gran parte de quienes viajan buscan algo económico, delicioso y tradicional de la comida típica de la ciudad.



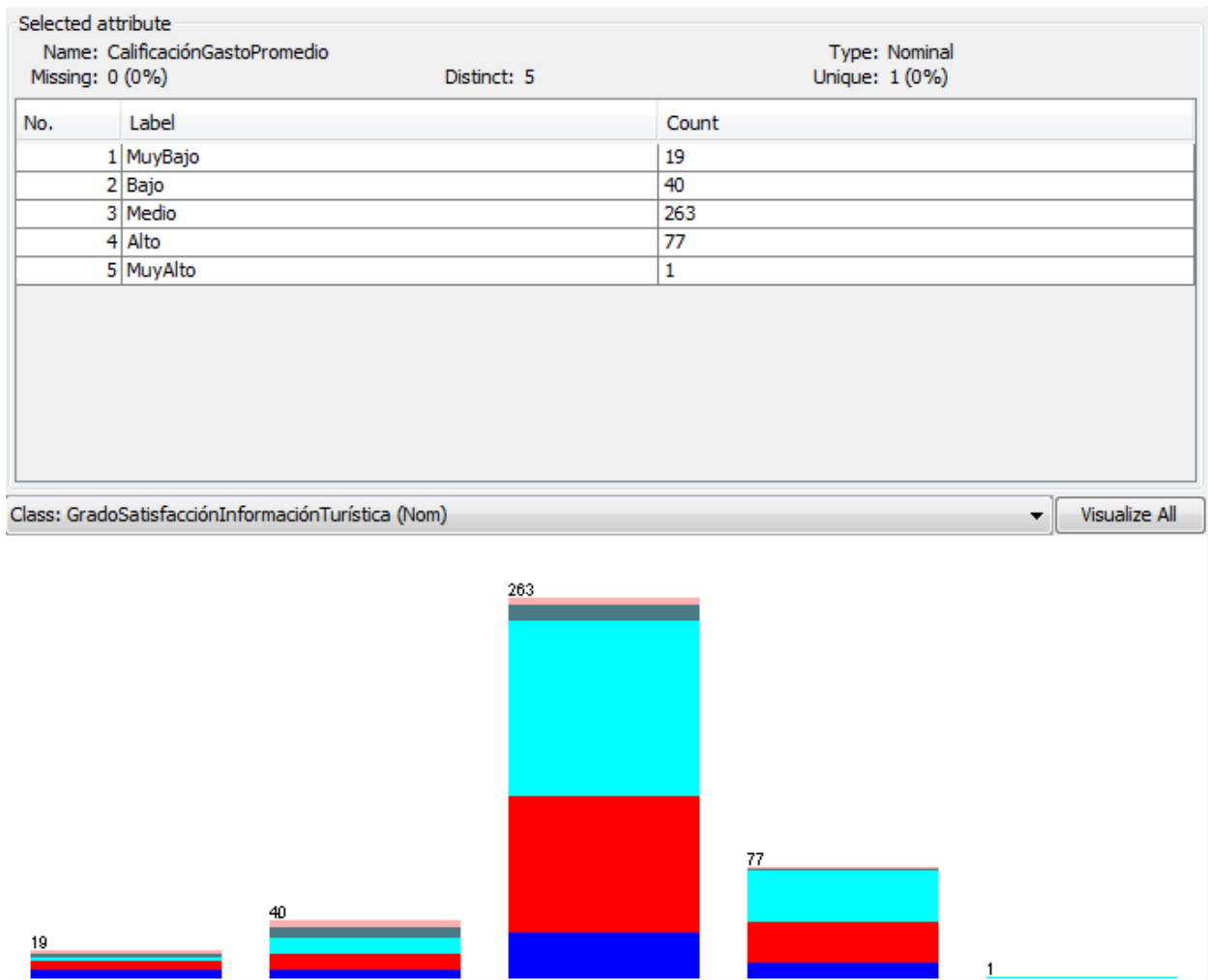
Φιγυρα 26. Ατριβυτο ΛυγαρΑλιμενταχι ίν δε λοσ ενχυεσταδοσ. . Αδαπταδο δε: Προχυεσο Οβτε νιδο πορ ελ σοφτωαρε ΩΕΚΑ.

Gran parte de los viajeros coinciden que el Gasto Promedio durante su viaje tiende a oscilar entre los \$500.000-\$700.000, representado con un 48.25%, cuyo dinero es utilizado para Transporte, Alimentación, Hospedaje y una que otra Compra.



Φιγυρα 27. Ατριβυτο ΓαστοΠρομεδιο δε λοσ ενχυεσταδοσ. Αδαπταδο δε: Προχεςο Οβτενιδο πο ρ ελ σοφτωαρε ΩΕΚΑ.

El Gasto Promedio antes analizado, para los viajeros es calificado como un Gasto Medio, ya que con la cantidad de dinero utilizada pueden realizarse diversas actividades, que cumplen completamente el itinerario programado del viaje. Esta calificación está representada por un 65.75% del total de la población encuestada.



Φιγυρα 28. Ατριβυτο Χαλιφιχαχι ΓνΓαστοΠρομεδιο δε λοσ ενχυεσταδος. . Αδαπαδο δε: Προ χεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

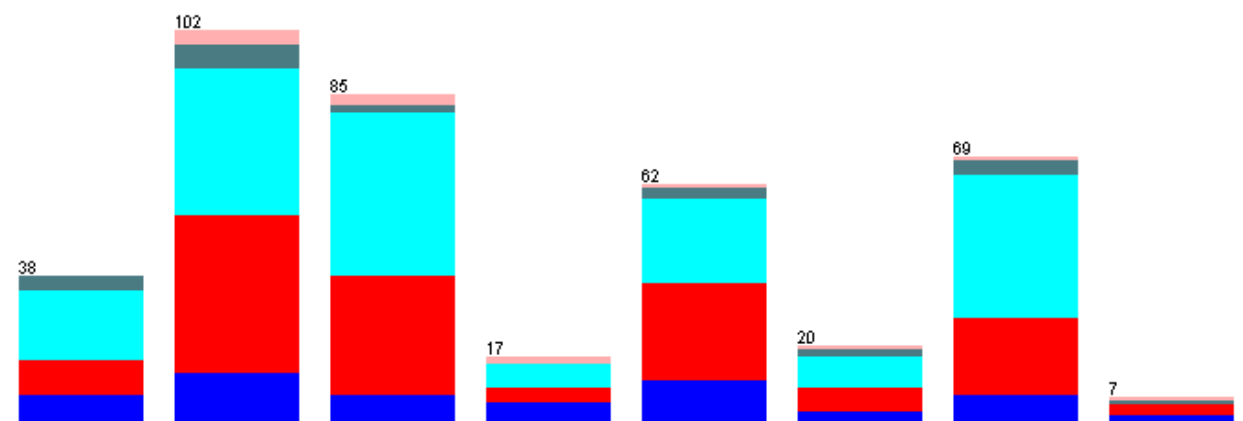
Los Aspectos que más impactaron de la ciudad a los encuestados son: Clima Agradable con un 25.5%, seguido de Ciudad Ruidosa con un 21.25%, y finalmente con Mucha Indigencia representado por un 17.25%.

Selected attribute

Name: AspectoqueImpactodelaCiudad Type: Nominal
 Missing: 0 (0%) Distinct: 8 Unique: 0 (0%)

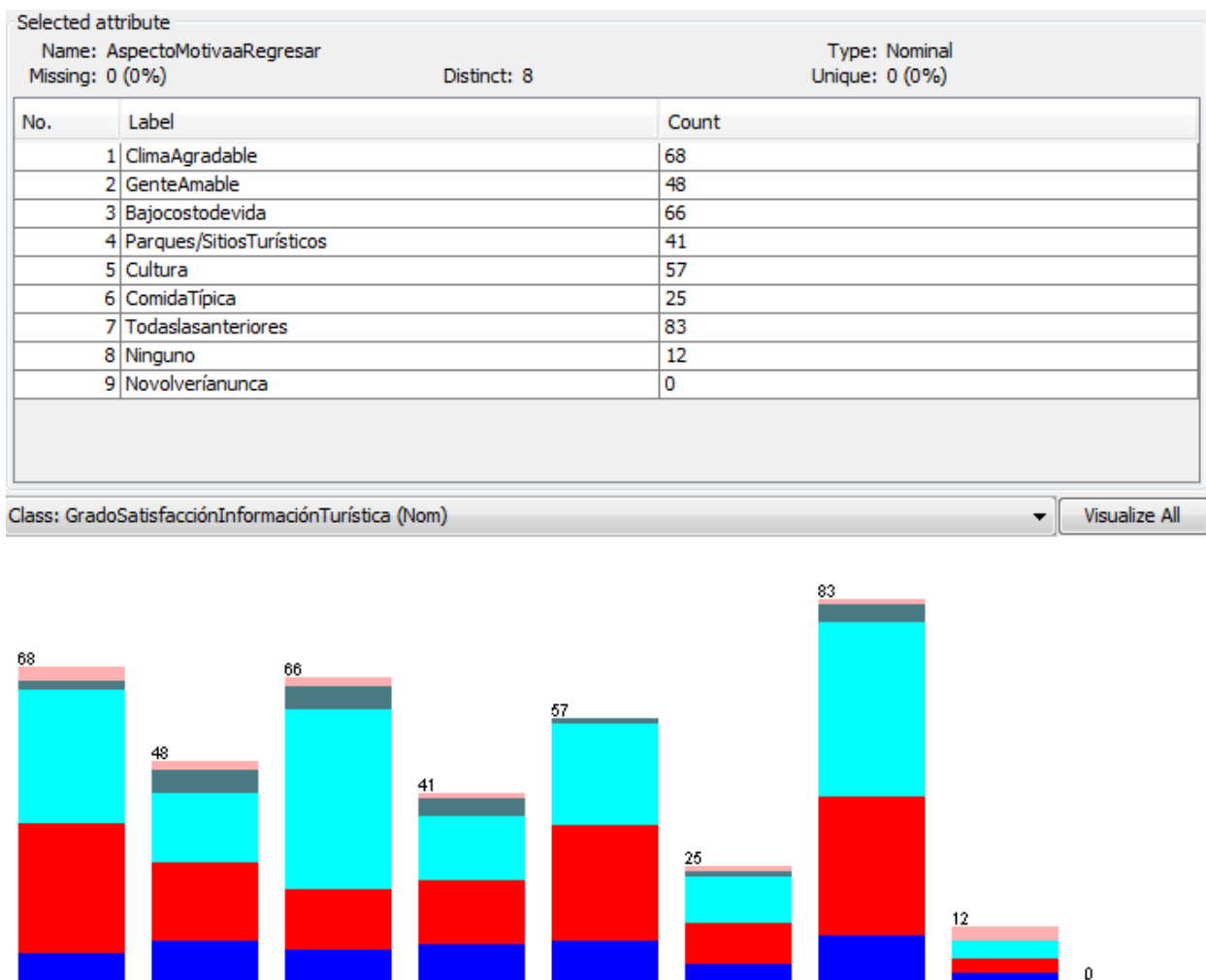
No.	Label	Count
1	AmbienteAgradable	38
2	ClimaAgradable	102
3	CiudadRuidosa	85
4	AltaContaminación	17
5	GenteAmable	62
6	CiudadSegura	20
7	MuchaIndigencia	69
8	Otro	7

Class: GradoSatisfacciónInformaciónTurística (Nom) Visualize All



Φιγυρα 29. Ατριβυτο ΑσπεχτοθυειμπαχτοδελαΧιυδαδ δε λοσ ενχυεσταδοσ. . Αδαπταδο δε: Πρ οχεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

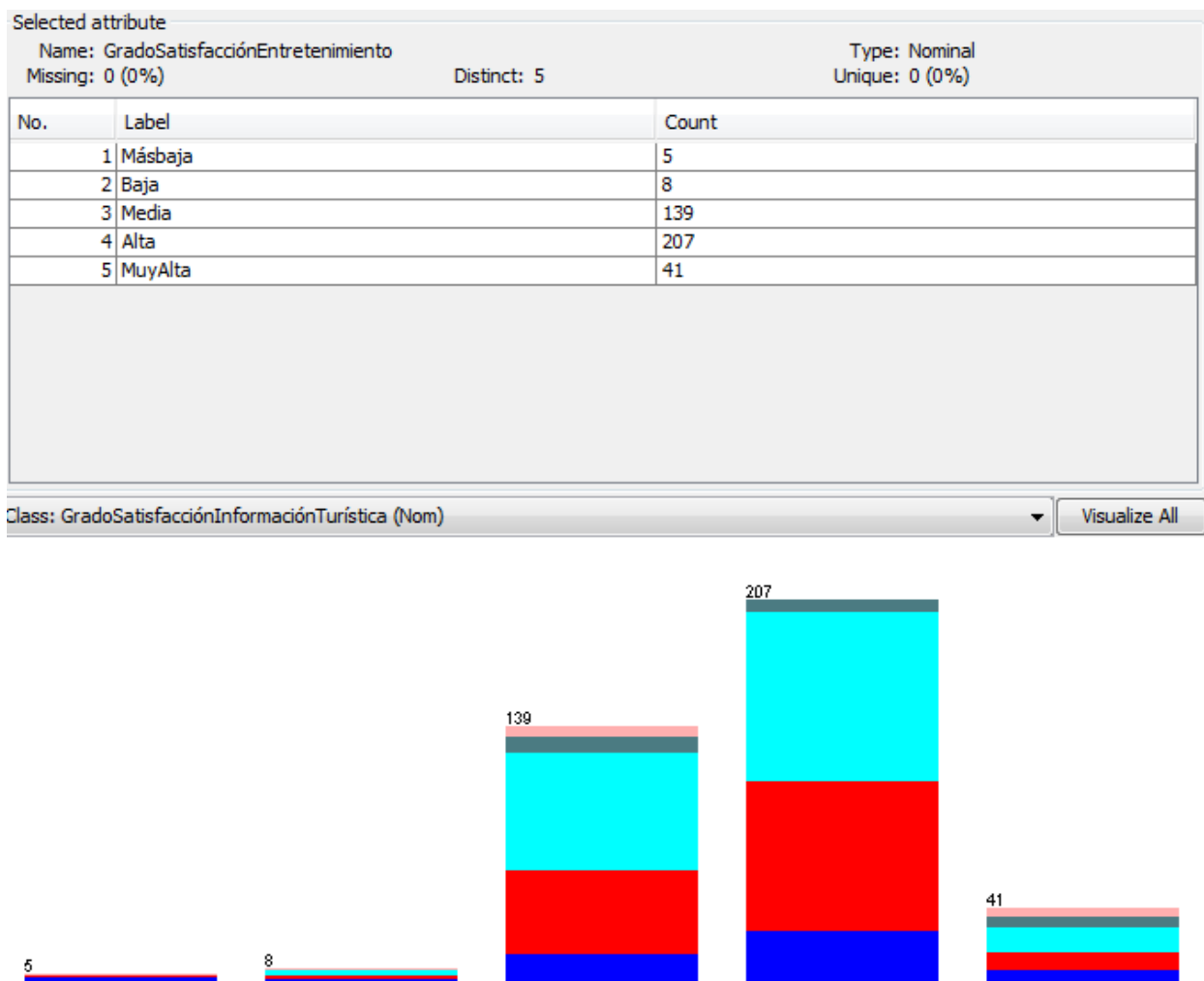
Entre los Aspectos más importantes por los cuales los viajeros estarían motivados a regresar a la ciudad, encontramos: Todas las anteriores (20.75%) por la cantidad de aspectos positivos que ofrece Bucaramanga, Clima Agradable (17%) y Bajo Costo de Vida (16.5%).



Φιγυρα 30. Ατριβυτο ΑσπεχτοΜοτιπιααΡεγρεσαρ δε λοσ ενχυεσταδοσ. . Αδαπταδο δε: Προχεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

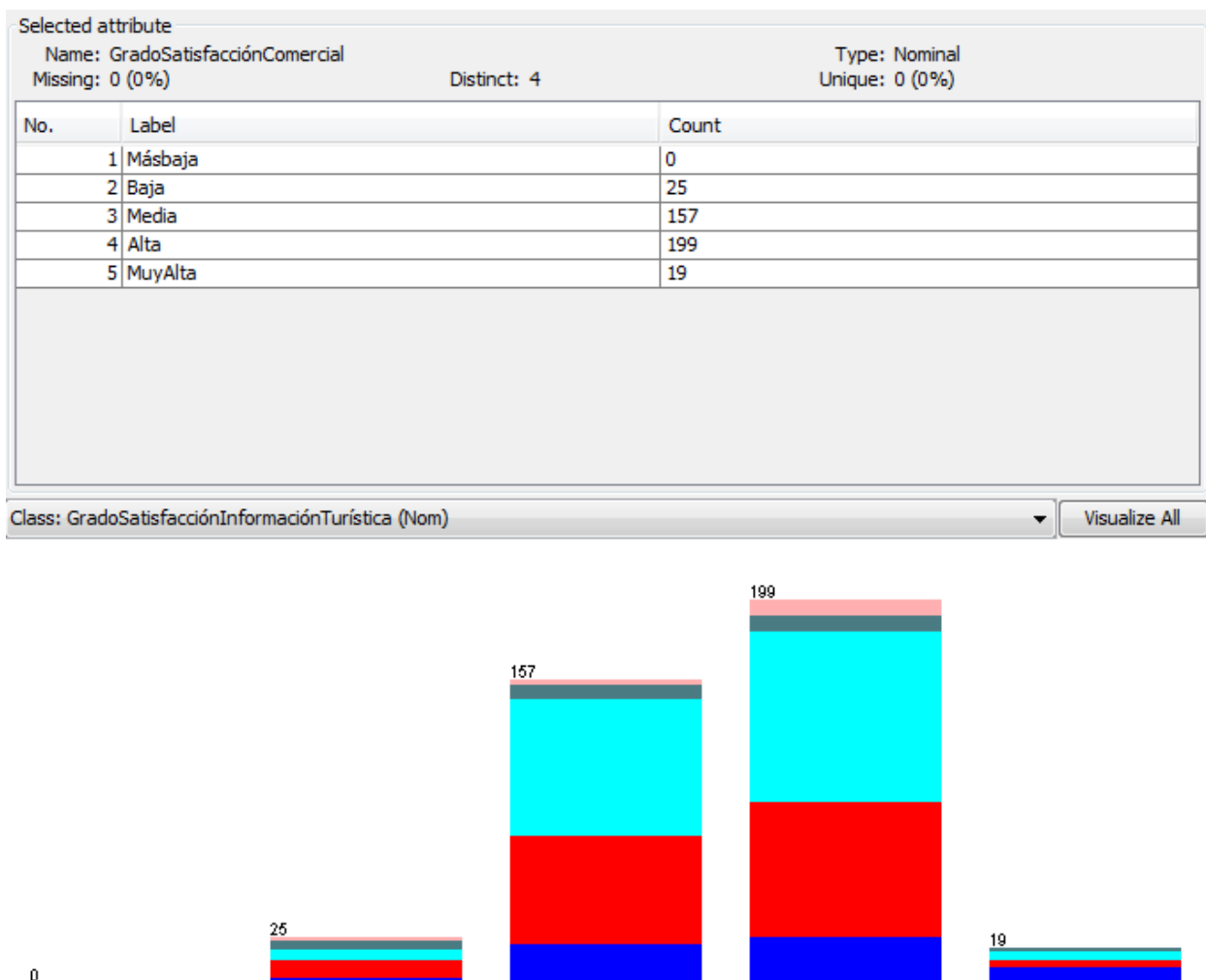
Para los viajeros el Grado de Satisfacción frente a ciertas ofertas ofrecidas por la ciudad son:

El Grado de Satisfacción en Ofertas de Entretenimiento es Alta con un 51.75% de la población encuestada, ya que Bucaramanga cuenta con lugares y espacios que permiten el desarrollo de actividades recreativas tanto para residentes como visitantes, y así disfrutar de las opciones de entretenimiento ofertados.



Φιγυρα 31. Ατριβυτο ΓραδοΣατισφαχχι (νΕντρναμμεντο δε λοσ ενχυεσταδοσ. . Αδαπαδο δε: Προχεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

El Grado de Satisfacción en Oferta Comercial es Alta con un 49.75% del total de viajeros encuestados, por la variedad de productos y servicios para la compra y venta de estos.



Φιγυρα 32. Ατριβυτο ΓραδοΣατισφαχχι ΊνΧομερχιαλ δε λοσ ενχυεσταδοσ. . Αδαπταδο δε: Προ χεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

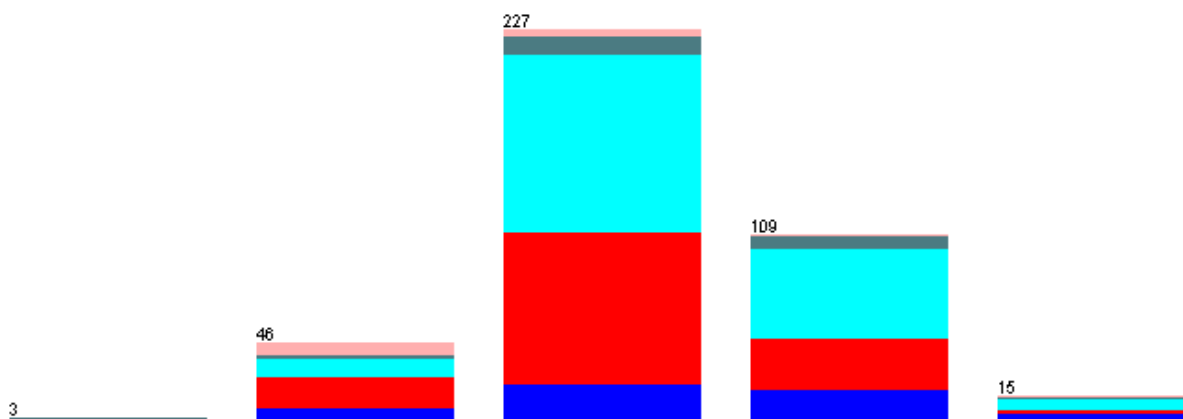
El Grado de Satisfacción en Oferta Cultural es Media con un 56.75% del total de viajeros encuestados, por ser considerada como Centro Cultural de Santander, al ofrecer diversos eventos formativos y de interés como herramienta para el desarrollo económico y social de la ciudad y la región.

Selected attribute

Name: GradoSatisfacciónCultural Type: Nominal
 Missing: 0 (0%) Distinct: 5 Unique: 0 (0%)

No.	Label	Count
1	Másbaja	3
2	Baja	46
3	Media	227
4	Alta	109
5	MuyAlta	15

Class: GradoSatisfacciónInformaciónTurística (Nom) Visualize All



Φιγυρα 33. Ατριβυτο ΓραδοΣατισφαχχι ΓνΧυλτυραλ δε λοσ ενχυεσταδοσ. . Αδαπταδο δε: Προχ εσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ..

El Grado de Satisfacción en Señalización Turística es Media con un 58.75%, pues los viajeros comparten que durante el tiempo que estuvieron en la ciudad existía muy poca visualización de señales o información turística, y si en ocasiones las observaban, su localización visual les impedía ver con claridad lo que indicaban.

Selected attribute

Name: GradoSatisfacciónSeñalizaciónTurística Type: Nominal
 Missing: 0 (0%) Distinct: 5 Unique: 0 (0%)

No.	Label	Count
1	Másbaja	11
2	Baja	86
3	Media	235
4	Alta	58
5	MuyAlta	10

Class: GradoSatisfacciónInformaciónTurística (Nom) Visualize All

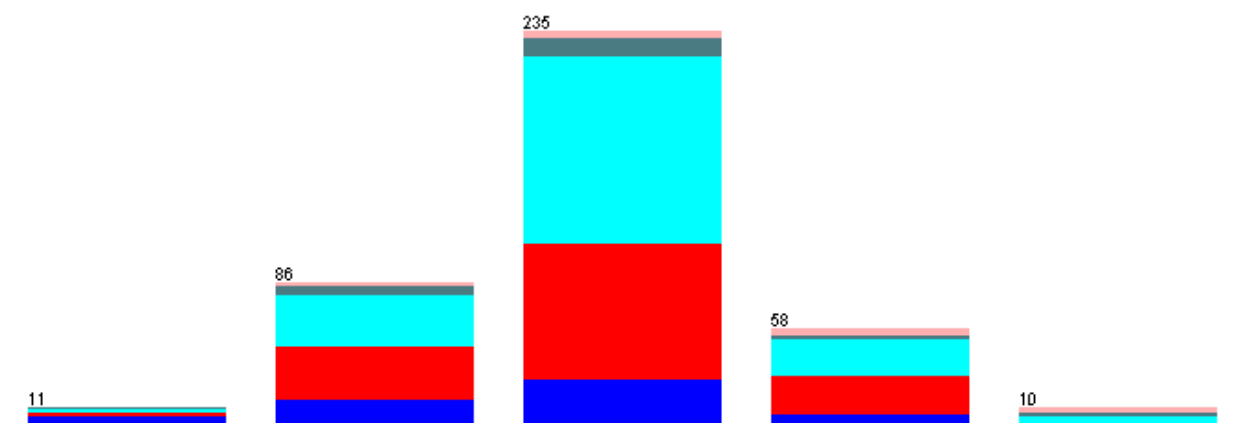


Figura 34. Atributo GradoSatisfacciónSeñalizaciónTurística (Nom) de los encuestados. Adaptado de: Proceso Obtenido por el software ODK.

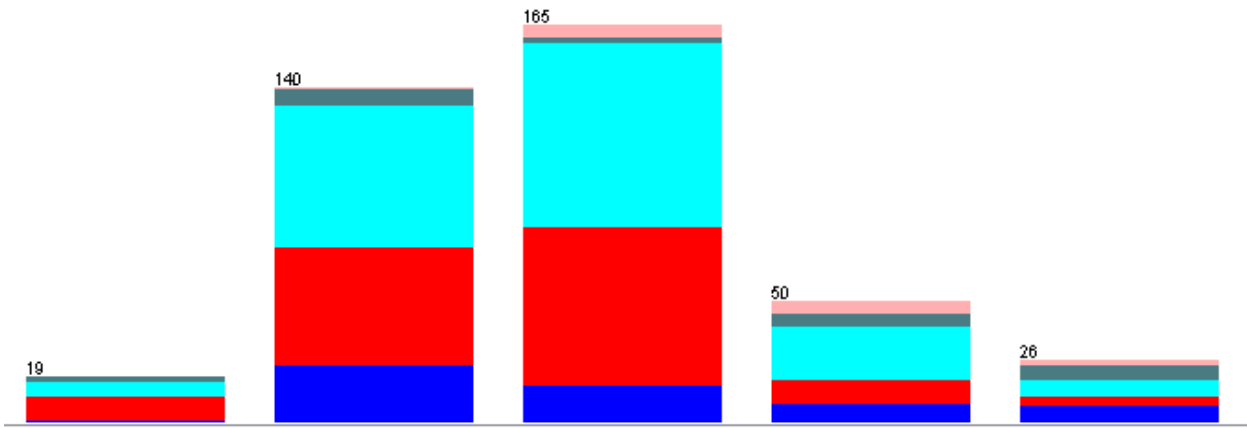
El Grado de Satisfacción en cuanto a Cultura Ciudadana es Media con un 41.25% del total de viajeros encuestados, ya que la mayoría coincide que a pesar que Bucaramanga cuenta con una diversidad cultural significativa, esto también es piedra de tropiezo para el desarrollo, pues aun no existen en dichas personas sentido de pertenencia por el lugar donde residen, y es por esto que se ve reflejado en una ciudad con un gran índice de basuras y un una creciente alarma de contaminación.

Selected attribute

Name: GradoSatisfacciónCulturaCiudadana Type: Nominal
 Missing: 0 (0%) Distinct: 5 Unique: 0 (0%)

No.	Label	Count
1	Másbaja	19
2	Baja	140
3	Media	165
4	Alta	50
5	MuyAlta	26

Class: GradoSatisfacciónInformaciónTurística (Nom) Visualize All



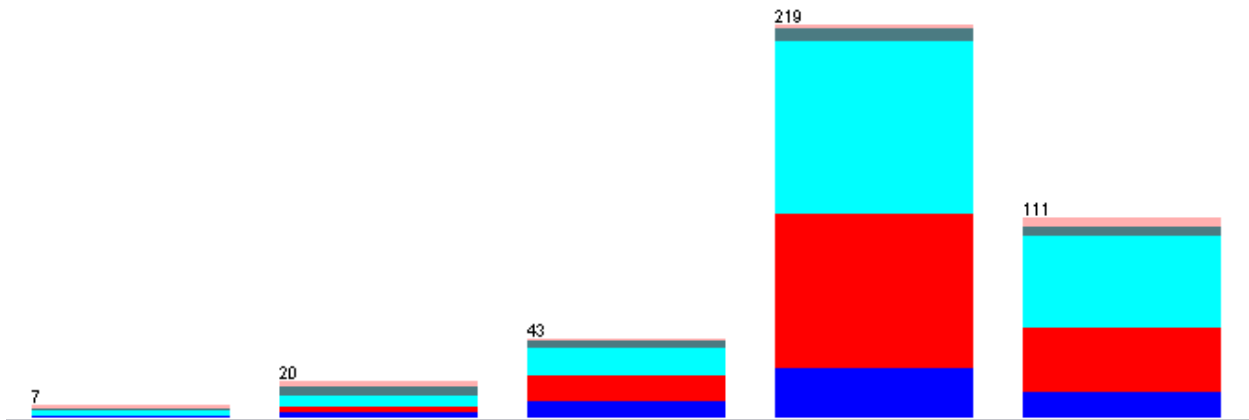
Φιγυρα 35. Ατριβυτο ΓραδοΣατισφαλλι ΊνΧυλτυραΧιυδααδανα δε λοσ ενχυεσταδος. . Αδαπτα δο δε: Προχεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

El Grado de Satisfacción en Ofertas de Alojamiento es Alta con un 54.75% al contar con variedad en Hoteles, Hostales, entre otros; que le ofrecen a los visitantes diversas opciones que se ajusten a su necesidad y presupuesto.

Selected attribute
 Name: GradoSatisfacciónAlojamiento Type: Nominal
 Missing: 0 (0%) Distinct: 5 Unique: 0 (0%)

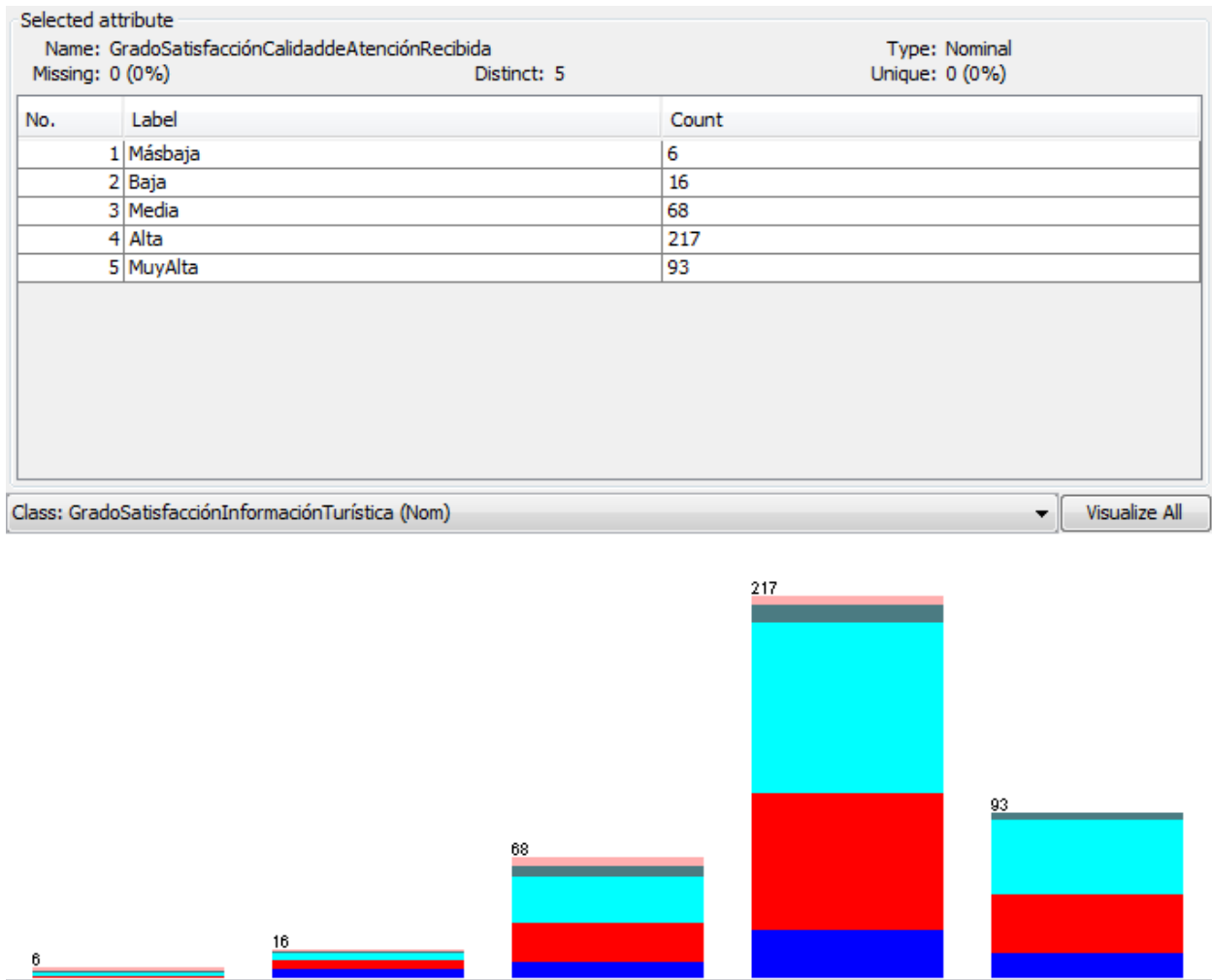
No.	Label	Count
1	Másbaja	7
2	Baja	20
3	Media	43
4	Alta	219
5	MuyAlta	111

Class: GradoSatisfacciónInformaciónTurística (Nom) Visualize All



Φιγυρα 36. Ατριβυτο ΓραδοΣατισφαχχι ἴναλοφαιμεντο δε λοσ ενχυεσταδοσ. . Αδαπαδο δε: Π ροχεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

El Grado de Satisfacción de la Calidad de la Atención Recibida es Alta con un 54.2%, donde los viajeros coinciden que en el lugar de alojamiento y a los sitios donde estuvieron durante su instancia, recibieron pro parte de las personas una cálido trato y excelente atención.



Φιγυρα 37. Ατριβυτο ΓραδοΣατισφαχχι ἸνΧαλιδαδΑτενχι ἸνΡεχιβιδα δε λοσ ενχυεσταδοσ . Αδ απταδο δε: Προχεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

El Grado de Satisfacción en la Calidad de Alojamiento es Alta con un 46.5% del total de viajeros encuestados, al estar gran parte de estos satisfechos y conformes por los servicios prestados en los Lugares de Alojamiento.

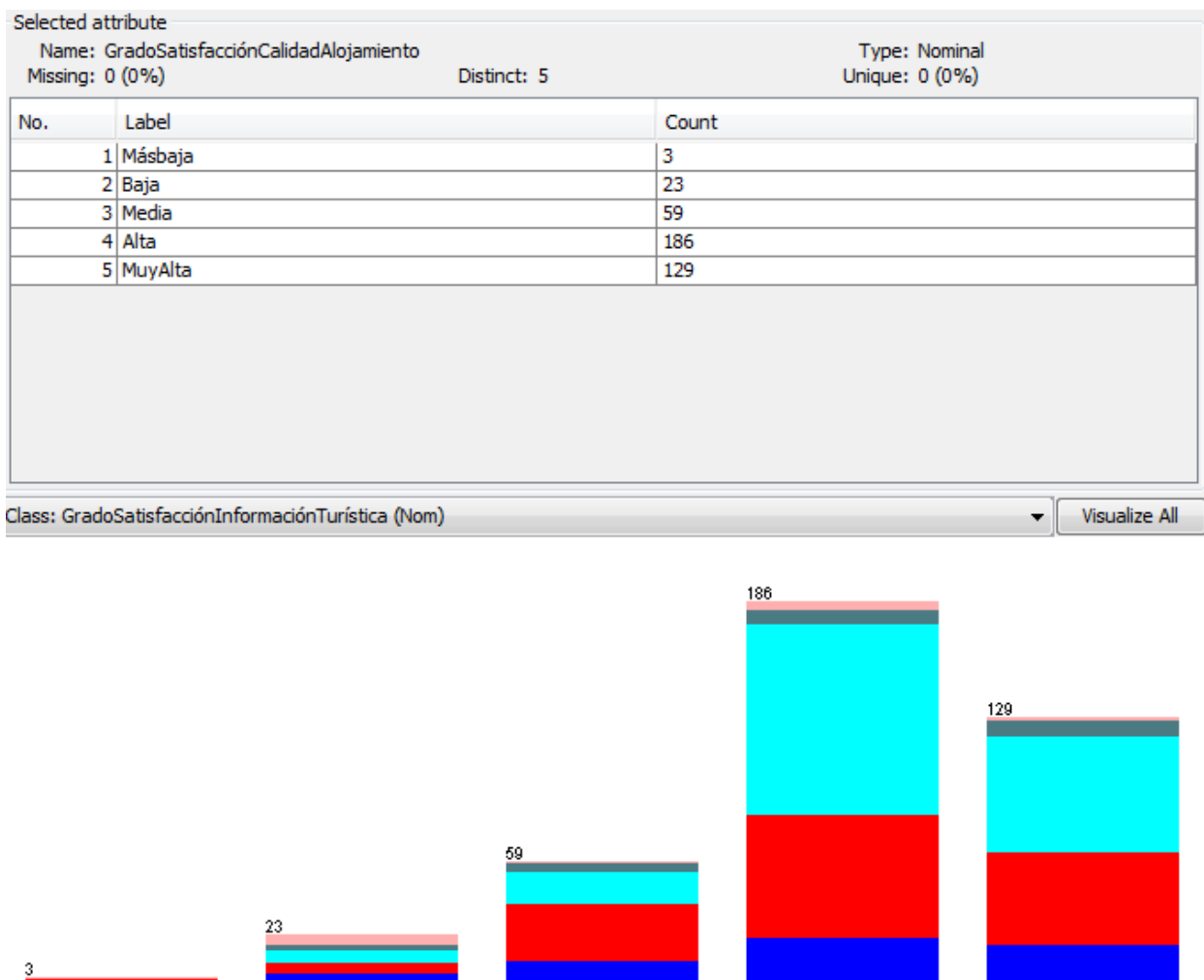
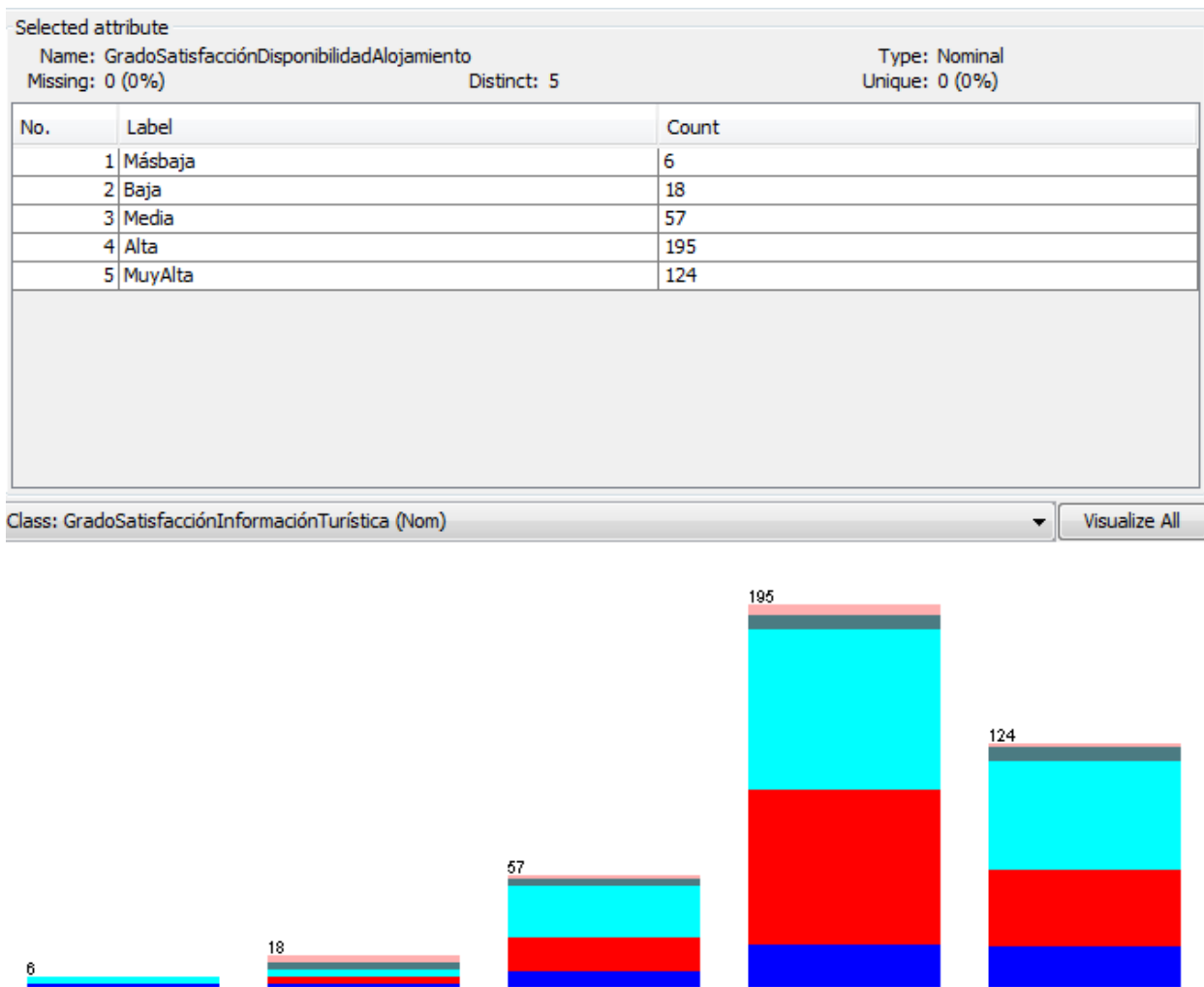


Figura 38. Atributo GradoSatisfacciónCalidadAlojamiento de los encuestados. . Adaptado de: Proceso Obtenido por el software WEGA.

El Grado de Satisfacción en la Disponibilidad de Alojamiento es Alta con un 48.75%, ya que gran parte de los encuestados reservaron con anticipación para no presentar problemas al momento de encontrar un lugar disponible para hospedarse, y quienes no pudieron reservar tenían ya conocimiento a que lugares llegar para encontrar una habitación disponible.



Φιγυρα 39. Ατριβυτο ΓραδοΣατισφαχχι ΓνΔισπονιβιλιδαΔΑλοφαμμεντο δε λοσ ενχυεσταδοσ. .
 Αδαπαδο δε: Προχεςο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

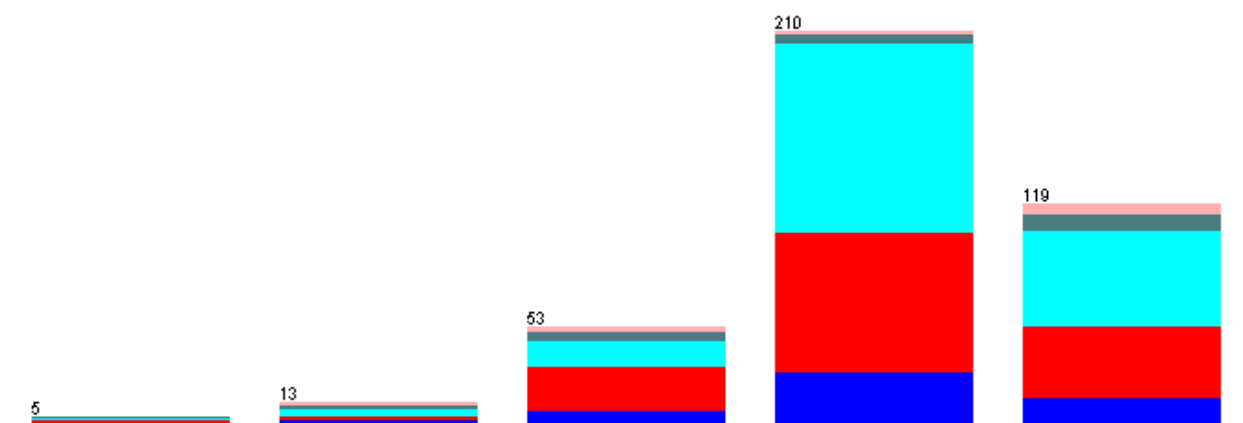
El Grado de Satisfacción en el Precio de Alojamiento es Alta con un 52.5% del total de viajeros encuestados, pues consideran que los precios que ofrecen los hoteles y demás lugares de alojamiento son adecuados y se ajustan a las necesidades de estos.

Selected attribute

Name: GradoSatisfacciónPrecioAlojamiento Type: Nominal
 Missing: 0 (0%) Distinct: 5 Unique: 0 (0%)

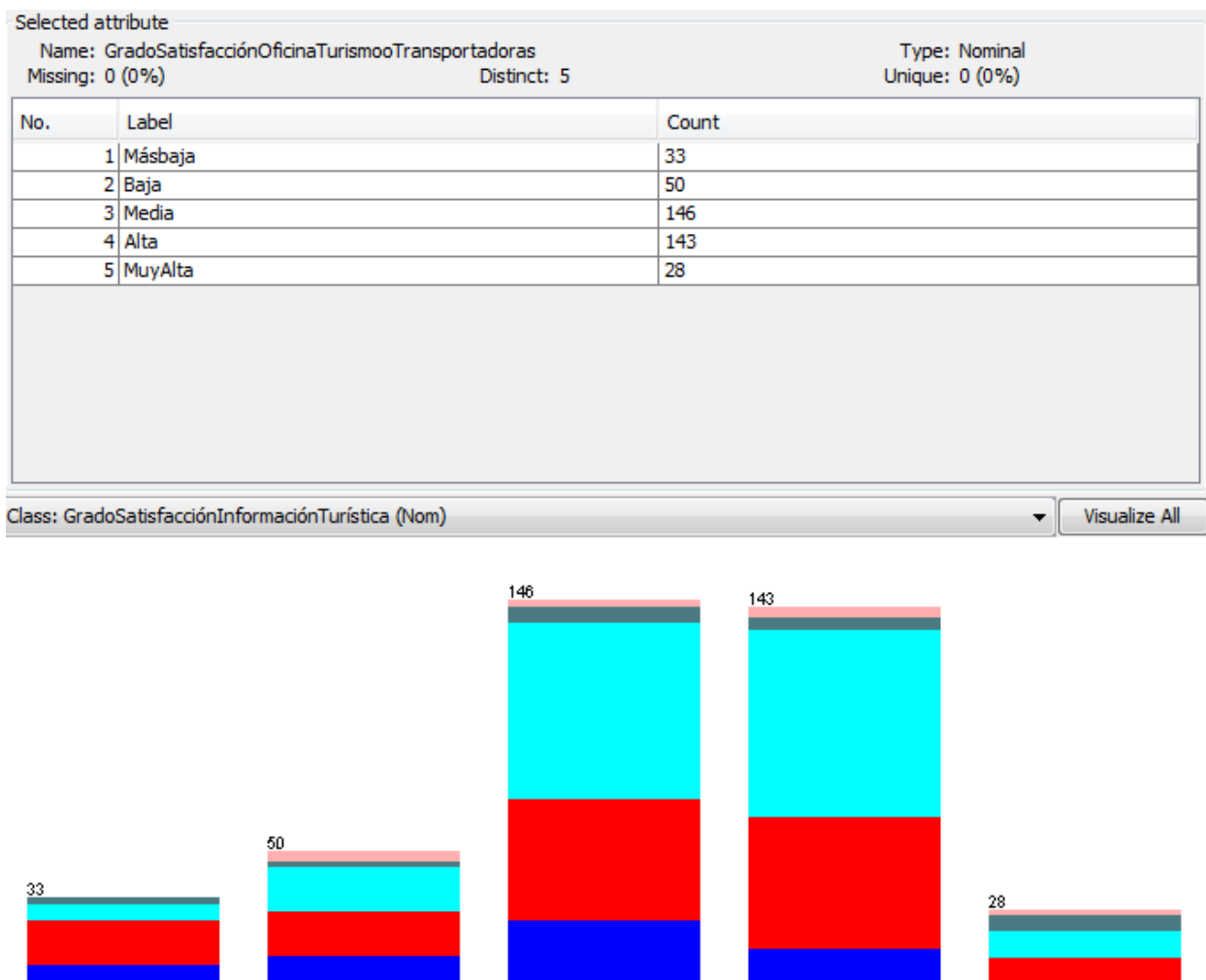
No.	Label	Count
1	Másbaja	5
2	Baja	13
3	Media	53
4	Alta	210
5	MuyAlta	119

Class: GradoSatisfacciónInformaciónTurística (Nom) Visualize All



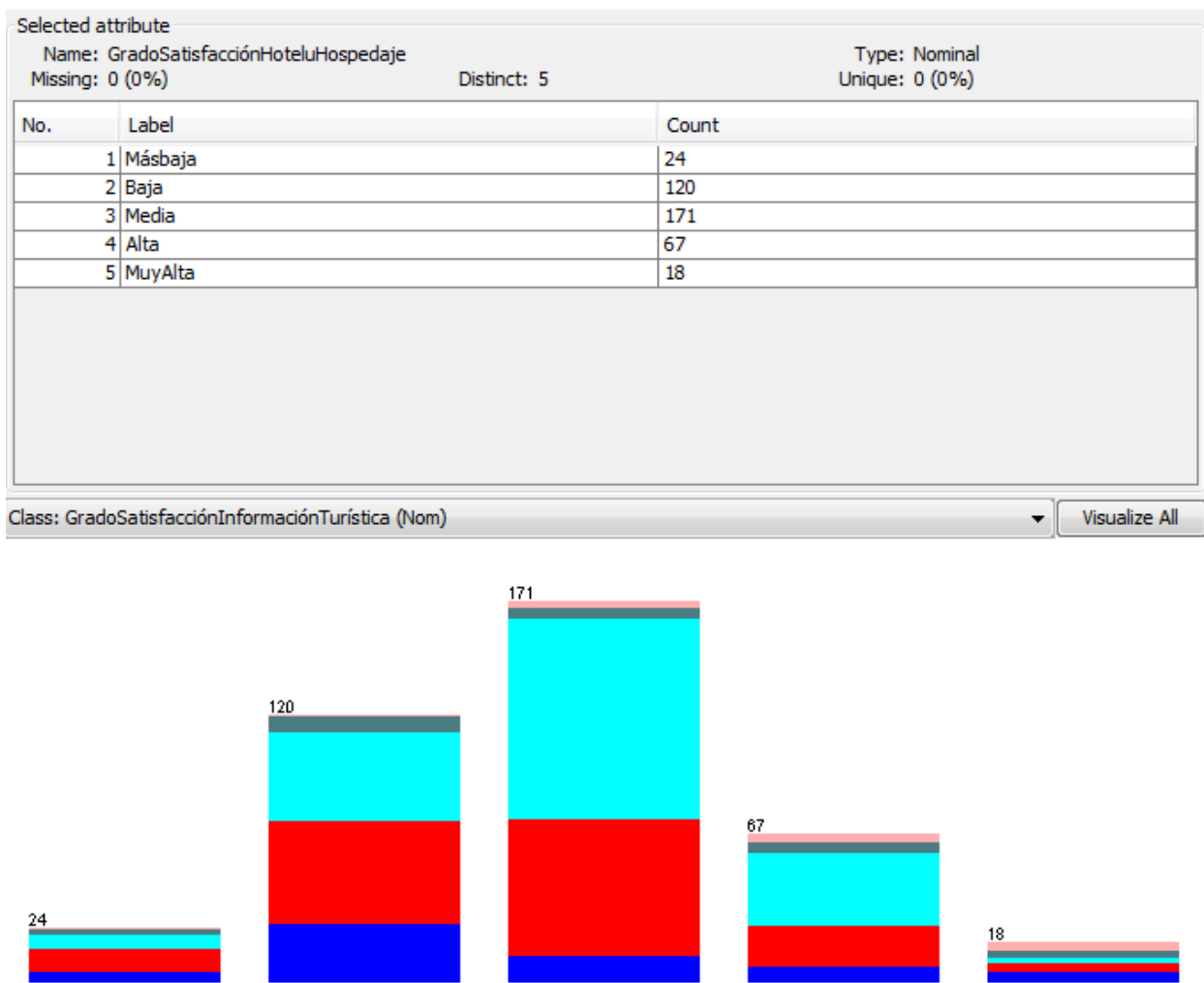
Φιγυρα 40. Ατριβυτο ΓραδοΣατισφαχχι ΓνΠρεχιοΑλοφαμμεντο δε λοσ ενχνεσταδος. . Αδαπταδ ο δε: Προχεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

El Grado de Satisfacción en Oficinas de Turismo o Empresas Transportadoras es Media con un 36.5% del total de viajeros encuestados, ya que aclaran que muchas de estas no cuentan con la información completa y adecuada para brindar a los viajeros, en temas de transporte, alojamiento, sitios turísticos y demás.



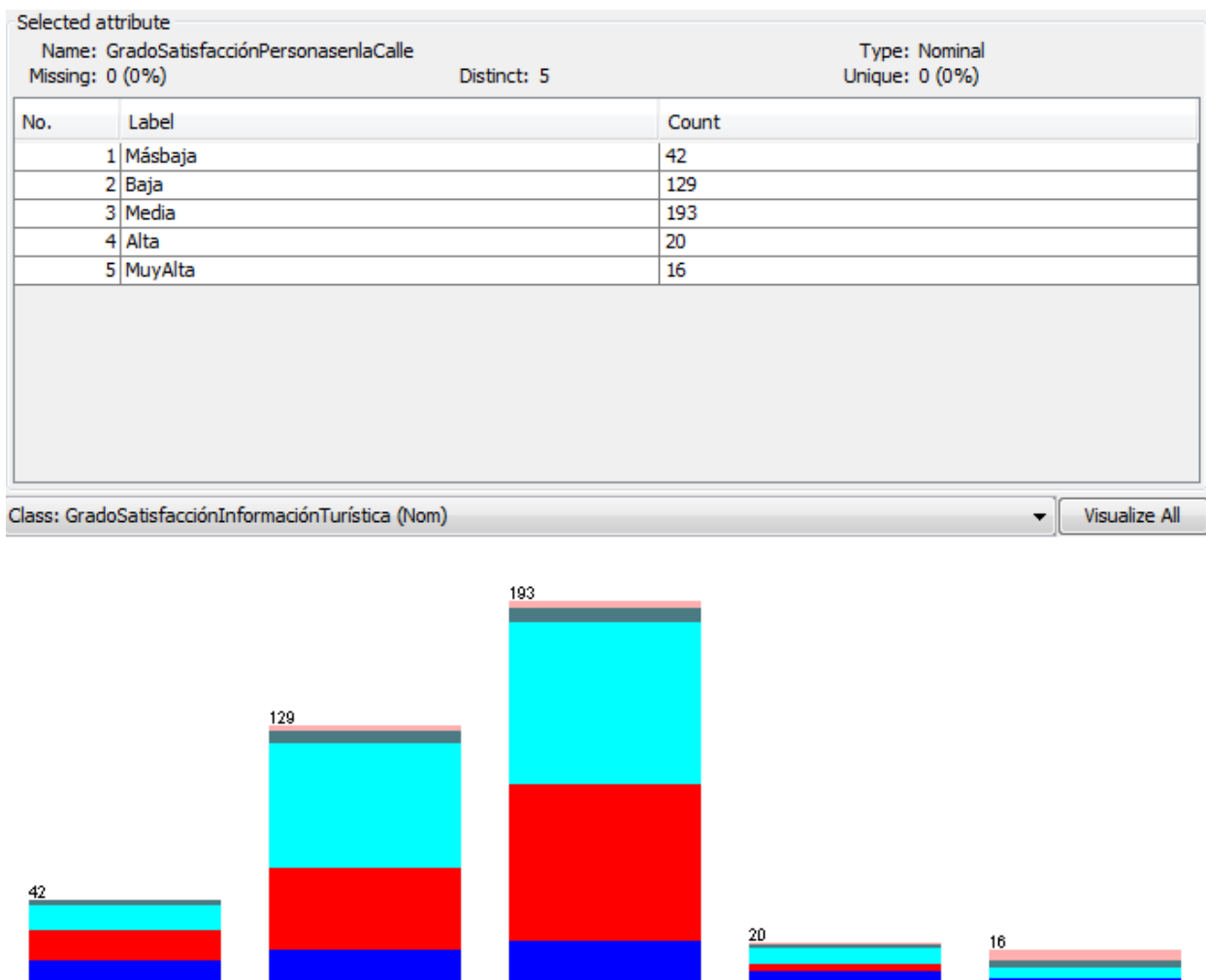
Φιγυρα 41. Ατριβυτο ΓραδοΣατισφαχχι ίνΟφιχίναΤυρισμοοΤρανσπορταδορασ δε λοσ ενχυεστ αδοσ. . Αδαπταδο δε: Προχεςο Οβτενίδο πορ ελ σοφτωαρε ΩΕΚΑ.

El Grado de Satisfacción en Hotel u Hospedaje es Media con un 42.75%, pues aunque gran parte coincide que la disponibilidad de alojamiento es alta, aun es necesario contar con más variedad y excelente atención por parte de quienes prestan dichos servicios.



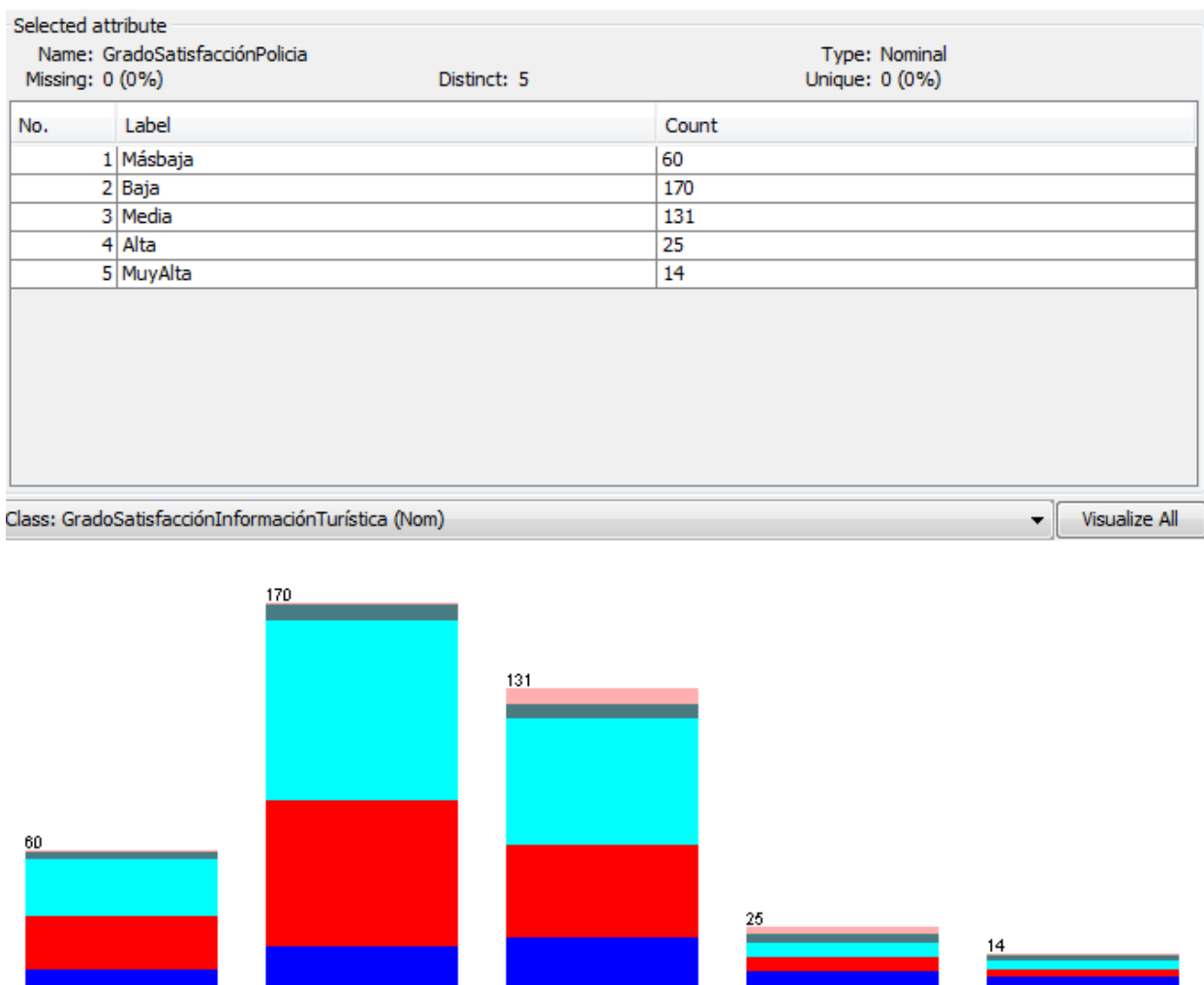
Φιγυρα 42. Ατριβυτο ΓραδοΣατισφαχχι (νΗοτελυΗοσπεδαφε δε λοσ ενχυε σταδοσ. Αδαπταδο δε: Προχεςο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

El Grado de Satisfacción de Personas en la Calle es Media con un 48.25%, debido a que ciertos viajeros afirman que muchas personas al momento de preguntarles respondían de manera amable, otros sencillamente decían desconocer la ciudad o simplemente ignorar a la persona.



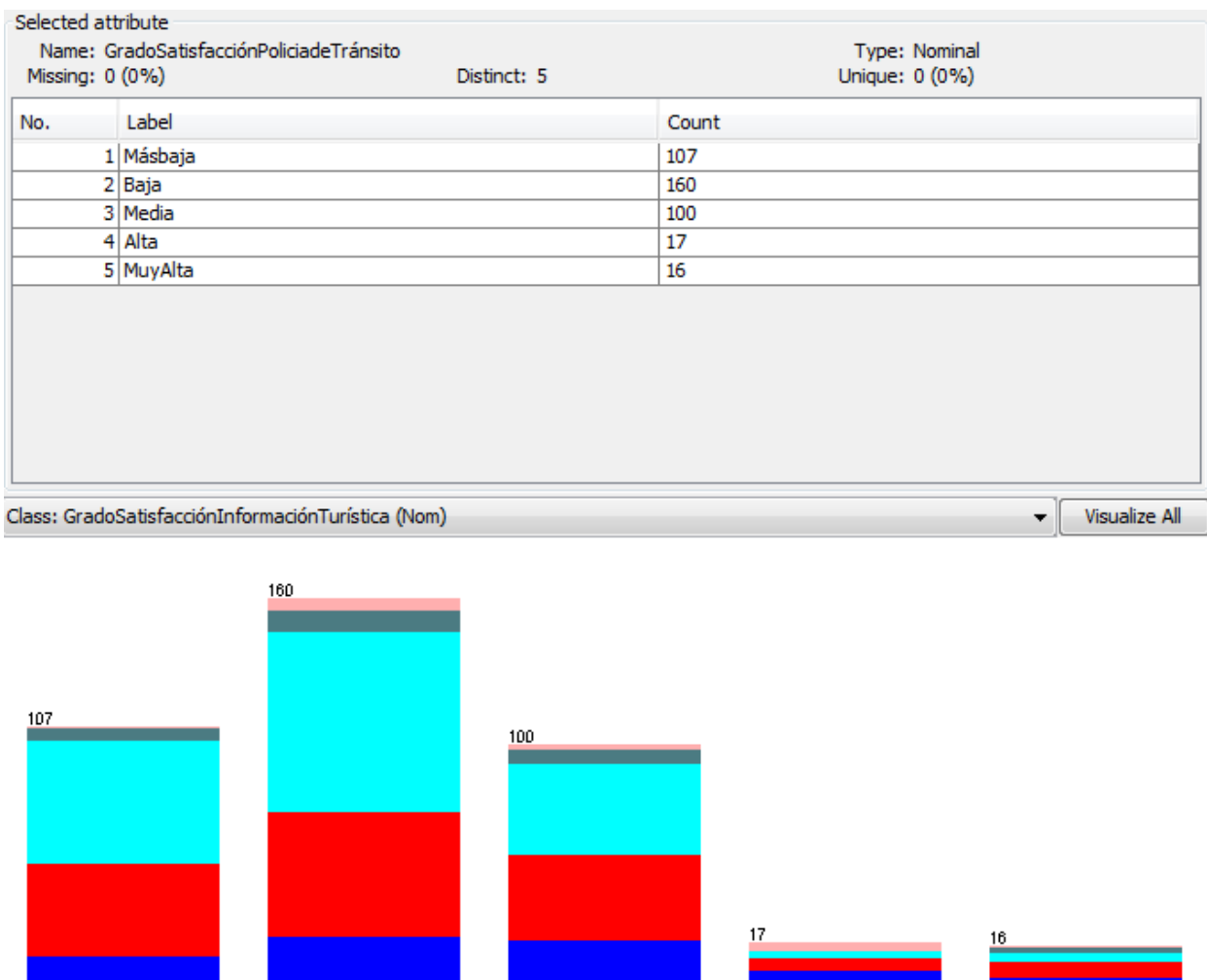
Φιγυρα 43. Ατριβυτο ΓραδοΣατισφαχχι ίνΠερσωναενλαΧαλλε δε λοσ ενχυεσταδοσ. . Αδαπαδο δε: Προχεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

El Grado de Satisfacción por el servicio prestado por parte de la Policía es Baja con un 42.5%, pues los viajeros afirman que durante el tiempo que estuvieron en la ciudad muy poco o casi nunca notaron la presencia de estos, lo cual provocaba que en los encuestados se sintiera un poco de inseguridad.



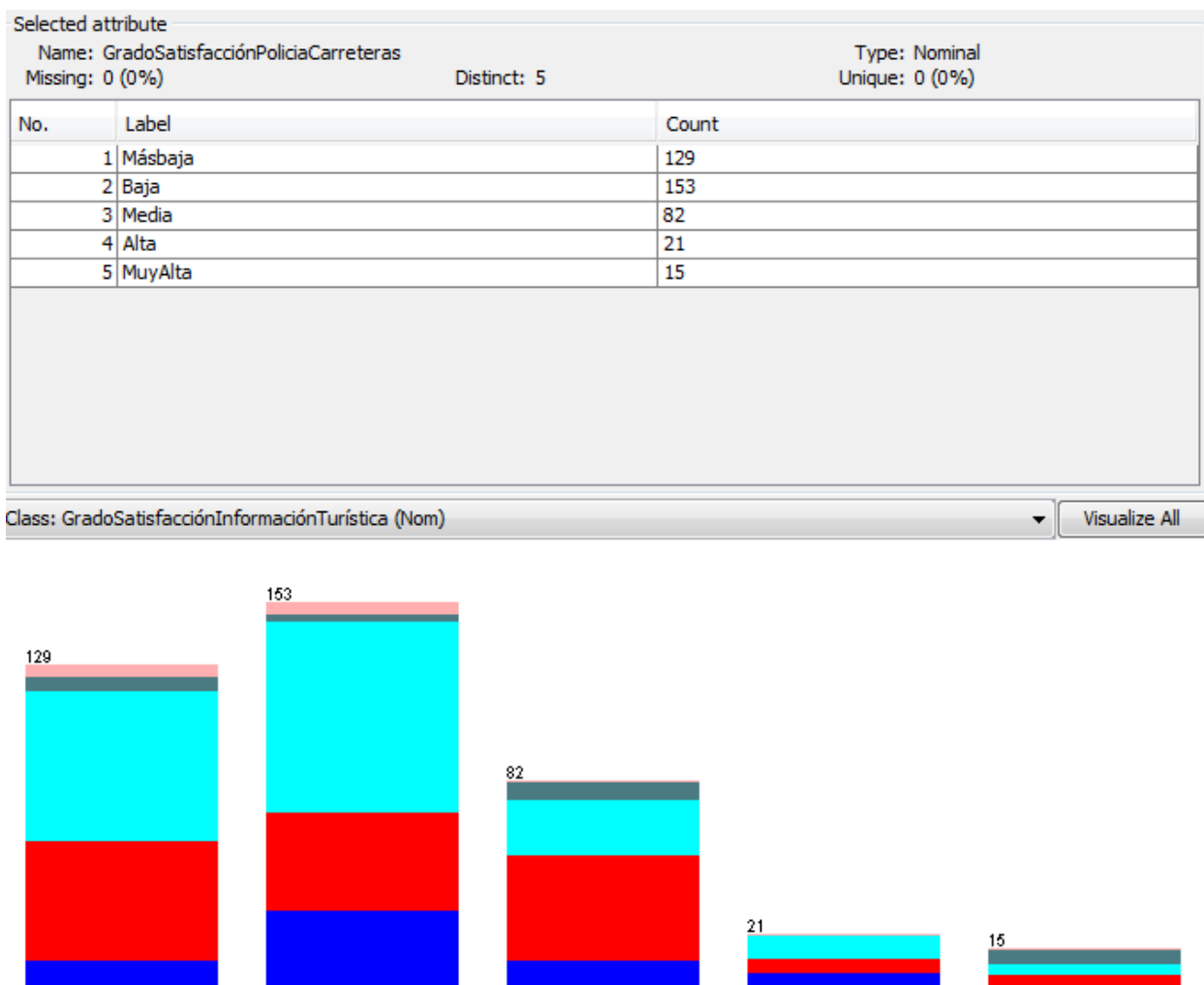
Φιγυρα 44. Ατριβυτο ΓραδοΣατισφαχχι ΓνΠολιχια δε λοσ ενχυεσταδοσ. . Αδαπταδο δε: Προχε σο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

El Grado de Satisfacción por el servicio prestado por parte de la Policía de Tránsito es Baja con un 40%, pues los viajeros afirman que durante la estadía en la ciudad en pocas ocasiones se pudo ver la presencia de estos brindando un poco de orden y cumplimiento de las normas de tránsito para la movilidad dentro de Bucaramanga.



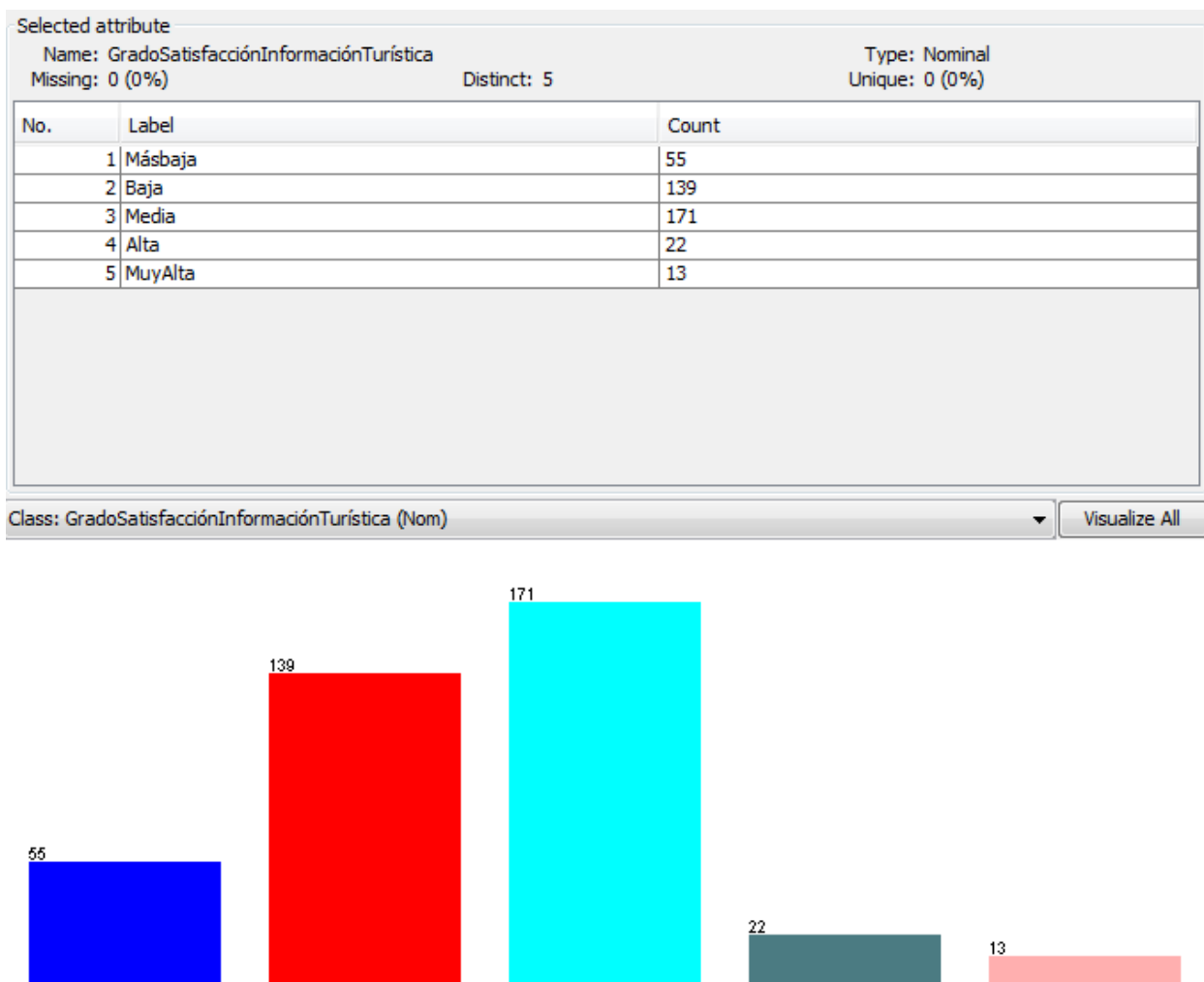
Φιγυρα 45. Ατριβυτο ΓραδοΣατισφαχχι (νΠολιχιαδεΤρ(νσιτο δε λοσ ενχυεσταδος. . Αδαπταδο δε: Προχεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

El Grado de Satisfacción por el servicio prestado por parte de la Policía de Carreteras es Baja con un 38.25%, pues los viajeros afirman que durante el viaje para llegar a la ciudad notaron que poco era el acompañamiento de estos en las rutas de acceso, para así controlar y brindar seguridad a los viajeros y conductores que arriban a Bucaramanga, y permitir que el viaje sea más seguro y confiable para aquellos que lo realicen.



Φιγυρα 46. Ατριβυτο ΓραδοΣατισφαχχι ΓνΠολιχιαΧαρρετερασ δε λοσ ενχυεσταδοσ. . Αδαπταδο δε: Προχεςο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

El Grado de Satisfacción por la Información Turística brindada a los viajeros es Media con un 42.75%, donde ellos afirman que la información a la que pudieron tener acceso era la que en pocas ocasiones brindaban las personas de la calle, algunas empresas transportadoras, anuncios publicitarios que se visualizaban en la ciudad o simplemente como ya conocían Bucaramanga, tenían conocimiento de lugares turísticos para visitar.



Φιγυρα 47. Ατριβυτο ΓραδοΣατισφαχχι (νΙνφορμαχι (νΤυρ)στιχα δε λοσ ενχυεσταδοσ. .Αδαπταδο δε: Προχεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

Analizando los resultados del proceso, observamos que la normalización realizada con anterioridad no permite alteración de los datos y que a su vez, que estén muy bien definidos en escala y cantidad.

La presente investigación agrupa datos de acuerdo al Método Clustering y no determinamos anticipadamente valores para margen de error en estadísticas. Además de no conocer la muestra total del país, ya que hay viajeros que ingresan por aire, otros por tierra, no sabemos, no sabemos

si todos los que diligenciaron la encuesta son realmente extranjeros, no sabemos a ciencia cierta su todos respondieron adecuadamente, entre otros aspectos.

Es por esto que tomamos datos completos, tomando diferentes variables en el Clustering para hacer experimentos, y de este modo decidir si analizar nuevamente todos los datos o tomar una muestra. Todo esto depende del análisis resultante y de las relaciones visibles en los histogramas.

De esta manera se excluyen todas las variables tenidas en cuenta para el análisis, dejando únicamente aquellas que nos aporten la información necesaria para dicha investigación. Estas variables son: “Edad”, “Género”, “CiudadOrigen”, “Ocupación”, “GastoPromedio”, “CalificaciónGastoPromedio”, “AspectoqueImpactodelaCiudad” y “AspectoMotivaaRegresar”.

```

Number of iterations: 3
Within cluster sum of squared errors: 1659.0
Missing values globally replaced with mean/mode

Cluster centroids:

Attribute                Full Data                Cluster#
                        (400)                    0
                        (208)                    1
                        (138)                    2
                        (54)
=====
Edad                    18-25Años                26-35Años                18-25Años                18-25Años
Género                  Femenino                 Masculino                 Femenino                 Masculino
CiudadOrigen            Barrancabermeja         Bogotá                   Barrancabermeja         Bogotá
Ocupación               Estudiante               Empleado                 Estudiante               Pensionado/Jubilado
GastoPromedio           $500.000-$700.000      $500.000-$700.000      $300.000-$500.000      $500.000-$700.000
CalificaciónGastoPromedio Medio                     Medio                     Medio                     Alto
AspectoqueImpactodelaCiudad ClimaAgradable          ClimaAgradable          MuchaIndigencia          CiudadRuidosa
AspectoMotivaaRegresar  Todaslasanteriores      Todaslasanteriores      Bajocostodevida         ClimaAgradable

Time taken to build model (full training data) : 0.11 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      208 ( 52%)
1      138 ( 35%)
2       54 ( 14%)
    
```

Φιγυρα 48. Ανάλισις δε Χλ/στερ τοταλεσ εσταδ'στιχοσ δε λοσ πιαφεροσ. . Αδαπταδο δε: Ρεσυλταδο Οβτενιδο δεσδε ελ σοφτωαρε ΩΕΚΑ.

Ya que lo datos obtenidos hacen referencia a personas con Edad, Género, Ciudad de Origen, Ocupación, Gasto Promedio, Calificación de Gasto Promedio, Aspecto que lo Impacto de la ciudad y Aspecto que lo motivaría a regresar; estos datos analizados contendrán varias tablas en las cuales se observa los datos analizados con sus frecuencias y una tabla de frecuencias relativas:

Tabla 43.

Edad viajeros que visitan la ciudad de Bucaramanga.

Edad				
	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Menos de 18 Años	37	9,3	9,3	9,3
18 - 25 Años	104	26,0	26,0	35,3
26 - 35 Años	86	21,5	21,5	56,8
36 - 45 Años	64	16,0	16,0	72,8
Válidos 46 - 55 Años	44	11,0	11,0	83,8
56 -65 Años	41	10,3	10,3	94,0
66 - 75 Años	17	4,3	4,3	98,3
Más de 75 Años	7	1,8	1,8	100,0
Total	400	100,0	100,0	

Nota: Resultado Obtenido desde el software SPSS. Adaptado de: Software SPSS.

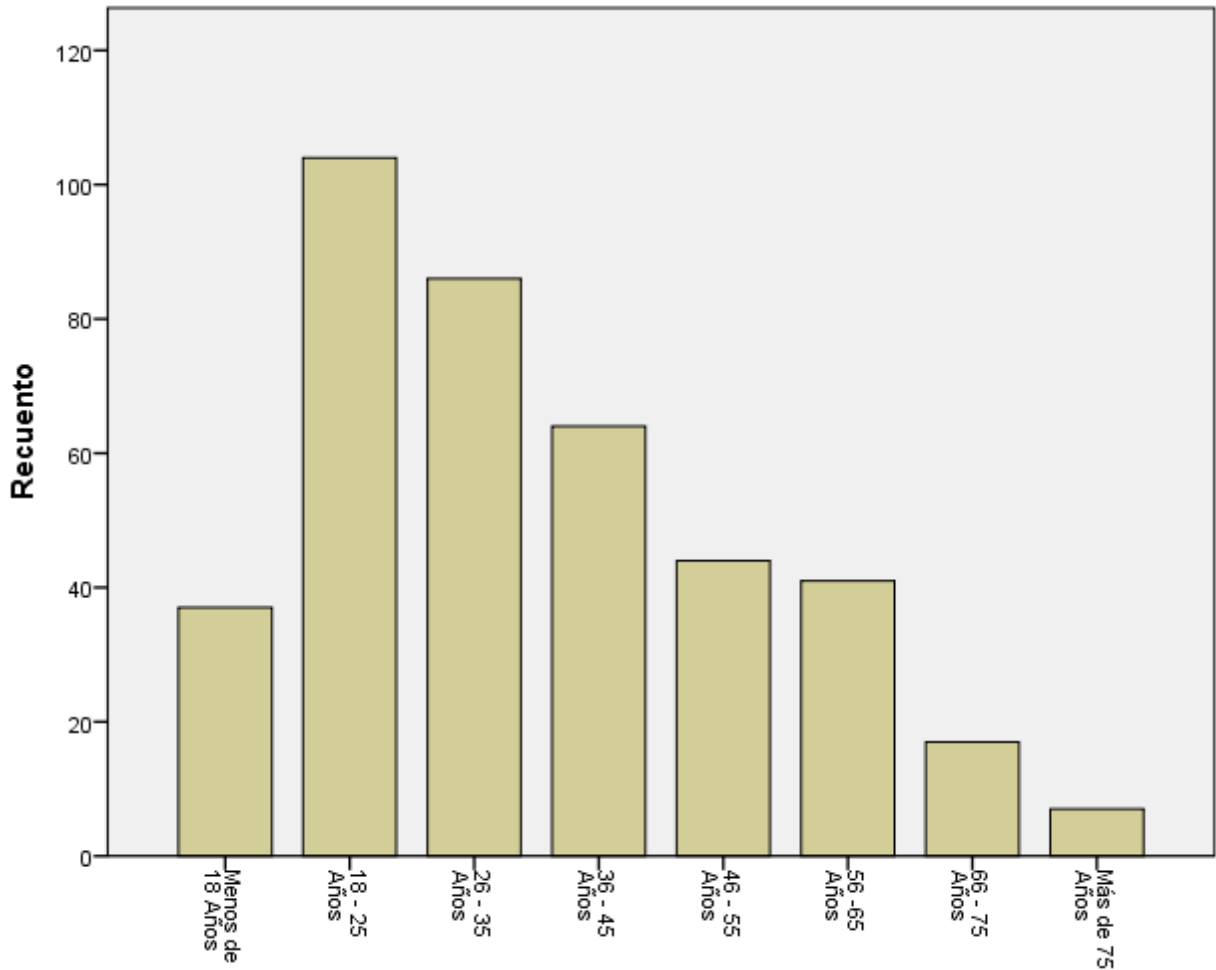


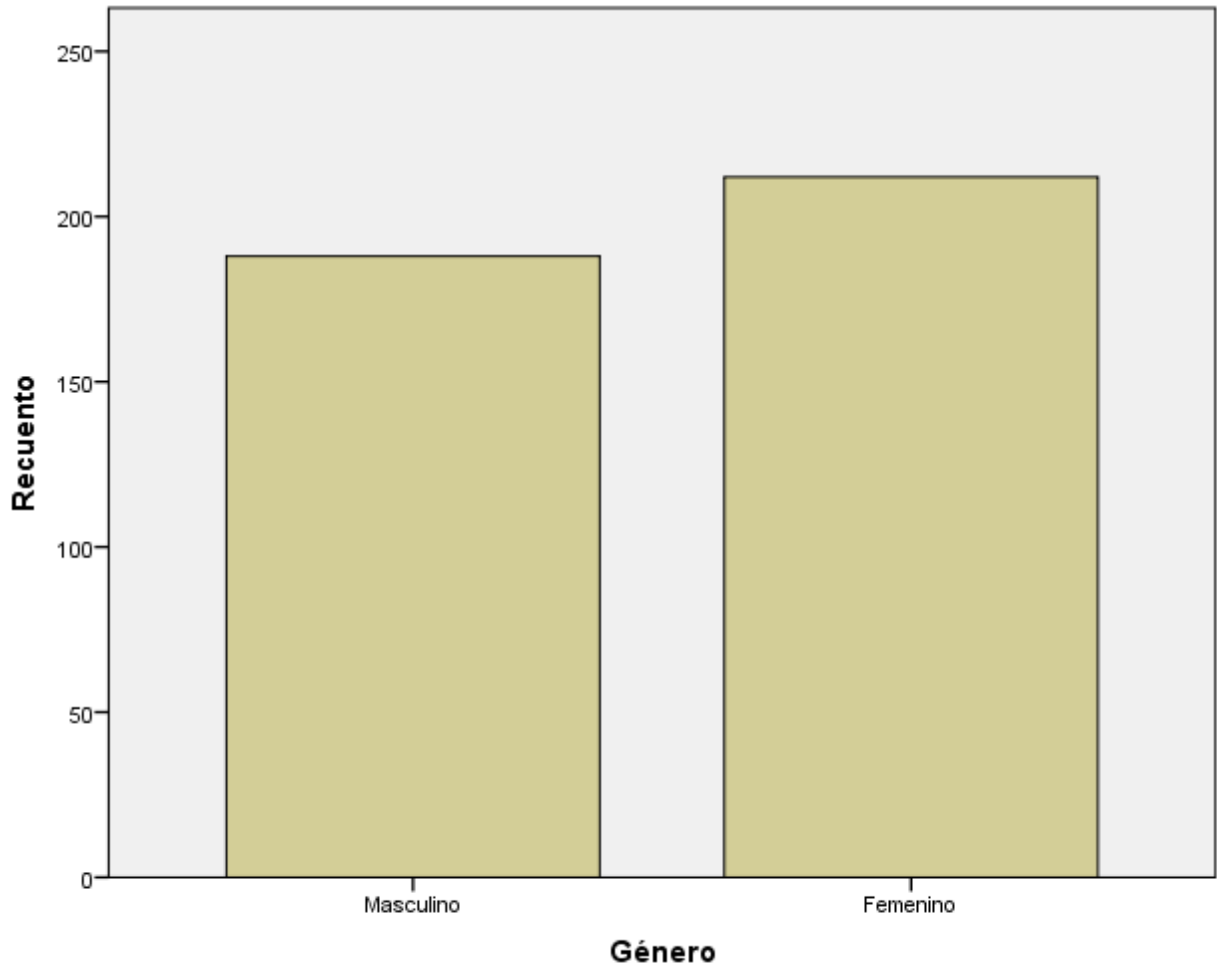
Figura 49. Edad piaferos θve πισιταν λα χιυδαδ δε Βυχαραμανγα. . Αδαπταδο δε: Ρεσυλαδ ο Οβτενιδο δεσδε ελ σοφτωαρε ΣΠΣΣ.

Tabla 44.

Género viajeros que visitan la ciudad de Bucaramanga.

Género				
	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Masculino	188	47,0	47,0	47,0
Válidos Femenino	212	53,0	53,0	100,0
Total	400	100,0	100,0	

Nota: Resultado Obtenido desde el software SPSS. Adaptado de: Software SPSS.



Φιγυρα 50. Γένερο πιαφεροσ θνε πισιταν λα χιυδαδ δε Βυχαραμανγα. . Αδαπταδο δε: Ρεσυλταδ ο Οβτενιδο δεσδε ελ σοφτωαρε ΣΠΣΣ.

Tabla 45.

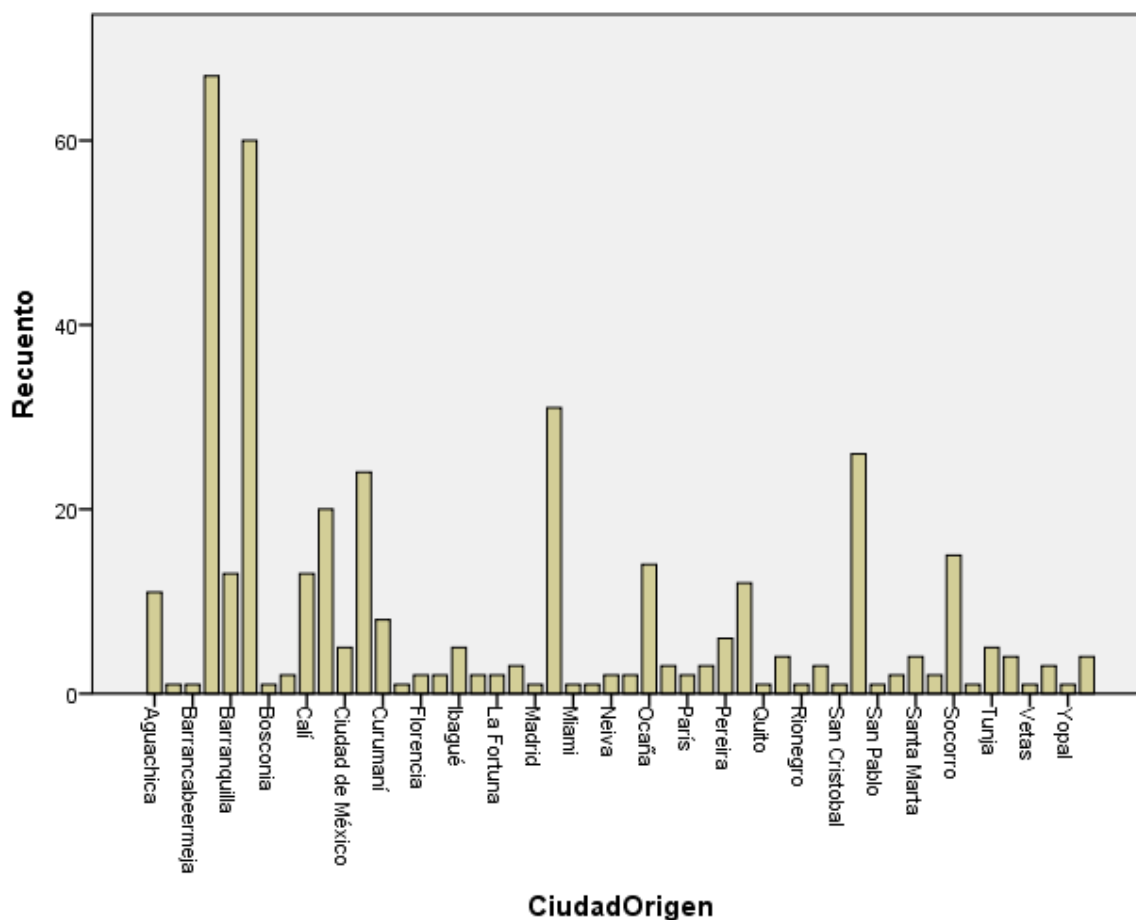
CiudadOrigen viajeros que visitan la ciudad de Bucaramanga.

CiudadOrigen				
	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos				
Aguachica	11	2,8	2,8	2,8
Arenal	1	,3	,3	3,0
Barrancabeermeja	1	,3	,3	3,3
Barrancabermeja	67	16,8	16,8	20,0
Barranquilla	13	3,3	3,3	23,3
Bogotá	60	15,0	15,0	38,3

Bosconia	1	,3	,3	38,5
Buenos Aires	2	,5	,5	39,0
Calí	13	3,3	3,3	42,3
Cartagena	20	5,0	5,0	47,3
Ciudad de México	5	1,3	1,3	48,5
Cúcuta	24	6,0	6,0	54,5
Curumaní	8	2,0	2,0	56,5
Duitama	1	,3	,3	56,8
Florencia	2	,5	,5	57,3
Guatapé	2	,5	,5	57,8
Ibagué	2	,5	,5	59,5
Ibagué	5	1,3	1,3	59,0
La Fortuna	2	,5	,5	60,0
La Guajira	3	,8	,8	60,8
Madrid	1	,3	,3	61,0
Medellín	31	7,8	7,8	68,8
Miami	1	,3	,3	69,0
Montería	1	,3	,3	69,3
Neiva	2	,5	,5	69,8
Nueva York	2	,5	,5	70,3
Ocaña	14	3,5	3,5	73,8
Pamplona	3	,8	,8	74,5
París	2	,5	,5	75,0
Pasto	3	,8	,8	75,8
Pereira	6	1,5	1,5	77,3
Puerto Wilches	12	3,0	3,0	80,3
Quito	1	,3	,3	80,5
Riohacha	4	1,0	1,0	81,5
Rionegro	1	,3	,3	81,8
Sabana de Torres	3	,8	,8	82,5
San Cristobal	1	,3	,3	82,8
San gil	26	6,5	6,5	89,3
San Pablo	1	,3	,3	89,5
San Vicente	2	,5	,5	90,0
Santa Marta	4	1,0	1,0	91,0

Sarabena	2	,5	,5	91,5
Socorro	15	3,8	3,8	95,3
Sopetran	1	,3	,3	95,5
Tunja	5	1,3	1,3	96,8
Valledupar	4	1,0	1,0	97,8
Vetas	1	,3	,3	98,0
Villavicencio	3	,8	,8	98,8
Yopal	1	,3	,3	99,0
Zapatoca	4	1,0	1,0	100,0
Total	400	100,0	100,0	

Nota: Resultado Obtenido desde el software SPSS. Adaptado de: Software SPSS.



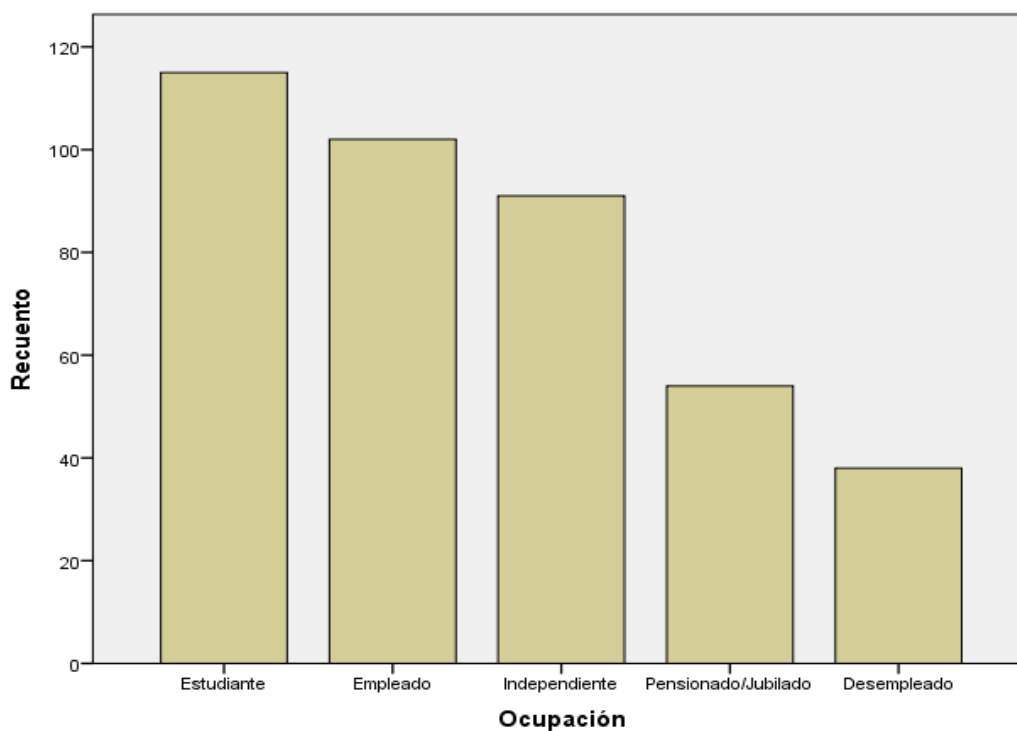
Φιγυρα 51. ΧιωδαδΟριγεν πιαφεροσ θυε πισιταν λα χιωδαδ δε Βυχαραμανγα. . Αδαπαδο δε: Ρεσυλταδο Οβτενιδο δεσδε ελ σοφτωαρε ΣΠΣΣ.

Tabla 46.

Ocupación viajeros que visitan la ciudad de Bucaramanga.

		Ocupación			
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Estudiante	115	28,8	28,8	28,8
	Empleado	102	25,5	25,5	54,3
	Independiente	91	22,8	22,8	77,0
	Pensionado/Jubilado	54	13,5	13,5	90,5
	Desempleado	38	9,5	9,5	100,0
	Total	400	100,0	100,0	

Nota: Resultado Obtenido desde el software SPSS. Adaptado de: Software SPSS.



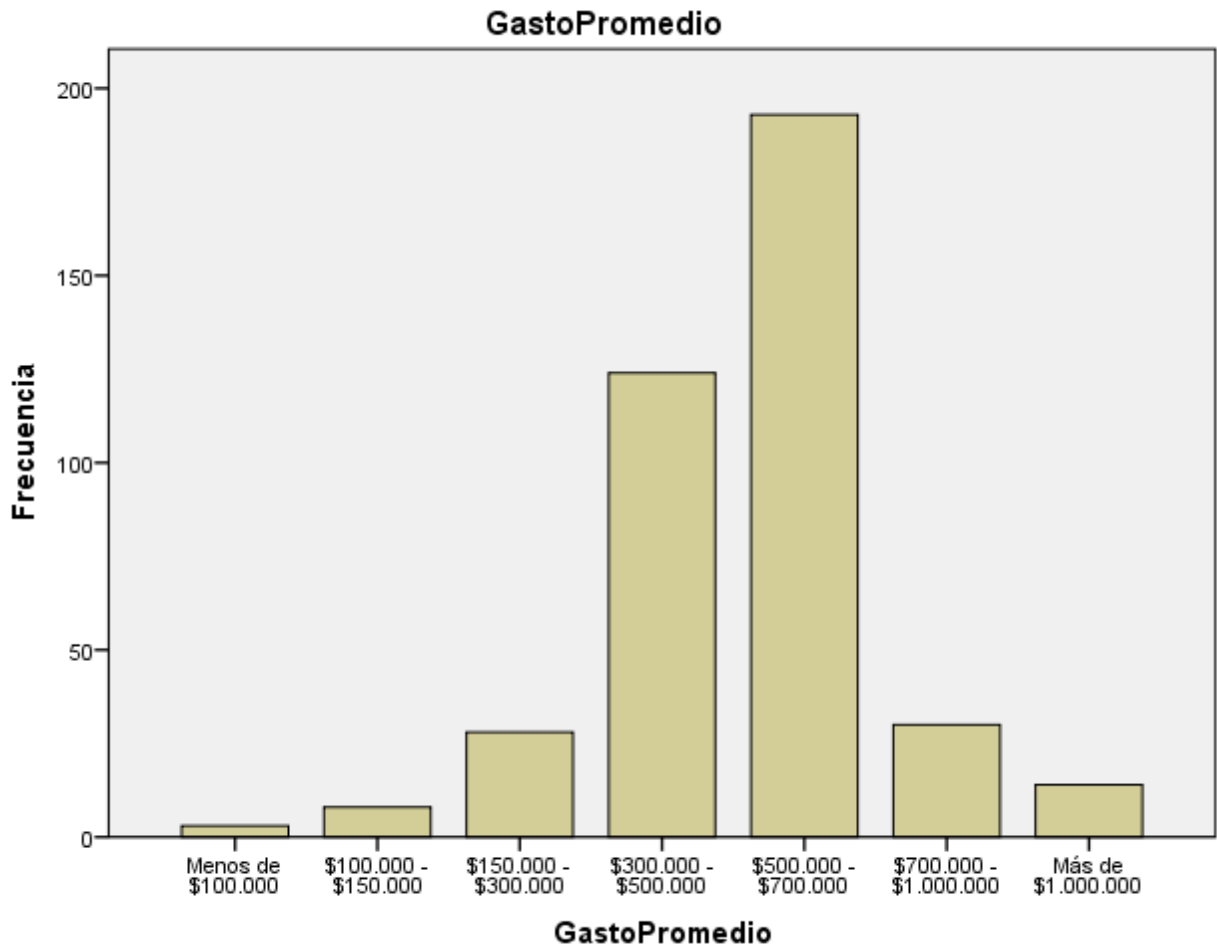
Φιγυρα 52. Οχυπαχι ίν πιαφεροσ θυε πισιταν λα χυδαδ δε Βυχαραμανγα. . Αδαπταδο δε: Ρε συλταδο Οβτενιδο δεσδε ελ σοφτωαρε ΣΠΣΣ.

Tabla 47.

GastoPromedio viajeros que visitan la ciudad de Bucaramanga.

GastoPromedio				
	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Menos de \$100.000	3	,8	,8	,8
\$100.000 - \$150.000	8	2,0	2,0	2,8
\$150.000 - \$300.000	28	7,0	7,0	9,8
\$300.000 - \$500.000	124	31,0	31,0	40,8
\$500.000 - \$700.000	193	48,3	48,3	89,0
\$700.000 - \$1.000.000	30	7,5	7,5	96,5
Más de \$1.000.000	14	3,5	3,5	100,0
Total	400	100,0	100,0	

Nota: Resultado Obtenido desde el software SPSS. Adaptado de: Software SPSS.



Φιγυρα 53. ΓαστοΠρομεδιο πιαφεροσ θυε πισιταν λα χιυδαδ δε Βυχαραμανγα. . Αδαπταδο δε: Ρεσυλταδο Οβτενιδο δεσδε ελ σοφτωαρε ΣΠΣΣ.

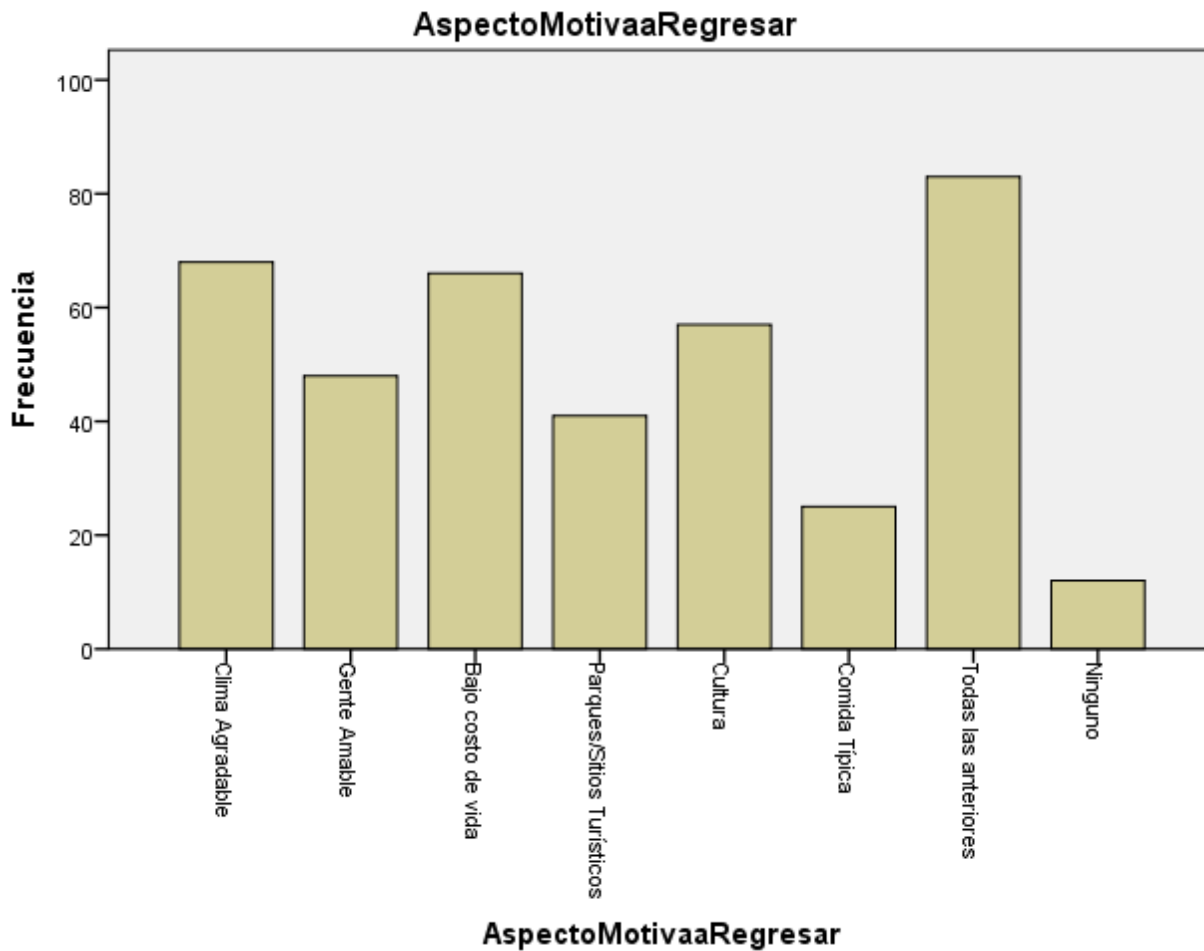
Tabla 48.

AspectoMotivaaRegresar viajeros que visitan la ciudad de Bucaramanga.

AspectoMotivaaRegresar					
	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado	
Válidos					

Comida Típica	25	6,3	6,3	76,3
Todas las anteriores	83	20,8	20,8	97,0
Ninguno	12	3,0	3,0	100,0
Total	400	100,0	100,0	

Nota: Resultado Obtenido desde el software SPSS. Adaptado de: Software SPSS.



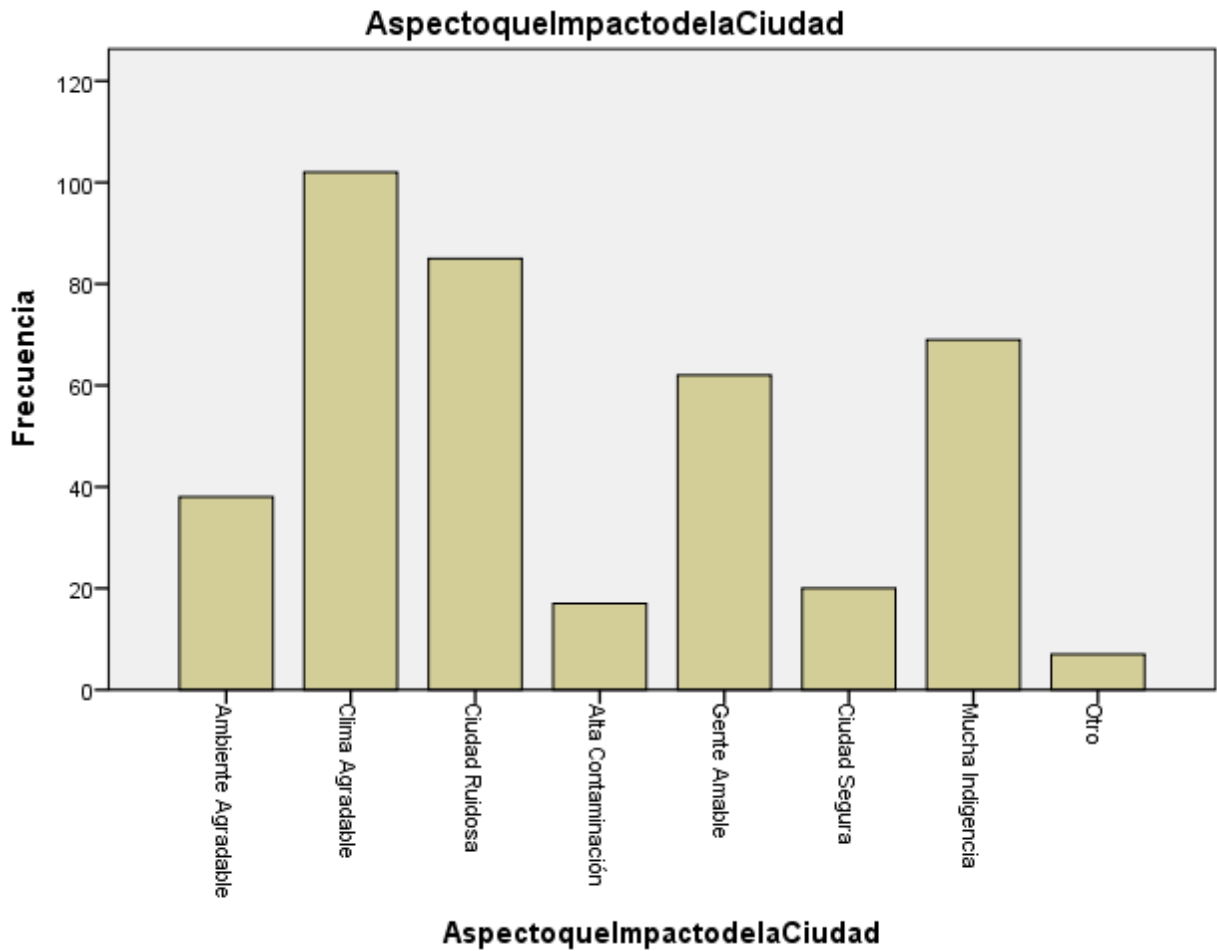
Φιγυρα 54. ΑσπεχτοΜοτιπιααΡεγρεσαρ πιαφεροσ θυε πισιταν λα χυδαδ δε Βυχαραμανγα. .
 Αδαπταδο δε: Ρεσυλταδο Οβτενιδο δεσδε ελ σοφτωαρε ΣΠΣΣ.

Tabla 49.

Aspecto que impacta de la Ciudad viajeros que visitan la ciudad de Bucaramanga.

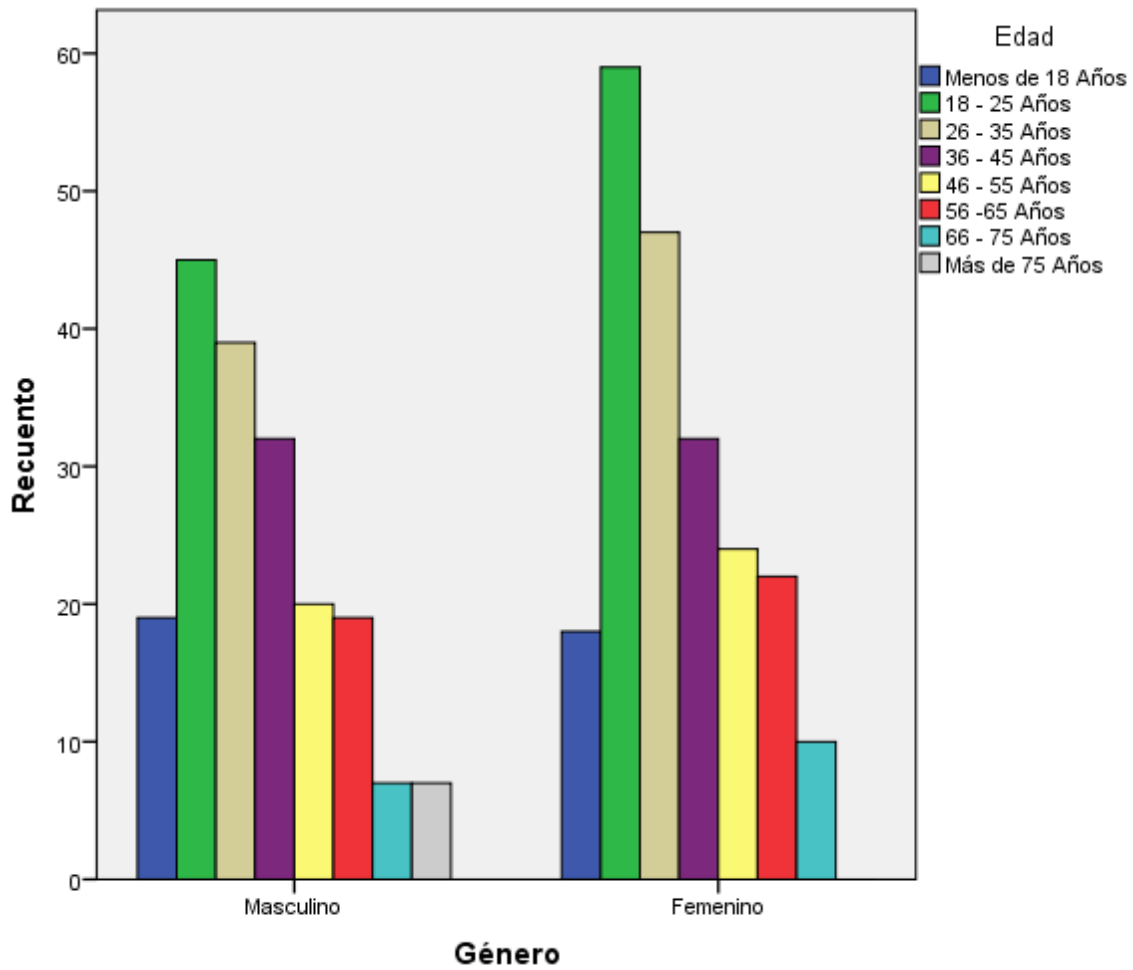
Aspecto que Impacta de la Ciudad				
	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Ambiente Agradable	38	9,5	9,5	9,5
Clima Agradable	102	25,5	25,5	35,0
Ciudad Ruidosa	85	21,3	21,3	56,3
Alta Contaminación	17	4,3	4,3	60,5
Válidos Gente Amable	62	15,5	15,5	76,0
Ciudad Segura	20	5,0	5,0	81,0
Mucha Indigencia	69	17,3	17,3	98,3
Otro	7	1,8	1,8	100,0
Total	400	100,0	100,0	

Nota: Resultado Obtenido desde el software SPSS. Adaptado de: Software SPSS.

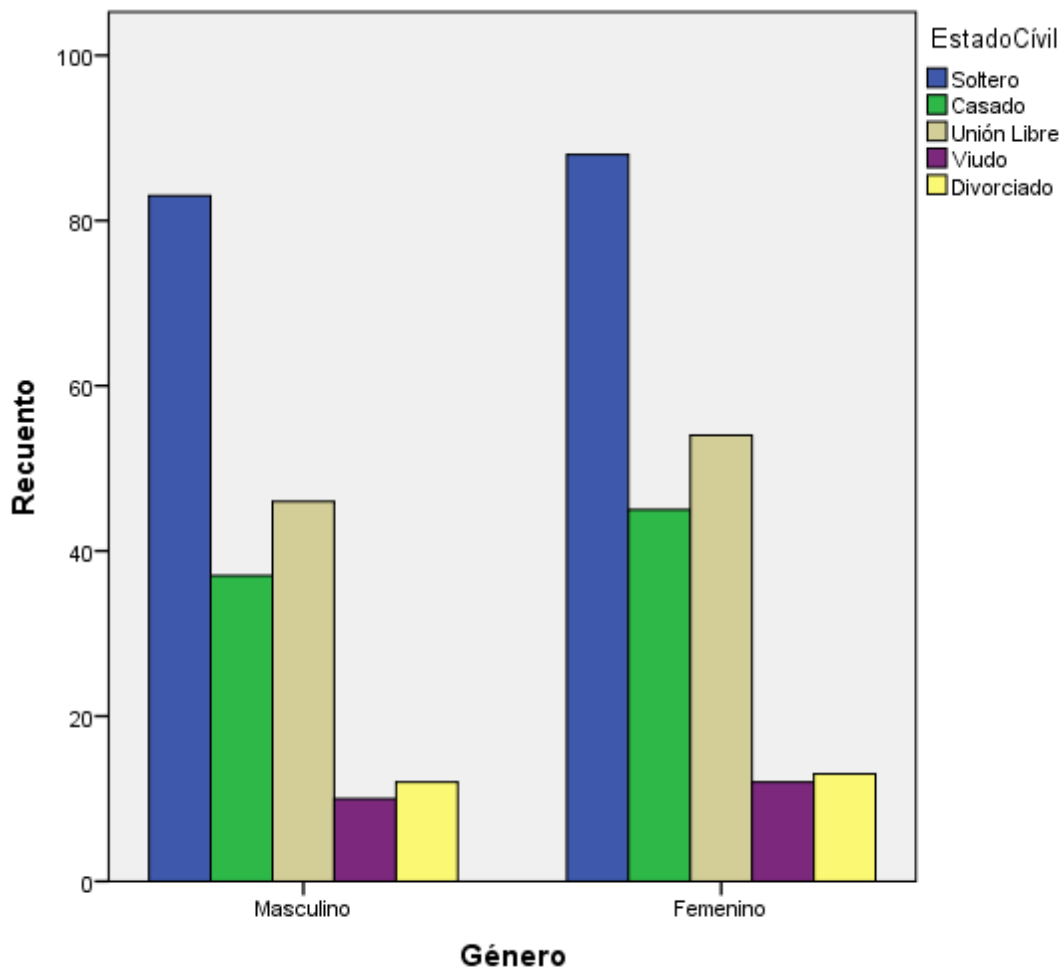


Φιγυρα 55. ΑσπεχτοθυεΙμπαχτοδελαΧιυδαδ πιαφεροσ θυε πισιταν λα χιυδαδ δε Βυχαραμανγα. . Αδαπταδο δε: Ρεσυλταδο Οβτενιδο δεσδε ελ σοφτωαρε ΣΠΣΣ.

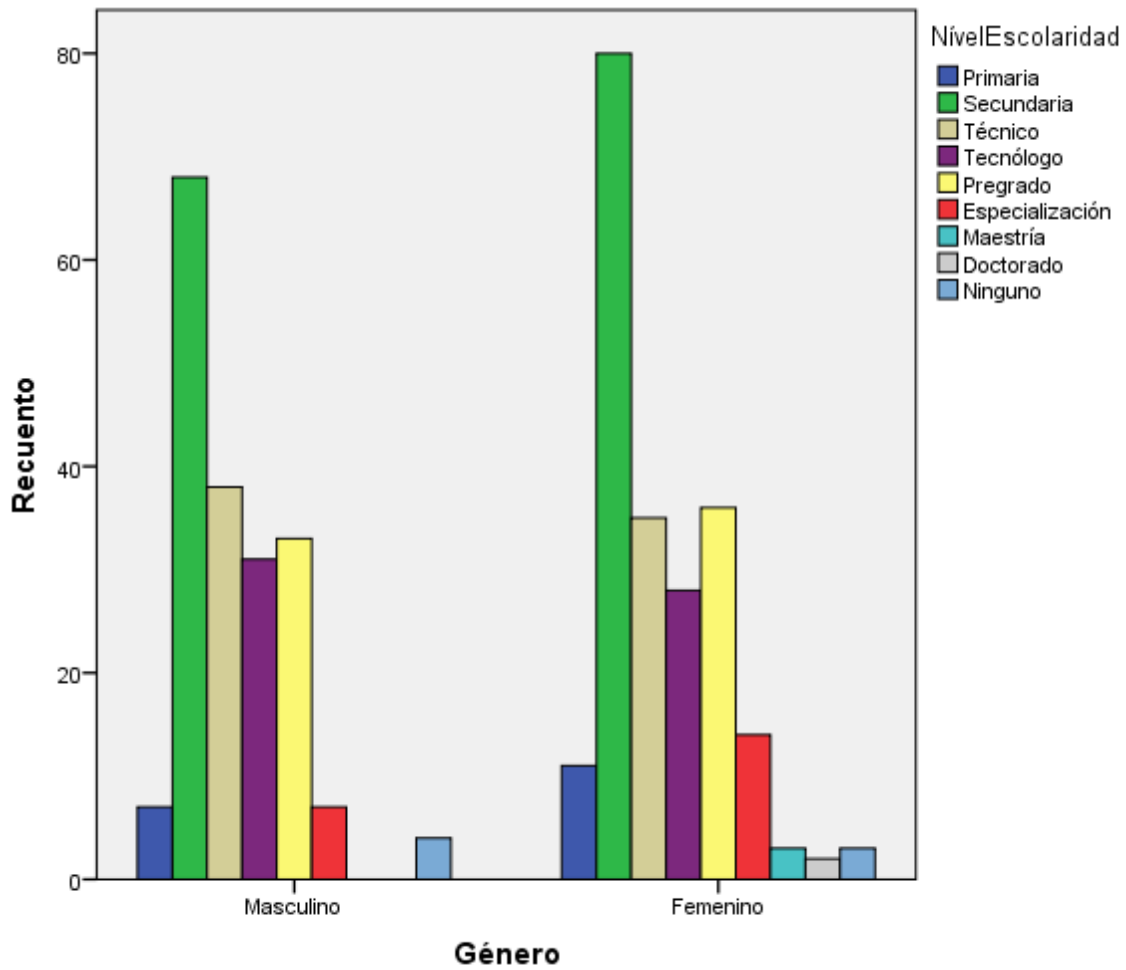
A continuación se describen varios Histogramas con dos variables, tomando en cuenta las consideradas anteriormente. Con estos diagramas tendremos más claridad sobre el perfilamiento y caracterización de los viajeros.



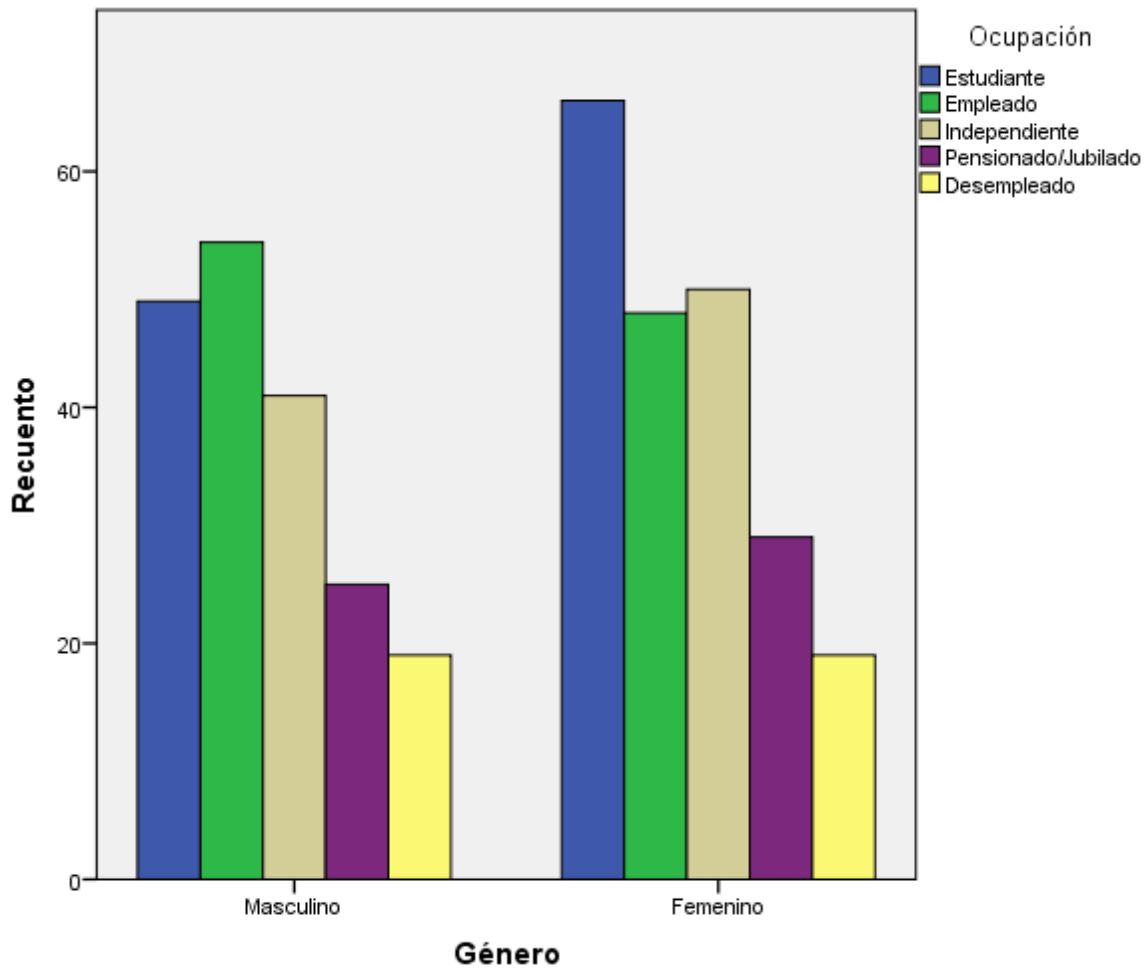
Φιγυρα 56. Ηιστογρμμ Γνερμ-Εδμδ.. Αδμπτμδμ δε: Ρεσυλτμδμ Οβτενιδμ δεσδε ελ σφτμωρε Σ ΠΣΣ.



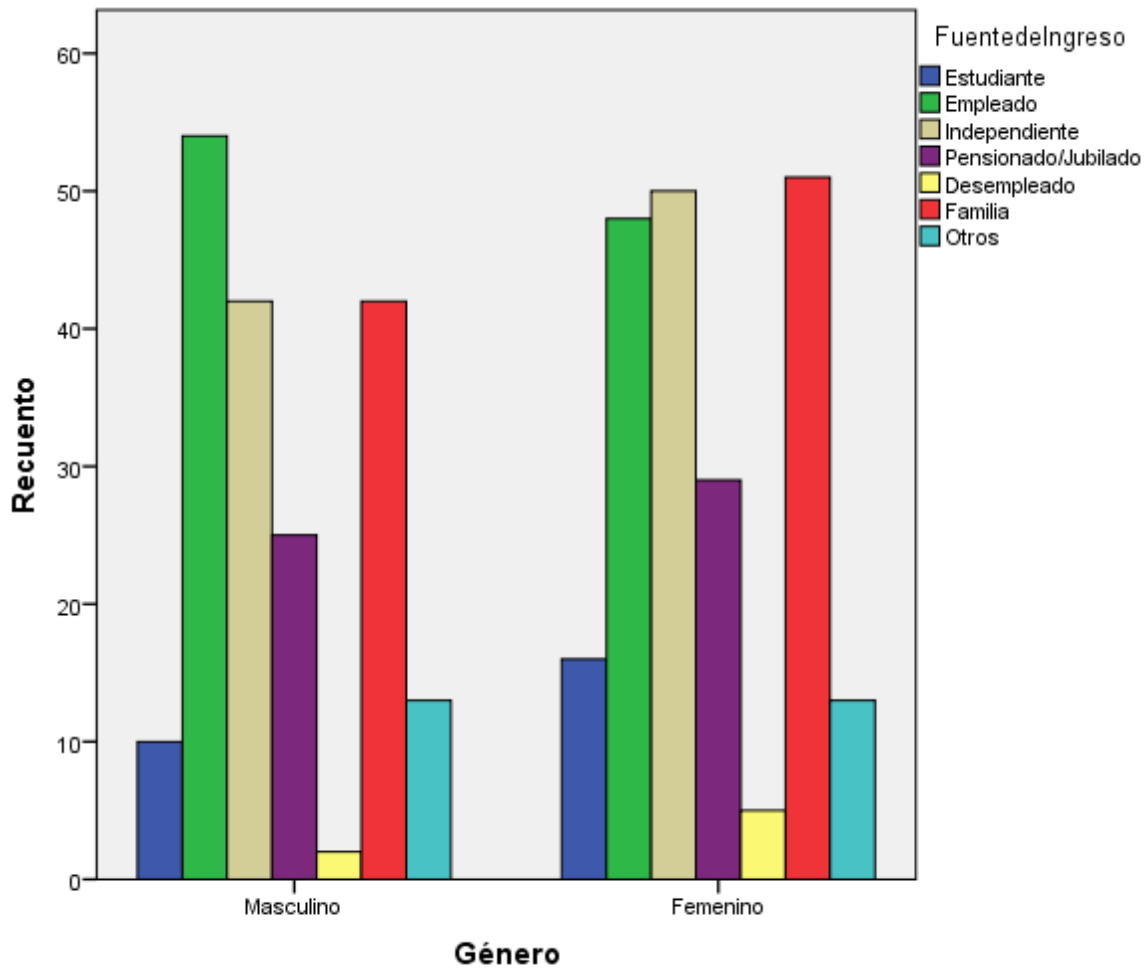
Φιγυρα 57. Ηιστογραιο Γένερο-ΕσταδοΧ΄πιλ.. Αδαπταδο δε: Ρεσυλταδο Οβτενιδο δεσδε ελ σοφτωαρε ΣΠΣΣ.



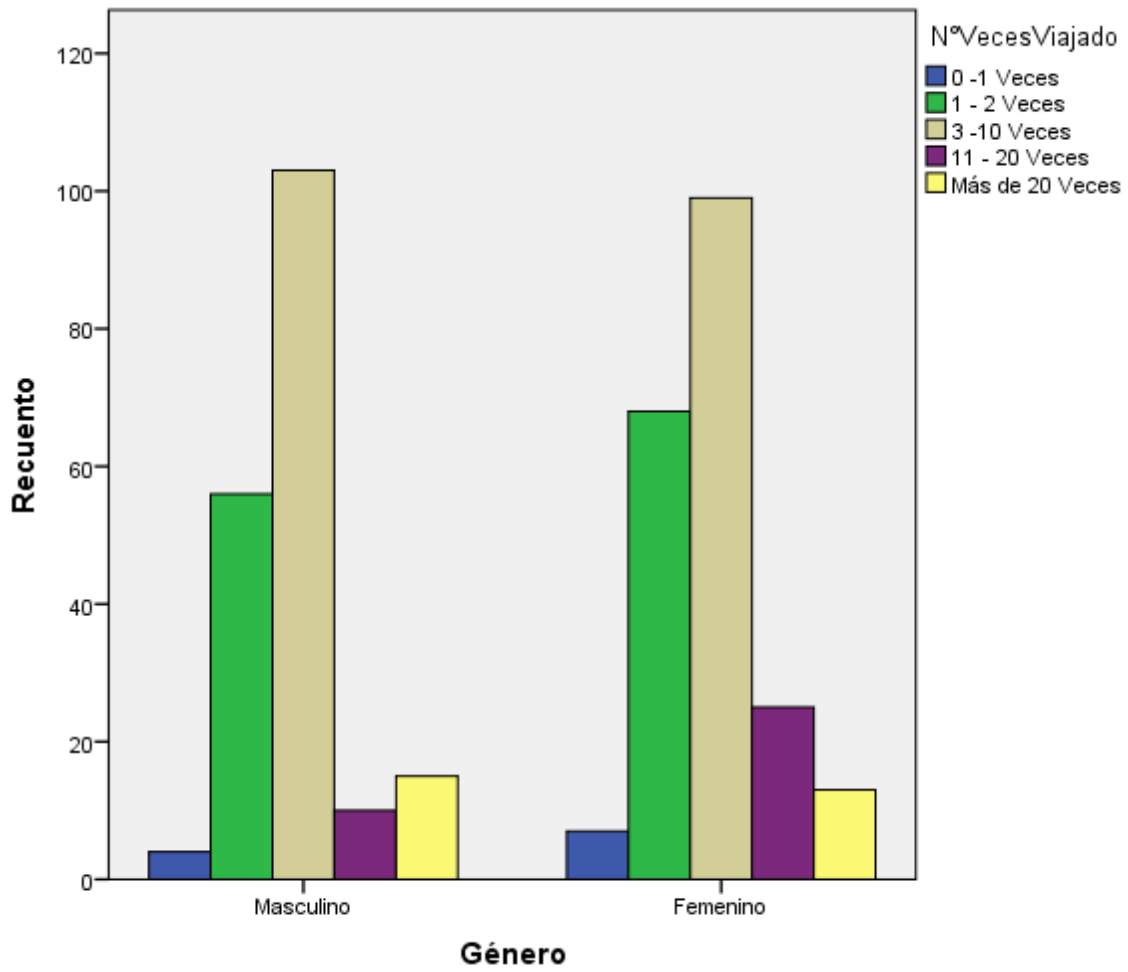
Φιγυρα 58. Ηιστογραμα Γ̅νερο-Νιπελεσχολαριδαδ. . Αδαπαδο δε: Ρεσυλαδο Οβτενιδο δεσδε ελ σοφτωαρε ΣΠΣΣ.



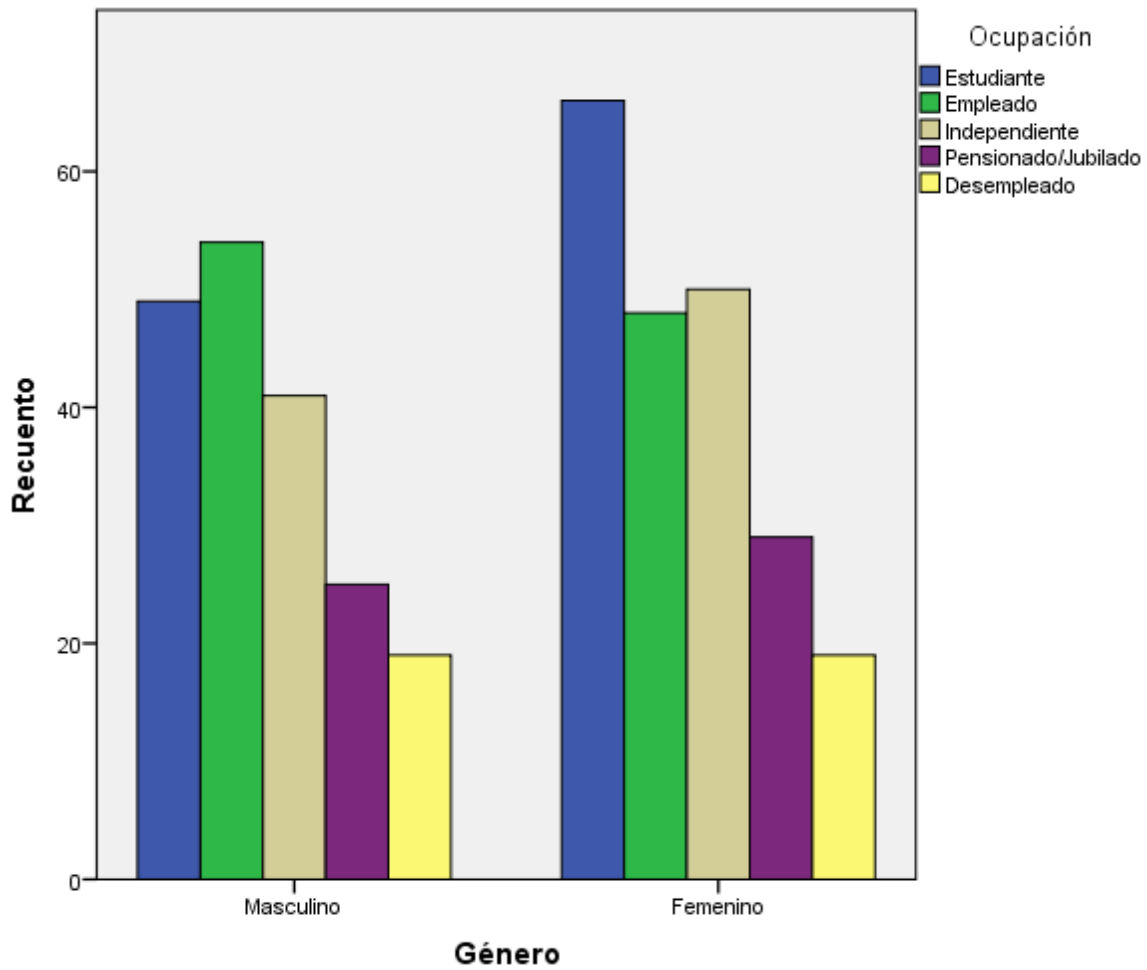
Φιγυρα 59. Ηιστογραμα Γ̅νερο-Οχυπαχι (ν. . Αδαπταδο δε: Ρεσυλταδο Οβτενιδο δεσδε ελ σοφτωαρε ΣΠΣΣ.



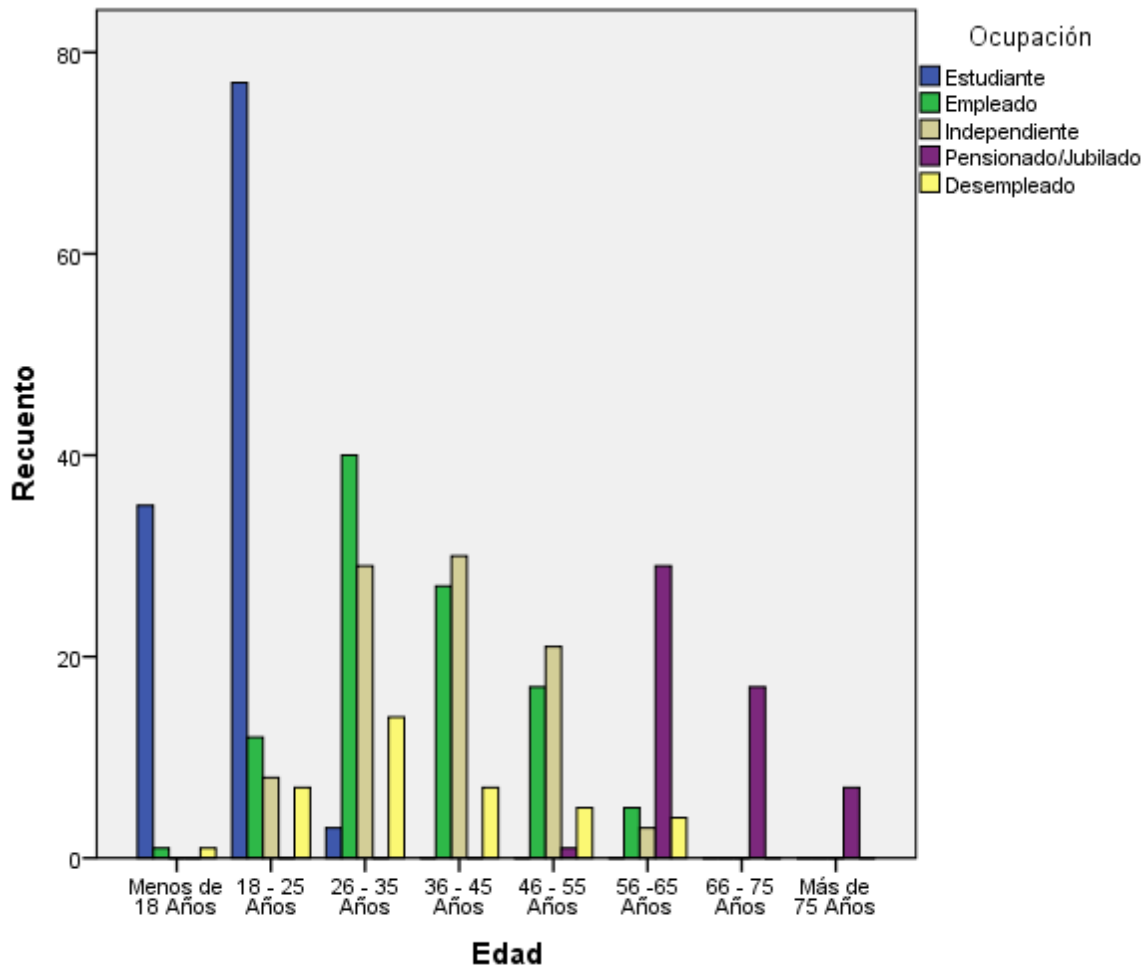
Φιγυρα 60. Ηιστογραμα Γ̅νερο-ΦυεντεδεΙνγρεσο. . Αδαπαδο δε: Ρεσυλαδο Οβτενιδο δεσδε ελ σοφτωαρε ΣΠΣΣ.



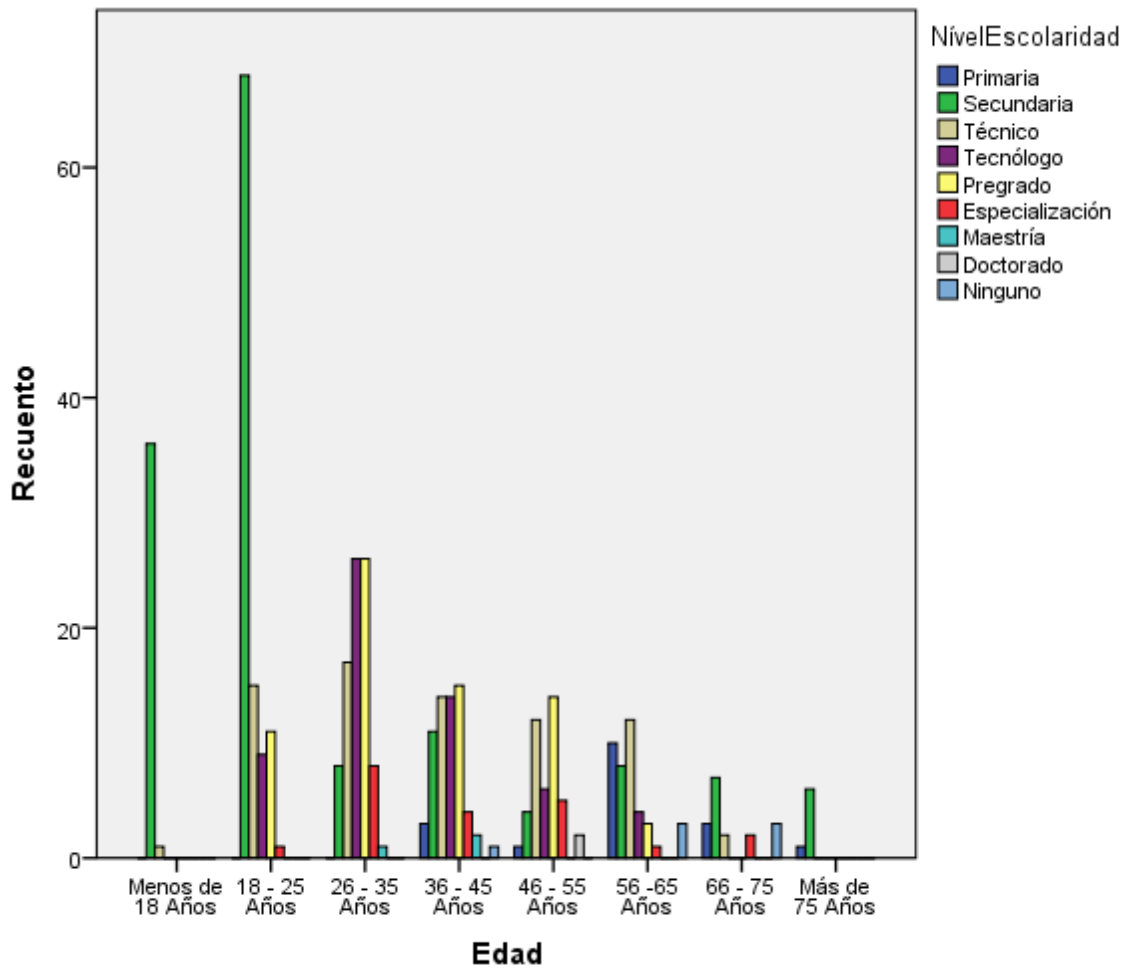
Φιγυρα 61. Ηιστογραμα Γ̅νερο–Ν̅ζεχεσσιαφαδο. . Αδαπαδο δε: Ρεσυλταδο Οβτενιδο δεσδε ελ σοφτωαρε ΣΠΣΣ.



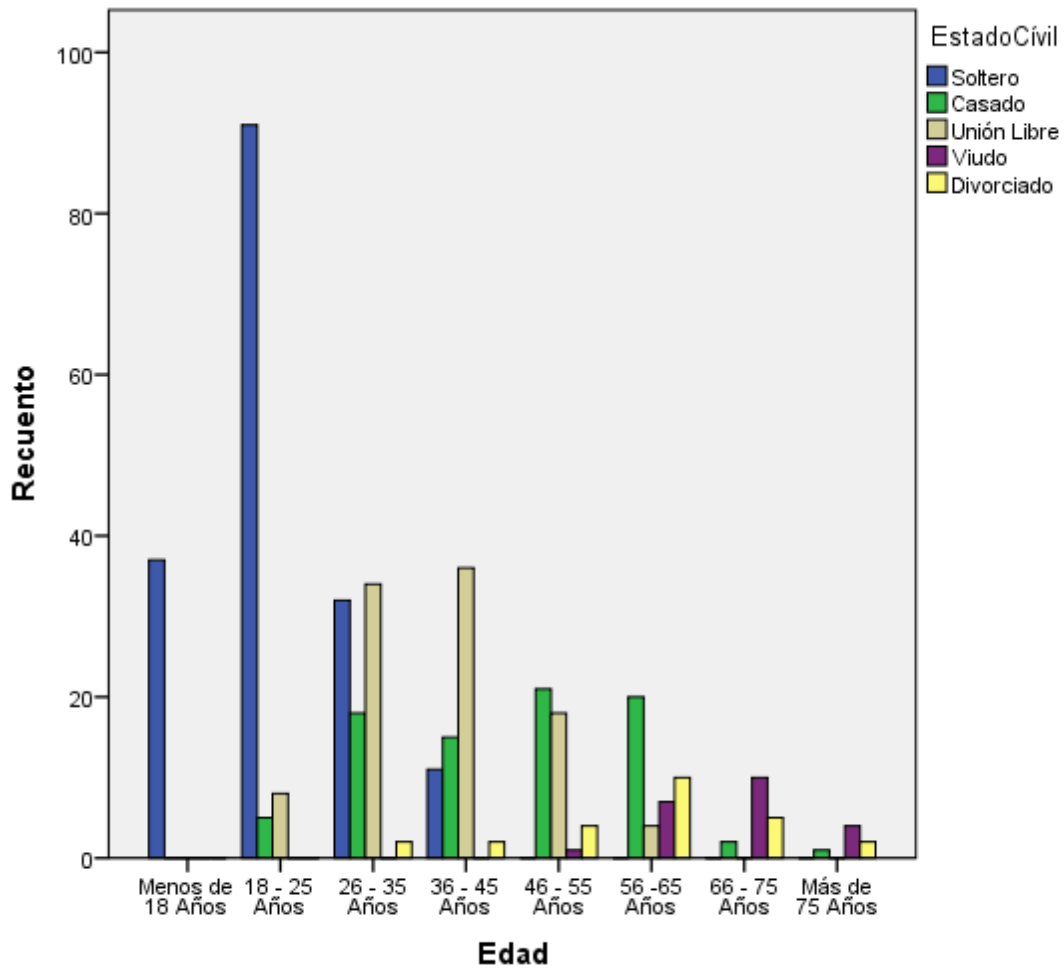
Φιγυρα 62. Ηιστογρραμα Γένερο-Οχυπαχι (ν. . Αδαπταδο δε: Ρεσυλταδο Οβτενιδο δεσδε ελ σοφτωαρε ΣΠΣΣ.



Φιγυρα 63. Ηιστογραιο Εδαδ-Οχυπαχι Γν.. Αδαπαδο δε: Ρεσυλταδο Οβτενιδο δεσδε ελ σοφτωα ρε ΣΠΣΣ.



Φιγυρα 64. Ηιστογραμα Εδαδ–ΝιπελΕσχολαριδαδ. . Αδαπταδο δε: Ρεσυλταδο Οβτενιδο δεσδε ελ σοφτωαρε ΣΠΣΣ.



Φιγυρα 65. Ηιστογραμα Εδαδ–ΕσταδοΧιπιλ. .Αδαπαδο δε: Ρεσυλταδο Οβτενιδο δεσδε ελ σοφτωαρε ΣΠΣΣ.

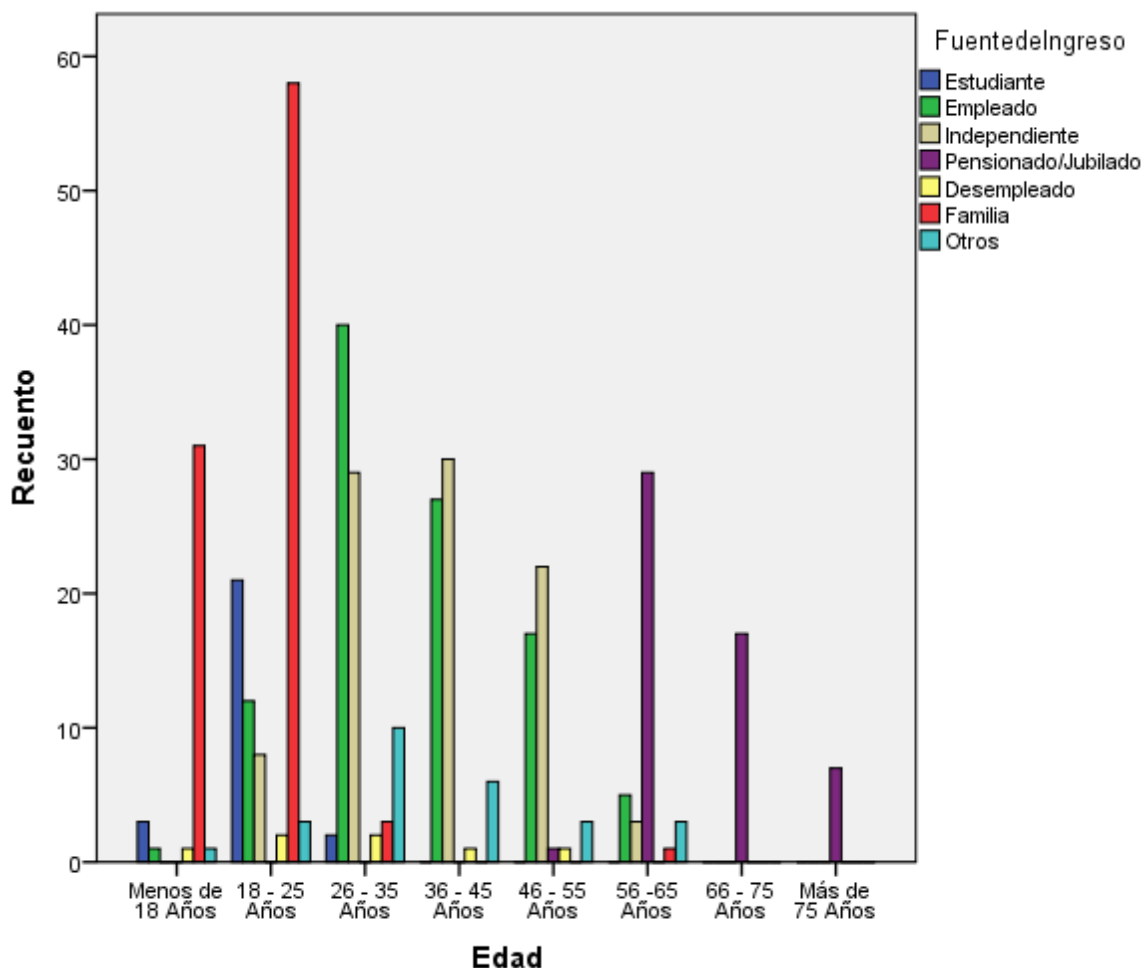


Figura 66. Histograma Edad- Fuente Ingreso. . Aadaptado de: Resultado Obtenido desde el software SPSS.

Se debe recordar que el análisis de minería de datos requiere de un conocimiento del sector en donde se realiza el estudio. A continuación se describen algunas observaciones con unas posibles causas:

1. La variable N°VecesViajado es importante (Ver Figura 63. *Histograma Género-N°VecesViajado*), pues tiene una variación marcada entre los viajeros. Por tanto es interesante analizar las características de las personas que regresan nuevamente a la ciudad. Se puede observar que existe una marcada tendencia en el N°VecesViajado, pues existen datos más extremos o algunos con pocos datos. Estos son:

- 3-10 Veces (Alto número de viaje)
- 0-1 Veces (Bajo número de viaje)

Teniendo en cuenta la información de la ciudad, esta tendencia nos puede indicar que aquellos que pertenecen al grupo de Bajo número de viajes no le interesa regresar o solo viajan por necesidad (Salud). Puede ser que en algún momento llegaran a la ciudad de paso y se quedaron durante cierto día.

El otro grupo de viajeros muy alto, es interesante analizarlo con Clustering ya que nos indicaría la tendencia de ese grupo de personas, argumentando que es posible que visiten familia, amigos, estén realizando negocios en la ciudad, busquen ofertas educativas, sean propietarios de una casa, apartamento o local comercial; o puede ser que les genere algún tipo de interés conocer a fondo la cultura y variedad turística en Bucaramanga.

2. Otra variable a analizar son las Fuentes de Ingreso y Ocupación (Ver Figura 62. *Histograma Género-Fuentes de Ingreso* y Figura 65. *Histograma Edad-Ocupación*). La ciudad se ve beneficiada con la visita y el gasto que realizan los viajeros. La descripción de las personas que visitan la ciudad depende de su fuente de ingreso, puesto esto nos ayuda a inferir aspectos que permiten el desarrollo de actividades para que estas personas regresen o den buenas referencias a otras personas, incentivando un incremento en las visitas a Bucaramanga.

Para esto se deben tener en cuenta ciertos aspectos:

- El mayor número de visitantes está en el grupo de Estudiantes.
- El menor grupo de visitantes son el grupo de Desempleados.

- El tercer grupo que más visita es el de Empleados, pero con una diferencia corta entre dicho grupo y el de Independientes.

La Edad, EstadoCivil, y NivelEscolaridad brindan al estudio datos importantes sobre el nivel socioeconómico, nivel de preparación académica y datos familiares de los viajeros (Ver Figura 66. *Histograma Edad-NivelEscolaridad*, Figura 67. *Histograma Edad-EstadoCivil* y Figura 68. *Histograma Edad-Fuente de Ingreso*). Observamos tendencias diferenciadoras entre las edades de los viajeros, y podemos inferir lo siguiente:

- Tenemos un grupo de personas que viven en Unión Libre entre los 36-45 Años.
- En el nivel de estudios encontramos que viajeros entre los 18-25 años cuentan con formación académica de Secundaria, y visitan la ciudad en busca de ofertas educativas para continuar con sus estudios.
- Las personas que más visitan la ciudad son jóvenes solteros en rangos de edad entre los 18-25 años.
- Una característica del grupo de casados es que tenemos personas entre los 46-55 años con esta estado civil, por lo tanto parece normal dicho grupo.
- La fuente de ingresos por edad es variable. El menor grupo es el de quienes reciben dinero por Otros (Ver Figura 68. *Histograma Edad-Fuente de Ingreso*); pocos menores de edad con ingresos de empleado, lo que contrasta y corresponde con la información encontrada en el histograma de ocupación (Ver Figura 65. *Histograma Edad-Ocupación*), el cual muestra un alto número de estudiantes menores de edad.

Proceso inicial de Clustering. El proceso de Clustering tiene como propósito crear los clúster adecuados para ser analizados y ajustarlos al modelo y/o los datos.

Tabla 50.

Clustering Validación Proceso inicial.

Number of iterations: 4
 Within cluster sum of squared errors: 1662.0
 Missing values globally replaced with mean/mode

Cluster centroids:

Attribute	Full Data (400)	Cluster#		
		0 (206)	1 (138)	2 (56)
Edad	18-25Años	26-35Años	18-25Años	36-45Años
Género	Femenino	Masculino	Femenino	Masculino
CiudadOrigen	Barrancabermeja	Bogotá	Barrancabermeja	Bogotá
ActividadPlaneada	Educación	Comercio	Educación	Salud
GastoPromedio	\$500.000-\$700.000	\$500.000-\$700.000	\$300.000-\$500.000	\$500.000-\$700.000
CalificaciónGastoPromedio	Medio	Medio	Medio	Alto
AspectoqueImpactodeLaCiudad	ClimaAgradable	ClimaAgradable	MuchaIndigencia	CiudadRuidosa
AspectoMotivaaRegresar	Todaslasanteriores	Todaslasanteriores	Bajocostodevida	Cultura

Time taken to build model (full training data) : 0.15 seconds

=== Model and evaluation on training set ===

Clustered Instances

0 206 (52%)
 1 138 (35%)
 2 56 (14%)

Nota: Resultado Obtenido desde el software WEKA. Adaptado de: Software WEKA.

Durante este proceso se determinó que al analizar solo estas variables se tiene una información más precisa, se eliminan aquellas inconsistencias entre los datos y se observan las relaciones entre los datos de cada clúster por reglas de asociación.

Al seleccionar como dato óptimo numClusters=3, se obtuvieron tres clúster, donde 206 de los encuestados, que representan el 52%, se ubican en el Clúster 0; 138 de los encuestados (35%) pertenecen al Clúster 1 y 56 de los viajeros encuestados (14%) al Clúster 2.

Para la obtención de este resultado, se utilizó el método “Simple K-Means”, y posterior a esto se visualizó las asignaciones de cada clúster de la siguiente manera:

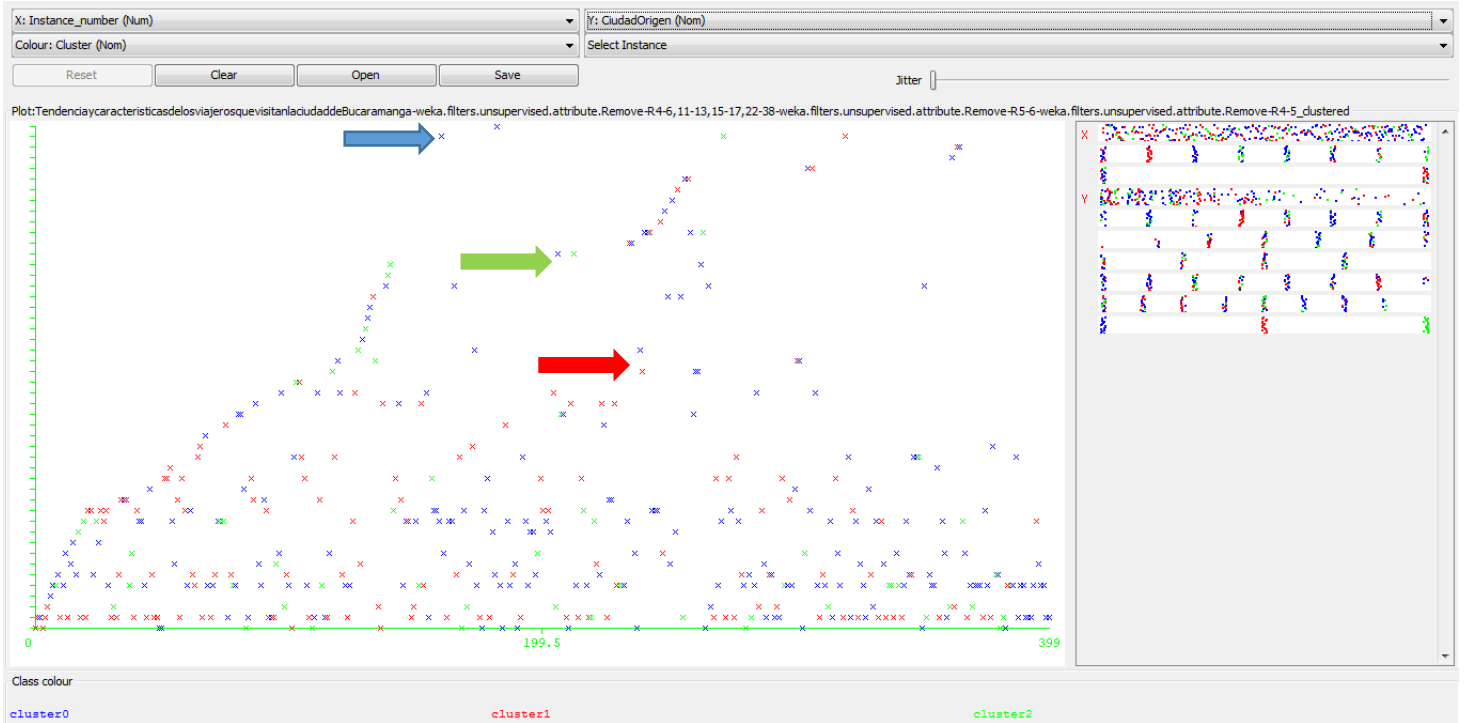


Figura 67. Asignaciones de cada clúster utilizando el software WEKA. Adaptado de: Proceso Obte video por el software WEKA.

Mediante el siguiente cuadro se puede visualizar de manera más detallada cada uno de los datos analizados, es decir, los utilizados para este ejercicio (los que están señalados por las flechas):

Plot : 09:51:02 - SimpleKMeans (Tendenciaycarac
Instance: 161

Clúster 0

Instance_number: 160.0
Edad: 36-45Años
Género: Femenino
CiudadOrigen: Pamplona
ActividadPlaneada: Comercio
GastoPromedio: \$500.000-\$700.000
CalificaciónGastoPromedio: Medio
AspectoqueImpactodelaCiudad: GenteAmable
AspectoMotivaaRegresar: GenteAmable
Cluster: cluster0


Φιγυρα 68. Ασιγναχιονεσ δελ χλ/στερ 0 υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. . Αδαπταδο δε: Προχεσ ο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

Plot : 09:51:02 - SimpleKMeans (Tendenciaycaracte
Instance: 240

Clúster 1

Instance_number: 239.0
Edad: 18-25Años
Género: Femenino
CiudadOrigen: Riohacha
ActividadPlaneada: Educación
GastoPromedio: \$700.000-\$1.000.000
CalificaciónGastoPromedio: Bajo
AspectoqueImpactodelaCiudad: CiudadRuidosa
AspectoMotivaaRegresar: Bajocostodevida
Cluster: cluster1

Φιγυρα 69. Ασιγναχιονεσ δελ χλ/στερ 1 υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. . Αδαπταδο δε: Προχεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

Clúster 2


```

Plot : 09:51:02 - SimpleKMeans (Tendenciaycarac
Instance: 141
    Instance_number: 140.0
    Edad: 18-25Años
    Género: Femenino
    CiudadOrigen: Guatapé
    ActividadPlaneada: Cultura
    GastoPromedio: $500.000-$700.000
    CalificaciónGastoPromedio: Alto
    AspectoqueImpactodelaCiudad: GenteAmable
    AspectoMotivaaRegresar: Cultura
    Cluster: cluster2

```

Φιγυρα 70. Ασιγναχιονεσ δελ χλ/στερ 2 υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. .Αδαπταδο δε: Προχεσ ο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

Segundo Proceso de Clustering

En este proceso se seleccionan los atributos necesarios para generar el clúster, para la etapa de Clustering de percepción y satisfacción. Para esto se seleccionaron los datos de hábitos (N°VecesViajado, MotivoViaje, N°DíasPermanencia, N°CompañeroViaje, LugarAlojamiento, ActividadPlaneada, MedioTrnsportedeLlegada, MedioTransporteenlaCiudad, LugarAlimentación).

Tabla 51.

Clustering Validación Segundo Proceso.

```

Cluster centroids:
Attribute                               Full Data          Cluster#
                                      (400)             0                 1                 2
                                      (163)             (134)             (103)
=====
N°VecesViajado                          3-10Veces         3-10Veces         3-10Veces         3-10Veces
MotivoViaje                              Educación          Comercio          Salud             Educación
N°DíasPermanencia                        1-2Días           1-2Días           1-2Días           Másde20Días
N°CompañeroViaje                         0-1Persona        0-1Persona        1-2Personas       0-1Persona
LugarAlojamiento                         Hotel              Hotel              Hotel              Otro
ActividadPlaneada                        Educación          Comercio          Salud             Educación
MedioTransportedeLlegada                  TransportePúblico TransportePúblico TransportePúblico TransportePúblico
MedioTransporteenlaCiudad                 TransportePúblico TransportePúblico TransportePúblico TransportePúblico
LugarAlimentación                         Restaurante        Restaurante        Restaurante        Otro

Time taken to build model (full training data) : 0.2 seconds

=== Model and evaluation on training set ===

Clustered Instances
0      163 ( 41%)
1      134 ( 34%)
2      103 ( 26%)
    
```

Nota: Resultado Obtenido desde el software WEKA. Adaptado de: Software WEKA.

Al seleccionar como dato óptimo numClusters=3, se obtuvieron tres clúster, donde 163 de los encuestados, que representan el 41%, su ubican en el Clúster 0; 134 de los encuestados (34%) pertenecen al Clúster 1 y 103 de los viajeros encuestados (26%) al Clúster 2.

Para la obtención de este resultado, se utilizó el método “Simple K-Means”, y posterior a esto se visualizó las asignaciones de cada clúster de la siguiente manera:



Φιγυρα 71. Ασιγναχιονεσ δε χαδα χλ /στερ υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. . Αδαπταδο δε: Προ χεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

Mediante el siguiente cuadro se puede visualizar de manera más detallada cada uno de los datos analizados, es decir, los utilizados para este ejercicio (los que están señalados por las flechas):

Plot : 10:22:57 - SimpleKMeans (Tendenciaycarac
Instance: 366
Instance_number: 365.0
N°VecesViajado: Másde20Veces
MotivoViaje: Cultura
N°DíasPermanencia: 11-20Días
N°CompañeroViaje: 3-10Personas
LugarAlojamiento: Hotel
ActividadPlaneada: Cultura
MedioTransportedeLlegada: TransportePúblico
MedioTransporteenlaCiudad: TransportePúblico
LugarAlimentación: Restaurante
Cluster: cluster0

Clúster 0
→

Φιγυρα 72. Ασιγναχιονεσ δελ χλ/στερ 0 υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. . Αδαπταδο δε: Προχεσ ο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

|Plot : 10:22:57 - SimpleKMeans (Tendenciaycarac
Clúster 1
→
ice_number: 219.0
N° vecesViajado: 11-20Veces
MotivoViaje: Salud
N°DíasPermanencia: 3-10Días
N°CompañeroViaje: 1-2Personas
LugarAlojamiento: Hotel
ActividadPlaneada: Salud
MedioTransportedeLlegada: TransportePúblico
MedioTransporteenlaCiudad: TransportePúblico
LugarAlimentación: Restaurante
Cluster: cluster1

Φιγυρα 73. Ασιγναχιονεσ δελ χλ/στερ 1 υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. . Αδαπταδο δε: Προχεσ ο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

Plot : 10:22:57 - SimpleKMeans (Tendenciaycara
 Instance: 167
 Instance_number: 166.0
 N°VecesViajado: 11-20Veces
 MotivoViaje: Educación
 Clúster 2
 N°DíasPermanencia: 11-20Días
 N°CompañeroViaje: 0-1Persona
 LugarAlojamiento: Otro
 ActividadPlaneada: Educación
 MedioTransportedeLlegada: TransportePúblico
 MedioTransporteenlaCiudad: TransportePúblico
 LugarAlimentación: Otro
 Cluster: cluster2

Φιγυρα 74. Ασιγναχιονεσ δελ χλ /στερ 2 υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. .Αδαπταδο δε: Προχεσ ο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ..

Tercer Proceso de Clustering

En este proceso se seleccionan los atributos necesarios para generar el clúster, para la etapa de Clustering de percepción y satisfacción. Para esto se seleccionaron los datos de hábitos (Fuente de Ingreso, Nivel Escolaridad, N° Veces Viajado, Motivo Viaje, N° Días Permanencia, N° Compañero Viaje, Lugar Alojamiento, Medio Transporte de Llegada, Medio Transporte en la Ciudad, Lugar Alimentación).

Tabla 52.

Clustering Validación Tercer Proceso.

Cluster centroids:

Attribute	Cluster#			
	Full Data (400)	0 (193)	1 (107)	2 (100)
NivelEscolaridad	Secundaria	Técnico	Secundaria	Secundaria
Fuente de Ingreso	Empleado	Empleado Pensionado/Jubilado		Familia
N°VecesViajado	3-10Veces	1-2Veces	3-10Veces	3-10Veces
MotivoViaje	Educación	Comercio	Salud	Educación
N°DíasPermanencia	1-2Días	1-2Días	1-2Días	Másde20Días
N°CompañeroViaje	0-1Persona	0-1Persona	1-2Personas	0-1Persona
LugarAlojamiento	Hotel	Hotel	Hotel	Otro
MedioTransportedeLlegada	TransportePúblico	TransportePúblico	TransportePúblico	TransportePúblico
MedioTransporteenlaCiudad	TransportePúblico	TransportePúblico	TransportePúblico	TransportePúblico
LugarAlimentación	Restaurante	Restaurante	Restaurante	Otro

Time taken to build model (full training data) : 0.05 seconds

=== Model and evaluation on training set ===

Clustered Instances

```

0      193 ( 48%)
1      107 ( 27%)
2      100 ( 25%)
    
```

Nota: Resultado Obtenido desde el software WEKA. Adaptado de: Software WEKA.

Al seleccionar como dato óptimo numClusters=3, se obtuvieron tres clúster, donde 193 de los encuestados, que representan el 48%, su ubican en el Clúster 0; 107 de los encuestados (27%) pertenecen al Clúster 1 y 100 de los viajeros encuestados (25%) al Clúster 2.

Para la obtención de este resultado, se utilizó el método “Simple K-Means”, y posterior a esto se visualizó las asignaciones de cada clúster de la siguiente manera:



Figura 75. Ασιγναχιονες δε χαδα χλ/στερ υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. . Αδαπταδο δε: Προ χεσο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

Mediante el siguiente cuadro se puede visualizar de manera más detallada cada uno de los datos analizados, es decir, los utilizados para este ejercicio (los que están señalados por las flechas):

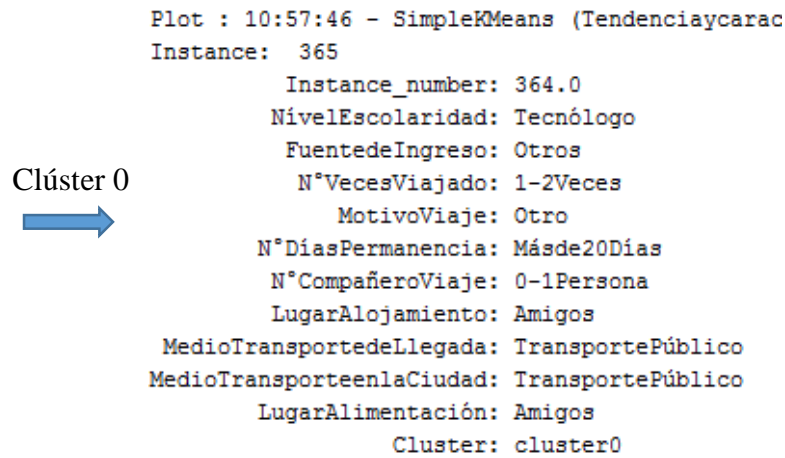



Figura 76. Ασιγναχιονες δελ χλ/στερ 0 υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. . Αδαπταδο δε: Προ χεσο ο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ

Plot : 10:57:46 - SimpleKMeans (Tendenciaycarac
Instance: 272

Clúster 1



```

Instance_number: 271.0
NivelEscolaridad: Primaria
Fuente de Ingreso: Pensionado/Jubilado
N°VecesViajado: 3-10Veces
MotivoViaje: Salud
N°DíasPermanencia: 1-2Días
N°CompañeroViaje: 1-2Personas
LugarAlojamiento: Hotel
MedioTransportedeLlegada: TransportePúblico
MedioTransporteenlaCiudad: TransportePúblico
LugarAlimentación: Restaurante
Cluster: cluster1

```

Φιγυρα 77. Ασιγναχιονεσ δελ χλ/στερ 1 υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. . Αδαπταδο δε: Προχεσ ο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

Plot : 10:57:46 - SimpleKMeans (Tendenciaycar
Instance: 213

Clúster 2


```

Instance_number: 212.0
NivelEscolaridad: Tecnólogo
Fuente de Ingreso: Otros
N°VecesViajado: 3-10Veces
MotivoViaje: Cultura
N°DíasPermanencia: 3-10Días
N°CompañeroViaje: 0-1Persona
LugarAlojamiento: Familia
MedioTransportedeLlegada: TransportePúblico
MedioTransporteenlaCiudad: Particular
LugarAlimentación: Familia
Cluster: cluster2

```

Φιγυρα 78. Ασιγναχιονεσ δελ χλ/στερ 2 υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. . Αδαπταδο δε: Προχεσ ο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

Cuarto Proceso de Clustering

En este proceso se seleccionan los atributos necesarios para generar el clúster, para la etapa de Clustering de percepción y satisfacción. Para esto se seleccionaron los datos de hábitos (Edad, EstadoCivil, Género, Fuente de Ingreso, NivelEscolaridad, N°VecesViajado, MotivoViaje,

NºDíasPermanencia, NºCompañeroViaje, LugarAlojamiento, ActividadPlaneada, MedioTrnsportedeLlegada, MedioTransporteenlaCiudad, LugarAlimentación).

Tabla 53.

Clustering Validación Cuarto Proceso.

Cluster centroids:

Attribute	Full Data (400)	Cluster#		
		0 (183)	1 (99)	2 (118)
Edad	18-25Años	26-35Años	56-65Años	18-25Años
Género	Femenino	Masculino	Femenino	Femenino
EstadoCivil	Soltero	UniónLibre	Casado	Soltero
NivelEscolaridad	Secundaria	Técnico	Secundaria	Secundaria
FuenteIngreso	Empleado	Empleado	Pensionado/Jubilado	Familia
NºVecesViajado	3-10Veces	3-10Veces	3-10Veces	3-10Veces
MotivoViaje	Educación	Comercio	Salud	Educación
NºDíasPermanencia	1-2Días	1-2Días	1-2Días	Másde20Días
NºCompañeroViaje	0-1Persona	0-1Persona	1-2Personas	0-1Persona
LugarAlojamiento	Hotel	Hotel	Hotel	Otro
ActividadPlaneada	Educación	Comercio	Salud	Educación
MedioTransportedeLlegada	TransportePúblico	TransportePúblico	TransportePúblico	TransportePúblico
MedioTransporteenlaCiudad	TransportePúblico	TransportePúblico	TransportePúblico	TransportePúblico
LugarAlimentación	Restaurante	Restaurante	Restaurante	Otro

Time taken to build model (full training data) : 0.17 seconds

=== Model and evaluation on training set ===

Clustered Instances

0	183 (46%)
1	99 (25%)
2	118 (30%)

Nota: Resultado Obtenido desde el software WEKA. Adaptado de: Software WEKA.

Al seleccionar como dato óptimo numClusters=3, se obtuvieron tres clúster, donde 183 de los encuestados, que representan el 46%, su ubican en el Clúster 0; 99 de los encuestados (25%) pertenecen al Clúster 1 y 118 de los viajeros encuestados (30%) al Clúster 2.

Para la obtención de este resultado, se utilizó el método “Simple K-Means”, y posterior a esto se visualizó las asignaciones de cada clúster de la siguiente manera:




Figura 79. Asignaciones de cada clúster utilizando el software WEKA. . Adaptado de: Προχέσο Οβτενίδο por el software WEKA.

Mediante el siguiente cuadro se puede visualizar de manera más detallada cada uno de los datos analizados, es decir, los utilizados para este ejercicio (los que están señalados por las flechas):

Plot : 12:33:54 - SimpleKMeans (Tendenciaycaract
 Instance: 248


Instance_number: 247.0
 Edad: 18-25Años
 Género: Masculino
 EstadoCívil: Soltero
 NivelEscolaridad: Tecnólogo
 FuentedeIngreso: Desempleado
 N°VecesViajado: 3-10Veces
 MotivoViaje: Otro
 N°DíasPermanencia: Másde20Días
 N°CompañeroViaje: 0-1Persona
 LugarAlojamiento: Hotel
 ActividadPlaneada: Otro
 MedioTransportedeLlegada: TransportePúblico
 MedioTransporteenlaCiudad: TransportePúblico
 LugarAlimentación: Restaurante
 Cluster: cluster0

Clúster 0


Φιγυρα 80. Ασιγναχιονεσ δελ χλ /στερ 0 υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. . Αδαπταδο δε: Προχεσ ο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.


Plot : 12:33:54 - SimpleKMeans (Tendenciaycaracte
 Instance: 193

Instance_number: 192.0
 Edad: Menosde18Años
 Género: Femenino
 EstadoCívil: Soltero
 NivelEscolaridad: Secundaria
 FuentedeIngreso: Familia
 N°VecesViajado: 1-2Veces
 MotivoViaje: Deportes
 N°DíasPermanencia: 1-2Días
 N°CompañeroViaje: 11-20Personas
 LugarAlojamiento: Hotel
 ActividadPlaneada: Deportes
 MedioTransportedeLlegada: TransportePúblico
 MedioTransporteenlaCiudad: TransportePúblico
 LugarAlimentación: ComidarápidaenlaCalle
 Cluster: cluster1

Clúster 1


Φιγυρα 81. Ασιγναχιονεσ δελ χλ /στερ 1 υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. . Αδαπταδο δε: Προχεσ ο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

Plot : 12:33:54 - SimpleKMeans (Tendenciaycaracteri:
 Instance: 362
 Instance_number: 361.0
 Edad: Menosde18Años
 Género: Masculino
 EstadoCivil: Soltero
 NivelEscolaridad: Secundaria
 FuentedeIngreso: Desempleado
 N°VecesViajado: 1-2Veces
 MotivoViaje: Otro
 N°DíasPermanencia: Másde20Días
 N°CompañeroViaje: 0-1Persona
 LugarAlojamiento: Otro
 ActividadPlaneada: Otro
 MedioTransportedeLlegada: TransportePúblico
 MedioTransporteenlaCiudad: Caminar
 LugarAlimentación: ComidarápidaenlaCalle
 Cluster: cluster2

Clúster 2


Φιγυρα 82. Ασιγναχιονες δελ χλ/στερ 2 υτιλιζανδο ελ σοφτωαρε ΩΕΚΑ. . Αδαπταδο δε: Προχεσ ο Οβτενιδο πορ ελ σοφτωαρε ΩΕΚΑ.

Evaluación del modelo

A partir de este punto se desarrolla el Objetivo 2: **Establecer relaciones y tendencias existentes entre las características de los viajeros y los aspectos que rodean su visita a la ciudad de Bucaramanga por medio de reglas de asociación.**

En el grafico se observa que la distribución de los clúster se da con diferencias en casi todas las variables, y en la que se aprecia que está marcada la diferencia de atributos como Edad, Género, EstadoCivil, NivelEscolaridad, FuentedeIngreso, ActividadPlaneada, MedioTransportedeLlegada, MedioTransporteenlaCiudad.

Best rules found:

1. FuentedeIngreso=6 ActividadPlaneada=4 MedioTransporteenlaCiudad=1 63 ==> EstadoCivil=1 62 conf:(0.98)
2. NivelEscolaridad=2 ActividadPlaneada=4 MedioTransportedeLlegada=2 MedioTransporteenlaCiudad=1 62 ==> EstadoCivil=1 61 conf:(0.98)
3. NivelEscolaridad=2 FuentedeIngreso=6 MedioTransporteenlaCiudad=1 71 ==> EstadoCivil=1 69 conf:(0.97)
4. ActividadPlaneada=4 MedioTransportedeLlegada=2 MedioTransporteenlaCiudad=1 71 ==> EstadoCivil=1 69 conf:(0.97)
5. NivelEscolaridad=2 ActividadPlaneada=4 MedioTransporteenlaCiudad=1 70 ==> EstadoCivil=1 68 conf:(0.97)
6. FuentedeIngreso=6 ActividadPlaneada=4 MedioTransportedeLlegada=2 66 ==> EstadoCivil=1 64 conf:(0.97)
7. NivelEscolaridad=2 FuentedeIngreso=6 MedioTransportedeLlegada=2 MedioTransporteenlaCiudad=1 64 ==> EstadoCivil=1 62 conf:(0.97)
8. NivelEscolaridad=2 FuentedeIngreso=6 ActividadPlaneada=4 MedioTransportedeLlegada=2 62 ==> EstadoCivil=1 60 conf:(0.97)
9. FuentedeIngreso=6 MedioTransporteenlaCiudad=1 79 ==> EstadoCivil=1 76 conf:(0.96)
10. NivelEscolaridad=2 FuentedeIngreso=6 MedioTransportedeLlegada=2 76 ==> EstadoCivil=1 73 conf:(0.96)
11. NivelEscolaridad=2 ActividadPlaneada=4 MedioTransportedeLlegada=2 76 ==> EstadoCivil=1 73 conf:(0.96)
12. FuentedeIngreso=6 ActividadPlaneada=4 74 ==> EstadoCivil=1 71 conf:(0.96)
13. NivelEscolaridad=2 FuentedeIngreso=6 ActividadPlaneada=4 70 ==> EstadoCivil=1 67 conf:(0.96)
14. FuentedeIngreso=6 MedioTransportedeLlegada=2 MedioTransporteenlaCiudad=1 70 ==> EstadoCivil=1 67 conf:(0.96)
15. NivelEscolaridad=2 FuentedeIngreso=6 85 ==> EstadoCivil=1 81 conf:(0.95)
16. FuentedeIngreso=6 MedioTransportedeLlegada=2 82 ==> EstadoCivil=1 78 conf:(0.95)
17. FuentedeIngreso=6 93 ==> EstadoCivil=1 88 conf:(0.95)
18. FuentedeIngreso=6 ActividadPlaneada=4 74 ==> NivelEscolaridad=2 70 conf:(0.95)
19. Edad=2 ActividadPlaneada=4 73 ==> EstadoCivil=1 69 conf:(0.95)
20. EstadoCivil=1 FuentedeIngreso=6 ActividadPlaneada=4 71 ==> NivelEscolaridad=2 67 conf:(0.94)

Figura 83. Regla de Asociación \hat{v} . Aadaptado de: Resultado Obtenido desde el software ΩEKA.

Reglas de Asociación: Se describen las primeras cinco reglas de asociación.

[FuentedeIngreso = Familia] \rightarrow [ActividadPlaneada = Educación, MedioTransporteenlaCiudad = Transporte Público, EstadoCivil = Soltero] (Confidence: 0,98).

[NivelEscolaridad = Secundaria] $\rightarrow \rightarrow$ [ActividadPlaneada = Educación, MedioTransportedeLlegada = Transporte Público, MedioTransporteenlaCiudad = Transporte Público, EstadoCivil = Soltero] (Confidence: 0,98).

[NivelEscolaridad = Secundaria] $\rightarrow \rightarrow$ [FuentedeIngreso = Familia, MedioTransporteenlaCiudad = Transporte Público, EstadoCivil = Soltero] (Confidence: 0,97).

[ActividadPlaneada = Educación] $\rightarrow \rightarrow$ MedioTransportedeLlegada = Transporte Público, MedioTransporteenlaCiudad = Transporte Público, EstadoCivil = Soltero] (Confidence: 0,97).

[NivelEscolaridad = Secundaria] $\rightarrow \rightarrow$ [ActividadPlaneada = Educación, MedioTransporteenlaCiudad = Transporte Público, EstadoCivil = Soltero] (Confidence: 0,97).

Análisis sobre los resultados del clúster del Proceso Inicial

- El grupo analizado está conformado por personas entre los 18-25 Años y los 36-45 Años.
- El clúster está conformado por solteros.
- El gasto Promedio esta entre los \$500.000 hasta \$1.000.000.
- Los viajeros consideran el Gasto Promedio durante su viaje como Bajo, Medio y Alto.
- Bucaramanga se caracteriza para estas personas como una ciudad con gente amable y con bajo costo de vida.

Análisis sobre los resultados del clúster del Segundo Proceso de Clustering

- Cuando tienen bajo nivel de permanencia coincide que tienen bajo nivel de acompañantes y bajo nivel de viajes anteriores a la ciudad.
- Hospedaje en Hotel y Otros.
- Cuando se quedan pocos días es debido a que llegan en Transporte Público, en la ciudad para transportarse optan por utilizar transporte público, seguramente de sus amigos o alquiler, y entre las actividades que realizan cuando viajan son Salud, Educación y Cultura.
- La visita a la ciudad se hace en grupos de personas y el tiempo de estadía es relativamente largo, utilizando Hoteles recomendados por amigos o familiares o encontrados por internet. Este grupo sería uno importante en el caso que se busque incentivar más sitios turísticos como actividades culturales para personas dentro del rango de edad más representativo, donde los agentes turísticos analicen y fomenten planes más acorde a sus gustos.

Análisis sobre los resultados del clúster del Tercer Proceso de Clustering

- Tenemos viajeros con todos los niveles de estudio.

- Las principales fuentes de ingreso en orden de cantidad son: Otros y Pensionado/Jubilado.
- Cuando tienen pocos días de permanencia coincide con el bajo número de acompañantes y el bajo número de veces que ha viajado a la ciudad.
- El tipo de hospedajes principales son: Hotel, Familia y Amigos.
- Los principales motivos de viaje son: Cultura, Salud y Otros.
- Cuando tienen un nivel de escolaridad elemental como primaria, se hospedan en hoteles.

Análisis sobre los resultados del clúster del Cuarto Proceso de Clustering

- Este grupo está conformado por personas entre menores de 18 años hasta los 25.
- El estado civil que tiene más viajero es Soltero.
- La principal fuente de ingreso es por Familia.
- El número de días de permanencia en la ciudad es Medio y Alto principalmente.
- El sitio de Hospedaje más utilizado es Hotel.
- Principalmente cuando el motivo de viaje es Educación, utilizan medio de transporte público y se hospedan en hoteles.

Es notable en este clúster una gran dependencia de la familia, en cuanto a Fuente de Ingreso, y se asume que los viajeros visitan la ciudad con familiares cercano.

Este grupo es interesante, ya que de acuerdo con una gran clasificación que realizaron para este tipo de nuevos jóvenes adultos el comercio tendría un potencial para la venta de artículos como juegos, videojuegos, actividades recreativas, es decir, actividades que podríamos tomar como de niños.

*A partir de este punto se desarrolla el Objetivo 3: **Determinar en los viajeros que visitan la ciudad de Bucaramanga el nivel de gastos, percepción y satisfacción luego de visitar la ciudad por medio de un análisis de clúster.***

Se seleccionan los atributos relacionados con la percepción, gasto y satisfacción del viajero (GastoPromedio, CalificaciónGastoPromedio, AspectoqueImpactodelaCiudad, AspectoMotivaaRegresar, GradoSatisfacción en ofertas de Entretenimeinto, Comercial, Cultural, SeñalozaciónTurística, CulturaCiudadana, Alojamiento, CalidadAtenciónRecibida, CalidadAlojamiento, DisponibilidadAlojamiento, PrecioAlojamiento, OficinaTurismooTransportadoras, HoteluHopedaje, PersonasenlaCalle, Policia, PoliciadeTránsito, PoliciaCarreteras, Información Turística).

Tabla 54.

Proceso para Análisis de la percepción.

Cluster centroids:				
Attribute	Full Data (400)	Cluster#		
		0 (248)	1 (152)	
GastoPromedio	\$500.000-\$700.000	\$500.000-\$700.000	\$300.000-\$500.000	
CalificaciónGastoPromedio	Medio	Medio	Medio	
AspectoqueImpactodelaCiudad	ClimaAgradable	ClimaAgradable	MuchaIndigencia	
AspectoMotivaaRegresar	Todaslasanteriores	Todaslasanteriores	Bajocostodevida	
GradoSatisfacciónEntretenimiento	Alta	Alta	Media	
GradoSatisfacciónComercial	Alta	Alta	Media	
GradoSatisfacciónCultural	Media	Media	Media	
GradoSatisfacciónSeñalizaciónTurística	Media	Media	Media	
GradoSatisfacciónCulturaCiudadana	Media	Media	Baja	
GradoSatisfacciónAlojamiento	Alta	Alta	Alta	
GradoSatisfacciónCalidaddeAtenciónRecibida	Alta	Alta	Alta	
GradoSatisfacciónCalidadAlojamiento	Alta	MuyAlta	Alta	
GradoSatisfacciónDisponibilidadAlojamiento	Alta	Alta	Alta	
GradoSatisfacciónPrecioAlojamiento	Alta	Alta	Alta	
GradoSatisfacciónOficinaTurismooTransportadoras	Media	Media	Alta	
GradoSatisfacciónHoteluHospedaje	Media	Baja	Media	
GradoSatisfacciónPersonasenlaCalle	Media	Media	Media	
GradoSatisfacciónPolicia	Baja	Media	Baja	
GradoSatisfacciónPolicia deTránsito	Baja	Baja	Másbaja	
GradoSatisfacciónPoliciaCarreteras	Baja	Baja	Másbaja	
GradoSatisfacciónInformaciónTurística	Media	Baja	Media	

Time taken to build model (full training data) : 0.17 seconds

=== Model and evaluation on training set ===

Clustered Instances

0 248 (62%)

1 152 (38%)

Nota: Resultado Obtenido desde el software WEKA. Adaptado de: Software WEKA.

- Se realiza la lectura de cada clúster. No se generan nuevos clúster porque requerimos conocer los grupos creados, es decir los 4 clúster para analizar cómo fue la percepción que tuvieron de su visita, de esta manera mejorar servicios, crear nuevos o eliminar servicios.

Caracterización del gasto promedio y percepción del clúster. El clúster indica que el gasto promedio es de \$500.000-\$700.000, considerado Medio.

- Las ofertas que se consideran altas son: Entretenimiento, Comercial, todo lo relacionado con Alojamiento.
- Las ofertas que se consideran bajas son: Policia, PoliciadeTránsito, PoliciaCarreteras e InformaciónTurística.
- Las ofertas que se consideran medias son: HoteluHospedaje, OficinadeTurismooTransportadoras y PersonasenlaCalle.

El Gasto Promedio durante el viaje está en una consideración Media, pues el costo de vida de la ciudad ha mantenido estable y se considera bajo respecto las demás ciudades del país.

Basado en esto se puede concluir que si hubo exceso de gastos durante el viaje en sí, bien sea en alimentación, alojamiento o transporte. Las ofertas que se consideran más altas son las ofertas de entretenimiento, comercial y alojamiento, y es posible que los viajeros hayan gastado gran parte del dinero en actividades como: casinos, apuestas, discotecas, bares, compra de ropa, electrodomésticos, comida costosa, entre otras.

El alojamiento está en un puesto intermedio, en la ciudad casi todos los hoteles se clasifican como 3 estrellas.

La principal característica de la ciudad respecto al impacto es: Clima Agradable

- Calificaciones positivas: Gente Amable, Clima Agradable, Ambiente Agradable.
- Calificaciones negativas: Mucha Indigencia, Ciudad Ruidosa.

Analizando las reglas y el consolidado, lo viajeros tienen una percepción de indigencia y ciudad ruidosa, ya que la conservación de los monumentos es muy baja y la afluencia vehicular es caótica en determinadas horas. Los temas de seguridad y acompañamiento por parte de la Policía, Policía de Tránsito, de Carretera es Baja y la información turística también es muy baja y se puede apreciar cuando visitamos dos parques, observamos poca vigilancia y muchas personas indigentes consumiendo sustancias psicoactivas, pocos letrados, vallas, pancartas y personal de turismo.

Calificación respecto a la información en la ciudad tenemos:

- Calificaciones positivas (Excelentes): Entretenimiento, Comercial y Alojamiento.
- Calificaciones negativas (Bajas): Policía, Policía de Tránsito, de Carretera y la información turística
- Calificaciones intermedias: HoteluHospedaje, OficinadeTurismooTransportadoras y PersonasenlaCalle.

Tenemos datos con bastantes viajeros en calificación alta e intermedia lo cual da una clara visión de conformidad en dichos aspectos. El alojamiento es regular pero como vimos anteriormente, la ciudad se caracteriza por hoteles de 3 estrellas; a información en las oficinas de turismo no es muy buena, ya que dichas oficinas funcionan únicamente en semana santa y son pequeñas y las personas encargadas solo entregan un volante informativo y pero no tienen conocimiento sobre esta.

Calificación principal de motivos para regresar a la ciudad:

- Calificaciones positivas: Todas las anteriores (clima, gente, cultura, etc.)

- Calificaciones negativas: Ninguna
- Calificaciones intermedias: Comida típica, parques y sitios turísticos.

Las calificaciones positivas y negativas no están al mismo nivel, siendo superiores las positivas que las negativas, observando que el clima es de los que más puntaje obtuvo.

Bucaramanga es un eje por el cual circulan personas de muchas ciudades, tal vez por esto sea difícil la cultura ciudadana.

Evaluación

Evaluación de los resultados. Los resultados obtenidos en el proceso de minería de datos nos brindan una clasificación del grupo de viajeros de acuerdo con validaciones internas de algoritmo de clustering K-Means.

Las diferentes pruebas realizadas a los clúster mediante los procesos y los resultados del Simple K-Means y algoritmo A priori, que nos permite validar la información recomendada para que los clúster sean más precisos.

Cada clúster nos brinda información general y específica de perfiles de viajero. Todos los análisis realizados en la etapa anterior están enfocados a los resultados de los algoritmos de Clustering y argumentando con las respectivas referencias las posibles causas de dicho comportamiento de los viajeros.

Los análisis realizados se dan ya que el analista conoce el objeto del negocio el cual es “La Ciudad de Bucaramanga”, conoce sus calles, su cultura, las dificultades diarias, las zonas de peligro, algunos de los comentarios que realizan las personas de otras regiones y de otros países. Lo anterior

no indica que este análisis es el único, por el contrario, brinda un marco de referencia de la ciudad para que los lectores realicen su propio análisis a la luz de los datos.

Los resultados obtenidos son satisfactorios. Se obtuvo 4 grupos en total donde se caracterizó a la población, de acuerdo a los clúster, concluyendo lo siguiente:

- Los resultados son claros y dan una descripción de los viajeros.
- Se han descubierto 4 perfiles de viajeros sobre los cuales mejorar todos los procesos de la ciudad y los procesos de los entes encargados del turismo y promoción de la ciudad.
- Dentro de dichos perfiles llama la atención de la gran cantidad de viajeros con diversos niveles de preparación académica.
- Otro aspecto que llama la atención es que comparando los problemas de la ciudad y los aspectos negativos de la visita por parte de los viajeros, estos han perdurado en el tiempo ya que gran parte de los encuestados han realizado varios viajes a la ciudad y desde hace tiempo atrás existen las mismas dificultades de años atrás.
- También llama la atención que a pesar de las dificultades, la ciudad de Bucaramanga sigue siendo unos de los destinos preferidos en la región.

Respecto a los datos, método de encuesta y las preguntas de la encuesta:

- Las preguntas seleccionadas para la encuesta fueron estudiadas cuidadosamente para dar una visión general de los viajeros.
- Como en toda encuesta, no podemos saber a ciencia cierta si los datos son reales o no.
- Algunas preguntas se pueden modificar de acuerdo a las características de la ciudad o la problemática en el momento de la recolección de los datos.

- Las preguntas fueron pertinentes para la ciudad debido a su idiosincrasia, problemas, vías de transporte, idioma, ubicación respecto a otras regiones y paisaje santandereano.

En términos comerciales los resultados son pertinentes porque:

- Da una perspectiva desde los viajeros a la ciudad.
- Nos muestra una clara tendencia de las problemáticas vividas diariamente las cuales se convierten el algo cotidiano para los que vivimos en la ciudad pero las cuales afectan el turismo.

Proceso de revisión: Como parte de una autoevaluación se tienen en cuenta los siguientes aspectos:

- El proceso de minería de datos utilizado resultó efectivo para analizar los datos nominales y numéricos.
- El proceso puede ser mejorado en su herramienta SPSS y WEKA para realizar el análisis en un solo momento, es decir, minería en una sola herramienta.
- La recolección de los datos y la limpieza se debería mejorar con el apoyo de los entes de control encargados del turismo en la región.
- El proceso realizado en la presente investigación se puede aplicar a los datos socioeconómicos o datos descriptivos de características en otras regiones o ciudades ya que es genérico para datos nominales, discretizados y normalizados.
- Se deberían aplicar las encuestas a grupos familiares y crear un nuevo campo que los identifique. De esta manera evitar segmentar la muestra, esto ya que muchas veces la persona que paga disfruta pero el viaje tiene una percepción diferente del aspecto económico con respecto a la persona que simplemente viaja, disfruta pero no gasta su propio dinero.

Despliegue

El despliegue de la aplicación puede ser realizado por cualquier persona interesada en analizar los datos de la ciudad o implementar dichos análisis a sus propias ciudades.

Debería ser parte del plan de cada alcaldía la cual aportaría bastante al mejoramiento de la ciudad y también en la innovación de productos y/o servicios para atraer a los turistas y mantener a los actuales que nos visitan.

El proyecto plantea una arquitectura inicial de análisis de datos para los viajeros, esta queda abierta para modificaciones, ejecuciones y análisis, ya que la gran cantidad de información recopilada es un insumo para cualquier analista de datos que quiera tratar la información y de esta manera poder contribuir al mejoramiento de la región.

Debe ser un propósito para el lector interesado, el continuar haciendo pruebas, análisis, conclusiones y argumentaciones.

Esta investigación da una referencia de cómo abordar un problema de minería de datos, para etapas posteriores se debería aplicar otros algoritmos de minería de datos los cuales permitan predecir resultados de los viajeros o predecir si un visitante regresará o no.

Análisis de Resultados

- En la ciudad tenemos muchos visitantes de un solo día, los cuales no pernoctan pero hacen diversas actividades, de acuerdo con los análisis realizados a los datos, se debe mejorar en la oferta cultural y así posiblemente el viajero se quedaría en la ciudad, dicha oferta debe estar a cargo de las personas encargadas de turismo o de los comerciantes de la región.

- La ciudad de Bucaramanga recibe diversidad de visitantes, hombres, mujeres, adultos, jóvenes y niños. Para cada grupo de personas hay algunos intereses en común como se pudo apreciar en los clúster, permitiendo segmentar dichos grupos para su análisis.
- Llama la atención que a la ciudad llegan personas de diversos niveles educativos con intereses en común. La ciudad podría aprovechar dichos conocimientos al utilizar el centro de convenciones o al crear una base de datos de los viajeros, determinar sus intereses, conocimientos y posibilidades de que realicen conferencias, charlas o eventos y de esta manera aportar al desarrollo de la región.
- Ya que nos visitan tantas personas independientes, se podrían generar mejoras en la ciudad para que dichas personas puedan hacer parte de los proyectos de la alcaldía, como por ejemplo, generar un establecimiento o portal web para que cualquier persona pueda participar de licitaciones de proyectos privados de la región, que permitan el mejoramiento económico y social de esta.
- Según el análisis realizado, de todos los Clúster en el que había jóvenes involucrados, el porcentaje de ingresos familiares era más del 40% en la población del clúster. Según este análisis dichos jóvenes visitan la ciudad para compras o por diversión, indicando un poder adquisitivo por parte de sus padres, ya que la mayoría tienen como ingreso el dinero familiar.
- El nivel de gasto estimado es bajo para las personas que visitan un solo día y vienen a realizar actividades para su Salud, mejorando un poco el aporte a restaurante y medios de transporte como una oportunidad de negocio para hacer ofertas, promociones o eventos programados con anterioridad.

- En la ciudad nos visitan muchas personas de edad pensionadas, por lo tanto se debe brindar seguridad y educar a los conductores del servicio público para mejorar las relaciones humanas, y de esta manera las personas se sentirían más seguras y utilizarían dicho transporte y mejorarían la economía de estas personas que prestan este servicio.
- En el caso del ruido en la ciudad, hay zonas identificadas con esta problemática, se deben implementar políticas más estrictas o dar una solución de forma tecnológica esta situación.

Aun así cabe recalcar que muchas personas regresan a la ciudad y dan buenas referencias. Parece ser que la amabilidad de la gran mayoría de personas y el clima tienen mayor peso frente a las cosas malas.

Conclusiones.

Los modelos de minería de datos aplicados a los datos propios del proceso y actividades características de la ciudad, sirven para que todos los visitantes y personas de la región. Además el trabajo realizado es importante para la ciudad de Bucaramanga como para el área metropolitana, ya que de esta manera, los lectores (interesados en el Turismo de la ciudad y región) podrán visualizar como los datos nos generaron información relevante que puede utilizarse para mejorar diversos aspectos que no eran notados a simple vista.

Las entidades encargadas del Turismo en la región cuentan con la infraestructura para enfocarse realmente en los viajeros y turistas, siendo estos un factor importante de negocio: viaje-ciudad-entidades de viaje-turismo. Un viajero contento recomienda la ciudad; un viajero que no fue bien recibido, es un viajero que no ve ni cree potencial la ciudad y lo más razonable es que no la recomiende a futuros y nuevos turistas.

El análisis realizado en el presente trabajo no es el único a desarrollarse, pues los datos recopilados aún tienen gran cantidad de información por deducir, extraer y experimentar. Esta investigación solo fue una parte potencial de la minería de datos enfocado a los viajeros.

La minería de datos posee muchas herramientas que permiten el desarrollo de diversos tipos de análisis, como utilizarlo en la web, en los portales públicos de las entidades interesadas de la ciudad y región que deseen hacer un análisis de información e identificar aquellos aspectos positivos y negativos que pueden publicar los viajeros y turistas en dichas páginas, y así mejorar progresivamente este servicio que se les ofrece a quienes nos visitan o planean hacerlo.

Realizando una comparación con la realidad de la ciudad y los datos recopilados en el análisis, podemos concluir ciertas cosas como:

- Aunque la indigencia se ha reducido un poco a comparación de años anteriores, para un viajero este tema resulta, en muchas ocasiones algo molesto.
- Según cifras obtenidas por el DANE, actualmente en la ciudad bajó la tasa de desempleo, pero es necesario que se implemente en la formación de los ciudadanos una preparación en idiomas, para seguir generando empleos y mejorar la imagen de la ciudad y la región.
- Un alto número de viajeros pernoctan en la ciudad, lo que plantea la hipótesis de la cultura popular, que Bucaramanga es una ciudad “*vividera*”, siendo así por su clima, ubicación geográfica y costo de vida. Por este motivo sería favorable promover la gestión por y para los ciudadanos y turistas, en pro del crecimiento económico, social y cultural de la ciudad y región.
- La percepción del gasto promedio durante el viaje es marcada, pues para algunos es Media, para otros es Alta, pero según el análisis realizado todo va de acuerdo con el número de permanencia en la ciudad y cantidad de dinero gastado durante ese tiempo, en diversas opciones planteadas en la encuesta como: compras, actividades culturales, turismo, educación, aventura, comercio, etc.
- Es importante estudiar que actividades enfocadas en la educación realizan los viajeros que llegan a la ciudad, ya que Bucaramanga cuenta con una amplia oferta educativa en educación superior para la región y el país, hasta nivel internacional al permitir ofertas de intercambio con otros países. Esto es planteado para que aquellas entidades interesadas en este tema realicen un estudio más profundo sobre dichos aspectos.

Recomendaciones.

Las recomendaciones para los lectores (entidades interesadas) serían las siguientes:

- Tener en cuenta que la minería de datos es un proceso que implica tener conocimiento del tema por parte del encargado de realizar el análisis, y de este tratamiento de la información saber que tanto nos pueden aportar a la investigación.
- Para las entidades públicas y privadas es necesario recomendarles el uso adecuado de los datos, a las encuestas y viajeros, pues como puede observarse en este trabajo, los datos recopilados brindan información valiosa para su análisis, y posterior a esto concluir y realizar recomendaciones que permitan el mejoramiento del tratamiento de la información.
- A las entidades encargadas de fomentar el Turismo en la ciudad y la región, esta investigación permite que tomen cartas en el asunto y publiciten de forma activa los lugares turísticos y actividades que Bucaramanga brinda para los viajeros, y así recopilar datos que mejoren sus estrategias de negocio.
- A quienes se encargan de analizar los datos (analistas) tengan muy claro que la minería de datos permite que la información analizada se debe a la implementación de técnicas no supervisadas que reflejan necesidades, estudian algoritmos y datos, y proporcionan información relevante y válida para cumplir con los objetivos del negocio.
- Para analista y programadores, se le recomienda profundizar en la extensión de la minería de datos, es decir, programar adecuadamente la herramienta y aprovechar los algoritmos matemáticos del software.

Las recomendaciones para continuar con la línea de investigación de minería de datos aplicada a los viajeros y turistas tenemos:

- Luego de realizar el análisis de clúster, se puede generar una segmentación de estos datos, de tal manera que se vayan creando subgrupos de viajeros y a cada subgrupo aplicarle nuevamente la segmentación, para que el analista divida los grupos de acuerdo a los requerimientos planteados.
- Encuestar y segmentar los viajeros teniendo en cuenta la edad, caracterizándolo por esta variable, realizando en este punto clustering con segmentación.
- Crear más intereses en la lista de actividades planeadas y realizadas, y crear dos nuevas categorías denominadas como actividades de interés y actividades que le gustaría realizar que no existen en la ciudad. Esta modificación permitirá realizar un nuevo estudio de reglas de asociación para determinar que tanto influye una variable a otra y validar los resultados.
- Realizar otro estudio en el análisis de regresión utilizando la herramienta DataSet que genere modelos para diversas preguntas que predigan las características de un viajero y determinar si este regresaría o no.
- Luego de realizar la predicción si un viajero regresaría a la ciudad o no, ejecutar un seguimiento al modelado para determinar el motivo de su regreso o no, y el margen de error de dicho análisis.
- Para trabajos posteriores, se plantea la implementación de una herramienta automatizada dentro de las entidades interesadas en temas de viajeros y turismo, para que el personal especializado se encargue de analizar los aspectos positivos y negativos y las opiniones que tienen las personas frente a esta temática.
- En caso que las entidades encargadas del turismo no implementen el uso de la herramienta antes mencionada, los analistas de datos pueden aprovechar el uso y beneficio que brinda la minería web, pues filtra los comentarios y observaciones que las personas tienen sobre

la ciudad, para realizar una segmentación de dicha información, con el fin de mejorar y crear nuevos productos y servicios en la ciudad y región para ofertarlos en viajeros y futuros turistas, y así permitir el crecimiento en esta ciudad, importante metrópoli santandereana.

Bibliografía.

- [1]. *ALCALDIA DE BOGOTA*. (26 de 07 de 1996). Obtenido de Ley 300 de 1996 Nivel Nacional:
<http://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=8634>
- [2]. *ALCALDIA DE BOGOTA*. (14 de 04 de 1997). Obtenido de Decreto 1075 de 1997 Nivel Nacional: <http://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=14696>
- [3]. *ALCALDIA DE BOGOTA*. (09 de 07 de 2010). Obtenido de Decreto 2438 de 2010 Nivel Nacional: <http://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=39997>
- [4]. *ANATO*. (2016). Obtenido de Resultados de Encuesta de Mitad del Año 2016:
<http://www.anato.org/sites/default/files/Resultados%20Encuesta%20de%20Mitad%20de%20A%C3%B1o%202016.compressed.pdf>
- [5]. *CONGRESO DE COLOMBIA*. (30 de 12 de 1993). Obtenido de LEY 105 DE 1993:
<file:///D:/Descargas/Ley105de1993.pdf>
- [6]. *CONGRESO DE COLOMBIA*. (20 de 12 de 1996). Obtenido de Ley 336 de 1996:
<file:///D:/Descargas/Ley336de1996.pdf>
- [7]. *DANE*. (2014). Obtenido de Cuenta Satélite de Turismo:
<https://www.dane.gov.co/index.php/estadisticas-por-tema/cuentas-nacionales/cuentas-satelite/cuentas-economicas-cuenta-satelite-de-turismo-cst>
- [8]. *DefiniciónABC*. (2007). Obtenido de Defición de Viaje:
<http://www.definicionabc.com/general/viaje.php>

- [9]. *DEFINICIONABC*. (2008). Obtenido de Definición de Características:
<http://www.definicionabc.com/general/caracteristicas.php>
- [10]. Enhancing Tourism Intermediaries with the Data Mining Process. (23 de 02 de 2012). 2012
International Conference on Information and Knowledge Management, 2-4. Obtenido de
<http://www.ipcsit.com/vol45/025-ICIKM2012-M0061.pdf>
- [11]. *FONTUR COLOMBIA*. (09 de 04 de 1997). Obtenido de DECRETO NUMERO 972 DE
1997 :
http://www.fontur.com.co/aym_document/aym_normatividad/1997/DECRETO_972_DE_1997.pdf
- [12]. Fuchs, M. (2009). Praxis der Wirtschaftsinformatik. En *Data mining in tourism (In German: „Data Mining im Tourismus“)* (págs. 73 - 81).
- [13]. Han, J., & Kamber, M. (2012). *DATA MINING*. Obtenido de
http://ccs1.hnue.edu.vn/hungtd/DM2012/DataMining_BOOK.pdf
- [14]. *ICONTEC INTERNACIONAL*. (16 de 02 de 2011). Obtenido de ISO 31000:
<https://mail.google.com/mail/u/0/#inbox/15874982468db307?projector=1>
- [15]. Information and Communication Technologies in Education. (2013). *ICES*, 393-403.
Obtenido de
http://s3.amazonaws.com/academia.edu.documents/41626919/Investigating_Students_Metaphor_Related_20160127-18528-eexyv.pdf?AWSAccessKeyId=AKIAJ56TQJRTWSMTNPEA&Expires=1479167361&S

signature=G2%2Bl8ktY%2BzglO4OgU%2B8uOtT24BY%3D&response-content-
disposition=inl

[16]. *Instituto Municipal de Cultura y Turismo de Bucaramanga (IMCUT)*. (2016). Obtenido de <http://www.imcut.gov.co/>

[17]. *MASADELANTE.COM*. (2009). Obtenido de Definición de Base de Datos : <http://www.masadelante.com/faqs/base-de-datos>

[18]. *MICROSOFT*. (2016). Obtenido de Concepto de Minería de Datos: <https://msdn.microsoft.com/es-es/library/ms174949.aspx>

[19]. *Min Comercio, Industria y Turismo*. (2015). Obtenido de Informe de Turismo: <http://www.mincit.gov.co/publicaciones.php?id=16590>

[20]. *Ministerio de Comercio, Industria y Turismo (MINCIT)*. (2014). Obtenido de Informe de Turismo Mayo 2014: <http://www.mincit.gov.co/descargar.php?id=71370>

[21]. *Ministerio de Comercio, Industria y Turismo*. (01 de 06 de 1994). Obtenido de DECRETO NUMERO 1095 DE 1994: <http://www.mincit.gov.co/loader.php?lServicio=Documentos&lFuncion=verPdf&id=57572&name=Decreto-1095-1994.pdf&prefijo=file>

[22]. *Ministerio de Comercio, Industria y Turismo*. (16 de 06 de 2003). Obtenido de SEGURIDAD TURÍSTICA: Reto Competitivo de Colombia (Plan Estratégico): <http://www.mincit.gov.co/minturismo/loader.php?lServicio=Documentos&lFuncion=verPdf&id=61781&name=PlanSeguridad.pdf&prefijo=file>

- [23]. *Ministerio de Comercio, Industria y Turismo*. (12 de 2009). Obtenido de COMPETITIVIDAD: El desafío para alcanzar un Turismo de clase mundial: <http://www.mincit.gov.co/minturismo/loader.php?lServicio=Documentos&lFuncion=verPdf&id=67&name=CompetitividadDesafio2009.pdf&prefijo=file>
- [24]. *Ministerio de Comercio, Industria y Turismo*. (2010). Obtenido de Histórico Informe de Turismo: <http://www.mincit.gov.co/publicaciones.php?id=32452>
- [25]. *MINISTERIO DE TRANSPORTE*. (25 de 02 de 2015). Obtenido de DECRETO 0348 DE 2015: <file:///D:/Descargas/DECRETO%200000348-2015.pdf>
- [26]. *ORGANIZACION MUNDIAL DEL TURISMO OMT*. (2008). Obtenido de Entender el turismo: Glosario Básico: <http://media.unwto.org/es/content/entender-el-turismo-glosario-basico>
- [27]. Pinot, M., Rojas Pulido, J. A., & Mora Martínez, C. A. (10 de 2013). *MEDICIÓN DEL TURISMO INTERNO EN COLOMBIA*. Obtenido de <http://www.eafit.edu.co/cec/congresos/move2013/papers/Medici%C3%B3n%20del%20turismo%20interno%20en%20Colombia%20experiencias%20y%20retos.pdf>
- [28]. *Portal del Municipio de Bucaramanga*. (2011). Obtenido de Boletines y Actualidad: Datos Generales: <http://old.bucaramanga.gov.co/Contenido.aspx?Param=9>
- [29]. *PRESIDENCIA DE LA REPÚBLICA*. (10 de 07 de 2012). Obtenido de <http://wsp.presidencia.gov.co/Normativa/Leyes/Documents/ley155810072012.pdf>
- [30]. Pujarl, A. K. (2011). *DATA MINING TECHNIQUES*. Obtenido de http://rungta.weebly.com/uploads/2/0/3/4/2034099/data_mining-arun_k_pujari.pdf

- [31]. *Red de Revistas Científicas de América Latina y el Caribe, España y Portugal*. (05 de 10 de 2011). Obtenido de Sistema de Información Científica Redalyc: <http://www.redalyc.org/home.oa>
- [32]. Rincon , D. E., & Bravo Mantilla, A. M. (01 de 2013). *Estudio de Competitividad en el Sector Turismo en Colombia*. Obtenido de <http://repository.urosario.edu.co/bitstream/handle/10336/4269/1020746197-2013.pdf;jsessionid=A99B615D06C827B0FB363128AE751E6C?sequence=5>
- [33]. Ritchie, R. J. (2002). A framework for an industry supported destination marketing information system. *Tourism Management*, 439 - 454.
- [34]. Rodríguez Pulgarín, E. A. (10 de 08 de 2011). *REVISTA VINCULANDO*. Obtenido de Evolución histórica del turismo: http://vinculando.org/vacaciones_viajes/turismo_sostenible/3_evolucion_historica_del_turismo.html
- [35]. Santos, M. F., & Azevedo, C. S. (2005). *RepositoriUm*. Obtenido de Preâmbulo [a] "Data mining : descoberta de conhecimento em bases de dados": <http://repositorium.sdum.uminho.pt/handle/1822/19136>
- [36]. *SGS COLOMBIA*. (2013). Obtenido de ISO 31000 - GESTIÓN DE RIESGOS: <http://www.sgs.co/es-ES/Health-Safety/Quality-Health-Safety-and-Environment/Risk-Assessment-and-Management/ISO-31000-Risk-Management/ISO-31000-Risk-Management.aspx>

- [37]. *Significado*. (27 de 05 de 2015). Obtenido de Significado de Tendencia:
<http://significado.net/tendencia/>
- [38]. *SINNEXUS*. (2009). Obtenido de Business Intelligence, Informática Estratégica:
http://www.sinnexus.com/business_intelligence/
- [39]. *Superintendencia de Industria y Comercio*. (2011). Obtenido de Estudios de Mercados,
Estudio de Agencias de Viajes en Colombia.:
http://www.sic.gov.co/drupal/masive/datos/estudios%20economicos/Documentos%20%20elaborados%20por%20la%20Delegatura%20de%20Protecci%C3%B3n%20de%20la%20Competencia/2012/Agencias_Viajes.pdf
- [40]. *Tecnologías de la Información*. (2016). Obtenido de Sistemas de Información de Gestión (SIG): Minería de Datos: <http://www.tecnologias-informacion.com/mineria-de-datos.html>
- [41]. Vallejos, S. (2006). *Trabajo de Adscripción: Minería de Datos*. Obtenido de http://exa.unne.edu.ar/informatica/SO/Mineria_Datos_Vallejos.pdf
- [42]. *Wikipedia*. (26 de 10 de 2016). Obtenido de ISO 31000:2009:
[43https://es.wikipedia.org/wiki/ISO_31000](https://es.wikipedia.org/wiki/ISO_31000)
- [43]. Zapata Suárez, A. (2015). *Análisis de Riesgo por Procesos basado en la norma iso 31000:2011*. Obtenido de <https://red.uao.edu.co/bitstream/10614/8029/1/T06032.pdf>

Apéndice.

Apéndice A. Modelo de la encuesta aplicada

ENCUESTA SOBRE LA CARACTERIZACION DE LOS VIAJEROS QUE VISITAN LA CIUDAD DE

OBJETIVO: El objetivo de esta encuesta es conocer y analizar las tendencias y características que tienen los viajeros al visitar la ciudad de Bucaramanga e identificar la percepción e interés que usted tiene sobre la oferta turística que ofrece la ciudad y los motivos que lo llevan a elegir este destino. Esta encuesta consta de 27 preguntas con única respuesta, y el tiempo de realización

2. Cuántos años tiene?

- a) Menos de 18 años
- b) Entre 18 - 25 años
- c) Entre 26 - 35 años
- d) Entre 36 - 45 años
- e) Entre 46 - 55 años
- f) Entre 56 - 65 años
- g) Entre 66 - 75 años
- h) Más de 75 años

3. Género

- a) Masculino
- b) Femenino

4.Cuál es su ciudad de origen?

5.Cuál es su país de origen?

6.Cuál es su estado civil?

- a) Soltero
- b) Casado
- c) Unión libre
- d) Viudo

- e) Divorciado

7.Cuál es su nivel de escolaridad?

- a) Primaria
- b) Secundaria
- c) Técnico
- d) Tecnólogo
- e) Pregrado
- f) Especialización
- g) Maestría
- h) Doctorado
- i) Ninguno

8.Cuál es su ocupación?

- a) Estudiante
- b) Empleado
- c) Independiente
- d) Pensionado/Jubilado
- e) Desempleado(a)

9.Cuál es su fuente de ingreso?

- a) Estudiante
- b) Empleado
- c) Independiente
- d) Pensionado/Jubilado
- e) Desempleado(a)
- f) Familia
- g) Otros

10. Número de veces que ha viajado a la ciudad de Bucaramanga

- a) 0 - 1
- b) 1 - 2
- c) 3 - 10
- d) 11 - 20
- e) Más de 20

Página 1

11. Motivo de viaje a la ciudad de Bucaramanga	
<input type="checkbox"/>	a) Turismo
<input type="checkbox"/>	b) Comercio
<input type="checkbox"/>	c) Aventura
<input type="checkbox"/>	d) Educación
<input type="checkbox"/>	e) Salud
<input type="checkbox"/>	f) Cultura
<input type="checkbox"/>	g) Deportes
<input type="checkbox"/>	h) Otro
12. Número de días de permanencia en la ciudad.	
<input type="checkbox"/>	a) 0 - 1
<input type="checkbox"/>	b) 1 - 2
<input type="checkbox"/>	c) 3 - 10
<input type="checkbox"/>	d) 11 - 20
<input type="checkbox"/>	e) Más de 20
13. Número de compañeros de viaje	
<input type="checkbox"/>	a) 0 - 1
<input type="checkbox"/>	b) 1 - 2
<input type="checkbox"/>	c) 3 - 10
<input type="checkbox"/>	d) 11 - 20
<input type="checkbox"/>	e) Más de 20
15. Cuál es el lugar de alojamiento durante el viaje?	
<input type="checkbox"/>	a) Hotel
<input type="checkbox"/>	d) Educación
<input type="checkbox"/>	e) Salud
<input type="checkbox"/>	f) Cultura
<input type="checkbox"/>	g) Deportes
<input type="checkbox"/>	h) Otro
17. Qué medio de transporte utilizó para llegar a la ciudad?	
<input type="checkbox"/>	a) Avión
<input type="checkbox"/>	b) Transporte público
<input type="checkbox"/>	c) Particular (Moto, Carro)
<input type="checkbox"/>	d) Otro
18. Qué medio de transporte utilizó para desplazarse en su estadía en la ciudad?	
<input type="checkbox"/>	a) Transporte público
<input type="checkbox"/>	b) Particular (Moto, Carro)
<input type="checkbox"/>	c) Caminar
<input type="checkbox"/>	d) Otro
19. Qué tipo de lugar de alimentación utilizó más al	
<input type="checkbox"/>	a) Restaurantes
<input type="checkbox"/>	b) Comida rápida en restaurantes
<input type="checkbox"/>	c) Comida rápida en la calle
<input type="checkbox"/>	d) Familia
<input type="checkbox"/>	e) Amigos
<input type="checkbox"/>	f) Otro
20. Cuál es el gasto estimado antes de viajar?	
<input type="checkbox"/>	a) Menos de \$100.000
<input type="checkbox"/>	b) \$100.000 - \$150.000
<input type="checkbox"/>	c) \$150.000 - \$300.000
<input type="checkbox"/>	d) \$300.000 - \$500.000
<input type="checkbox"/>	e) \$500.000 - \$700.000
<input type="checkbox"/>	f) \$700.000 - \$1'000.000
<input type="checkbox"/>	g) Más de \$1'000.000

21. Cómo califica el gasto que se generó durante su visita?	
<input type="checkbox"/>	a) Muy bajo
<input type="checkbox"/>	b) Bajo
<input type="checkbox"/>	c) Medio
<input type="checkbox"/>	d) Alto
<input type="checkbox"/>	e) Muy alto
22. Cuál de los siguientes aspectos lo impactó más cuando llegó a la ciudad?	
<input type="checkbox"/>	a) Ambiente agradable
<input type="checkbox"/>	b) Clima agradable
<input type="checkbox"/>	c) Ciudad ruidosa
<input type="checkbox"/>	d) Alta contaminación
<input type="checkbox"/>	e) Gente amable
<input type="checkbox"/>	f) Ciudad segura
<input type="checkbox"/>	g) Mucha indigencia
<input type="checkbox"/>	h) Otro
23. Qué aspecto lo motivaría a regresar a la ciudad?	
<input type="checkbox"/>	a) Clima agradable
<input type="checkbox"/>	b) Gente amable
<input type="checkbox"/>	c) Bajo costo de vida
<input type="checkbox"/>	d) Parques/Sitios turísticos
<input type="checkbox"/>	e) Cultura
<input type="checkbox"/>	f) Comida Típica
<input type="checkbox"/>	g) Todas las anteriores
<input type="checkbox"/>	h) Ninguno
<input type="checkbox"/>	i) No volvería nunca
24. Califique de 1 a 5 el grado de satisfacción a los siguientes aspectos sobre la ciudad, siendo 1 la más baja y 5 la más alta.	
<input type="checkbox"/>	a) Oferta de entretenimiento
<input type="checkbox"/>	b) Oferta comercial
<input type="checkbox"/>	c) Oferta cultural
<input type="checkbox"/>	d) Señalización turística
<input type="checkbox"/>	e) Cultura ciudadana
25. Califique de 1 a 5 el grado de satisfacción sobre el alojamiento en la ciudad, siendo 1 la más baja y 5 la más alta.	
<input type="checkbox"/>	a) Oferta de alojamiento
<input type="checkbox"/>	b) Calidad en la atención recibida
<input type="checkbox"/>	c) Calidad del alojamiento
<input type="checkbox"/>	d) Disponibilidad de los sitios de alojamiento
<input type="checkbox"/>	e) Precio de alojamiento
26. Califique de 1 a 5 el grado de satisfacción sobre la información turística brindada por diferentes sectores, siendo 1 la más baja y 5 la más alta.	
<input type="checkbox"/>	a) Información en oficina de turismo/empresas transportadoras
<input type="checkbox"/>	b) Información en el hotel
<input type="checkbox"/>	c) Información por personas de la calle
<input type="checkbox"/>	d) Información por parte de la policía
<input type="checkbox"/>	e) Información por parte de la policía de tránsito
<input type="checkbox"/>	f) Información por parte de la policía de carreteras
<input type="checkbox"/>	g) Información turística en vallas, carteles, folletos, volantes, etc.