



Desarrollo de un índice, a través de métodos multivariados, para la evaluación de la satisfacción turística en el Valle del Cauca

Diana Marcela Mosquera

Universidad del valle
Facultad de Ingeniería, Escuela de Estadística
Cali, Valle del Cauca
2024

Desarrollo de un índice, a través de métodos multivariados, para la evaluación de la satisfacción turística en el Valle del Cauca

Por

Diana Marcela Mosquera Mosquera

Trabajo de grado presentado como requisito parcial para optar por el título de:

Estadístico(a)

Director:

PhD. Rafael Tovar Cuevas

Codirector:

Mg. Harold Humberto Rivas

Universidad del valle
Facultad de Ingeniería, Escuela de Estadística
Cali, Valle del Cauca
2024

Agradecimientos

En primer lugar, quiero agradecer a Dios por esta oportunidad. A mi madre, por su incesante apoyo, su ejemplo de superación, resiliencia, y por enseñarme que sí se pueden cumplir los sueños por más difíciles que parezcan. Agradezco también a la familia Mosquera, mi preciada familia, por su tenacidad y el acompañamiento constante; sus consejos fraternales han sido, son y siempre serán un pilar en mi crecimiento integral.

Agradezco a mi director de tesis, Rafael Tovar, por sus correcciones, su paciencia y sus palabras precisas. De igual manera, agradezco a todos los docentes cuya pasión y vocación me guiaron hacia los resultados de este trabajo de grado, transmitiéndome sus valiosos conocimientos. Quiero destacar al profesor Andrés Felipe Ochoa por su sincera disposición para ayudar y sus significativas contribuciones que mejoraron este trabajo.

Un agradecimiento especial al SITUR y a su Director, quien también es mi cotutor, Harold Rivas, por la oportunidad de explorar un tema tan significativo para mí. Gracias por abrir las puertas de la entidad y brindarme tan apreciable asesoría.

Finalmente, quiero agradecer a mis compañeros, especialmente a mi incansable compañera de batallas académicas, Daniela, y a Isaac. Gracias por ser luz y un sostén en todo este proceso.

Resumen

Este estudio se centra en la exploración de variables latentes en el ámbito turístico, con un énfasis particular en el desarrollo de un índice para medir la Satisfacción Turística en el Valle del Cauca. El propósito principal de este trabajo es presentar una herramienta, entre otras existentes, para que el Sistema de Información Turística de la región, pueda proporcionar información valiosa a los actores e interesados, destacando oportunidades de mejora y éxitos en la gestión.

Los datos utilizados provienen del instrumento de medición 'Receptivo', suministrado por el SITUR (Sistema de Información Turística del Valle del Cauca). La investigación se basó en 5.101 registros correspondientes a individuos que visitaron el departamento en calidad de turistas. Se evaluaron variables relacionadas con su percepción en aspectos como seguridad, limpieza, hospitalidad, así como su experiencia en restaurantes y alojamientos y visita a lugares específicos.

El proceso de investigación abarcó varias etapas. Inicialmente, se realizó un análisis exploratorio de los datos, seguido de una validación del instrumento mediante un Análisis Factorial Exploratorio (AFE). Posteriormente, se construyeron tres índices utilizando las metodologías de Análisis Factorial Múltiple (AFM), Análisis de Correspondencias Múltiples (ACM) y Análisis de Componentes Principales para Datos Mixtos (FAMD). Finalmente, se eligió el índice con menor varianza. Los resultados se consolidaron en un tablero de visualización en LookerStudio.

Palabras clave: Análisis Factorial Múltiple, Análisis de Componentes Principales para Datos Mixtos, Análisis de Correspondencias Múltiples, Análisis Factorial Exploratorio, Turismo, Índice, Modelo Reflectivo, Variable Latente, Satisfacción Turística.

Abstract

This study focuses on the exploration of latent variables within the tourism field, with a particular emphasis on the development of a index to measure Tourism Satisfaction in Valle del Cauca. The main purpose of this work is to present a tool, among other existing tools, so that the Valle del Cauca Tourism Information System can provide valuable information to actors and stakeholders, highlighting opportunities for improvement and management successes.

The data used came from the 'Inbound' measurement instrument, provided by SITUR (Sistema de Información Turística del Valle del Cauca). The research was based on 5,101 records corresponding to individuals who visited the department as tourists. Variables related to their perception of aspects such as safety, cleanliness, hospitality, as well as their experience in restaurants and lodgings and visits to specific places were evaluated.

The research process involved several stages: initially, an exploratory analysis of the data was carried out, followed by a validation of the instrument through an Exploratory Factor Analysis (EFA). Subsequently, an indicator was constructed using Multiple Factor Analysis (MFA), Multiple Correspondence Analysis (MCA) and Principal Component Analysis for Mixed Data (FAMD) methodologies, the results of which were consolidated in a visualization dashboard.

Key words: Multiple Factor Analysis, Principal Component Analysis for Mixed Data, Multiple Correspondence Analysis, Exploratory Factor Analysis, Tourism, Index, Reflective Model, Latent Variable, Tourism Satisfaction.

Contenido

Abstract	IV
1 Introducción	2
1.1 Planteamiento del problema	3
1.1.1 Problema contextual	3
1.1.2 Problema estadístico	4
1.2 Justificación	4
1.3 Pregunta de investigación	5
1.4 Objetivos	5
2 Antecedentes	6
3 Marco teórico	9
3.1 Marco teórico contextual	9
3.1.1 Relevancia del turismo en Colombia	9
3.1.2 Turismo	10
3.1.3 Tipos de turismo	10
3.1.4 Diversificación de la oferta turística en el Valle del Cauca	12
3.1.5 Los Sistemas de Información Turística	12
3.1.6 CITUR	13
3.1.7 Sistema de Información Turística del Valle del Cauca	13
3.1.8 Turismo Receptivo	14
3.1.9 Satisfacción del cliente	14
3.1.10 Satisfacción turística	14
3.2 Marco teórico estadístico	15
3.2.1 Variables observables	15
3.2.2 Variables latentes y modelos de variables latentes	15
3.2.3 Modelo formativo	16
3.2.4 Modelo unidimensional	17
3.2.5 Análisis Factorial Exploratorio (AFE)	18
3.2.6 Matriz de correlación heterogénea	20
3.2.7 KMO (Kaiser-Meyer-Olkin)	21
3.2.8 Rotación oblicua <i>Promax</i>	22
3.2.9 Análisis de Componentes Principales (ACP)	22

3.2.10	Análisis de Correspondencias Múltiples (ACM)	24
3.2.11	Análisis Factorial Multiple (AFM)	26
3.2.12	Índice o medida compuesta	27
3.2.13	Validación criterio del índice	27
3.2.14	Factor tamaño	28
3.2.15	Bootstrap o remuestreo	28
3.2.16	Análisis de Clúster Jerárquico	29
4	Metodología	30
4.1	Origen de los datos	30
4.2	Procesamiento de los datos	31
4.3	Análisis exploratorio	32
4.3.1	Categorización del gasto	32
4.3.2	Evaluación de asociación entre variables	34
4.4	Evaluación de la validez del instrumento	34
4.4.1	Análisis Factorial Exploratorio - AFE	36
4.4.2	Ajuste del AFE	36
4.5	Construcción de los índices generales	37
4.5.1	Construcción de un índice con ACM	37
4.5.2	Construcción de un índice con ACP para datos mixtos	40
4.5.3	Construcción de un índice con AFM	43
4.6	Construcción de índices para restaurantes y alojamientos	46
4.7	Transformación rango de valores del índice	47
4.7.1	Comparación de los índices con estadísticos básicos	48
4.8	Desempeño de los índices	48
4.9	Puntos de corte para el índice seleccionado	48
4.10	Presentación de resultados para SITUR	49
5	Resultados	50
5.1	Análisis exploratorio	50
5.1.1	Percepción de la calidad en restaurantes y alojamientos	53
5.2	Evaluación del instrumento de medición	55
5.3	Comparación de los índices generales propuestos ($ISTV_i$)	59
5.3.1	Ecuación del índice general seleccionado	61
5.3.2	Puntos de corte para el índice general seleccionado	62
5.3.3	Nivel de satisfacción turística según el perfil del turista	63
5.4	Desarrollo de índices para restaurantes y alojamientos	66
5.4.1	Puntos de corte para ISTV-R y ISTV-H	66
5.5	Visualización de resultados	67
6	Conclusiones y trabajos futuros	68

7	Anexos	71
7.1	Tablas	72
7.2	Código usado	76
	Bibliografía	77

Lista de Figuras

3-1	Representación de las dos primeras componentes principales en un caso práctico.	23
3-2	Matriz de datos expresada en forma de tabla disjunta	25
4-1	Diagrama de flujo de la metodología	30
4-2	Método del codo para definir el número óptimo de clusters	33
4-3	Estratificación del gasto turístico en clusters definidos	33
4-4	Gráfico de sedimentación para elección de factores	35
4-5	Gráfico de sedimentación para el ACM	38
4-6	Gráfico de contribución de las variables al primer eje factorial en el ACM . .	38
4-7	Gráfico de sedimentación para el FAMD	40
4-8	Gráfico de contribución de las variables al primer eje factorial en el FAMD .	41
4-9	Círculo de correlaciones - FAMD	41
4-10	Gráfico de contribución de las variables cuantitativas al primer eje factorial .	44
4-11	Gráfico de contribución de las variables cuantitativas al primer eje factorial .	44
4-12	Gráfico de contribución de las variables cuantitativas al primer eje factorial .	45
4-13	Círculo de correlaciones - AFM	45
5-1	Gráfico de sedimentación para elección de factores	56
5-2	Resultados del análisis de clúster	57
5-3	Histogramas de los índice propuestos	60
5-4	Histograma de los puntajes obtenidos para el Índice de Satisfacción Turística del Valle del Cauca (ISTV) y sus puntos de corte	62
5-5	Valores promedios del ISTV por departamento en Colombia	65
5-6	Boxplot de los índices de satisfacción en restaurantes y hoteles	66

Lista de Tablas

3-1	Interpretación de los valores del índice KMO	21
4-1	Definición de rangos para determinar el tipo de calificación	34
4-2	Estadísticas descriptivas de las subíndices transformadas	42
5-1	Distribución de visitantes al Valle del Cauca por departamento de origen (julio 2017 - enero 2023)	51
5-2	Distribución por edades de los turistas según su procedencia	53
5-3	Definición de los rangos para el gasto turístico total	54
5-4	Calificaciones de alojamiento por categoría de Gasto	55
5-5	Calificaciones de restaurantes por categoría de Gasto	55
5-6	Factores y su contribución a la Varianza	57
5-7	Dimensiones y variables observables	58
5-8	Rango de valores teóricos de los índices	59
5-9	Estadísticas descriptivas de los índice transformados	59
5-10	Evaluación empírica del desempeño de los índice propuestos	60
5-11	Suma de pesos relativos de cada factor	61
5-12	Rangos mínimos y máximos para el nivel de satisfacción turística	62
5-13	Clasificación de los turistas según el Índice de Satisfacción Turística (ISTV) en función de variables demográficas	63
5-14	Clasificación de los turistas según el Índice de Satisfacción Turística (ISTV)	64
5-15	Clasificación de los turistas según el Índice de Satisfacción Turística (ISTV) y la intención de revisita	65
5-16	Estadísticas descriptivas de los índice de restaurante y alojamiento	66
5-17	Rangos mínimos y máximos para el nivel de satisfacción turística en hoteles	67
5-18	Puntos de corte para el nivel de satisfacción turística en restaurantes	67
7-1	Codificación de variables observables relacionadas con actividades	72
7-2	Codificación de variables observables relacionada con servicios turísticos	73
7-3	Codificación de variables observables relacionadas con el ambiente y la intención de volver	73
7-4	Motivos de viaje de turistas colombianos por departamento	74
7-5	Motivos de viaje de turistas extranjeros por país (excluyendo a Colombia)	75
7-6	Valores de KMO para los ítems	76

1 Introducción

El presente trabajo de investigación, tiene como propósito la creación de una medida sintética que le permita al Sistema de Información Turístico Regional (SITUR) conocer el nivel de satisfacción de los turistas que llegan al Valle del Cauca. Además, busca evaluar todos los posibles factores relacionados con la complacencia de cada uno de los viajeros (Ministerio de Comercio, 2022). La satisfacción del turista es un indicador clave en la competitividad de los destinos turísticos, ya que influye directamente en la probabilidad de repetición de visita y la recomendación del destino a otros potenciales turistas.

El turismo, como fenómeno global en constante crecimiento, desempeña un papel significativo en la sociedad contemporánea. A medida que el mundo se vuelve más interconectado y accesible, las personas buscan cada vez más experiencias y destinos turísticos que les permitan escapar de su rutina diaria, descubrir nuevas culturas y sumergirse en entornos distintos. Este crecimiento del turismo global, presenta tanto oportunidades como desafíos para los destinos, que deben adaptarse para satisfacer las crecientes y cambiantes expectativas de los viajeros.

En el territorio colombiano, el turismo tiene una gran importancia debido a su impacto económico y social (Ministerio de Comercio, Industria y Turismo de Colombia, 2022). Este sector genera, entre otros beneficios, la creación de nuevos empleos, el incremento de la producción, la optimización de recursos públicos, el mejoramiento de infraestructuras y la elaboración de políticas públicas. Además, incentiva al sector a generar y promocionar nuevos productos turísticos, estimula la creación de nuevos prestadores de servicios en el gremio y dinamiza la economía local (Brida et al., 2011). La evaluación de la satisfacción del turista permite identificar áreas de mejora y diseñar estrategias más efectivas para el desarrollo turístico sostenible.

Cabe destacar que, en el año 2022, los servicios turísticos aportaron aproximadamente 45 billones de pesos al PIB del país, según estimaciones del Ministerio de Comercio, lo que representa un incremento del 14.8 % en comparación con el año 2021 (Ministerio de Comercio, Industria y Turismo de Colombia, 2022). En este contexto, es fundamental centrar la atención en la evaluación y el perfeccionamiento de los servicios y productos ofrecidos en el sector (Ministerio de Comercio, 2022).

1.1. Planteamiento del problema

Los Sistemas de Información Turística Regional (SITUR) son plataformas financiadas e implementadas por el Gobierno colombiano con el fin de recopilar, gestionar y difundir información turística a nivel nacional. Su creación se remonta al año 2014, y su finalidad era generar información primaria, diferente a la obtenida a partir de otras entidades como el DANE, COTELCO, DIMAR, la Aerocivil y Migración Colombia, entre otras, para que los estamentos que intervienen en el turismo, como restaurantes, hoteles y agencias, tuvieran una idea clara de lo que estaba sucediendo realmente en el sector. Estos sistemas fueron ejecutados en diferentes departamentos a lo largo del país, como parte de una estrategia integral de desarrollo turístico (Cano, 2021).

Sin embargo, en el caso particular del departamento del Valle del Cauca, la financiación se mantuvo únicamente durante los dos primeros años de funcionamiento (Cano, 2021). Esta situación ha planteado varios desafíos en el sector y, como consecuencia, condujo a la búsqueda de subvenciones privadas por parte de los actores involucrados en la cadena productiva del turismo.

1.1.1. Problema contextual

La dependencia de la financiación privada ha llevado al SITUR del Valle del Cauca a adoptar una actitud de cuidado en el manejo de los recursos, priorizando la eficiencia y la optimización de los gastos. La reducción de los fondos disponibles ha resultado en una limitación significativa en la capacidad para generar datos consistentes y realizar análisis profundos. En consecuencia, la administración de SITUR del Valle del Cauca se vio impulsada a buscar herramientas que les permitieran obtener información clara y precisa para medir acertadamente los niveles de satisfacción de los turistas; medidas que, en asocio con otros indicadores existentes como los económicos, les servirán para tomar decisiones relacionadas con el sector, acciones que permitan fortalecer el turismo en el Valle.

Para una comprensión más profunda de los problemas financieros que afectan a los sistemas turísticos, es importante referirse a estudios como el de Cano (2021), quien discute la importancia de una gestión adecuada de los recursos financieros en el turismo para asegurar su sostenibilidad y crecimiento. La falta de financiación adecuada no solo impacta la capacidad operativa del SITUR, sino también su habilidad para adaptarse a las demandas cambiantes del sector y contribuir efectivamente al desarrollo turístico regional.

1.1.2. Problema estadístico

Teniendo en cuenta que la satisfacción es una variable latente, resulta imprescindible identificar el conjunto de variables observables, a partir de los ítems del instrumento de medición *Receptivo*, que permitan captar la presencia o ausencia de dicha variable en un individuo (Corbetta, 2003). En el caso de esta investigación, estas variables observables se refieren a los turistas que llegan al Valle del Cauca.

Para abordar este reto, se propone la creación de una medida resumen que proporcione una perspectiva más clara de la satisfacción de los turistas que visitan el departamento. Las medidas individuales ofrecen información detallada sobre aspectos específicos; sin embargo, carecen de una representación general que permita comprender integralmente el problema.

En esencia, se plantea como un desafío definir cuál es la metodología estadística multivariada más adecuada para la medición de la satisfacción de los turistas que arriban al Valle del Cauca. Es necesario considerar diversas técnicas multivariadas, como el Análisis de Correspondencias Múltiples (ACM), el Análisis de Componentes Principales (ACP) y el Análisis Factorial Múltiple (AFM), para determinar cuál de ellas puede captar mejor la varianza de las variables observables y representar de manera precisa la variable latente estudiada.

1.2. Justificación

La implementación de los SITUR representó un avance en la gestión de información turística a nivel nacional, proporcionando una estructura organizada para la recopilación y el análisis de datos del sector (Cano, 2021). Sin embargo, la falta de financiamiento sostenido afectó tanto la continuidad como la calidad de las mediciones. Esta limitación impidió realizar una evaluación precisa del impacto del turismo y restringió la capacidad de las autoridades para tomar decisiones estratégicas informadas, fundamentales para el desarrollo y la competitividad del sector.

El turismo es una actividad económica esencial que impulsa el crecimiento de las regiones al fomentar la creación de empleos y el desarrollo de infraestructura. En el caso del Valle del Cauca, la región posee una oferta turística variada que abarca desde paisajes naturales hasta sitios culturales y patrimoniales de gran relevancia. Esta diversidad no solo atrae a turistas nacionales e internacionales, sino que también genera divisas y aporta beneficios a la economía regional (Gobernación del Valle del Cauca, 2023).

Dada la importancia del turismo para el desarrollo económico del Valle del Cauca, resulta relevante definir la metodología estadística con técnicas multivariadas que garantice que

el índice desarrollado refleje adecuadamente la experiencia de los turistas y proporcione información valiosa que, apoyada en otras herramientas, sirva para la toma de decisiones en el ámbito turístico del Valle del Cauca.

Contar con datos precisos y actualizados permitirá a las autoridades tomar decisiones informadas y oportunas, mejorando su capacidad de respuesta ante los desafíos y oportunidades del turismo. Un SITUR robusto y bien financiado es esencial para maximizar los beneficios económicos y sociales del turismo en el Valle del Cauca, con el fin de asegurar un crecimiento sostenible y equilibrado del sector.

1.3. Pregunta de investigación

Teniendo en cuenta el anterior razonamiento, se plantea la siguiente pregunta de investigación: ¿Cuál es la metodología de análisis multivariado más adecuada para construir un índice que permita medir la satisfacción de los turistas que visitan el Valle del Cauca?

1.4. Objetivos

Con el propósito de responder a la pregunta de investigación, se establecen los siguientes objetivos:

Objetivo general

Desarrollar un índice, a través de métodos estadísticos multivariados, que permita evaluar el nivel de satisfacción general de los turistas que visitan el Valle del Cauca.

Objetivos específicos

- Desarrollar un índice mediante metodologías multivariadas que refleje la satisfacción de los viajeros con el servicio de alojamiento.
- Construir un índice estadístico que mida la satisfacción de los viajeros con el servicio de restaurantes y alimentos.
- Implementar un tablero de datos que proporcione al SITUR una plataforma organizada, monitoreada y actualizada para acceder a la información recolectada.

2 Antecedentes

Con el propósito de enriquecer el contexto y establecer un marco comparativo con los enfoques utilizados a nivel nacional e internacional para la medición de la satisfacción turística, se proporcionan a continuación antecedentes relevantes:

En el Valle del Cauca, se ha adoptado la metodología propuesta por la OMT en su documento de referencia *Indicators of Sustainable Development for Tourism Destinations: A Guidebook* (2005). Este enfoque ha evolucionado hasta convertirse en una herramienta esencial para medir la satisfacción de los clientes en el sector turístico, abordando áreas clave como la calidad de los servicios, la preservación del patrimonio cultural y natural, la gestión de los recursos naturales, la participación comunitaria y el impacto socioeconómico del turismo. La implementación de estos indicadores permite a las autoridades y a los actores turísticos tomar decisiones informadas para asegurar un desarrollo sostenible y una experiencia satisfactoria tanto para los visitantes como para la comunidad local.

Rivas Cano (2021) se enfoca en una evaluación técnica y financiera del SITUR Valle, presentándose como una respuesta a los desafíos económicos derivados de la falta de financiamiento gubernamental. Este estudio ofrece una alternativa valiosa para la toma de decisiones en la cadena productiva del turismo en el Valle del Cauca, abordando aspectos como la disponibilidad presupuestal, la demanda de indicadores turísticos y una evaluación minuciosa de los componentes técnicos del Sistema de Información Turística. Este documento contiene el conjunto de indicadores mínimos necesarios para el funcionamiento del SITUR, por lo que es relevante para el trabajo vigente.

A nivel nacional, destacan varias contribuciones significativas en el ámbito de la medición de la satisfacción turística. Entre ellas, se encuentra el trabajo realizado por **Otero Gómez et al. (2016)**, quienes llevaron a cabo un análisis correlacional cuantitativo utilizando datos de turistas nacionales, extranjeros y residentes. Este estudio, procesado con la herramienta estadística SPSS V.18, reveló niveles de satisfacción entre buenos y muy buenos en la ciudad analizada.

Los hallazgos de Otero y sus colaboradores no solo proporcionan una valiosa perspectiva sobre el nivel de satisfacción experimentado por diferentes *categorías de turistas*, sino

que también ofrecen un punto de partida sólido para la implementación de estrategias de mejora en el sector turístico a nivel nacional. La identificación de rutas de acción específicas derivadas del análisis correlacional permite a los responsables de la planificación turística y a los actores relevantes enfocar sus esfuerzos en áreas que contribuirán de manera efectiva a elevar la experiencia turística general. En Villavicencio, la satisfacción general manifestada entre residentes y turistas en actividades de recreo y ocio fluctúa entre *buena y muy buena*. Sin embargo, los principales actores de la oferta turística, junto con su infraestructura y servicios complementarios, deben realizar acciones más decididas para mejorar la oferta. Esta debe alinearse con las expectativas del consumidor turístico y ofrecer productos que generen una percepción de calidad. De esta manera, se logrará que la experiencia del turista sea gratificante y cumpla con las expectativas del visitante, sin depender únicamente del aprovechamiento de la riqueza natural de la región.

Castañeda (2021) realizó un estudio exhaustivo sobre la calidad percibida del servicio, la satisfacción y la intención de revisita de los turistas en destinos turísticos de Boyacá, utilizando técnicas avanzadas como el *Análisis Factorial Exploratorio (AFE)*, el *Análisis Factorial Confirmatorio (AFC)* y las *Ecuaciones Estructurales (SEM)*. Los resultados indicaron que la calidad percibida del servicio influye positivamente en la satisfacción del turista y que esta satisfacción, a su vez, influye positivamente en la intención de visitar el destino.

El análisis reveló que las dimensiones de la calidad del servicio con mayor impacto fueron las prácticas de sostenibilidad, actividades, accesibilidad, limpieza, seguridad, idioma de comunicación, hospitalidad, transporte local y alojamiento. Además, se encontró que las motivaciones de los turistas nacionales para viajar por impulso y atracción afectaron significativamente su satisfacción general y su disposición a recomendar y volver a visitar los destinos turísticos.

En conclusión, el estudio de Contreras ofrece una evaluación integral de las relaciones y dinámicas en la experiencia turística, proporcionando información valiosa para la gestión y mejora del sector turístico en Boyacá.

A nivel internacional, se destaca la investigación de **Jiang and Peng (2022)** sobre la medición de la satisfacción turística en Daegu, Corea del Sur, quienes desarrollaron un índice de satisfacción del turista utilizando un *modelo de ecuaciones estructurales*, lo que permitió una evaluación detallada de las relaciones y dinámicas en la experiencia turística.

Las conclusiones de Jiang y Peng subrayan la ausencia de diferencias significativas en la satisfacción de los turistas en función de factores demográficos como género y edad,

desafiando suposiciones previas. Además, identificaron que los factores más influyentes en la satisfacción incluyen la *calidad percibida del servicio, la percepción del valor y la lealtad del turista*. Este enfoque destaca la importancia de métodos analíticos avanzados para obtener resultados precisos en la evaluación de la satisfacción turística.

La investigación no solo enriquece la comprensión teórica de la satisfacción turística, sino que también ofrece un modelo metodológico valioso para investigadores y profesionales del turismo a nivel internacional. La aplicación exitosa de ecuaciones estructurales demuestra la relevancia de la innovación metodológica, sugiriendo que adoptar enfoques avanzados puede llevar a una comprensión más profunda de los factores que influyen en la satisfacción de los turistas.

Taborga (2016) realizó una recopilación bibliográfica comparando los *modelos formativo, reflexivo* y de antecedentes en la evaluación del servicio de docencia. Este trabajo discute la *utilidad de los modelos proyectivos y formativos multidimensionales*. El autor utilizó técnicas como el Análisis Factorial Exploratorio (*AFE*), el Análisis Factorial Confirmatorio (*AFC*) y el Modelado de Ecuaciones Estructurales (*SEM*), encontrando que el modelo de antecedentes, aunque con poder predictivo similar, tiene mejores índices de ajuste que los modelos reflexivo y formativo.

La investigación de Taborga (2016) proporciona una evaluación integral de los modelos de medición de calidad del servicio en la docencia, destacando la superioridad del modelo de antecedentes para la calidad del servicio de enseñanza. Este estudio se apoya en trabajos previos de J. Joseph Cronin and Taylor (1992) y Taylor and Baker (1994), quienes identificaron que la calidad del servicio es un antecedente significativo de la satisfacción y tiene una influencia notable en las intenciones de compra futuras. Además, Valdivieso Taborga menciona el modelo de calidad de servicio percibido y satisfacción desarrollado por Spreng and Mackoy (1996), integrando estos enfoques para ofrecer una visión comprensiva del impacto de la calidad del servicio en la satisfacción del usuario.

3 Marco teórico

3.1. Marco teórico contextual

Antes de abordar los aspectos estadísticos relevantes para este trabajo, es importante aclarar algunos conceptos de tipo contextual tratados en este documento.

3.1.1. Relevancia del turismo en Colombia

Desde tiempos inmemoriales, las personas se desplazaban a lugares distintos a sus residencias para asistir a certámenes deportivos, negociar productos, presenciar coronaciones o asistir a funerales. Estos viajes implicaban pernoctar en otros lugares, lo que les permitía conocer aspectos muy diferentes a los propios, como la alimentación, los paisajes, los bailes y los modos de comportamiento (Encyclopaedia Britannica, 2024).

Así nació el turismo, entendido como la actividad de desplazarse a un lugar diferente al de residencia habitual o actividades comerciales, ya sea a nivel local, regional, nacional o internacional. Esta actividad permite interactuar con otras personas, conocer y explorar manifestaciones de culturas diferentes (comida, bailes, artes, etc.) y visitar lugares, generalmente con fines de recreación u otros motivos (Tourism Notes, 2024).

A lo largo de los siglos, el turismo se consolidó como una verdadera actividad económica, permitiendo a los países obtener recursos para financiar gastos nacionales, satisfacer las necesidades básicas de la población y crecer económicamente, ya que genera empleo y divisas. Hoy en día, el turismo es considerado una fuente de ingresos tan importante como la agricultura, la ganadería y la minería, entre otros sectores (World Tourism Organization, 2024).

En Colombia, el sector turístico ha sido fundamental para el desarrollo del país. Esto se refleja en el Plan Sectorial de Turismo 2018-2022, donde se describe al turismo como “el nuevo petróleo de Colombia”. Este sector es considerado una opción viable y rentable para el desarrollo sostenible, potenciando la competitividad del país basada en su diversidad biológica y cultural. El turismo genera divisas, crea empleo y mejora las condiciones de vida de la población, además de promover la transformación regional y la reconciliación nacional

(Ministerio de Comercio, 2018).

A nivel global, el turismo es una de las actividades económicas más importantes, representando el 30 % de las exportaciones en el comercio de servicios y generando uno de cada 11 empleos, contribuyendo con el 10 % del PIB mundial (Organización Mundial del Turismo, 2016). Según la OMT, en 2017 se movilizaron 1.332 millones de viajeros a nivel mundial, y se proyecta un crecimiento anual del 3.8 % hasta el 2030. Para destinos turísticos en economías emergentes como Colombia, este crecimiento ha sido superior al promedio global, con una variación promedio anual del 4.8 % entre 2005 y 2017 (Organización Mundial del Turismo, 2018).

El Foro Económico Mundial destaca que, aunque tradicionalmente el desarrollo económico de los países se ha basado en el sector industrial, cada vez más estados estructuran modelos económicos donde el turismo es el principal dinamizador. En estos países, el crecimiento se sustenta en el sector servicios, donde el turismo tiene un efecto multiplicador significativo debido a la llegada de visitantes y su gasto en el territorio Foro Económico Mundial (2017).

3.1.2. Turismo

Según la Organización Mundial del Turismo (2023), organismo especializado de las Naciones Unidas, “el turismo es un fenómeno social, cultural y económico que supone el desplazamiento de personas a países o lugares fuera de su entorno habitual por motivos personales, profesionales o de negocios. Esas personas se denominan viajeros (que pueden ser turistas o excursionistas; residentes o no residentes) y el turismo abarca sus actividades, algunas de las cuales suponen un gasto turístico.”

3.1.3. Tipos de turismo

El turismo se divide en varias categorías, entre las cuales se encuentran el turismo de ocio, gastronómico, de aventura, cultural y formativo. La clasificación depende de las actividades que se realicen durante el desplazamiento. Estas actividades pueden incluir la compra de productos, la visita a museos, iglesias y galerías de arte, así como la exploración de ruinas, castillos y lugares históricos o arqueológicos. Además, el turismo puede involucrar caminatas por praderas y entornos ecológicos, visitas a parques naturales y ríos, o simplemente disfrutar de una playa y la cultura que la rodea.

Según lo indicado por Pérez Porto and Gardey (2022), los tipos de turismo pueden clasificarse de la siguiente manera:

Turismo de compras

Se caracteriza por estar organizado con el propósito de recorrer centros comerciales y otros espacios donde el viajero puede realizar compras.

Turismo cultural

Se centra en visitas a museos, exposiciones y otros lugares que permiten al viajero conocer más sobre la historia y la vida cultural del destino.

Turismo cultural

Se refiere a cualquier actividad turística que se desarrolla en un entorno rural, cerca de áreas naturales y litorales, permitiendo actividades relacionadas con el trabajo del sector rural. Es un tipo de turismo opuesto al masificado, común en ciudades costeras, y requiere un entorno endógeno sostenido por una comunidad humana y ambiental. Este tipo de turismo es respetuoso con el medio ambiente e intenta integrarse con otras actividades turísticas, aprovechando los recursos locales.

Turismo formativo

Tiene como objetivo ofrecer al viajero conocimientos sobre una materia específica, que es el motivo principal de su viaje.

Turismo gastronómico

Se enfoca en que los viajeros conozcan y degusten las comidas autóctonas del lugar, participando en actividades relacionadas con la cocina local.

Agroturismo

consiste en actividades en entornos rurales donde los turistas se alojan y pueden participar en las tareas propias del lugar, colaborando en la restauración de granjas o cortijos.

Ecoturismo

Es un tipo de turismo que se realiza en áreas naturales protegidas. Los residentes locales suelen recibir a los turistas y les muestran el entorno, priorizando la preservación de la naturaleza.

Turismo de aventura

Implica una serie de actividades en un entorno rural, diseñadas para ofrecer sensaciones de descubrimiento y poner a prueba los límites de supervivencia de los turistas. Se realiza en espacios que han sido escasamente explotados turísticamente.

3.1.4. Diversificación de la oferta turística en el Valle del Cauca

El Valle del Cauca, nuestro departamento, posee un gran potencial turístico. En el, se puede disfrutar de una diversidad de paisajes, cañaverales, ríos, sitios de interés, playas, museos e iglesias. Además, es posible realizar diversas actividades como canotaje, paracaidismo, pesca, entre otras.

El departamento cuenta con 42 municipios, cada uno con una rica diversidad de fauna y flora, lo que permite descubrir múltiples y variadas rutas y lugares para visitar. Entre las rutas más conocidas se encuentran:

- Hacienda El Paraíso
- Laguna de Sonso
- La basílica de los Milagros de Buga
- La ciudad de Santiago de Cali
- Parque Nacional Natural Farallones de Cali
- Museo Rayo
- Lago Calima
- La observación de ballenas en el Pacífico colombiano

3.1.5. Los Sistemas de Información Turística

Para obtener información relevante sobre el turismo, se crearon en el país los Sistemas de Información Turística, conocidos como SITUR. Su objetivo principal es recopilar datos para determinar el comportamiento del sector turístico en el país. Estos sistemas operan a nivel regional y proporcionan un soporte esencial para obtener información directa y específica sobre el turismo en cada una de las regiones del país.

3.1.6. CITUR

En 2016, el gobierno nacional colombiano, a través del Ministerio de Comercio, Industria y Turismo (MinCIT), creó el CITUR, o Centro de Información Turística de Colombia, con el objetivo de recopilar información sobre el turismo a nivel nacional, en lugar de regional.

En su lanzamiento, se destacó que este instrumento tecnológico sería seguro, formal, con reglas claras, capacitado, sostenible y responsable. Permitirá obtener, en un solo portal, información veraz y útil, proporcionando soluciones para maximizar el potencial turístico mediante indicadores como el registro nacional de turismo, empleo, flujos turísticos, tráfico aéreo y terrestre.

El CITUR se encargaría de realizar un análisis completo de las cifras nacionales, incluyendo las divisas generadas por el turismo y el crecimiento del sector. Se nutriría no solo de los SITUR, sino que también colaboraría con las regiones, empresarios y otras organizaciones para obtener información precisa, cuantitativa y cualitativa sobre el volumen, composición y dinamismo de las fuerzas productivas, la influencia del turismo en la economía nacional, y el número de empleos generados.

Esta plataforma fue reemplazada por la herramienta Portal de Información Turística de Colombia - PorTuColombia. Según el MinCIT, esta nueva plataforma asegura la transparencia de la información principal y atiende las recomendaciones de la evaluación del sistema estadístico nacional.

3.1.7. Sistema de Información Turística del Valle del Cauca

El Valle del Cauca, al igual que otros departamentos de Colombia, cuenta con un sistema de información turística conocido como SITUR. Según la página oficial del SITUR Valle del Cauca, este sistema se encarga de generar información estadística que orienta las decisiones de los diversos sectores involucrados en el turismo, tanto a nivel nacional como regional, con un enfoque especial en el Valle del Cauca. Para ello, han desarrollado y adaptado metodologías específicas para la recolección y análisis de datos, ajustándose a las necesidades de cada operador turístico y de las entidades encargadas de promover destinos turísticos (Sistema de Información Turística Regional del Valle del Cauca (2024)).

El SITUR Valle del Cauca fue creado en 2017 con el objetivo de observar y analizar el turismo en los 42 municipios del departamento. Este sistema se diseñó para recolectar datos cruciales sobre la oferta y la demanda del sector turístico, proporcionando información valiosa que facilita una planificación y gestión más efectivas de los recursos turísticos del Valle del Cauca.

3.1.8. Turismo Receptivo

El turismo receptivo se refiere al análisis de los viajeros que ingresan a una región o país, permitiendo identificar las actividades que realizan, sus gastos y otros aspectos importantes para medir el comportamiento del turismo en nuestro departamento.

Según la Organización Mundial del Turismo (OMT), el turismo receptivo, o turismo de entrada, incluye a todas las personas no residentes que viajan a una región específica con el objetivo de disfrutar, divertirse o participar en actividades particulares. Este tipo de turismo, al atraer visitantes de otras localidades o países, se convierte en un componente esencial de la industria turística, con un impacto significativo en el desarrollo económico y cultural de la región receptora Organización Mundial del Turismo (2023).

Recibir a estos turistas genera un flujo de ingresos derivados de la inversión en servicios turísticos, alojamientos y actividades recreativas. Además, enriquece culturalmente a la región mediante el intercambio de experiencias y perspectivas entre los visitantes y la comunidad local.

3.1.9. Satisfacción del cliente

La satisfacción del cliente se define como una evaluación post-consumo que compara la calidad percibida del producto o servicio con las expectativas del consumidor. Según Philip Kotler, profesor de marketing, la satisfacción es el nivel del estado de ánimo de una persona que resulta de comparar el rendimiento percibido de un producto o servicio con sus expectativas (Kotler and Keller, 2016).

3.1.10. Satisfacción turística

La satisfacción turística se considera un elemento clave en la construcción de la imagen y la competitividad de un destino (Otero et al., 2017). Según la definición de Kotler and Keller (2006), la satisfacción es el sentimiento que surge al comparar el desempeño experimentado con las expectativas de una persona. Sin embargo, el nivel de desempeño puede variar según el grado de expectativas del consumidor, las cuales, de acuerdo con Berry and Parasuraman (1991), se dividen en dos niveles: el nivel adecuado y el nivel deseado.

En otras palabras, la satisfacción turística puede asociarse, por un lado, al concepto de producto básico, donde el turista recibe un beneficio esencial para cubrir una necesidad fundamental, y por otro lado, al concepto de producto esperado, que consiste en un conjunto de atributos reales que marcan la diferencia en comparación con los productos ofrecidos por la competencia. Este enfoque permite entender cómo la satisfacción del turista no solo depende

del cumplimiento de necesidades básicas, sino también de la capacidad del destino para ofrecer experiencias y atributos diferenciadores que superen las expectativas del visitante.

3.2. Marco teórico estadístico

Conforme al problema estadístico que se pretende resolver en este trabajo de grado, es necesario definir los elementos involucrados en el modelo que busca darle respuesta. Este enfoque implica una comprensión profunda de las metodologías estadísticas y los **conceptos teóricos que fundamentan el análisis**.

3.2.1. Variables observables

Bollen (1989) describe las variables observables como aquellas que se pueden medir directamente y que representan constructos no observables o latentes. Estas variables permiten medir los constructos latentes mediante la observación y el análisis de sus indicadores.

3.2.2. Variables latentes y modelos de variables latentes

Tal y como se mencionó en la sección 1.1.2, en este trabajo nos interesa cuantificar la Satisfacción Turística de los viajeros que llegan al Valle del Cauca. Para ello, es fundamental entender que las **variables latentes** o **factores** son variables que no se pueden observar directamente, pero que se infieren a partir de otras variables que sí se pueden medir. Estas variables no observables representan factores subyacentes que influyen en las variables observables (Bartholomew and Knott, 1999).

Por ejemplo, consideremos la satisfacción del cliente como una variable latente. No es posible medir la satisfacción directamente, pero se puede inferir a partir de diferentes aspectos observables, tales como (Perron and Gillespie, 2015):

- **Calidad del servicio:** ¿Cómo evalúan los clientes la atención recibida?
- **Precio:** ¿Cómo perciben los clientes los precios en relación con el servicio recibido?
- **Recomendación:** ¿Estarían dispuestos a recomendar el servicio a otros?

Cada una de estas variables observables proporciona información sobre la satisfacción del cliente. Al analizar estas observaciones, podemos inferir la variable latente de **satisfacción** (Bollen, 1989).

En **términos matemáticos**, la relación entre las variables observables $\mathbf{x} = (x_1, x_2, \dots, x_p)$ y la variable latente $\mathbf{y}$ se puede modelar utilizando técnicas estadísticas como el análisis factorial o los modelos de ecuaciones estructurales (Bartholomew et al., 2011). Estos modelos permiten expresar la distribución conjunta de $\mathbf{x}$ y $\mathbf{y}$ para realizar inferencias sobre $\mathbf{y}$ basadas en $\mathbf{x}$.

Antes de incluir las variables latentes en el modelo, las variables observables ($\mathbf{x} = (x_1, x_2, \dots, x_p)$) pueden estar correlacionadas entre sí. Esta correlación es lo que se busca explicar mediante las variables latentes. Al introducir una variable latente $\mathbf{y}$, que actúa como un factor subyacente que influye en todas las variables observables, se modela la relación entre las variables observables y la variable latente (Belalcazar, 2023).

En el contexto de un **modelo de variables latentes**, se distinguen dos componentes principales: el modelo estructural o modelo interno y el modelo de medida o modelo externo (Belalcazar, 2023).

El **modelo estructural o modelo interno** se enfoca en las relaciones entre variables latentes, es decir, variables que no se pueden observar directamente pero que se asume que existen y subyacen a las variables observables. Este modelo examina cómo las variables latentes se relacionan y afectan entre sí, proporcionando una comprensión de las conexiones teóricas entre distintos conceptos latentes. Según Hair et al. (2017) y Bollen (1989), el modelo estructural es esencial para explorar estas conexiones hipotéticas (Bollen, 1989; Hair et al., 2017).

El propósito del **modelo de medida** es evaluar la validez y fiabilidad de las mediciones de las variables latentes. Según Hair et al. (2017), un modelo de medida bien especificado asegura que los indicadores seleccionados representen adecuadamente los constructos latentes, lo que es esencial para evitar sesgos y errores en la inferencia de las relaciones entre las variables latentes.

El modelo de medida se clasifica a su vez en dos tipos: el modelo reflectivo, que asume que las variables observables son manifestaciones de la variable latente, y el modelo formativo, que asume que las variables observables causan la variable latente (Allen and Seaman, 2007; Boari and Nai Ruscone, 2015; Hair et al., 2017).

3.2.3. Modelo formativo

Este modelo se basa en la premisa de que cada indicador captura una dimensión específica y esencial del constructo latente. En un modelo formativo, los indicadores no son

intercambiables; cada uno representa una dimensión distinta del constructo latente. La omisión o sustitución de un indicador implica la pérdida de información crucial sobre el constructo (Miranda Zapata, 2012).

A diferencia de los modelos reflectivos, en los modelos formativos no es necesario que los indicadores estén altamente correlacionados entre sí. Los indicadores pueden estar débilmente correlacionados o incluso no correlacionados, ya que cada uno captura un aspecto diferente del constructo ((Mazziotta and Pareto, 2013)). Cada indicador aporta una parte única y específica al constructo latente. Por lo tanto, la identificación y selección correcta de los indicadores es fundamental para la validez del modelo (Diamantopoulos and Winklhofer, 2001).

La variable latente y en un modelo formativo se define mediante la siguiente ecuación (Miranda Zapata, 2012):

$$y = \sum_p x_p w_p + \epsilon \quad (3-1)$$

Donde:

- y es la variable latente.
- x_p son los indicadores observables.
- w_p son los coeficientes que capturan el efecto de cada indicador en la variable latente.
- ϵ representa el error de medición.

3.2.4. Modelo unidimensional

En el ámbito del análisis factorial, según el nivel de abstracción descrito, los modelos que incluyen variables latentes pueden dividirse en cuatro categorías principales (Belalcazar, 2023; Brown, 2015; Dunn and McCray, 2020; Thurstone, 1944). Según un orden ascendente de complejidad estos son: modelos unidimensionales, modelos de factores correlacionados, modelos de alto orden y modelos bifactoriales.

Este trabajo se centra en el **modelo de alto orden**, aplicable cuando se tienen dos o más variables latentes o factores que son afectados por un factor o variable latente global. Este factor global relaciona dichos factores con los subgrupos de variables observables que se derivan de ellos. Las variables observables pueden describirse a partir de su relación con los factores locales específicos o en conjunto con el factor global (Dunn and McCray, 2020).

3.2.5. Análisis Factorial Exploratorio (AFE)

El Análisis Factorial Exploratorio (AFE) es una técnica estadística utilizada para identificar la estructura subyacente en un conjunto de datos multivariados. Según Tabachnick and Fidell (2013), el objetivo del AFE es descubrir las relaciones entre las variables observadas y reducir la dimensionalidad del conjunto de datos al agrupar variables correlacionadas en factores subyacentes comunes.

El AFE parte de la premisa de que la variabilidad en las variables observadas puede ser explicada por un número menor de factores latentes, que no son directamente observables. Estos factores representan dimensiones ocultas que subyacen en los datos y ayudan a simplificar y entender la estructura de las relaciones entre las variables. Además, este tipo de análisis factorial permite identificar patrones o estructuras en los datos, reducir la cantidad de variables necesarias para describir un fenómeno e interpretar mejor datos complejos (Fabrigar and Wegener, 2011).

Matemáticamente, el AFE está conformado por (Bartholomew and Knott, 1999; Fabrigar and Wegener, 2011):

$$\mathbf{X} = \mu + \mathbf{A}\mathbf{F} + \epsilon \quad (3-2)$$

Donde:

- $\mathbf{X}$ es una matriz $n \times p$ de datos observados, con n observaciones y p variables.
- μ es el vector de medias.
- $\mathbf{A}$ es una matriz $p \times k$ de cargas factoriales (α_{jk} indica la importancia que tiene el ítem j en el factor k).
- $\mathbf{F}$ es una matriz $k \times n$ de puntuaciones factoriales (f_{ik} es la puntuación del i -ésimo individuo en el k -ésimo factor).
- ϵ es una matriz $n \times p$ de términos de error, que representa la variabilidad no explicada por los factores latentes.

Adicionalmente, se tienen como supuestos del modelo (Bartholomew and Knott, 1999; Fabrigar and Wegener, 2011; Harman, 1976):

- Los factores comunes ($\mathbf{F}$) tienen media cero, varianza uno y no están correlacionados entre sí:

$$\mathbf{E}(\mathbf{F}) = 0, \quad \text{Cov}(\mathbf{F}) = \mathbf{I}_k \quad (3-3)$$

- Los factores únicos (ϵ) tienen media cero y no están correlacionados entre sí ni con los factores comunes:

$$\mathbf{E}(\epsilon) = 0, \quad \text{Cov}(\epsilon) = \mathbf{\Psi}, \quad \mathbf{E}(\mathbf{F}\epsilon) = 0 \quad (3-4)$$

Donde $\mathbf{\Psi}$ es una matriz diagonal.

Métodos para evaluar el análisis factorial

Para evaluar el ajuste del modelo factorial en el análisis factorial exploratorio (AFE), se puede utilizar el Root Mean Square of the Residuals (RMSR) y en análisis de las comunialidades.

El *RMSR*, conocido en español como el *Error Cuadrático Medio de los Residuales (ECMR)*, es una medida utilizada para evaluar el ajuste de un modelo factorial en el análisis factorial exploratorio (AFE). El ECMR se enfoca en la discrepancia entre la matriz de correlación observada y la matriz de correlación reproducida por el modelo factorial.

Se calcula según se enuncia en la Ecuación 3-5.

$$\text{RMSR} = \sqrt{\frac{\sum_{i < j} (r_{ij} - \hat{r}_{ij})^2}{\frac{p(p-1)}{2}}} \quad (3-5)$$

Donde:

- r_{ij} es el valor de la correlación observada entre las variables i y j .
- $\hat{r}_{ij}$ es el valor de la correlación reproducida entre las variables i y j según el modelo factorial.
- p es el número total de variables.

Por otra parte, las *comunialidades* representan la proporción de la varianza de cada variable que es explicada por los factores comunes.

Son calculadas según la Ecuación 3-6 (Bartholomew and Knott, 1999).

$$h_i^2 = \sum_{j=1}^m l_{ij}^2 \quad (3-6)$$

Donde:

- h_i^2 es la comunalidad de la variable i .
- l_{ij} es la carga factorial de la variable i en el factor j .
- m es el número de factores retenidos.

Y pueden ser interpretadas como (Harman, 1976):

- Valores altos de comunalidades, cercanos a 1, indican que la mayor parte de la varianza de la variable está explicada por los factores.
- Valores bajos de comunalidades, cercanos a cero, indican que la variable no está bien explicada por los factores.

3.2.6. Matriz de correlación heterogénea

Una matriz de correlación heterogénea incluye diferentes tipos de correlaciones, adaptadas para manejar las relaciones entre distintos tipos de variables, como variables continuas, ordinales y categóricas. Para variables continuas, se utiliza comúnmente la correlación de Pearson, siempre y cuando se cumpla el supuesto de normalidad de las variables involucradas. En el caso de variables ordinales, se emplean correlaciones policóricas o poliseriales, dependiendo de la combinación de variables involucradas. A continuación, se definen las correlaciones policóricas y poliseriales debido a su relevancia en este trabajo:

- **Correlación policórica para dos variables ordinales:**

Se estima utilizando modelos de máxima verosimilitud, que consideran la probabilidad conjunta de las categorías observadas dadas las variables latentes subyacentes (Olsson, 1979).

$$r_{\text{polychoric}} = \frac{\sum_{i,j} \phi^{-1}(P_{ij})}{\sqrt{\sum_i \phi^{-1}(P_{i.}) \sum_j \phi^{-1}(P_{.j})}} \quad (3-7)$$

Donde ϕ^{-1} es la función inversa de la distribución normal acumulativa y P_{ij} es la probabilidad conjunta de las categorías i y j .

- **Correlación poliserial** (para una variable ordinal y una continua): Similarmente, se asume que la variable ordinal es una manifestación discreta de una variable latente continua. La correlación poliserial $r_{\text{polyserial}}$ se estima mediante un modelo de máxima verosimilitud que relaciona la variable continua con la variable latente subyacente a la variable ordinal (Olsson et al., 1982).

$$r_{\text{polyserial}} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} \quad (3-8)$$

Donde x_i son los valores observados de la variable continua, y_i son las puntuaciones latentes esperadas, $\bar{x}$ y $\bar{y}$ son las medias de x_i y y_i , respectivamente.

3.2.7. KMO (Kaiser-Meyer-Olkin)

El índice KMO proporciona una medida de la adecuación de los datos para el análisis factorial, evaluando la proporción de la varianza en sus variables que podría ser causada por factores subyacentes comunes (Kaiser and Rice, 1974). Este índice ayuda a determinar si la muestra es lo suficientemente grande y adecuada para extraer factores significativos.

El coeficiente de KMO (Kaiser-Meyer-Olkin) se define como (Kaiser and Rice, 1974):

$$\text{KMO} = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} q_{ij}^2} \quad (3-9)$$

Donde:

- r_{ij} es el coeficiente de correlación entre las variables i y j .
- q_{ij} es el coeficiente de correlación parcial entre las variables i y j .

Los valores del índice KMO varían entre 0 y 1, los valores más altos indican que los datos son más adecuados para el análisis factorial. La interpretación de los valores del índice KMO es la siguiente (Kaiser and Rice, 1974):

Tabla 3-1: Interpretación de los valores del índice KMO

KMO	Interpretación
$\geq 0,90$	Excelentes
De 0,80 a 0,90	Muy buenos
De 0,70 a 0,80	Aceptables
De 0,60 a 0,70	Regulares
De 0,50 a 0,60	Pobre
$< 0,50$	Inacceptables

3.2.8. Rotación oblicua *Promax*

La rotación *Promax* es un método de rotación factorial en dos etapas: primero se realiza una rotación ortogonal, seguida de una transformación oblicua de la matriz de cargas factoriales $\mathbf{L}$ (Child, 2006). Este método maximiza la varianza de los elementos al cuadrado en una columna de la matriz de factores, simplificando así la estructura de las cargas factoriales.

Después de la rotación ortogonal inicial, se eleva la matriz de cargas factoriales a una potencia k . Esto da más peso a las cargas mayores y reduce el peso de las cargas menores. La matriz resultante se define como (Hendrickson and White, 1964):

$$\mathbf{L}^* = \mathbf{L}^{(k)} \quad (3-10)$$

Donde cada elemento l_{ij} de la matriz $\mathbf{L}$ se eleva a la potencia k :

$$l_{ij}^* = l_{ij}^k \quad (3-11)$$

3.2.9. Análisis de Componentes Principales (ACP)

Dentro de los métodos en el análisis multivariado, el ACP es una de las técnicas más populares y tiene como objetivo explicar la mayor parte de la variabilidad total observada en un conjunto de variables cuantitativas con el menor número de componentes posibles (Lebart et al., 1995).

Análisis de Componentes Principales tradicional

El análisis de Componentes Principales, tal y como lo introdujo Karl Pearson en 1901 e integró Hotelling (1933), considera un conjunto de datos inicial con $\mathbf{N}$ cantidad de registros y $\mathbf{P}$ cantidad de variables, con el objetivo de reducir el número de dimensiones necesarias para representar cada dato, buscando los $\mathbf{K}$ ($1 \leq \mathbf{K} \leq \mathbf{P}$) componentes principales. Estos componentes principales se calculan de forma iterativa, buscando cada vez el eje que explica la mayor inercia posible, siendo ortogonal a los componentes principales anteriores, tal y como se muestra en la Figura 3-1.

Es posible describir los datos a través de los $\mathbf{K}$ componentes principales. Aunque se vuelve más compleja su interpretación, permite conservar la mayor cantidad de información posible del conjunto de datos original, con un número muy reducido de variables. Estos componentes son la representación de nuevos ejes principales que representan y reducen, sin mantener ningún tipo de hipótesis o modelo particular el comportamiento de los datos, no obstante, con el fin de hacer inferencias certeras es importante mantener el concepto de normalidad (Blaufuks, 2021; González Rojas, 2014)

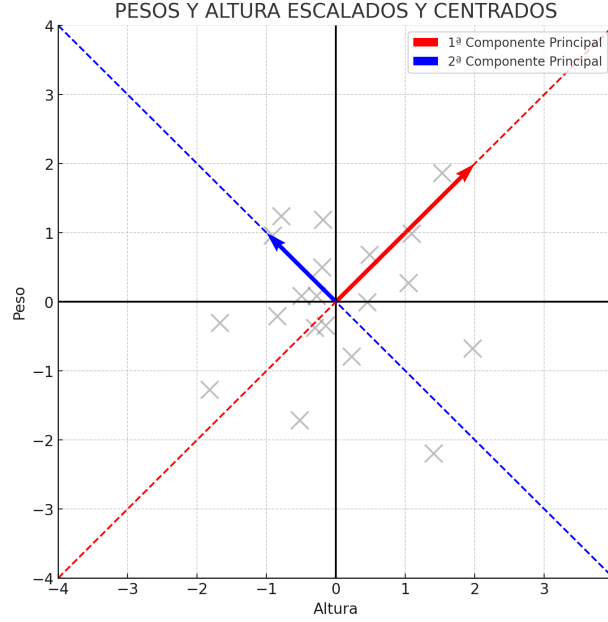


Figura 3-1: Representación de las dos primeras componentes principales en un caso práctico.

Análisis de Componentes Principales para datos mixtos (FAMD)

Pese a que el Análisis de Componentes Principales (ACP) es muy efectivo con datos numéricos continuos, este método no puede procesar datos cualitativos o categóricos (Blaufuks, 2021). Por esta razón, es importante introducir el Análisis de Componentes Principales para Datos Mixtos (FAMD), que tiene en cuenta tanto las variables numéricas como las categóricas, dándoles a cada una de ellas una importancia similar en la construcción de los componentes.

Si intentamos usar ACP en datos categóricos codificados, el resultado no es ideal. La razón es que el peso o la importancia que se le da a una variable categórica dependerá del número de categorías que tenga y de la frecuencia de cada una. Esto hace que sea difícil darle la misma importancia a todas las variables en el análisis. El objetivo del FAMD es dar el mismo peso a todas las variables, sin importar si son numéricas o categóricas. Para lograr esto, FAMD busca componentes que maximicen la correlación (r^2) para variables numéricas y una medida similar (η^2), tal y como se muestra en la Ecuación 3-12, para variables categóricas. Al hacer esto, se asegura que tanto las variables numéricas como las categóricas contribuyan de manera equilibrada a los componentes principales.

$$\eta^2(A, Y) = \frac{\sum_{m \in M_A} \sum_{k \in K_m} (\bar{y}_{*,m} - \bar{y})^2}{\sum_{m \in M_A} \sum_{k \in K_m} (y_{k,m} - \bar{y})^2} = \frac{SSE_{inter}}{SSE_{total}}, \quad (3-12)$$

donde:

- M_A : Modalidades de A
- K_m : Registros que toman la modalidad m
- $\bar{y}_{*,m}$: Media de Y para los registros K_m
- $\bar{y}$: Media global de Y

El numerador de η^2 mide la variación entre las modalidades respecto a la media global y el denominador mide la variación total en los datos. η^2 mide la proporción de la variabilidad total en la variable continua $\mathbf{Y}$ que puede explicarse por la variable categórica $\mathbf{A}$. Un valor de η^2 más alto indica una mayor relación entre la variable categórica y la variable continua.

A las variables numéricas se les asigna un peso constante $p_{\text{num}} = 1$. Esto significa que todas las variables numéricas tienen el mismo peso en el cálculo de la inercia proyectada. Por otra parte, a las variables categóricas se les asigna un peso proporcional a su probabilidad. La probabilidad μ_m de una modalidad es la proporción de datos que tienen esa modalidad en particular. Al dividir las columnas de las modalidades categóricas por esta probabilidad, ajustamos su contribución a la inercia proyectada de manera que refleje su rareza o frecuencia.

De esta manera, se construye cada componente principal C_k maximizando r^2 para variables numéricas y η^2 para variables categóricas. El vector w y, por ende, la componente principal C_k que maximiza la suma de las correlaciones cuadradas para las variables numéricas y las asociaciones cuadradas para las variables categóricas, se encuentran tal y como se describe en la Ecuación 3-13.

$$C_k = \arg \max_{\|w\|=1 \text{ con } w \perp \text{vect}(C_1, \dots, C_{k-1})} \left(\sum_{j=1}^P r^2(w, v_j) + \sum_{j=P+1}^{P+Q} \eta^2(w, v_j) \right) \quad (3-13)$$

Tener presente que, w debe ser un vector unitario (de longitud 1) y debe ser ortogonal a los componentes principales anteriores.

3.2.10. Análisis de Correspondencias Múltiples (ACM)

El ACM tiene como objetivo estudiar relaciones entre variables nominales expresadas en forma de tablas disyuntivas o tablas de Burt con información binaria. Así pues, para una sola variable nominal con j categorías, la tabla disjunta se expresa de forma binaria como se muestra en la Figura 3-2 (Belalcazar, 2023).

Partiendo de la Figura 3-2, supongamos que tenemos una variable categórica x_j con n observaciones $(i_1, i_2, \dots, i_n)$ y c_j categorías $(c_1, c_2, \dots, c_j)$.

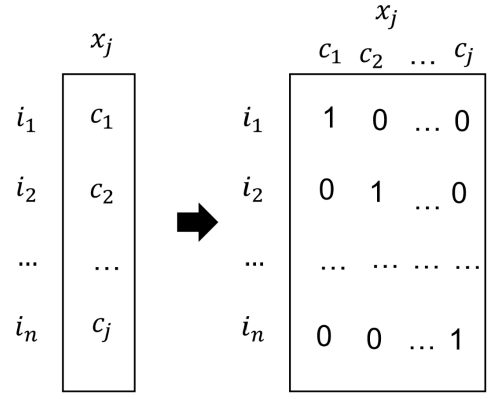


Figura 3-2: Matriz de datos expresada en forma de tabla disjunta

Convertimos la variable categórica x_j en una matriz de indicadores $\mathbf{Z}$. La matriz de indicadores $\mathbf{Z}$ tiene n filas (una por cada observación) y c_j columnas (una por cada categoría).

La variable original se puede representar como:

$$\mathbf{x}_j = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_j \end{pmatrix} \quad (3-14)$$

La matriz de indicadores $\mathbf{Z}$ correspondiente es:

$$\mathbf{Z} = \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1c_j} \\ z_{21} & z_{22} & \cdots & z_{2c_j} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{nc_j} \end{pmatrix} \quad (3-15)$$

Donde:

$$z_{ik} = \begin{cases} 1 & \text{si la observación } i \text{ pertenece a la categoría } c_k \\ 0 & \text{en caso contrario} \end{cases}$$

Después de construir la matriz de indicadores $\mathbf{Z}$, la Matriz de Burt $\mathbf{B}$ se construye como:

$$\mathbf{B} = \mathbf{Z}^\top \mathbf{Z} \quad (3-16)$$

Donde $\mathbf{B}$ es una matriz simétrica que muestra las frecuencias de co-ocurrencia de cada par de categorías. En la tabla de Burt, cada fila representa una observación y cada columna representa una categoría de las variables categóricas. De este modo, se asigna 1 a la presencia de la característica en la variable y 0 a la ausencia de la misma. (Ochoa Muñoz, 2018)

Se aplica SVD (descomposición de valores singulares) a la matriz de Burt para obtener las coordenadas principales.

$$\mathbf{B} = \mathbf{U}\mathbf{D}\mathbf{U}^\top \quad (3-17)$$

- $\mathbf{U}$ son los vectores singulares izquierdos.
- $\mathbf{D}$ es la matriz diagonal de valores singulares.

Las coordenadas principales de las observaciones y las categorías se obtienen multiplicando $\mathbf{U}$ por la raíz cuadrada de $\mathbf{D}$. La inercia total en ACM es una medida de la variabilidad de los datos y se calcula como la suma de los valores singulares al cuadrado, como se presenta en la Ecuación 3-18.

$$\text{Inercia total} = \sum_i d_i^2 \quad (3-18)$$

En cuestión de distancias, los individuos que aparecen cerca se parecen porque asumen las mismas modalidades de diferentes variables y es así como las modalidades de variables diferentes se consideran asociadas porque son asumidas por los mismos individuos.

3.2.11. Análisis Factorial Multiple (AFM)

El Análisis Factorial Múltiple (AFM) fue desarrollado por Escofier y Pagès (1992) como una estrategia para explicar la mayor parte de la variabilidad total observada en un conjunto de datos compuesto por variables de naturaleza cualitativa, cuantitativa o mixta, reteniendo el menor número posible de componentes. El AFM es una extensión tanto del Análisis en Componentes Principales (ACP) como del Análisis de Correspondencias Múltiples (ACM). En este método, se considera la contribución de todos los grupos activos de variables para definir la distancia entre individuos (Pagès, 2004).

La cantidad de variables en cada grupo puede variar, es decir, un grupo puede tener más variables que otro, y el tipo de variable, ya sea cualitativa o cuantitativa, puede diferir entre grupos. Sin embargo, las variables **deben ser de la misma naturaleza dentro de un grupo determinado** (Abdi and Williams, 2010).

Para completar este análisis, el AFM realiza tanto un análisis separado como un análisis global (Belalcazar, 2023):

En el análisis separado, se obtiene el primer valor propio a partir de la realización de un ACP para el grupo o tabla de variables cuantitativas y un ACM para el conjunto de variables cualitativas (Becerra Avella, 2010; Escofier and Pagès, 1992). En el análisis global, se construye la matriz de pesos para cada una de las tablas y se lleva a cabo un ACP ponderado sobre los conjuntos de datos.

3.2.12. Índice o medida compuesta

El índice es una medida compuesta utilizada para evaluar constructos complejos (Rojas, 2008). Esta se define mediante una función $f(x)$ que representa adecuadamente la medición de una variable latente (Belalcazar, 2023).

La función más comúnmente utilizada es una suma ponderada o combinación lineal de variables observables, como se muestra en la Ecuación 3-19. En esta ecuación, x_j representa un conjunto de variables observables y β_j son los pesos estimados para cada ítem, los cuales indican la contribución de cada variable observable en la construcción de la variable latente y .

$$\text{Índice} = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_j x_j \quad (3-19)$$

3.2.13. Validación criterio del índice

La validación de un índice es crucial para garantizar que mide con precisión y consistencia el constructo de interés. Según Hernández Sampieri et al. (2014), la validez de un índice se puede evaluar en términos de validez de contenido, validez de criterio y validez de constructo (Hernández Sampieri et al., 2014).

Validez de contenido

La validez de contenido se refiere a la medida en que los ítems de un índice representan adecuadamente todas las dimensiones y aspectos del constructo que se pretende medir. Para asegurar esta validez, es fundamental un proceso riguroso de revisión por expertos en el área. Estos expertos evalúan la relevancia y representatividad de cada ítem, garantizando que el índice cubra todos los aspectos del constructo de interés (Hernández Sampieri et al., 2014). Otros autores, como Haynes et al. (1995), también enfatizan **la importancia de la revisión por expertos para asegurar la validez de contenido**. (Haynes et al., 1995)

Validez de criterio

La validez de criterio se refiere a la efectividad de un índice para predecir o correlacionarse con otras medidas que se consideran válidas. Esta se subdivide en validez concurrente y validez predictiva. La validez concurrente se evalúa comparando el índice con otros instrumentos administrados simultáneamente, mientras que la validez predictiva se evalúa por la capacidad del índice para predecir resultados futuros (Hernández Sampieri et al., 2014). Según Cronbach y Meehl (1955), la validez de criterio es esencial para determinar la utilidad práctica de un índice. (Cronbach and Meehl, 1955)

Validez de constructo

La validez de constructo se refiere a la medida en que un índice realmente mide el constructo teórico que pretende evaluar. Para validar un constructo, es necesario que los resultados del índice correlacionen con otras medidas teóricamente relacionadas y no correlacionen con aquellas que no están relacionadas. La validación de constructo puede realizarse mediante análisis factoriales, los cuales permiten identificar la estructura subyacente de las relaciones entre los ítems (Hernández Sampieri et al., 2014). Nunnally y Bernstein (1994) destacan la importancia de la validez de constructo en el desarrollo de instrumentos de medida. (Nunnally and Bernstein, 1994)

3.2.14. Factor tamaño

El factor tamaño se manifiesta cuando los vectores propios del primer eje factorial presentan el mismo signo, lo cual indica que las variables apuntan en la misma dirección. En consecuencia, la primera componente principal se genera debido a la alta correlación existente entre todas las variables, creando un conjunto homogéneo de vectores. (Aluja Banet and Morineau, 1999)

Bajo estas condiciones, el primer componente principal se utiliza para construir el índice (variable compuesta) que distingue entre individuos con valores altos y bajos en todas las variables analizadas. Esto implica que un valor alto en una variable probablemente se asocie con valores altos en las demás variables.

3.2.15. Bootstrap o remuestreo

El *bootstrap* es una técnica estadística de remuestreo que permite estimar la distribución de un estadístico a partir de una muestra de datos. Esta técnica es especialmente útil cuando no se pueden hacer suposiciones sobre la distribución de los datos o cuando el tamaño de la muestra es pequeño (Efron, 1979).

1. **Muestra Original:** Se toma una muestra original de tamaño n de una población.

2. **Remuestreo con Reemplazo:** A partir de esta muestra original, se generan múltiples muestras *bootstrap* de tamaño n , seleccionando datos al azar con reemplazo.
3. **Cálculo del estadístico:** Para cada una de las muestras *bootstrap*, se calcula el estadístico de interés.
4. **Estimación de la distribución:** A partir de los valores del estadístico obtenidos en los diferentes remuestreos, se estima la distribución del mismo.

3.2.16. Análisis de Clúster Jerárquico

El análisis de clúster jerárquico es una técnica que permite agrupar variables o individuos en clústeres basados en la similitud de sus características, utilizando las cargas factoriales obtenidas. El procedimiento se desarrolla siguiendo estos pasos:

1. Se seleccionan las cargas factoriales.
2. Se calcula la matriz de distancias entre las variables utilizando dichas cargas factoriales. Para ello, se utiliza la distancia euclidiana, definida en la Ecuación 3-20:

$$d_{ij} = \sqrt{\sum_{k=1}^m (l_{ik} - l_{jk})^2} \quad (3-20)$$

donde l_{ik} y l_{jk} son las cargas factoriales de las variables i y j en el factor k , y m es el número de factores.

3. Se realiza un análisis de clúster jerárquico utilizando un método de agrupamiento, como por ejemplo, el método *Ward*. El método de *Ward* minimiza la varianza total dentro de cada clúster. En cada paso, se combinan los dos clústeres que resultan en la menor varianza intraclúster. La varianza intraclúster está definida en la Ecuación 3-21:

$$\text{Varianza intraclúster} = \sum_{i \in D_k} \|l_i - \bar{l}_k\|^2 \quad (3-21)$$

donde D_k es el k -ésimo clúster, l_i es el vector de cargas factoriales de la variable i , y $\bar{l}_k$ es el vector de medias de las cargas factoriales del k -ésimo clúster.

4 Metodología

Este capítulo detalla la metodología empleada en este estudio, con el objetivo de proporcionar una comprensión clara y precisa de los procedimientos seguidos para alcanzar los resultados obtenidos. Se presenta el origen y la depuración de los datos, el análisis exploratorio, la validación del instrumento y la construcción de los índices. Con el objetivo de presentar de manera lineal la metodología utilizada, se muestra el diagrama de flujo de la metodología (Figura 4-1).



Figura 4-1: Diagrama de flujo de la metodología

4.1. Origen de los datos

Los datos empleados en este proyecto fueron recolectados a través de un instrumento específico conocido como *Encuesta de Turismo Receptivo*. Para la recolección de esta información, se utilizó la herramienta *KOBO Collect*, llevada a cabo por un grupo de

monitores cuidadosamente seleccionados por el Sistema de Información Turística Regional. Inicialmente, la base de datos contenía un total de 35.600 registros. Posteriormente, después de un meticuloso proceso de depuración y análisis de los datos, el número de registros válidos se redujo a 5.101. Este conjunto de datos final abarca un período de tiempo que va desde el 1 de julio de 2017 hasta el 1 de enero de 2023, proporcionando una visión comprensiva del turismo receptivo en la región durante dicho intervalo.

4.2. Procesamiento de los datos

Con el objetivo de establecer una base de datos robusta para el análisis de información y garantizar la integridad de la variabilidad de los datos, se llevó a cabo una depuración manual de los registros, complementada posteriormente con la intervención del software Python. Durante este proceso, se eliminaron los registros considerados inválidos, es decir, aquellos que no se ajustaban a las escalas de medición o al tipo de datos establecido. Principalmente, se realizaron las siguientes modificaciones en los datos:

- Eliminación de las columnas que contenían información sensible (nombre, correo, teléfono, etc.) para garantizar la seguridad y el anonimato del procedimiento.
- Eliminación de registros que tenían como respuesta palabras en lugar de la escala numérica solicitada.
- Descarte de registros con calificaciones fuera del rango numérico establecido.
- Remoción de registros de personas con edades menores a 16 años (tope establecido por el SITUR) o con valores atípicos, como 250 años.
- Supresión de registros pertenecientes a encuestas de otros departamentos, es decir, que no evaluaban propiamente la visita al Valle del Cauca.
- Descarte de la variable “tipo de medición”, ya que en todos los registros tenía la misma respuesta.
- Corrección y homogenización de los nombres de países y municipios.
- Verificación de la consistencia entre el motivo de viaje general y el motivo de viaje específico.
- Se verificaron los datos originales de la edad de los turistas y su clasificación en los rangos propuestos por el SITUR. Si la edad no correspondía al rango establecido, se ajustó el dato para que coincidiera. Este mismo procedimiento se aplicó a las variables de “Cantidad de noches en el departamento” y “Cantidad de municipios visitados”.
- Exclusión de registros con datos faltantes.

Los detalles completos de estas modificaciones están documentados en el informe titulado *Modificaciones a la base de datos*, que será entregado al SITUR. Este informe tiene como propósito proporcionar información detallada sobre las correcciones realizadas, permitiendo a SITUR comprender y abordar posibles mejoras en la recopilación de datos en futuras instancias. Asimismo, se sugiere capacitar a los monitores designados para optimizar la calidad y consistencia de los datos recopilados en próximas oportunidades.

4.3. Análisis exploratorio

Con el fin de obtener una idea inicial de los datos, se realizó un análisis descriptivo a partir de tablas y el cálculo de estadísticos básicos.

4.3.1. Categorización del gasto

Se identificó la necesidad de categorizar la variable *Tipo de gasto* con el fin de proporcionar al SITUR una herramienta para definir el gasto de los turistas que visitan el departamento. Esta variable se refiere a cuánto gastaron los turistas durante su estancia y es crucial para realizar tablas cruzadas que ofrezcan información inicial sobre la posible relación entre el tipo de gasto y la calificación que los turistas dieron a restaurantes y alojamientos. La categorización de la variable mencionada se llevó a cabo bajo el siguiente criterio:

$$\min_S \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2 \quad (4-1)$$

Donde $S = \{S_1, S_2, \dots, S_k\}$ es una partición del conjunto de datos y μ_i es la media de los puntos en S_i .

Debido a la asimetría de los datos y la imposibilidad teórica para definir las categorías de gasto a partir de los cuartiles, se utilizó el método de K-means (análisis de clúster) para establecer dichas categorías, resolviendo el problema de minimización tal como se muestra en la Ecuación 4-1.

Este método facilitó la división de las 5.101 observaciones en cuatro clústeres óptimos, definidos a través del método del codo, como se ilustra en la Figura 4-2. La suma de cuadrados se redujo significativamente hasta alcanzar el cuarto clúster. A partir de este punto, la reducción se estabilizó, indicando que un incremento en el número de clústeres no resultaría en mejoras significativas en la partición de los datos.

Finalmente, cada observación fue asignada al clúster con la media más cercana, con el fin de minimizar la variabilidad interna y agrupar los datos de manera más eficiente.

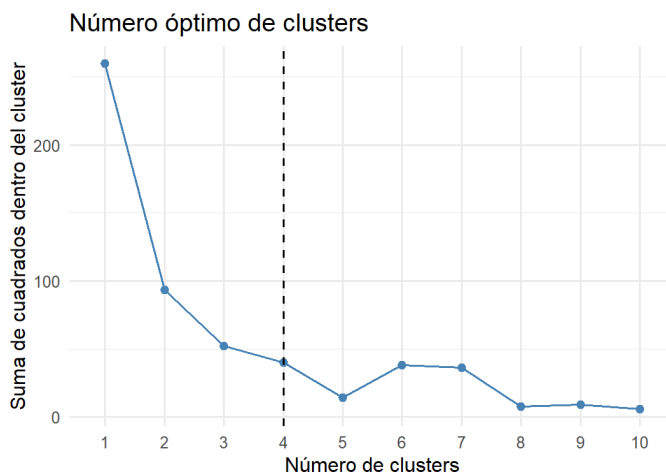


Figura 4-2: Método del codo para definir el número óptimo de clusters

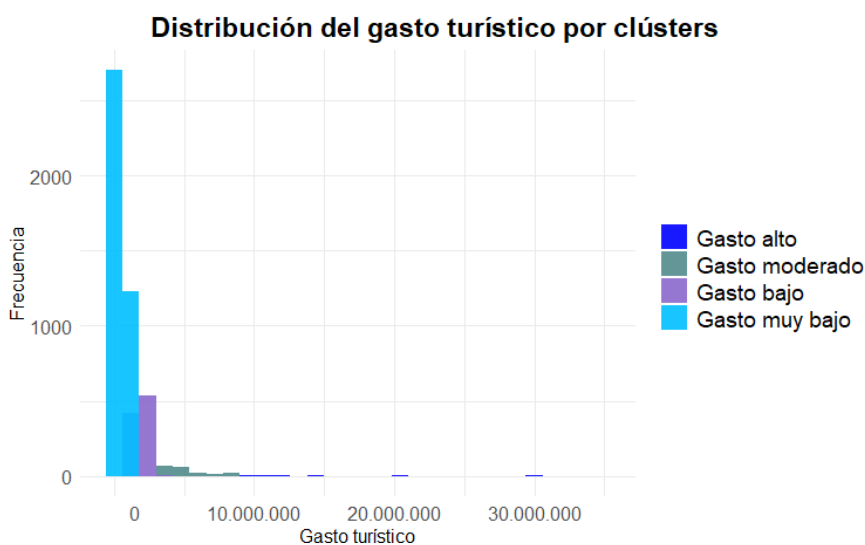


Figura 4-3: Estratificación del gasto turístico en clusters definidos

La clasificación del gasto turístico, apoyada por la Figura 4-3, se distribuye en los siguientes grupos:

- **Gasto bajo:** Representado por el cluster 4, este grupo contiene la mayor cantidad de observaciones, con un total de 4.096 y un gasto promedio de aproximadamente \$571.422 COP. Los montos en este cluster oscilan entre \$0 y \$1'440.000 COP.
- **Gasto moderado:** Definido por el cluster 3, incluye 921 observaciones con un gasto promedio de \$2'402.161 COP. Los gastos en este nivel varían de \$1'400.001 a \$5'500.000 COP.
- **Gasto alto:** Conformado por el cluster 2, este grupo incluye 78 observaciones y

un gasto promedio de \$8'700.641 COP, con gastos que oscilan entre \$5'500.001 y \$15'000.000 COP.

- **Gasto muy alto:** El cluster 1, siendo el más pequeño con sólo 6 observaciones, presenta un gasto promedio de \$27'500.000 COP. Los montos en este grupo se encuentran entre \$15,000.001 y \$35'000.000 COP.

4.3.2. Evaluación de asociación entre variables

Para evaluar la posible relación entre la variable *Tipo de gasto general*, categorizada en la subsección anterior, y la calificación asignada por los turistas tanto a restaurantes como a hospedajes, se llevó a cabo una prueba Chi-cuadrado para cada una de estas dos tablas de contingencia. Las calificaciones para estos sectores se clasifican según los criterios establecidos en la Tabla 4-1. Estos criterios de calificación fueron proporcionados por el SITUR y se originan en los rangos establecidos por el *Consejo Nacional de Estadísticas de Turismo para el CITUR-SITUR*, una entidad que ya no existe, según informó el Director de la regional Valle.

Tabla 4-1: Definición de rangos para determinar el tipo de calificación

Tipo de calificación	Rango	
	Mín.	Máx.
Mala calificación	1	5,99
Calificación regular	6	7,99
Buena calificación	8	10

Para llevar a cabo cada prueba, primero se calculó el estadístico de χ^2 considerando las frecuencias esperadas. Luego, se determinaron los grados de libertad asociados con la prueba y, finalmente, se calculó el valor p para la prueba Chi-cuadrado de cada tabla de contingencia (Minitab Support, 2024).

4.4. Evaluación de la validez del instrumento

Los datos empleados en esta investigación fueron extraídos de un instrumento de medición diseñado por Cano (2021) para su tesis titulada *Evaluación técnica y financiera del sistema de información turístico regional Situr para la toma de decisiones de los actores de la cadena productiva del turismo en el departamento del Valle del Cauca*. Cabe destacar que no se tuvo participación directa en el diseño o implementación del mencionado instrumento; los datos fueron proporcionados ya compilados y listos para su análisis.

Ante esta circunstancia, es imperativo entender y verificar la estructura del instrumento empleado. En consecuencia, se implementó un *Análisis Factorial Exploratorio (AFE)* con el objetivo de discernir las dimensiones fundamentales del instrumento y asegurar su pertinencia y eficacia para los objetivos de este estudio. El concepto del AFE se explican en la sección 4.4.1. Se tuvieron en cuenta los siguientes detalles:

Matriz de correlación heterogénea: Teniendo en cuenta que el *Análisis Factorial Exploratorio (AFE)* se utiliza comúnmente y preferentemente con variables continuas, se abordó el desafío de incluir variables nominales y ordinales en la base de datos mediante el uso de una matriz de correlación heterogénea, tal como se definió en la Subsección 3.2.6. Esta matriz fue adecuada porque permitió calcular diferentes tipos de correlaciones dependiendo de la naturaleza de las variables involucradas: policórica para dos ordinales, de Pearson para dos continuas y poliserial para una ordinal y una continua.

Evaluación del KMO (Kaiser-Meyer-Olkin): La tabla 7-6 mostró los resultados del índice de adecuación muestral de Kaiser-Meyer-Olkin (KMO) para el conjunto de datos analizado. El valor general del KMO fue de 0.92, lo que indicó una buena adecuación muestral para realizar el análisis factorial. Además, se presentaron los valores de KMO para cada ítem.

Gráfico de sedimentación

Se realizó un gráfico de sedimentación, también conocido como *scree plot*, para proporcionar una guía sobre la cantidad de factores que se debían retener en el análisis factorial. Este gráfico fue de ayuda para identificar el punto donde la varianza explicada por los factores adicionales comenzaba a disminuir de manera significativa.

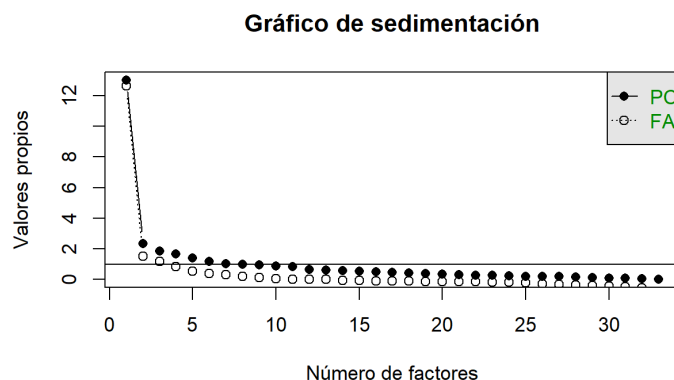


Figura 4-4: Gráfico de sedimentación para elección de factores

A partir de la interpretación de la gráfica de sedimentación, se decidió retener tres factores al identificar el punto de inflexión, donde cada factor adicional explicaba una proporción

significativamente menor de la varianza total. Esta decisión estuvo respaldada por la regla de Kaiser, ya que el valor propio del tercer factor superó el umbral de uno, lo que sugirió que estos factores eran significativos y debían ser retenidos.

4.4.1. Análisis Factorial Exploratorio - AFE

4.4.2. Ajuste del AFE

Después de evaluar la adecuación de los datos y determinar la cantidad de factores a retener, se procedió a realizar un Análisis Factorial Exploratorio (AFE) utilizando la función *fa* del software estadístico R y empleando el método de mínimos residuos (Harman and Jones, 1966). Se consideraron los siguientes aspectos:

Elección de una rotación oblicua

Se optó por implementar la rotación *Promax*, basada en el supuesto de que los nuevos factores están correlacionados (Méndez Martínez and Rondón Sepúlveda, 2012). En el contexto del análisis factorial exploratorio (AFE) aplicado al turismo, es probable que los diferentes factores o dimensiones influyan entre sí o compartan causas subyacentes comunes. Por ejemplo, las expectativas de los turistas pueden llevarlos a evaluar varios aspectos de manera similar. Un visitante que percibe un mal estado de las vías podría tener una percepción negativa general de la actividad turística, afectando simultáneamente múltiples dimensiones de su experiencia.

Aunque el modelo formativo de variables latentes no requiere que los indicadores estén altamente correlacionados entre sí, la naturaleza del fenómeno turístico sugiere que las dimensiones subyacentes a estas percepciones pueden estar interrelacionadas. La rotación oblicua, como *Promax*, permite captar estas posibles correlaciones entre factores latentes, reflejando mejor la realidad de cómo se forman las percepciones turísticas.

El uso de la rotación oblicua no contradice la premisa del modelo formativo, sino que complementa la interpretación de los datos al permitir que los factores latentes deriven de una estructura correlacionada, lo cual puede ser más representativo de la complejidad inherente a las experiencias turísticas.

Evaluación del AFE

El valor obtenido para el Root Mean Square of the Residuals (RMSR) fue de 0.04, lo cual es relativamente bajo y sugiere que el modelo presenta un buen ajuste. Además, el índice de ajuste basado en los valores fuera de la diagonal alcanzó un valor de 0.99, lo que indica un buen ajuste del modelo a las correlaciones observadas entre las variables.

Conformación de grupos

Después de realizar y analizar el Análisis Factorial Exploratorio (AFE), se llevó a cabo un análisis de clúster jerárquico con el objetivo de agrupar las variables en tres factores distintos.

4.5. Construcción de los índices generales

Para la construcción de los índices, se proponen tres metodologías diferentes. La primera utiliza el Análisis de Correspondencias Múltiples (ACM), la segunda aplica el Análisis de Componentes Principales para datos mixtos (FAMD), y la tercera emplea el Análisis Factorial Múltiple (AFM).

4.5.1. Construcción de un índice con ACM

Para evaluar si el Análisis de Correspondencias Múltiples, como se describe en la sesión 3.2.10, es una opción adecuada para desarrollar el *Índice de Satisfacción Turística*, se llevó a cabo el siguiente procedimiento:

Recodificación de las variables

Para adaptar las variables al análisis en el Análisis de Correspondencias Múltiples (ACM), se recodifican las que están en escala ordinal a una escala nominal. Siguiendo los lineamientos del Sistema de Información Turística (SITUR), reflejados en la Tabla 4-1, se redefinen las calificaciones como categóricas.

Utilizando la metodología de codificación *one-hot*, resulta en una base de datos con 95 variables tipo *dummy*, donde cada columna indica la presencia o ausencia de una categoría específica. Por ejemplo, los componentes de restaurante y alojamiento se dividieron en tres columnas *dummy* cada uno, correspondientes a alojamiento bueno, regular y malo, respectivamente.

Es crucial considerar que esta propuesta metodológica incluye la recodificación de variables, un proceso que podría comprometer la validez de interpretación de los resultados. Las variables ordinales, por su naturaleza, encapsulan información valiosa sobre el orden o rango inherente a sus categorías; información que inevitablemente se pierde al ser transformadas a un formato nominal (0/1). Dicha transformación homogeneiza todas las categorías, tratándolas como si fueran equidistantes entre sí, y por ende, omite cualquier jerarquía o gradación natural que las mismas podrían manifestar (MacCallum et al., 2002). Este razonamiento se tendrá en cuenta al momento de evaluar los resultados del índice y su habilidad para explicar la Satisfacción Turística.

Análisis del modelo

Se evalúa el ACM ajustado examinando el porcentaje de inercia explicada por los ejes (4-5) y la contribución de cada variable al modelo (4-6).

Se observa que el primer eje factorial contribuye con solo el 28,6% de la varianza, mientras que el segundo aporta un 8,7%, valores que se consideran bajos. Esto podría atribuirse a las modificaciones realizadas en las variables. Destacan particularmente en su contribución la variable ‘Calificación de platos en restaurantes’ (en la categoría de calificación ‘buena’) y ‘Evaluación del precio del alojamiento’ (también en la categoría de calificación ‘buena’).

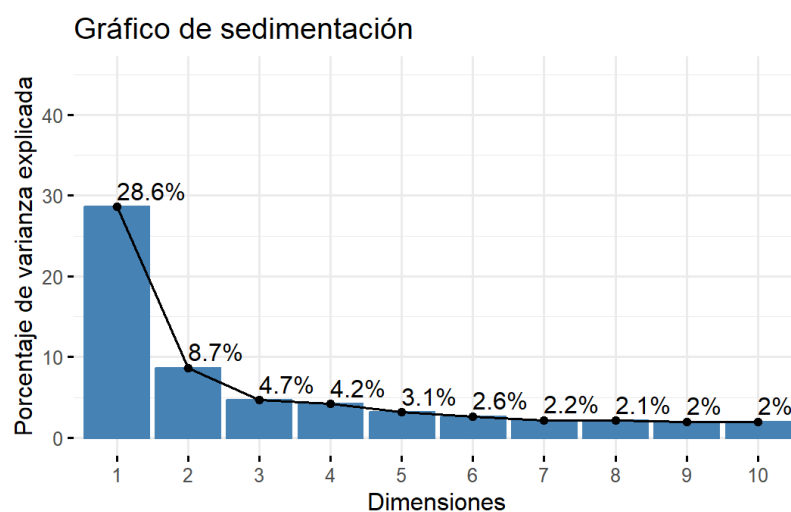


Figura 4-5: Gráfico de sedimentación para el ACM

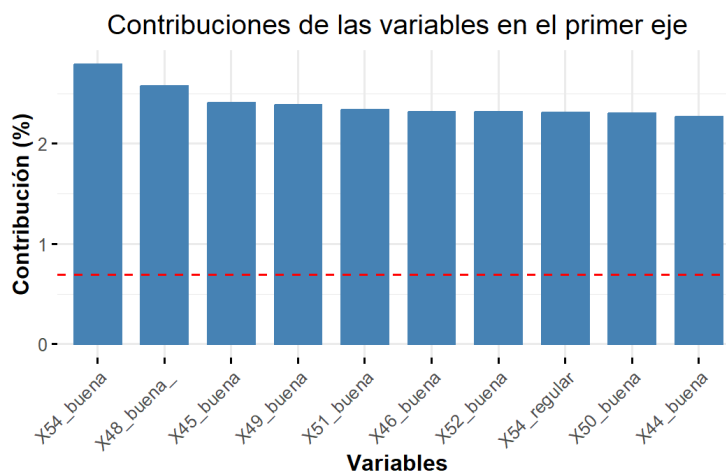


Figura 4-6: Gráfico de contribución de las variables al primer eje factorial en el ACM

Desarrollo del índice

Para el desarrollo del índice $ISTV_1$, se tuvo en cuenta la teoría existente para la construcción de índices con ACM (Ambapour, 2020).

Primero, se extraen las coordenadas de las variables en el espacio reducido definido por los componentes principales obtenidos del ACM. Usando estas coordenadas y para cada categoría c_k , se calcula su peso W_{c_k} . Este peso se ajusta con el primer valor propio (Ecuación 4-2)

$$W_{c_k} = \frac{F_{c_k1}}{\sqrt{l_1}} \quad (4-2)$$

Para la Ecuación 4-2 se define:

F_{c_k} es la coordenada de la categoría c_k en el primer eje del ACM y l_1 es el primer valor propio del ACM.

Por otra parte, $\mathbf{Z}$ es la matriz de respuestas codificadas, la cual es de tamaño $n \times q$. Cada entrada Z_{ic_k} en la matriz que indica si la observación i pertenece a la categoría c_k (1 para sí, 0 para no). Para esta se define un indicador binario I_{c_k} con la que se desarrolla el índice.

Dado que se tienen variables binarias originales, como las actividades que el turista realizó o no y ordinales convertidas en categóricas para el calculo del índice con esta metodología, se debe dar un manejo diferencial.

Para las variables binarias originales se calculó:

$$\text{Subíndice binarios} = \sum_{j=1}^9 W_{c_j} I_{c_j} \quad (4-3)$$

Para las variables recodificadas en varias categorías se tiene:

$$\text{Subíndices recodificadas} = \sum_{k=10}^{72} \sum_{j=1}^{J_k} W_{c_k} I_{c_k} \quad (4-4)$$

Donde:

- J_k es el número de categorías para la variable k
- W_{c_k} es el peso de la categoría c_k

- I_{c_k} es la presencia (1) o ausencia (0) de la categoría c_k para la observación i

De esta forma, el índice **ISTV₁** será:

ISTV₁ = Subíndices binarios + Puntaje recodificadas

4.5.2. Construcción de un índice con ACP para datos mixtos

Adecuación del modelo

Para evaluar si el Análisis Factorial de Datos Mixtos (FAMD), definido en la sesión 3.2.9, es una herramienta adecuada para construir el índice de *satisfacción turística*, se consideraron los siguientes aspectos:

1. Se verificó que los autovalores fueran mayores a 1.
2. Se comprobó el porcentaje de varianza explicado por los primeros componentes, asegurando que el primer componente explique un porcentaje alto de la varianza, seguido por el segundo componente y así sucesivamente.
3. Se visualizaron las relaciones entre las variables y los componentes principales.
4. Se verificaron las contribuciones de las variables y la calidad de la representación de las variables e individuos en el espacio de los componentes principales (Co^2).

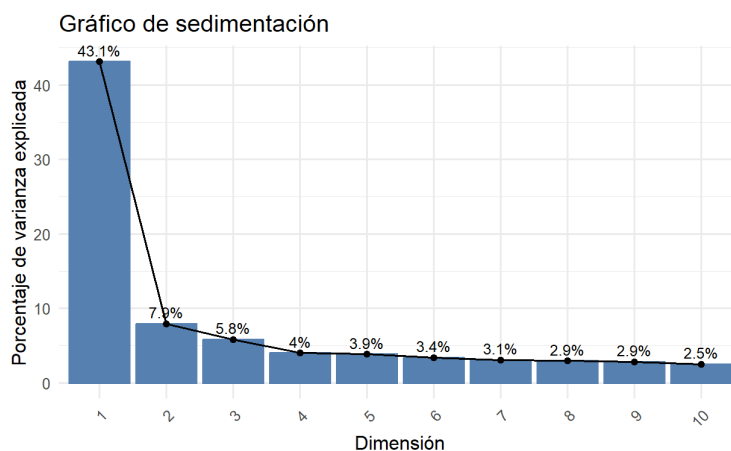


Figura 4-7: Gráfico de sedimentación para el FAMD

En las Figuras 4-8 y 4-7, se observa que el primer eje factorial del análisis realizado mediante el Análisis Factorial de Datos Mixtos (FAMD) explica el 43.1 % de la varianza

total. Adicionalmente, se destaca que la variable con mayor contribución a este eje es el componente restaurantes.

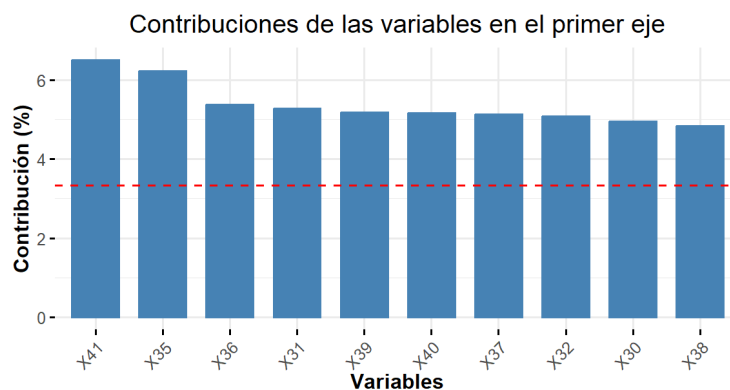


Figura 4-8: Gráfico de contribución de las variables al primer eje factorial en el FAMD

En la **Figura 4-9** se puede confirmar la presencia del factor tamaño para el modelo ajustado con FAMD (Subfigura 4.9(b)). Los vectores del primer eje factorial tienen el mismo signo, lo que indica que las variables están correlacionadas y apuntan en la misma dirección, sugiriendo que la primera componente principal es adecuada para la construcción del ISTV. En la Subfigura 4.9(a) se observa que las variables relacionadas con el componente de alojamiento (X35) y el componente de restaurantes (X41) tienen una buena representación en el espacio bidimensional, lo que implica que están bien explicadas por ambos componentes.

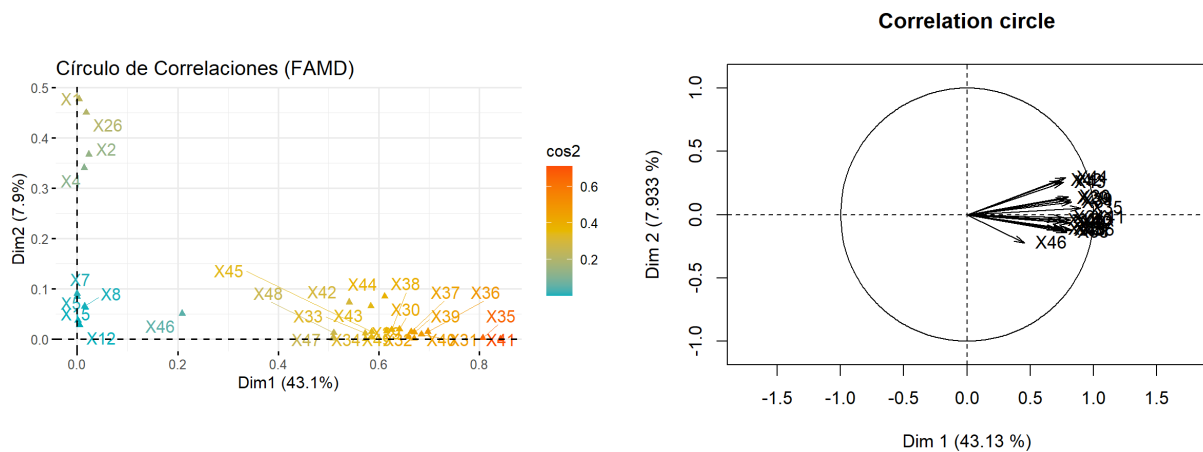


Figura 4-9: Círculo de correlaciones - FAMD

Desarrollo de subíndices

Seguido a la realización del ACP con datos mixtos, se plantean los subíndices por grupos, el resumen se encuentra en la Tabla 4-2.

Tabla 4-2: Estadísticas descriptivas de las subíndices transformadas

Comparación de las subíndices con FAMD						
Índice	Mín.	Máx.	Media	Mediana	Desviación estándar	C.V
SUBISTV₁	0	100	40,80	39,42	17,17	42,09
SUBISTV₂	0	100	86,67	88,91	13,36	15,41
SUBISTV₃	0	100	81,41	84,13	16,09	19,76

Desarrollo del índice

Para el desarrollo del índice con FAMD se sigue la siguiente metodología:

Primero, se visualizan las cargas de las variables cuantitativas y las contribuciones de las variables cualitativas. De manera siguiente, para cada grupo de variables, se calcula un peso de grupo (α_k) y las contribuciones de cada ítem j en el respectivo grupo k (β_{kj}^*).

El peso de cada grupo k se determina mediante la suma de los pesos de sus respectivos ítems (β_{kj}), tal y como se muestra en la Ecuación 5-11

$$\alpha_k = \sum_{j=1}^{m_k} \beta_{kj} \quad (4-5)$$

Donde:

- α_k es el peso del grupo k .
- m_k es el número de ítems en el grupo k .
- β_{kj} es el peso del ítem j en el grupo k .

Las contribuciones de cada ítem (β_{kj}^*) se obtienen dividiendo cada peso (β_{kj}) por el peso asignado a cada grupo k (α_k):

$$\beta_{kj}^* = \frac{\beta_{kj}}{\alpha_k} \quad (4-6)$$

Para calcular el *Índice de Satisfacción Turística* para cada grupo, se utiliza una suma ponderada de los ítems dentro de cada grupo, teniendo en cuenta el peso asignado al grupo

al cual corresponden.

Para los grupos 1 y 2, que contienen variables cuantitativas, el índice se calcula como se define en la Ecuación 4-7

$$\text{Subíndice}_{\text{grupo}} = \sum_j \left(X_{\text{quant}_{ij}} \times \beta_{kj}^* \right) \quad (4-7)$$

Para las cuales, $X_{\text{quant}_{ij}}$ es el valor de la variable cuantitativa j para el individuo i . Además grupo toma los valores de 1 y 2 (Calidad de servicios turísticos e infraestructura y servicios municipales).

Teniendo en cuenta que el tercer grupo está conformado por variables nominales, el índice del grupo uno se calcula como:

$$\text{subíndice}_3 = \sum_j \left(\text{Indicador}(X_{\text{qual}_{ij}} = \text{nivel}) \times \beta_{1j}^* \right) \quad (4-8)$$

Teniendo definidos todos los *subíndices*, se procede a calcular el $ISTV_2$ con la metodología de FAMD, teniendo en cuenta también los pesos de cada grupo, tal y como se define en la Ecuación 4-9.

$$ISTV_2 = \alpha_1 \cdot \text{Subíndice}_1 + \alpha_2 \cdot \text{Subíndice}_2 + \alpha_3 \cdot \text{Subíndice}_3 \quad (4-9)$$

Para la cual, la función *Indicador* devuelve 1 si la variable cualitativa $X_{\text{qual}_{ij}}$ está en el nivel específico y 0 en caso contrario.

4.5.3. Construcción de un índice con AFM

Adecuación del modelo

Para evaluar la adecuación del método estadístico empleado en la construcción de la $ISTV_3$, se analizó la contribución de los primeros componentes principales, así como la contribución específica de cada variable al modelo. Esto incluyó la revisión de los valores propios asociados a los componentes para determinar su relevancia en la explicación de la varianza del conjunto de datos, y el examen de las cargas factoriales para entender el peso y la influencia de cada variable en los componentes seleccionados.

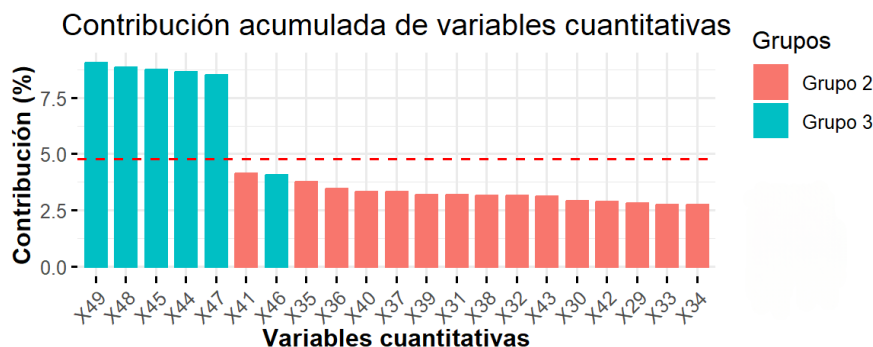


Figura 4-10: Gráfico de contribución de las variables cuantitativas al primer eje factorial

Como se observa en la Figura 4-10, las variables relacionadas con la valoración de la visita al departamento (X49), el componente de recompra (X48), la evaluación de la hospitalidad en los municipios (X45), el componente ambiental (X44) y la evaluación de la seguridad (X47) son las que más contribuyen al primer componente factorial. Estas variables, al tener una contribución alta, indican que son importantes para explicar la varianza dentro del modelo. Su fuerte relación con el primer eje sugiere que influyen de manera decisiva en la percepción global del destino turístico, demostrando su relevancia en la construcción del índice.

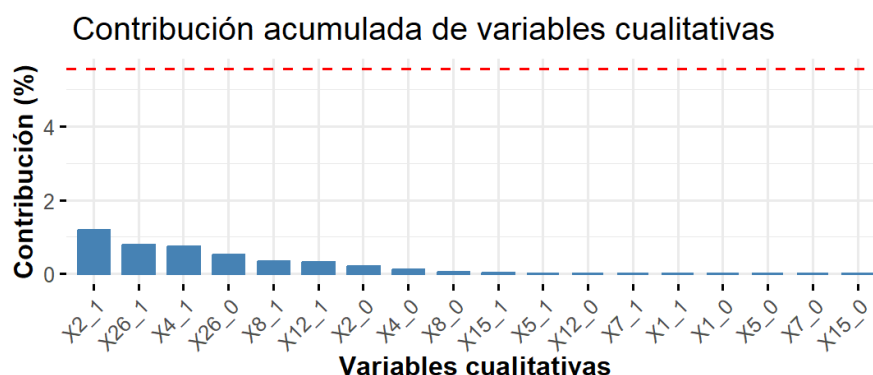


Figura 4-11: Gráfico de contribución de las variables cuantitativas al primer eje factorial

En el caso de las variables cualitativas, la contribución es mucho menor. En algunas, aunque positiva, es muy cercana a cero y, por lo tanto, insignificante (Figura 4-11). Las variables que más contribuyen son la visita a museos, casas de la cultura e iglesias; la realización de actividades religiosas; y la asistencia a playas, cuando estas actividades se realizaron (es decir, cuando $X_i = 1$). Por esta razón, se decidió mantener solo estas variables y retirar las demás variables cualitativas de la construcción del índice.

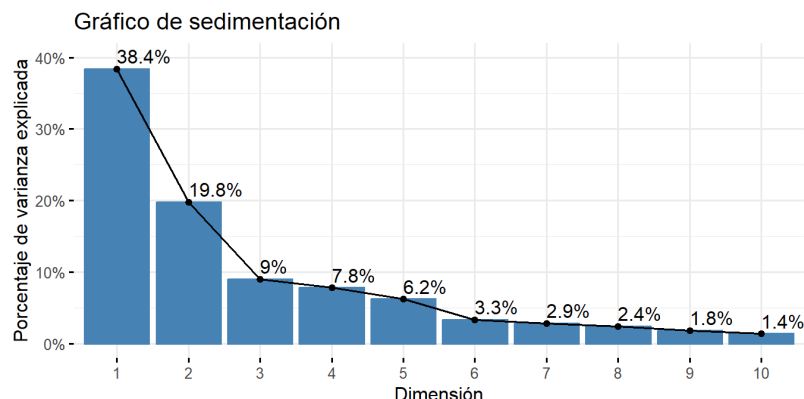


Figura 4-12: Gráfico de contribución de las variables cuantitativas al primer eje factorial

Después de eliminar las variables no significativas en la contribución a la construcción del índice, se evaluó el porcentaje de varianza explicado por cada componente. La Figura 4-12 muestra que la primera componente explica el 38.4% de la inercia total, seguida de la segunda componente que explica el 19.8%. En conjunto, estas dos componentes explican el 58.2% de la inercia, lo que significa que más de la mitad de la variabilidad presente en los datos originales puede ser capturada y representada por estas dos primeras componentes. Esto indica que las dos primeras componentes son bastante efectivas para resumir la información contenida en las variables originales.

En la Figura 4-13, se confirma la presencia del factor tamaño, lo que valida el uso de la primera componente principal para la construcción del índice. Además, se observa que variables como el componente restaurante (X41), el componente alojamiento (X35) y la calidad de los platos servidos en restaurantes (X36) están fuertemente correlacionadas con los dos primeros componentes. Esto sugiere que estas variables contribuyen de manera significativa a la explicación de la variabilidad total de los datos.

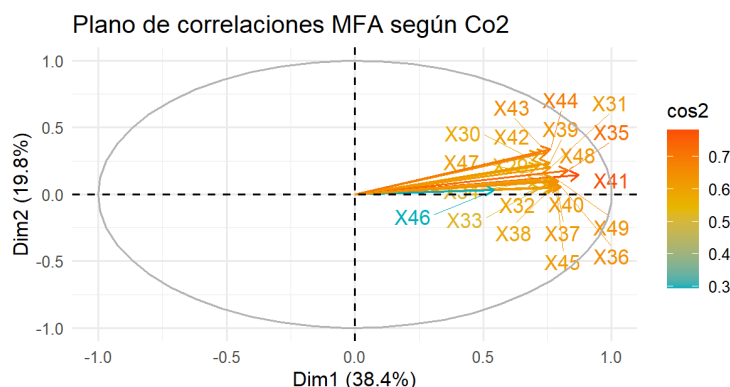


Figura 4-13: Circulo de correlaciones - AFM

Desarrollo del índice

Para el desarrollo del índice, se extrajo la primera coordenada ajustada del vector propio, v'_1 , que se calcula según la Ecuación 4-10.

$$v'_1 = \frac{v_1}{\sqrt{l}} \quad (4-10)$$

Los pesos relativos en el análisis se calcularon conforme a la Ecuación 4-11. Esta garantiza que la contribución de cada variable sea evaluada de manera equitativa.

$$\beta_k = \frac{v_k}{\sum_{i=1}^n v_i} \quad (4-11)$$

Posteriormente Se calcula el índice teniendo en cuenta la Ecuación 4-12.

$$\text{ISTV}_3 = \sum_{k=1}^{30} X_{ik} \beta_k \quad (4-12)$$

Donde β_k es el peso relativo para la k -ésima variable cuantitativa o binaria y X_{ik} es el valor de cada respuesta del individuo i .

4.6. Construcción de índices para restaurantes y alojamientos

Dada la necesidad de evaluar de manera separada la satisfacción en restaurantes y alojamientos, se diseñaron índices específicos para cada sector utilizando un enfoque basado en componentes principales. Este método permite condensar la información proporcionada por varias variables en un único índice representativo para cada categoría.

De manera inicial, se extrajeron las coordenadas de las variables para el primer componente para el método multivariado que haya resultado más eficiente en la construcción del ISTV.

Los índices para cada categoría se construyeron calculando las sumas ponderadas de las cargas factoriales y los valores correspondientes de las variables en el conjunto de datos. Estas cargas factoriales representan la contribución de cada variable a la variabilidad explicada por el primer componente principal, considerando que este componente capta la mayor parte de la información relevante.

Las variables relacionadas a cada grupo de interés son:

Variables alojamiento = $\{X_{29}, X_{30}, X_{31}, X_{32}, X_{33}, X_{34}, X_{35}\}$

Variables restaurantes = $\{X_{36}, X_{37}, X_{38}, X_{39}, X_{40}, X_{41}\}$

Para el sector de restaurantes, la *Índice de Satisfacción del Turista con los Restaurantes (ISTV-R)* se define como muestra en la 4-13.

$$ISTV-R_i = \sum_{k=36}^{41} X_{ik}\beta_k \quad (4-13)$$

Donde $ISTV-R_i$ es el índice de satisfacción para restaurantes para la observación i , X_{ik} son los valores de las variables para restaurantes, y β_k son las cargas factoriales correspondientes.

Para el sector de alojamientos, el *Índice de Satisfacción del Turista con los Hospedajes (ISTV-H)* se define de manera similar:

$$ISTV-H_i = \sum_{k=29}^{35} X_{ik}\beta_k \quad (4-14)$$

Donde $ISTV-H_i$ es el índice de satisfacción para alojamientos u hospedajes para la observación i , X_{ik} son los valores de las variables para alojamientos, y β_k son las cargas factoriales correspondientes.

4.7. Transformación rango de valores del índice

Se transformó el rango de los valores obtenidos en los índices para facilitar su interpretación y permitir comparaciones. Siguiendo los principios establecidos por Becerra Avella (2010), se llevó a cabo la siguiente transformación matemática:

$$ISTV_i^* = \frac{ISTV_{\text{Observado}} - ISTV_{\text{Mín}}}{ISTV_{\text{Máx}} - ISTV_{\text{Mín}}} \times 100 \quad (4-15)$$

En esta ecuación, $ISTV$ representó el valor original del índice obtenido para cada participante utilizando el método multivariado seleccionado. El valor transformado, $ISTV_i^*$, se ajustó para estar comprendido entre 0 y 100, lo cual estandarizó los resultados a un rango comúnmente aceptado y facilitó su interpretación.

Los términos $ISTV_{M\acute{a}x}$ y $ISTV_{M\acute{i}n}$ correspondieron, respectivamente, a los valores máximo y mínimo teóricos que pudo alcanzar el índice, de acuerdo con la metodología empleada, ya sea Análisis Factorial Múltiple de Correspondencias (AFMC), Análisis de Múltiples Correspondencias (AMC) o Análisis Factorial Múltiple (AFM).

4.7.1. Comparación de los índices con estadísticos básicos

Una vez transformados los valores de los índices $ISTV_i$, se realizó una comparación inicial utilizando estadísticos descriptivos básicos. Estos incluyeron el puntaje mínimo transformado, el máximo, la media, la mediana, el coeficiente de variación y la desviación estándar.

4.8. Desempeño de los índices

El desempeño de los índices fue evaluado mediante una estimación de su varianza.

Para ello, se empleó un procedimiento de muestreo con reemplazo sobre los n individuos de cada índice, replicando el proceso de construcción del índice 10.000 veces utilizando el método multivariado correspondiente. Esto permitió obtener 10.000 estimaciones de $\hat{\theta}$, basadas en los valores del Índice de Satisfacción Turística del Valle del Cauca (ISTV) calculados en cada iteración.

El promedio de estas 10.000 estimaciones se denota como el parámetro θ , y la varianza de dichas estimaciones se calcula mediante la siguiente fórmula:

$$\text{Var}(\hat{\theta}) = \frac{1}{n-1} \sum_{i=1}^n (\hat{\theta}_i - \bar{\hat{\theta}})^2$$

El estimador se considera insesgado si la media de $\hat{\theta}$ es aproximadamente igual al promedio de las estimaciones obtenidas. El mejor índice fue seleccionada basándose en el menor valor de PECM obtenido.

4.9. Puntos de corte para el índice seleccionado

Después de seleccionar el mejor *Índice de Satisfacción Turística para los visitantes del Valle del Cauca*, se propone clasificar a los individuos en tres categorías: Satisfacción baja, moderada y alta. Para esto, se emplean los métodos de K-means y bootstrap con el fin de identificar grupos de turistas con características de satisfacción similares.

Inicialmente, se llevó a cabo un remuestreo con reemplazo de las puntuaciones obtenidas por cada turista, manteniendo el mismo tamaño de muestra que los datos originales. Este procedimiento se repitió 10,000 veces para estimar el ISTV de cada turista en cada una de estas iteraciones.

Posteriormente, cada una de las 10.000 muestras generadas fue analizada mediante el método de K-means para categorizar la satisfacción turística en tres grupos distintos. Para cada grupo, se definieron límites inferiores y superiores basados en los valores obtenidos de la ISTV.

Se calcularon los valores mínimos y máximos de la ISTV para los turistas agrupados en cada nivel de satisfacción. Además, se estimaron los valores centrales de los rangos mínimos y máximos para cada uno de los tres grupos.

Finalmente, se establecieron los puntos de corte entre las categorías de satisfacción. El punto de corte entre la satisfacción baja y regular se definió como el valor máximo dentro del rango de satisfacción baja, y de manera similar, el punto de corte entre la satisfacción regular y alta se determinó como el valor máximo en el rango de satisfacción regular.

4.10. Presentación de resultados para SITUR

Se ha desarrollado un tablero en LookerStudio para presentar los índices construidos. Los datos generados a partir de los puntajes de los índices se han importado a un repositorio web privado, lo que permite organizar y visualizar los resultados de manera efectiva. Para explorar esta información visualmente, se invita a consultar el tablero correspondiente en la sección de resultados.

5 Resultados

En este capítulo, se presentan los resultados derivados de la metodología descrita en la sección anterior. Inicialmente, se realiza un análisis exploratorio de los datos para comprender mejor las características del conjunto de datos. Posteriormente, se lleva a cabo un *Análisis Factorial Exploratorio (AFE)* con el objetivo de validar el instrumento de medición utilizado.

Adicionalmente, se comparan los índices construidos utilizando las metodologías de: *Análisis de Correspondencias Múltiples (ACM)*, *Análisis Factorial Múltiple de Datos Mixtos (FAMD)* y *Análisis Factorial Múltiple (AFM)*. También se proporcionan detalles sobre los índices para evaluar la satisfacción del viajero con los servicios de restaurante y alojamiento.

Finalmente, se clasifican los turistas según variables demográficas y perfiles de viaje, basándose en los puntajes obtenidos en los índices mencionados.

5.1. Análisis exploratorio

Antes de proceder con la elaboración de los índices, resulta crucial identificar y describir a los encuestados que forman parte de la base de datos destinada a medir la satisfacción turística en el Valle del Cauca. Con este fin, se analizaron las variables de categorización disponibles para entender mejor los factores que influyen en la experiencia turística en la región.

Este análisis determinó que, de un total de 5.101 encuestados, 2.517 se identifican como femeninos, 2.569 como masculinos, y 10 eligieron una categoría de género diferente. Se encontraron 5 registros sin información de género. Asimismo, se registraron 23 nacionalidades distintas, predominando Colombia, Ecuador, Estados Unidos y España. Cabe destacar que 4.644 de los encuestados, equivalentes al 96,8 % del total, son colombianos, es decir, personas que residen en diversos municipios de Colombia y que viajaron al Valle del Cauca por motivos turísticos sin intención de establecerse de manera permanente.

Dado que la mayoría de los turistas en el Valle del Cauca son colombianos, resulta relevante determinar de qué departamentos provienen con mayor frecuencia. Para ello, se analiza la siguiente tabla:

Tabla 5-1: Distribución de visitantes al Valle del Cauca por departamento de origen (julio 2017 - enero 2023)

Departamento	Cantidad de encuestados
Valle del Cauca	2458
Cundinamarca	479
Antioquia	340
Risaralda	294
Quindío	199
Cauca	197
Caldas	187
Nariño	117
Tolima	85
Atlántico	42
Huila	27
Chocó	26
Meta	25
Caquetá	22
Santander	22
Putumayo	21
Boyacá	21
Norte de Santander	19
Bolívar	11
Córdoba	9
Magdalena	8
Sucre	8
La Guajira	6
San Andrés y Providencia	5
Cesar	5
Casanare	4
Arauca	4
Guaviare	1
Guainía	1
Amazonas	1

Como se muestra en la Tabla **5-1**, la mayoría de los viajeros encuestados son del Valle del Cauca y se trasladan a otros municipios dentro del mismo departamento. Les siguen los turistas provenientes de Cundinamarca, Risaralda y Antioquia, que se destacan como las principales fuentes de tráfico turístico. Esta distribución suscita una interrogante importante: ¿qué motiva a estas personas a visitar otros departamentos o municipios,

especialmente si están dentro del Valle del Cauca? Profundizar en las motivaciones de estos viajes podría proporcionar una visión más detallada sobre las dinámicas turísticas. Por esta razón, es fundamental analizar las razones de viaje según el departamento de origen de los turistas colombianos.

Según la Tabla **7-4**, la mayoría de los turistas encuestados se dirigieron al departamento por razones personales. Estas incluyen motivos como educación, tratamientos médicos o cirugías estéticas, visitas a amigos o familiares, y actividades de ocio como vacaciones o entretenimiento. Es relevante señalar que el estudio abarca datos de 30 de los 32 departamentos del país, con la excepción de Vichada y Vaupés, de los cuales no se dispone de registros. Además, hay 3 registros que se clasifican como datos faltantes, en los cuales no se pudo determinar el departamento de origen.

Asimismo, se realizó un análisis similar con turistas extranjeros para identificar sus principales motivaciones para visitar el departamento del Valle del Cauca, como se presenta en la Tabla **7-5**. Se encontró que tanto turistas locales como extranjeros visitan el departamento principalmente por motivos personales. En el caso de los turistas extranjeros, los datos se distribuyen de la siguiente manera: 100 visitaron por vacaciones o actividades recreativas (1,96 %); 12 asistieron a eventos culturales (0,23 %); 21 visitaron amigos o familiares (0,41 %); 3 participaron en eventos deportivos (0,05 %); 9 realizaron viajes de negocios (0,17 %); 3 exploraron la flora y fauna local (0,05 %); 1 buscó servicios de salud y bienestar (0,02 %); 1 pasó brevemente por el territorio (0,02 %); y 1 vino por motivos educativos. Además, hay 2 registros sin especificar el motivo del viaje.

En cuanto a los visitantes locales, 3.505 personas (68,71 %) eligieron el Valle del Cauca principalmente para vacacionar. Se contabilizaron 572 visitas para reunirse con familiares o amigos (11,21 %), y 85 visitas estuvieron vinculadas a eventos deportivos (1,66 %). La espiritualidad atrajo a 186 visitantes (3,64 %), mientras que 17 asistieron a eventos culturales (0,33 %). El turismo de negocios fue la elección de 94 personas (1,84 %), y el turismo de naturaleza captó el interés de 87 visitantes (1,70 %). Las conferencias y ferias comerciales atrajeron a 20 visitantes (0,39 %). Adicionalmente, 7 turistas viajaron por motivos educativos o de formación (0,13 %). La categoría 'Otro' incluyó a 25 personas (0,49 %), y 8 turistas optaron por el turismo orientado a la salud y bienestar (0,15 %). Por último, 34 personas estaban simplemente de paso por el departamento (0,66 %). Cuatro registros no cuentan con información sobre el motivo del viaje.

Por otro lado, además de analizar las preferencias de los viajeros para visitar el departamento del Valle, es relevante conocer la edad de las personas que reportan dichas visitas, la cual se detalla en la Tabla **5-2**.

Tabla 5-2: Distribución por edades de los turistas según su procedencia

Rango	Colombia				Otros países	
	Valle del Cauca	%	Otros departamentos	%	Cantidad	%
15-26	509	9,97	358	7,01	14	0,27
27-38	852	16,7	814	15,9	58	1,13
39-50	422	8,27	490	9,60	48	0,94
51-62	211	4,13	238	4,66	18	0,35
Más de 63	98	1,92	74	1,45	5	0,09

La mayoría de los viajeros se sitúa en el rango de edad de 27 a 38 años, una tendencia que se mantiene constante en todas las clasificaciones analizadas: originarios del Valle del Cauca, quienes constituyen la mayoría de los turistas, así como aquellos provenientes de los otros 28 departamentos con datos disponibles y los visitantes de países extranjeros. Además, la edad más frecuentemente registrada en la base de datos es 35 años. Esto sugiere que las actividades ofrecidas en el Valle del Cauca están alineadas con las expectativas de este grupo etario. Sin embargo, también representa una oportunidad para diversificar los servicios hacia otros grupos de edad y potencialmente incrementar el turismo.

5.1.1. Percepción de la calidad en restaurantes y alojamientos

Para obtener una visión preliminar del grado de conformidad de los turistas que visitan el departamento, se llevó a cabo un análisis de las variables contenidas en las tablas 5-4 y 5-5.

Se analizó el gasto de los turistas en la región, estableciendo rangos de gastos a partir de datos recolectados. Se observó un gasto promedio de aproximadamente \$1'057.946 COP, mientras que la mediana fue de \$600,000 COP. Esta diferencia significativa sugiere una distribución asimétrica de los gastos. Se registró un gasto mínimo de \$0 COP, lo cual es coherente dado que algunos turistas reportaron no haber incurrido en gastos durante su estancia. Por otro lado, el gasto máximo alcanzó los \$35'000.000 COP. La desviación estándar de los gastos fue de 1'621.979 COP, reflejando una amplia variabilidad en la cantidad de dinero que los turistas gastan en la región.

La categorización del gasto se realizó en la metodología (Sección 4.3.1), quedando definida de la siguiente manera como se muestra en la Tabla 5-3. Asimismo, las clasificaciones de las calificaciones se definieron en dicha sección.

Se realizaron comparaciones entre los tipos de gasto y las calificaciones generales asignadas por los turistas a los alojamientos y restaurantes, como se muestra en la Tabla 5-4. Los resultados indican que la mayoría de las calificaciones son 'Buenas' en todos los grupos

Tabla 5-3: Definición de los rangos para el gasto turístico total

Tipo de gasto	Rango	
	Mín.	Máx.
Gasto bajo	0	1.440.000
Gasto moderado	1.440.001	5.500.000
Gasto alto	5.500.001	15.000.000
Gasto muy alto	15.000.001	35.000.000

de gasto, reflejando una percepción generalmente positiva del alojamiento, que no varía significativamente con el nivel de gasto. Notablemente, el grupo de gasto ‘Bajo’ registra una concentración más alta de calificaciones positivas, sugiriendo que un gasto menor tiende a correlacionarse con experiencias de alojamiento más satisfactorias. Además, es interesante observar que las proporciones de calificaciones negativas son consistentemente bajas en todos los niveles de gasto y no muestran mucho incremento con los cambios en el gasto. Esto podría indicar que factores distintos al gasto tienen un mayor impacto en las experiencias negativas.

Se calculó la desviación estándar de las calificaciones de los alojamientos para cada grupo de gasto reportado por los turistas. Los resultados muestran que la variabilidad de las calificaciones es relativamente baja: 0,369 para el grupo de ‘Bajo gasto’, 0,300 para ‘Mediano gasto’ y 0,317 para ‘Alto gasto’. Estos valores sugieren que las experiencias de alojamiento son generalmente consistentes, aunque se observa una mayor variación en las experiencias entre aquellos que gastan menos.

Para continuar con el análisis, se aplicó una prueba de Chi-cuadrado para evaluar la posible asociación entre las variables categóricas ‘Tipo de gasto’ y ‘Tipo de calificación para alojamiento’, tal como se presenta en la Tabla 5-4. La hipótesis nula de esta prueba asume que no existe asociación entre estas dos variables.

Para llevar a cabo la prueba, primero se calculó el estadístico Chi-cuadrado, que mide la discrepancia entre las frecuencias observadas y las esperadas. El valor obtenido del estadístico fue de 6,37.

El valor p obtenido fue de aproximadamente 0.383, indicando que no hay suficiente evidencia para rechazar la hipótesis nula a un nivel de significancia de 0.05. Esto significa que no se encontró suficiente evidencia de una asociación significativa entre el ‘Tipo de gasto’ y el ‘Tipo de calificación para alojamiento’ al nivel de significancia del 5%.

La Tabla 5-5 muestra la distribución de las calificaciones de restaurantes segmentadas por categorías de gasto de los clientes. Es notorio que la mayoría de las calificaciones son positivas

Tabla 5-4: Calificaciones de alojamiento por categoría de Gasto

Tipo de gasto	Tipo de calificación para alojamiento					
	Bueno	%	Regular	%	Malo	%
Gasto bajo	3.397	66,5	633	12,4	65	1,27
Gasto moderado	770	15,0	135	2,64	16	0,31
Gasto alto	64	1,25	4	0,07	2	0,03
Gasto muy alto	12	0,23	3	0,05	0	0,00

(‘Bueno’), especialmente en las categorías de gasto más bajo, donde el 68,2 % de los clientes han valorado positivamente su experiencia. Esta tendencia disminuye significativamente a medida que aumenta el nivel de gasto, aunque sigue siendo predominantemente positiva incluso en los niveles más altos de gasto.

Una observación particularmente relevante es que las calificaciones negativas (‘Malo’) son bajas, no superando el 0,82 % en el gasto más bajo y están completamente ausentes en los niveles de gasto alto y muy alto. Esto puede sugerir que, a medida que aumenta el gasto, la satisfacción con la calidad del servicio o el producto ofrecido por los restaurantes tiende a ser mayor o al menos no resulta en experiencias negativas para los clientes.

De igual forma, se aplicó la prueba de Chi-cuadrado para explorar una posible asociación entre las variables categóricas ‘Tipo de gasto’ y ‘Calificación de calidad de restaurante’. El análisis reveló un estadístico de Chi-cuadrado de 11.01 y un valor p de 0.0881. Dado que el nivel de significancia es $\alpha = 0,05$, los resultados sugieren que, según los datos disponibles, no existe una relación estadísticamente significativa entre ambas variables.

Tabla 5-5: Calificaciones de restaurantes por categoría de Gasto

Tipo de gasto	Tipo de calificación para restaurante					
	Bueno	%	Regular	%	Malo	%
Gasto bajo	3.479	68,2	574	11,2	42	0,82
Gasto moderado	780	15,2	123	2,41	18	0,35
Gasto alto	64	1,25	6	0,11	0	0,00
Gasto muy alto	15	0,29	0	0,00	0	0,00

5.2. Evaluación del instrumento de medición

Con el fin de identificar las dimensiones que componen el instrumento de medición utilizado, se llevó a cabo un Análisis Factorial Exploratorio (AFE), utilizando el método de mínimos

residuos con rotación Promax. Es relevante anotar que, en una primera versión de este análisis, se hizo la exclusión de las siguientes variables debido a su poca contribución en el modelo: X_3 (Visita haciendas o casas históricas), X_6 (Realiza avistamiento de aves), X_9 (Visita a parques temáticos o parque de atracciones), X_{10} (Compra de bienes de consumo duradero), X_{11} (Realizar inversión o una reunión de negocio), X_{14} (Asiste a conferencias/congresos), X_{16} (Visita a fábricas o haciendas), X_{17} (Visita a casinos y otros sitios de juegos de azar), X_{18} (Visita a ríos, lagos o cascadas), X_{19} (Visita miradores paisajísticos), X_{20} (Visita zoológicos), X_{21} (Realiza proceso del café), X_{22} (Realiza proceso del azúcar dulce), X_{23} (Realiza actividades de aventura), X_{24} (Realiza senderismo o caminatas), X_{25} (Visita reservas o parques), X_{27} (Realiza otra actividad), X_{51} (Gasto alto), X_{52} (Gasto moderado), X_{53} (Gasto muy alto), X_{13} (Asiste a competencias o muestras deportivas), X_{54} (Intención de revisita sí) y X_{55} (Intención de revisita no sabe).

En la figura 5-1, se observa que las calificaciones del trato de personal en restaurantes y alojamientos, componente restaurante, componente alojamiento y evaluación de la calidad de estos servicios se agrupan en el lado derecho del eje X, indicando una fuerte representación por el primer factor. Las variables como componente ambiental, componente recompra, evaluación de los municipios (conservación y limpieza) y estado de las carreteras se agrupan en la parte superior del eje Y, sugiriendo una fuerte representación por el segundo factor. Finalmente, las variables relacionadas con actividades como recorrer calles, visitar museos y realizar actividades religiosas se encuentran cerca del origen, lo que podría indicar que no están bien representadas por ninguno de los dos primeros factores, sino por un tercer factor.

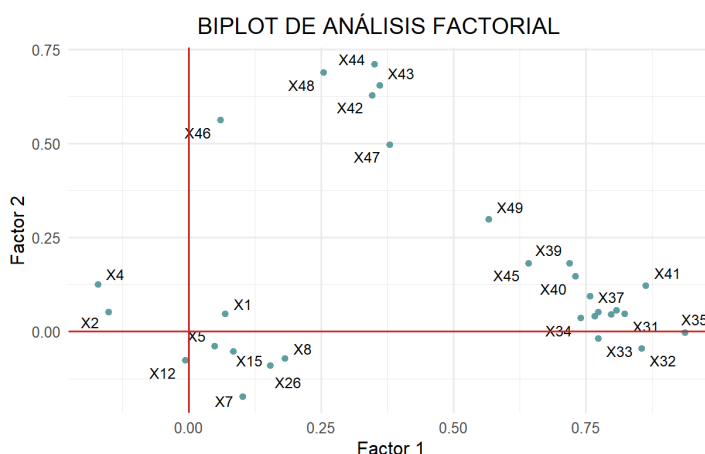


Figura 5-1: Gráfico de sedimentación para elección de factores

Dadas las características de cada grupo de variables, se interpretaron los factores como tres dimensiones que nos ayudan a entender la variable latente de satisfacción turística. Estas dimensiones se describen en la Tabla 5-6.

El factor de *Calidad de servicios turísticos* explica el 35 % de la varianza y se refiere a la percepción de los turistas sobre la calidad de los servicios proporcionados en los destinos turísticos, incluyendo la atención del personal y la oferta gastronómica, entre otros aspectos. Este factor tiene un SSL de 10.57, indicando una alta contribución al modelo. El factor de *Infraestructura y servicios municipales* explica el 12 % de la varianza y agrupa variables relacionadas con la infraestructura y los servicios ofrecidos por los municipios, como la limpieza, la conservación de lugares y el estado de las carreteras, con un SSL de 3.46, reflejando una contribución moderada. Por último, el factor de *Actividades turísticas y recreativas* explica el 6 % de la varianza y engloba las actividades recreativas y turísticas disponibles en el destino, como visitas a museos, parques y eventos culturales, con un SSL de 1.73, indicando una contribución más modesta en comparación con los otros factores.

Tabla 5-6: Factores y su contribución a la Varianza

Factor	Interpretación	Varianza (%)	SSL
1	Calidad de servicios turísticos	35	10,57
2	Infraestructura y servicios municipales	12	3,46
3	Actividades turísticas y recreativas	6	1,73

Posteriormente, se realizó un análisis de clúster jerárquico utilizando las cargas factoriales derivadas del análisis factorial exploratorio previo para agrupar variables relacionadas. Se eligió el método de Ward (Subsección 3.2.16) debido a su eficacia en minimizar la varianza total dentro de cada clúster, facilitando así una segregación precisa de las variables según sus patrones de correlación comunes. Este procedimiento resultó en la formación de tres clústeres distintos, como se muestra en la Tabla 5-6 y la Figura 5-2. Cada clúster representa una dimensión subyacente que no solo es clara, en términos de la distinción y cohesión entre las variables agrupadas, sino también teóricamente relevante, reflejando la construcción de constructos.

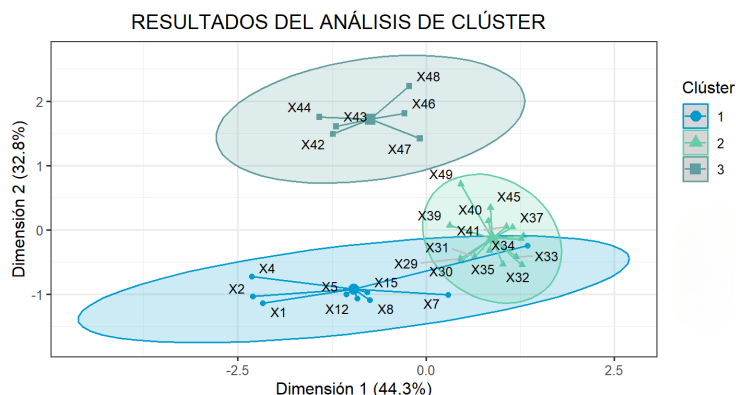


Figura 5-2: Resultados del análisis de clúster

En la Figura 5-2, las variables del *Clúster 1* muestran una dispersión considerable, lo que sugiere una diversidad en la percepción de las actividades turísticas y recreativas por parte de los turistas. Esta variabilidad podría estar relacionada con cómo diferentes expectativas individuales de entretenimiento y cultura influyen en la experiencia turística. Adicionalmente, la Tabla 5-7 presenta las variables que componen cada factor, incluyendo sus respectivas cargas factoriales, permitiendo apreciar la contribución de cada una al factor correspondiente.

Tabla 5-7: Dimensiones y variables observables

Factor	Ítem	Carga factorial
1	Estado del edificio de su alojamiento	0,77
	Estado de higiene y limpieza del alojamiento	0,80
	Estado de muebles de su alojamiento	0,82
	Calidad del trato personal en su alojamiento	0,85
	Calidad de la alimentación en su alojamiento	0,77
	Evaluación del precio de su alojamiento	0,74
	Componente alojamiento	0,93
	Calidad de los platos servidos en restaurantes	0,81
	Calidad oferta gastronómica en restaurantes	0,76
	Calidad del trato personal en restaurantes	0,77
	Estado de higiene y limpieza en restaurantes	0,72
	Evaluación del precio en restaurantes	0,73
	Componente restaurante	0,86
2	Evaluación de hospitalidad en los municipios	0,64
	Valoración visita al departamento	0,57
	Evaluación de limpieza y aseo en los municipios visitados	0,35
	Evaluación de limpieza y conservación de lugares en municipios	0,36
	Componente ambiental	0,35
	Evaluación estado de carreteras	0,56
3	Evaluación de la seguridad	0,38
	Componente recompra	0,25
	Recorre calles y parques del casco urbano	0,59
	Visita a museos, casas de la cultura, iglesias y monumentos	0,48
	Actividades religiosas	0,44
	Asistencia a espectáculos artísticos y festivales	0,21
	Asistencia a avistamiento de ballenas	-0,17
	Visita a discotecas bares y karaokes	0,20
	Práctica de deportes	0,12
	Asistencia a fiestas y ferias de pueblos o municipios	0,14
	Visita playas	-0,56

5.3. Comparación de los índices generales propuestos ($ISTV_i$)

Se construyeron tres índices ($ISTV_i$), como se describe en la Sección 4 de metodología. En la Tabla 5-8, se presentan los valores teóricos que estos índices pueden tomar.

Tabla 5-8: Rango de valores teóricos de los índices

$ISTV_i$	Metodología	Mín. teórico	Mediana	Máx. teórico
$ISTV_1$	ACM	-15.01	-0.191	76,097
$ISTV_2$	FAMD	44,78	147,21	163,94
$ISTV_3$	AFM	43,95	143,41	160,06

Cada índice propuesto, ya sea construido mediante el *Análisis de Componentes Principales para datos Mixtos*, *Análisis Factorial Múltiple* o *Análisis de Correspondencias Múltiples* se elaboró considerando el grupo de actividades turísticas y recreativas como variables categóricas. En contraste, los factores locales de calidad de servicios turísticos e infraestructura y servicios municipales se mantuvieron como variables numéricas.

Para mejorar la interpretación de los índices y garantizar la comparabilidad, estos fueron transformados. En la Tabla 5-9 se presentan las estadísticas descriptivas de los tres índices transformados, $ISTV_1$, $ISTV_2$ y $ISTV_3$, utilizadas para evaluar la satisfacción turística general en el departamento del Valle del Cauca, Colombia. Además, en la Figura 5-3 se muestra el histograma de cada índice para apreciar su distribución.

Tabla 5-9: Estadísticas descriptivas de los índice transformados

Comparación de los índice transformados						
Índice	Mín.	Máx.	Media	Mediana	Desviación estándar	C.V
$ISTV_1$	0	100	20,01	25,15	15,15	60,24
$ISTV_2$	0	100	83,73	85,96	14,25	17,02
$ISTV_3$	0	100	83,21	85,75	14,34	17,24

En la **Figura 5-3** se observa que los puntajes del índice construido con ACM van en sentido contrario a los índices construidos con FAMD y AFM, pareciendo ser más bajos. Esta inversión en los puntajes del ACM es probablemente una consecuencia directa de la recodificación de variables ordinales en binarias.

Los pesos negativos en el Análisis de Correspondencias Múltiples (ACM) se deben a la manera en que las coordenadas factoriales de las categorías están distribuidas en el espacio factorial. Las coordenadas factoriales en el ACM se calculan de modo que las categorías con

patrones similares de respuesta se sitúan cerca unas de otras en el espacio reducido. Estas coordenadas representan la posición de cada categoría en los ejes principales del espacio factorial y pueden ser positivas o negativas dependiendo de la relación de las categorías con el eje de variabilidad principal. Las categorías en extremos opuestos de un eje tendrán coordenadas con signos opuestos.

Como resultado, el índice se interpreta de manera opuesta en comparación con los índices obtenidos por los métodos de FAMd y AFM.

Los resultados indican que los índice transformados $ISTV_2$ y $ISTV_3$ presentan características estadísticas similares, lo que sugiere una comparabilidad consistente entre ambas. En contraste, los resultados para el índice $ISTV_1$ difieren significativamente. Por esta razón, se decide descartar este índice.

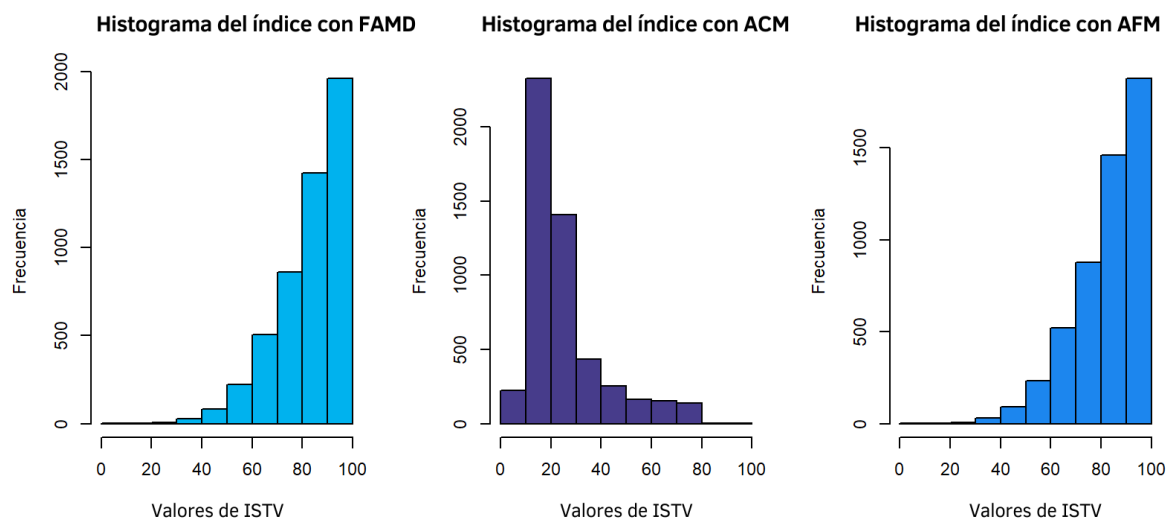


Figura 5-3: Histogramas de los índice propuestos

Teniendo en cuenta estas similitudes, se procede a realizar el remuestreo bootstrap, como se especifica en la sección de metodología 4.8, para estimar la distribución de la mediana y calcular su error estándar.

Tabla 5-10: Evaluación empírica del desempeño de los índice propuestos

Índice	SE	Varianza
$ISTV_2$	0,024	0,0227
$ISTV_3$	0,043	0,0153

Basado en estos criterios, el índice $ISTV_2$ parece ser la mejor opción. Tiene una media

ligeramente superior, y más importante aún, tiene un error estándar considerablemente menores que **ISTV₃**, indicando que es más precisa y confiable para la evaluación de la *Satisfacción Turística General* en el departamento azucarero.

Es importante destacar que la primera componente del *Análisis Factorial de Datos Mixtos (FAMD)*, el método utilizado para calcular el índice **ISTV₂**, explica un 43,1 % de la varianza total. Este valor refleja la relevancia de esta componente en la captura de la variabilidad principal dentro del conjunto de datos, sugiriendo que encapsula factores críticos que influyen en las dimensiones subyacentes del modelo.

5.3.1. Ecuación del índice general seleccionado

Con el objetivo de lograr una conceptualización más sólida de la escala seleccionada, se describen las ecuaciones utilizadas para construir el índice. En primer lugar, se determinaron los valores de alpha (α), es decir, los pesos relativos de cada factor para cada dimensión identificada en el Análisis Factorial Exploratorio (AFE), los cuales conforman cada subíndice. Estos valores se presentan en la Tabla 5-11.

Tabla 5-11: Suma de pesos relativos de cada factor

Factor	Interpretación	α_i
1	Calidad de servicios turísticos	12,16
2	Infraestructura y servicios municipales	4,163
3	Actividades turísticas y recreativas	-0,20

Las ecuaciones de los factores se definen como:

$$\text{Subíndice}_1 = \beta_{1,1}^* X_{1,1} + \beta_{1,2}^* X_{1,2} + \cdots + \beta_{1,15}^* X_{1,15} \quad (5-1)$$

$$\text{Subíndice}_2 = \beta_{2,1}^* X_{2,1} + \beta_{2,2}^* X_{2,2} + \cdots + \beta_{2,6}^* X_{2,6} \quad (5-2)$$

$$\text{Subíndice}_3 = \beta_{3,1}^* X_{3,1} + \beta_{3,2}^* X_{3,2} + \cdots + \beta_{3,9}^* X_{3,9} \quad (5-3)$$

Por lo tanto, la ecuación general del **ISTV** seleccionado es:

$$\text{ISTV} = 12,16 \times \text{Subíndice}_1 + 4,163 \times \text{Subíndice}_2 - (0,20 \times \text{Subíndice}_3) \quad (5-4)$$

La Ecuación 5-4 sugiere que un aumento en las actividades turísticas está asociado con efectos negativos en la satisfacción del turista.

5.3.2. Puntos de corte para el índice general seleccionado

Para categorizar la satisfacción turística en el Valle del Cauca en tres niveles distintos: baja, moderada y alta, se emplearon los métodos de K-means y Bootstrap. Estos métodos ofrecen una manera robusta y precisa de identificar los umbrales clave en el Índice de Satisfacción Turística (ISTV). Los resultados obtenidos son los siguientes:

Tabla 5-12: Rangos mínimos y máximos para el nivel de satisfacción turística

Nivel de satisfacción	Mín.	Máx.
Satisfacción baja	0	70,30
Satisfacción regular	70,30	88,64
Satisfacción alta	88,64	100

Según los puntos de corte presentados en la **Tabla 5-11**, los turistas se clasificaron según su nivel de satisfacción durante la visita al departamento. Un total de 2,132 turistas (41.8 %) tuvieron un nivel de satisfacción turística alta, 2,101 turistas (41.1 %) tuvieron una satisfacción moderada, y 868 turistas (17.0 %) tuvieron una satisfacción baja. Estos resultados se pueden observar gráficamente en la **Figura 5-4**, que presenta el histograma del Índice de Satisfacción Turística (ISTV) con sus respectivos puntos de corte.

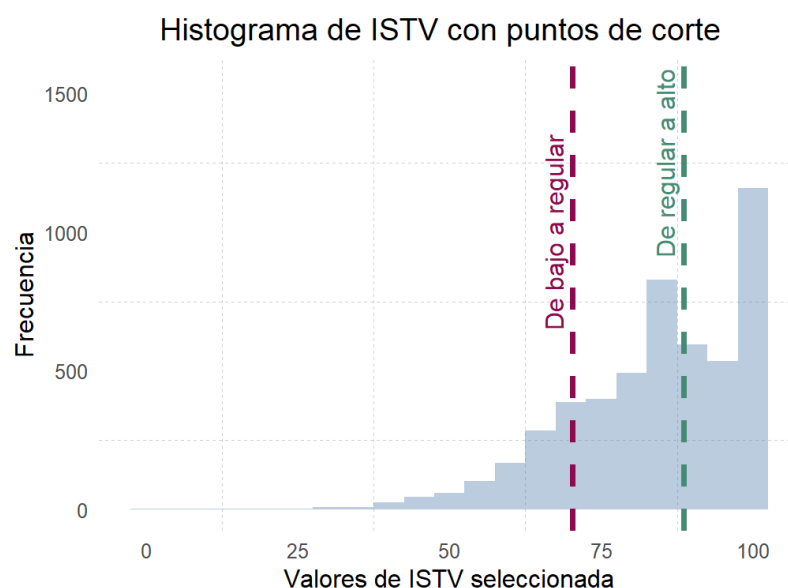


Figura 5-4: Histograma de los puntajes obtenidos para el Índice de Satisfacción Turística del Valle del Cauca (ISTV) y sus puntos de corte

5.3.3. Nivel de satisfacción turística según el perfil del turista

Se clasificó a los viajeros teniendo en cuenta distintas variables y el puntaje obtenido. La Tabla 5-13 presenta la clasificación de los turistas según las variables demográficas y los niveles o puntajes obtenidos en el Índice de Satisfacción Turística del Valle del Cauca. Es relevante notar que los hombres y mujeres presentan una distribución de la satisfacción muy similar. Las personas con edades entre 15 y 26 años, y aquellas mayores de 63 años, suelen tener niveles de satisfacción más altos respecto al total de turistas en su categoría. La satisfacción para el grupo de adultos jóvenes (27-38 años) es moderada, al igual que para adultos entre 39 y 50 años.

Un dato destacable es que los extranjeros aparentan tener un nivel de satisfacción más bajo que los nacionales; la distribución de los puntajes de satisfacción para los extranjeros está sesgada hacia la izquierda, es decir, el 70.2 % presentaron una satisfacción entre moderada y baja, en contraste con los nacionales, de los cuales el 84.1 % presentaron una satisfacción entre alta y moderada.

Tabla 5-13: Clasificación de los turistas según el Índice de Satisfacción Turística (ISTV) en función de variables demográficas

Variable		ISTV					
		Alta	%	Moderada	%	Baja	%
Género	Femenino	1063	42,2 %	1055	15,9 %	99	41,9 %
	Masculino	1065	41,5 %	1040	18,1 %	464	40,5 %
	Otro	2	20,0 %	4	40,0 %	4	40,0 %
	NA	5					
Edad	15 - 26	401	43,6 %	385	41,9 %	133	14,5 %
	27 - 38	692	38,1 %	795	43,8 %	330	18,2 %
	39 - 50	379	37,4 %	430	42,4 %	204	20,1 %
	51 - 62	211	42,4 %	198	39,8 %	89	17,9 %
	Más de 63	90	44,8 %	66	32,8 %	45	22,4 %
	NA	653					
Tipo de procedencia	Extranjero	45	29,8 %	59	39,1 %	47	31,1 %
	Nacional	1985	42,7 %	1924	41,4 %	735	15,8 %
	NA	306					

Por otra parte, la Tabla 5-14 clasifica a los turistas que llegaron al Valle del Cauca de acuerdo con los puntajes obtenidos en el Índice de Satisfacción Turística del Valle del Cauca (ISTV) y las variables asociadas al tipo de viaje. Se observa que el 83,7 % de los turistas que viajaron por motivos personales —tales como educación, salud, visitas a familiares o amigos, y vacaciones— registraron niveles de satisfacción que variaron entre altos y moderados. De

forma similar, para aquellos que viajaron por motivos profesionales —incluyendo asistencia a eventos culturales y deportivos, reuniones de negocios y conferencias—, estas categorías de satisfacción alcanzaron el 80,5 %.

La mayoría de las personas que se alojaron en establecimientos comerciales clasificaron su satisfacción como moderada. Mientras tanto, la mayoría de aquellos que se hospedaron en alojamientos no comerciales, como casas de familiares y amigos o cabañas de vacaciones, presentaron una mayor satisfacción general durante el viaje.

El análisis revela una variabilidad en la satisfacción turística dependiendo de la duración de la estancia. Los turistas que se quedan entre 6 y 8 noches reportan un índice alto de satisfacción, del 51,8 %, mientras que aquellos que se quedan más de 10 noches experimentan un aumento en la insatisfacción, elevándose hasta el 48,3 %. Por otro lado, estancias de menos de 6 noches también muestran menores niveles de satisfacción, sugiriendo que períodos demasiado breves no permiten una experiencia turística plenamente satisfactoria.

Tabla 5-14: Clasificación de los turistas según el Índice de Satisfacción Turística (ISTV)

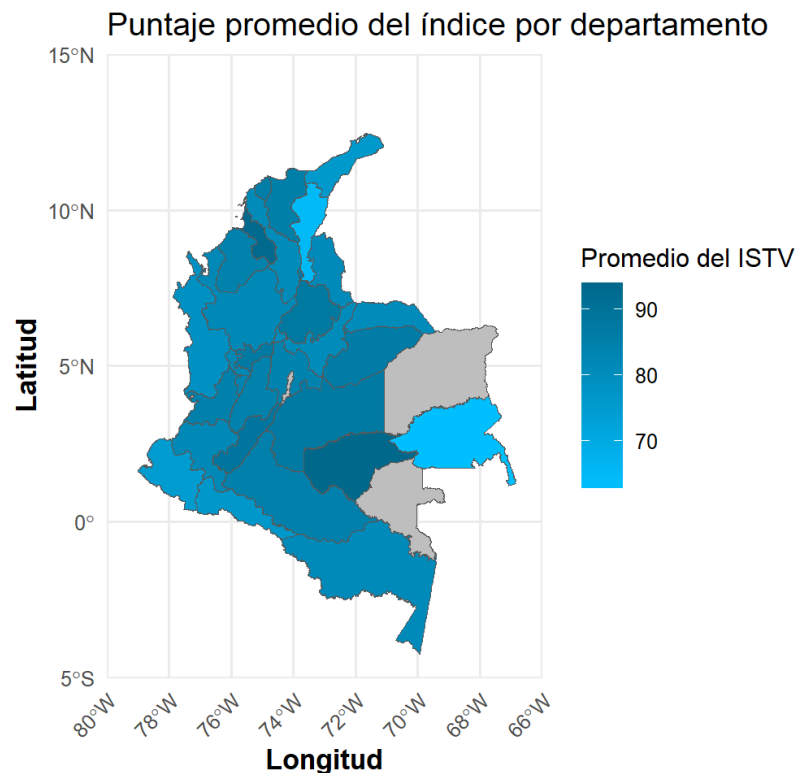
Variable		ISTV					
		Alta	%	Moderada	%	Baja	%
Motivo de viaje	Personales	1901	42,8 %	1815	40,9 %	726	16,3 %
	Profesionales	186	34,4 %	249	46,1 %	105	19,4 %
	Otro	27	31,4 %	29	33,7 %	30	34,9 %
	NA	33					
Tipo de alojamiento	Comercial	1581	39,6 %	1702	42,6 %	712	17,8 %
	No comercial	426	54,5 %	254	32,5 %	102	13,0 %
	Otro	4	36,4 %	2	18,2 %	5	45,5 %
	NA	313					
Noches en el departamento	0 a 2	897	34,5 %	1221	46,9 %	485	18,6 %
	3 a 5	833	49,6 %	661	39,4 %	185	11,0 %
	6 a 8	142	51,8 %	89	32,5 %	43	15,7 %
	9 a 10	12	16,0 %	34	45,3 %	29	38,7 %
	Más de 10	46	26,1 %	45	25,6 %	85	48,3 %
	NA	294					

Como era de esperarse, la mayoría de los turistas encuestados que manifestaron su deseo de volver al departamento fueron clasificados con un nivel de satisfacción alto según la ISTV (Tabla 5-15). En cambio, aquellos que manifestaron que no volverían fueron clasificados mayoritariamente con niveles de satisfacción baja y moderada.

Tabla 5-15: Clasificación de los turistas según el Índice de Satisfacción Turística (ISTV) y la intención de revisita

Variable		ISTV					
		Alta	%	Moderada	%	Baja	%
Intención de revisita	Sí	2128	42,1 %	2087	41,3 %	838	16,6 %
	No	2	22,2 %	3	22,2 %	4	44,4 %
	No estoy seguro	2	5,30 %	10	26,3 %	26	68,4 %
	NA	1					

En la Figura 5-5 se puede observar que, en la muestra de datos analizada, los turistas con mayor satisfacción promedio provienen de los departamentos de Guaviare (94.0), Sucre (93.1) y Huila (89.0). En contraste, los viajeros de los departamentos de Guainía (62.91), Cesar (64.58) y Nariño (74.99) presentaron una baja satisfacción turística. Enfocarse en la satisfacción por departamento puede ser esencial para determinar si esta está relacionada con los prestadores de servicios turísticos, como las agencias que operan en dichas regiones. Los departamentos en color gris no contienen datos. Esto incluye a Bogotá D.C., cuyos datos se han incorporado en el análisis de Cundinamarca.

**Figura 5-5:** Valores promedios del ISTV por departamento en Colombia

5.4. Desarrollo de índices para restaurantes y alojamientos

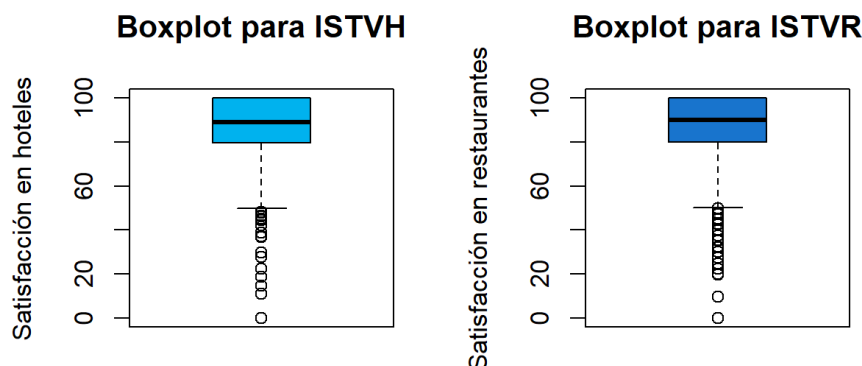
Con base en el Índice de Satisfacción Turística General (ISTV) seleccionado, se desarrollaron dos índices específicos: uno para medir la satisfacción percibida de los turistas en restaurantes (ISTV-R) y otro para medir la satisfacción en hoteles (ISTV-H). Los resultados obtenidos son los siguientes:

Tabla 5-16: Estadísticas descriptivas de los índice de restaurante y alojamiento

Estadísticas descriptivas para índice de restaurantes y alojamiento						
Índice	Mín.	Máx.	Media	Mediana	Desviación estándar	C.V
ISTV – H	0	100	87,71	88,89	12,66	14,43
ISTV – R	0	100	87,55	89,99	13,27	15,16

La Tabla 5-16 presenta las estadísticas descriptivas de los índice ISTV-H (Satisfacción en hospedajes o alojamientos) y ISTV-R (Satisfacción en restaurantes). Las medidas de los puntajes asignados a los turistas para ambos sectores son altas y bastante uniformes. La ligera variabilidad observada, explicada por la desviación estándar y los bajos coeficientes de variación, refuerza la idea de que, en general, los turistas están satisfechos tanto con los alojamientos como con los restaurantes.

Figura 5-6: Boxplot de los índices de satisfacción en restaurantes y hoteles



5.4.1. Puntos de corte para ISTV-R y ISTV-H

Al igual que en la construcción del Índice de Satisfacción Turística general (ISTV), los índices para medir la satisfacción en restaurantes y hoteles se categorizarán en tres niveles distintos: baja, moderada y alta. Para esta clasificación, se emplearon los métodos de K-means y Bootstrap.

Se clasificaron a los 5.101 turistas encuestados según los puntos de corte registrados en la Tabla **5-17**. De esta manera, se determinó que 2.209 turistas (43,3 %) tuvieron una satisfacción alta con los lugares donde se hospedaron, 2.176 turistas (42,65 %) tuvieron una satisfacción moderada y 716 turistas (14,03 %) tuvieron una satisfacción baja.

Tabla 5-17: Rangos mínimos y máximos para el nivel de satisfacción turística en hoteles

Nivel de satisfacción	Mín.	Máx.
Satisfacción baja	0	74,40
Satisfacción regular	74,40	91,38
Satisfacción alta	91,38	100

De forma análoga, la Tabla **5-18** contiene los puntos de corte para la satisfacción de los turistas con los restaurantes. Se encontró que 2.447 turistas (47,97 %) tuvieron una satisfacción alta con el servicio de alimentos, 1.634 turistas (32,03 %) tuvieron una satisfacción moderada y 1.020 turistas (20,00 %) fueron clasificados con una satisfacción baja.

Tabla 5-18: Puntos de corte para el nivel de satisfacción turística en restaurantes

Nivel de satisfacción	Mín.	Máx.
Satisfacción baja	0	75,50
Satisfacción regular	75,50	91,51
Satisfacción alta	91,51	100

5.5. Visualización de resultados

Para presentar de manera más funcional, un resumen de los resultados desarrollados en este trabajo, se realizó un tablero de datos con ayuda de la herramienta LookerStudio, a la que se puede acceder en el siguiente enlace: [Tablero para SITUR](#)

6 Conclusiones y trabajos futuros

A partir del Análisis Factorial Exploratorio, que también sirvió para validar el instrumento de medición, se evidenció que el modelo formativo de alto orden fue adecuado para describir la relación entre las variables del estudio. Se identificaron tres factores principales: Calidad de servicios turísticos, Infraestructura y servicios municipales, y Actividades turísticas y recreativas. Estos factores explican de manera adecuada un factor de segundo orden, la Satisfacción turística, que es el objetivo principal de este estudio. Estos hallazgos fueron respaldados y validados por la opinión de los expertos de SITUR.

Se determinó que la metodología estadística más adecuada para la construcción de los Índices de Satisfacción Turística del Valle del Cauca, tanto en su versión general como en las específicas para Restaurantes y Alojamientos, es el Análisis de Componentes Principales para Datos Mixtos (FAMD). El índice construido con esta metodología demostró un desempeño superior en términos de su capacidad para explicar la varianza.

Se constató que la metodología de Análisis Factorial Múltiple (AFM) genera puntajes muy similares a los del índice construido con FAMD, lo que indica que también es una buena alternativa para construir los Índices de Satisfacción Turística del Valle del Cauca, aunque con un rendimiento ligeramente inferior.

Los puntos de corte utilizados en el análisis resultaron ser altos, reflejando que, en general, las puntuaciones de los datos originales presentan valores elevados. Esto sugiere que la mayoría de los turistas reportan experiencias positivas, lo que establece un estándar elevado para medir la satisfacción y destaca la necesidad de mantener y superar estos niveles para seguir mejorando la oferta turística.

El Análisis de Correspondencias Múltiples (ACM) no resultó ser una estrategia efectiva para la construcción del Índice de Satisfacción Turística del Valle del Cauca (ISTV). La consistencia y el desempeño de los puntajes obtenidos no fueron óptimos, y la varianza explicada por la primera componente fue demasiado baja (28.6%). Esto se debió a una recodificación de las variables en su escala original, lo que resultó en la pérdida de características importantes de los datos.

Se observó que la distribución de la Satisfacción Turística General es similar entre mujeres y hombres, con la mayoría de las puntuaciones clasificadas como altas, 42,2% y 41,5%

respectivamente.

Por grupos de edad, los mayores de 63 años mostraron el mayor porcentaje de puntuaciones bajas, con un 22,4 %, mientras que los grupos de 27 a 38 años y de 39 a 50 años presentaron una alta proporción de puntuaciones regulares. Este hallazgo sugiere la necesidad de enfocarse en estos grupos para comprender su percepción turística y desarrollar estrategias que se alineen con sus preferencias.

El 70,2 % de los puntajes de turistas extranjeros se clasificaron entre medianos y bajos, lo que indica que pocos viajeros internacionales están completamente satisfechos con los servicios ofrecidos en el departamento. En contraste, los locales (residentes en Colombia) mostraron muy pocas puntuaciones bajas, no más del 15,8 %. Esto podría señalar que las expectativas de los turistas extranjeros son más altas que las de los locales, subrayando la importancia de revisar y potencialmente elevar los estándares del turismo local.

Las personas alojadas en establecimientos comerciales mostraron, en su mayoría, una satisfacción media o regular con el destino, representando el 42,6 % de las respuestas. Cercanamente, el 39,6 % de los huéspedes calificaron su satisfacción como alta, lo que resalta oportunidades significativas de mejora en la calidad del servicio de hospedaje.

Curiosamente, la mayoría de los turistas que se alojaron hasta 2 noches reportaron una satisfacción media (46,9 %). La satisfacción aumentó entre quienes se quedaron de 3 a 5 noches y de 6 a 8 noches, con tasas del 49,6 % y 51,8 % respectivamente. Sin embargo, para aquellos que extendieron su estadía a 9 o 10 noches, la satisfacción generalmente disminuyó a media o regular (45,3 %), con un aumento notable en las calificaciones bajas, alcanzando el 38,7 %. Los turistas que se quedaron más de 10 noches experimentaron predominantemente bajas satisfacciones, con un 48,3 %. Esta tendencia indica que la satisfacción decrece a medida que los turistas prolongan su visita, lo que representa un desafío crítico para los proveedores de servicios turísticos.

Entre los visitantes que respondieron que ‘no están seguros de volver’ al ser consultados sobre su intención de revisita, el 68,45 % fue clasificado con baja satisfacción. Por otro lado, de aquellos que afirmaron que sí volverían, el 16,6 % recibió una puntuación baja en satisfacción. A pesar de estas experiencias menos positivas, estos viajeros aún contemplan la posibilidad de regresar, resaltando la necesidad de desarrollar estrategias efectivas que incentiven su retorno.

Trabajos futuros

Para futuros trabajos dentro de la entidad turística, se recomienda implementar un plan de mejoramiento del instrumento ‘Receptivo’ con el objetivo de optimizar el proceso de recolección de datos, garantizando la confiabilidad de los mismos y minimizando la pérdida de información para obtener resultados más consistentes. Asimismo, sería interesante desarrollar un instrumento diseñado específicamente para obtener variables categóricas o nominales, permitiendo experimentar con la metodología de Análisis de Correspondencias Múltiples (ACM) sin necesidad de realizar transformaciones.

Para futuros trabajos que busquen explicar la relación entre variables latentes a partir de variables observables en el sector del turismo, se recomienda optar por la metodología de **Ecuaciones Estructurales**.

7 Anexos

7.1. Tablas

Reasignación de nombres de las variables utilizadas en el análisis

Variable observable	Codificación
Recorrer calles y parques del casco urbano	X1
Visita a museos, casas de la cultura, iglesias y monumentos	X2
Visita haciendas o casas históricas	X3
Realiza actividades religiosas	X4
Asistencia a espectáculos artísticos y festivales	X5
Realiza avistamiento de aves	X6
Realiza avistamiento de ballenas	X7
Visita a discotecas, bares o karaokes	X8
Visita a parques temáticos o parques de atracciones	X9
Compra de bienes de consumo duradero	X10
Realizar inversión o una reunión de negocio	X11
Práctica de deportes	X12
Asiste a competencias o muestras deportivas	X13
Asiste a conferencias/congresos	X14
Asiste a fiestas y ferias de pueblos y municipios	X15
Visita a fábricas o haciendas	X16
Visita a casinos y otros sitios de juegos de azar	X17
Visita a ríos, lagos o cascadas	X18
Visita miradores paisajísticos	X19
Visita zoológicos	X20
Realiza proceso del café	X21
Realiza proceso del azúcar/dulce	X22
Realiza actividades de aventura (parapente, bungee, etc.)	X23
Realiza senderismo o caminatas	X24
Visita reservas o parques	X25
Asiste a playas	X26
Realiza otra actividad	X27

Tabla 7-1: Codificación de variables observables relacionadas con actividades

Reasignación de nombres de las variables de servicios turísticos utilizadas en el análisis

Variable observable	Codificación
Gasto turístico en pesos colombianos	X28
Estado del edificio de su alojamiento	X29
Estado de higiene y limpieza de su alojamiento	X30
Estado de muebles de su alojamiento	X31
Calidad del trato personal en su alojamiento	X32
Calidad de la alimentación en su alojamiento	X33
Evaluación del precio de su alojamiento	X34
Componente alojamiento	X35
Calidad de los platos servidos en restaurantes	X36
Calidad oferta gastronómica en restaurantes	X37
Calidad del trato personal en restaurantes	X38
Estado de higiene y limpieza en restaurantes	X39
Evaluación del precio en restaurantes	X40
Componente restaurante	X41

Tabla 7-2: Codificación de variables observables relacionada con servicios turísticos

Reasignación de nombres de las variables relacionadas con el ambiente e intención de revisita utilizadas en el análisis

Variable observable	Codificación
Evaluación de limpieza y aseo en los municipios	X42
Evaluación de limpieza y conservación de municipios	X43
Componente ambiental	X44
Evaluación de hospitalidad en los municipios	X45
Evaluación estado de carreteras	X46
Evaluación de la seguridad	X47
Componente recompra	X48
Valoración visita al departamento	X49
Intención de revisita	X50

Tabla 7-3: Codificación de variables observables relacionadas con el ambiente y la intención de volver

Motivos generales de viaje para turistas colombianos y extranjeros

Tabla 7-4: Motivos de viaje de turistas colombianos por departamento

Departamento	Motivos personales		Motivos profesionales		Otro	
	Cantidad	%	Cantidad	%	Cantidad	%
Valle del Cauca	2263	45.45	180	3.61	14	0.28
Cundinamarca	395	7.93	74	1.49	8	0.16
Antioquia	260	5.22	73	1.47	7	0.14
Risaralda	268	5.38	18	0.36	8	0.16
Quindío	186	3.73	12	0.24	1	0.02
Cauca	170	3.41	20	0.40	7	0.14
Caldas	156	3.13	28	0.56	3	0.06
Nariño	89	1.79	19	0.38	9	0.18
Tolima	76	1.53	9	0.18	0	0.00
Atlántico	41	0.82	0	0.00	1	0.02
Huila	18	0.36	9	0.18	0	0.00
Chocó	21	0.42	5	0.10	0	0.00
Meta	19	0.38	3	0.06	3	0.06
Caquetá	19	0.38	3	0.06	0	0.00
Santander	18	0.36	4	0.08	0	0.00
Putumayo	18	0.36	2	0.04	1	0.02
Boyacá	18	0.36	3	0.06	0	0.00
Norte de Santander	12	0.24	7	0.14	0	0.00
Bolívar	10	0.20	1	0.02	0	0.00
Córdoba	5	0.10	4	0.08	0	0.00
Magdalena	7	0.14	1	0.02	0	0.00
Sucre	4	0.08	4	0.08	0	0.00
La Guajira	2	0.04	2	0.04	2	0.04
San Andrés y Providencia	5	0.10	0	0.00	0	0.00
Cesar	3	0.06	0	0.00	2	0.04
Casanare	4	0.08	0	0.00	0	0.00
Arauca	4	0.08	0	0.00	0	0.00
Guaviare	0	0.00	1	0.02	0	0.00
Guainía	0	0.00	1	0.02	0	0.00
Amazonas	1	0.02	0	0.00	0	0.00

Tabla 7-5: Motivos de viaje de turistas extranjeros por país (excluyendo a Colombia)

País	Motivos personales		Motivos profesionales		Otro	
	Cantidad	%	Cantidad	%	Cantidad	%
Ecuador	30	0.59	1	0.02	0	0.00
Estados Unidos	19	0.37	6	0.12	2	0.04
España	17	0.33	2	0.04	1	0.02
Argentina	10	0.20	1	0.02	1	0.02
Chile	7	0.14	2	0.04	2	0.04
Francia	6	0.12	2	0.04	0	0.00
Perú	5	0.10	2	0.04	0	0.00
Venezuela	5	0.10	1	0.02	0	0.00
Guayana Francesa	5	0.10	0	0.00	0	0.00
República Dominicana	4	0.08	0	0.00	0	0.00
Suiza	4	0.08	0	0.00	0	0.00
Uruguay	2	0.04	0	0.00	0	0.00
Italia	0	0.00	1	0.02	1	0.02
Costa Rica	1	0.02	0	0.00	1	0.02
Aruba	2	0.04	0	0.00	0	0.00
República Checa	2	0.04	0	0.00	0	0.00
Puerto Rico	0	0.00	0	0.00	1	0.02
Grecia	1	0.02	0	0.00	0	0.00
Alemania	1	0.02	0	0.00	0	0.00
Islas Malvinas	0	0.00	0	0.00	1	0.02
México	1	0.02	0	0.00	0	0.00
Suriname	1	0.02	0	0.00	0	0.00

Valores de KMO para cada variable

Tabla 7-6: Valores de KMO para los ítems

Codificación	KMO	Descripción de la variable
X1	0.70	Recorrer calles y parques del casco urbano
X2	0.85	Visita a museos, casas de la cultura, iglesias y monumentos
X4	0.81	Realiza actividades religiosas
X5	0.80	Asistencia a espectáculos artísticos y festivales
X7	0.66	Realiza avistamiento de ballenas
X8	0.69	Visita a discotecas, bares o karaokes
X12	0.70	Practica de deportes
X15	0.72	Asiste a fiestas y ferias de pueblos municipios
X26	0.72	Asiste a playas
X29	0.97	Estado del edificio de su alojamiento
X30	0.96	Estado de higiene y limpieza de su alojamiento
X31	0.97	Estado de muebles de su alojamiento
X32	0.97	Calidad del trato personal en su alojamiento
X33	0.96	Calidad de la alimentación en su alojamiento
X34	0.96	Evaluación del precio de su alojamiento
X35	0.90	Componente alojamiento
X36	0.97	Calidad de los platos servidos en restaurantes
X37	0.96	Calidad oferta gastronómica en restaurantes
X38	0.96	Calidad del trato personal en restaurantes
X39	0.95	Estado de higiene y limpieza en restaurantes
X40	0.96	Evaluación del precio en restaurantes
X41	0.89	Componente restaurante
X42	0.85	Evaluación de limpieza y aseo en los municipios
X43	0.86	Evaluación de limpieza y conservación en municipios
X44	0.83	Componente ambiental
X45	0.98	Evaluación de hospitalidad en los municipios
X46	0.83	Evaluación estado de carreteras
X47	0.93	Evaluación de la seguridad
X48	0.88	Componente recompra
X49	0.98	Valoración visita al departamento

7.2. Código usado

El código en R utilizado para desarrollar los índices está disponible en un repositorio de GitHub. Puede acceder a él haciendo clic [aquí](#).

Bibliografía

- Abdi, H. and Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4):433–459. Citations: 6,246.
- Allen, I. E. and Seaman, C. A. (2007). Likert scales and data analyses. *Quality Progress*, 40(7):64–65.
- Aluja Banet, T. and Morineau, A. (1999). *Aprender de los datos: el análisis de componentes principales: una aproximación desde el data mining*. Universidad de Barcelona.
- Ambapour, S. (2020). Using multiple correspondence analysis to measure multidimensional poverty in congo. *Journal of Data Analysis and Information Processing*, 8(4).
- Bartholomew, D., Knott, M., and Moustaki, I. (2011). *Latent Variable Models and Factor Analysis: A Unified Approach*. John Wiley & Sons.
- Bartholomew, D. J. and Knott, M. (1999). *Latent Variable Models and Factor Analysis*. Arnold, London.
- Becerra Avella, M. O. (2010). Comparación del análisis factorial múltiple y del análisis en componentes principales para datos cualitativos, en la construcción de índices. Master’s thesis, Universidad Nacional de Colombia, Departamento de Estadística, Facultad de Ciencias.
- Belalcazar, M. (2023). Evaluación del desempeño estadístico de escalas construidas para representar la expresión de variables latentes mediante análisis multivariados. Tesis de maestría, Escuela de Estadística, Facultad de Ingeniería, Maestría en Estadística. Director: Jose Rafael Tovar Cuevas, PhD. Codirector: Víctor Manuel González Rojas, PhD.
- Berry, L. L. and Parasuraman, A. (1991). *Marketing Services: Competing Through Quality*. Free Press.
- Blaufuks, W. (2021). Famd: How to generalize pca to categorical and numerical data. *Towards Data Science*. Published in Towards Data Science, May 25, 2021.
- Boari, G. and Nai Ruscone, M. (2015). A procedure simulating likert scale item responses. *Electronic Journal of Applied Statistical Analysis*, 8(3):288–297.

- Bollen, K. (1989). *Structural Equations with Latent Variables*, volume 210. John Wiley & Sons.
- Brida, J. G., Pulina, M., Riaño, E., and Sánchez-Casado, N. (2011). El impacto económico del turismo en Colombia. *Revista de Estudios Económicos y Empresariales*, 22:45–60.
- Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research*. Guilford Press, New York.
- Cano, H. H. R. (2021). Trabajo de grado: Entregado como requisito parcial para optar al título de magíster en políticas públicas. Master's thesis, Universidad del Valle, Facultad de Administración, Maestría en Políticas Públicas, Santiago de Cali, Colombia. Tutor: Raúl Andrés Tabarquino Muñoz, Ph.D. en Administración con énfasis en Gobierno y Política Pública.
- Castañeda, E. D. C. (2021). *Calidad percibida del servicio, satisfacción e intención de revisita del turista en destinos turísticos de Boyacá, Colombia*. PhD thesis, Universidad EAN, Facultad de Administración, Finanzas y Ciencias Económicas, Bogotá, Colombia. Doctorado en Gestión, Directores: Gabriel Alfonso Pérez Cifuentes, Ph.D., Enrique Edgardo Gilles Romero, Ph.D.
- Child, D. T. (2006). *The Essentials of Factor Analysis*. Continuum, 3rd edition.
- Corbetta, P. (2003). *Social Research: Theory, Methods and Techniques*. Sage.
- Cronbach, L. J. and Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4):281–302.
- Diamantopoulos, A. and Winklhofer, H. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research*, 38:269–277.
- Dunn, T. J. and McCray, G. (2020). *The Place of the Second-Order Factor Model in Confirmatory Factor Analysis*. Routledge, New York.
- Efron, B. (1979). *Bootstrap Methods: Another Look at the Jackknife*, volume 7.
- Encyclopaedia Britannica (2024). Tourism. <https://www.britannica.com/topic/tourism>. Accedido: 2024-09-04.
- Escofier, B. and Pagès, J. (1992). *Análisis factoriales simples y múltiples: objetivos, métodos e interpretación*. Universidad del País Vasco, España.
- Fabrigar, L. R. and Wegener, D. T. (2011). *Implementing and Interpreting Exploratory Factor Analysis*. Oxford Academic, online edition edition. Accessed: 2024-09-04.

- Foro Económico Mundial (2017). *Informe sobre la competitividad en viajes y turismo*. FEM.
- Gobernación del Valle del Cauca (2023). Turismo movió más de 2000 millones de dólares en la economía vallecaucana. Accessed: 2024-09-04.
- González Rojas, V. M. (2014). *Análisis conjunto de múltiples tablas de datos mixtos mediante PLS*. PhD thesis, UPC, Departamento de Estadística e Investigación Operativa.
- Hair, J. F. J., Matthews, L. M., Matthews, R. L., and Sarstedt, M. (2017). Pls-sem or cb-sem: updated guidelines on which method to use. *International Journal of Multivariate Data Analysis*, 1(2):107–123.
- Harman, H. H. (1976). *Modern Factor Analysis*. University of Chicago Press, 3rd edition.
- Harman, H. H. and Jones, W. H. (1966). Factor analysis by minimizing residuals (minres). *Psychometrika*, 31(4):351–368.
- Haynes, S. N., Lenchman, R. P., and Steinberg, S. W. G. (1995). A comprehensive review of construct validity: Advances in instrumentation theory. *Psychological Assessment*, 7(3):238–247.
- Hendrickson, A. E. and White, P. O. (1964). Promax: A quick method for rotation to oblique simple structure. *British Journal of Statistical Psychology*, 17(1):65–70.
- Hernández Sampieri, R., Fernández Collado, C., and Baptista Lucio, P. (2014). *Metodología de la Investigación*. McGraw-Hill.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6):417–441.
- J. Joseph Cronin, J. and Taylor, S. A. (1992). Measuring service quality: A reexamination and extension. *Journal of Marketing*, 56(3):55–68.
- Jiang, X. and Peng, Y. (2022). Analysis of tourist satisfaction index based on structural equation model. *Journal of Sensors*, 2022:1–12.
- Kaiser, H. F. and Rice, J. (1974). Little jiffy, mark iv. *Educational and Psychological Measurement*, 34:111 – 117.
- Kotler, P. and Keller, K. L. (2006). *Marketing Management*. Pearson Prentice Hall, 12th edition.
- Kotler, P. and Keller, K. L. (2016). *Marketing Management*. Pearson.
- Lebart, L., Morineau, A., and Piron, M. (1995). *Statistique exploratoire multidimensionnelle*, volume 3. Dunod, Paris.

- MacCallum, R. C., Zhang, S., Preacher, K. J., and Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7(1):19–40.
- Mazziotta, M. and Pareto, A. (2013). Methods for constructing composite indices: One for all or all for one. *Rivista Italiana di Economia Demografia e Statistica*, 67(2):67–80.
- Ministerio de Comercio, I. y. T. (2022). Informe anual de turismo 2022. Disponible en <https://www.mincit.gov.co>.
- Ministerio de Comercio, I. y. T. d. C. (2018). *Plan Sectorial de Turismo 2018-2022*. Ministerio de Comercio, Industria y Turismo de Colombia.
- Ministerio de Comercio, Industria y Turismo de Colombia (2022). Los servicios turísticos aportarían \$45 billones al pib de 2022. <https://www.mincit.gov.co/prensa/noticias/turismo/servicios-turisticos-aportarian-45-billones-al-pib>. Consultado el 09 de marzo de 2024.
- Minitab Support (2024). Tablas de contingencia. <https://support.minitab.com/es-mx/minitab/help-and-how-to/statistics/tables/supporting-topics/basics/contingency-tables/>. Consultado el 26 de marzo de 2024.
- Miranda Zapata, E. D. (2012). *Comparación de la eficiencia en la estimación de parámetros entre análisis de estructura de covarianzas, modelo de rutas mediante mínimos cuadrados parciales y análisis generalizado de componentes estructurales: un estudio de simulación*. Tesis doctoral, Universidad Autónoma de Madrid. Dirigida por: Miguel Ángel Ruiz Díaz. Tribunal: Rafael San Martín Castellanos (Presidente), Francisco José Abad García (Secretario), Jesús Varela Mallou (Vocal), Jesús Privado Zamorano (Vocal), Gerardo Prieto Adánez (Vocal).
- Méndez Martínez, C. and Rondón Sepúlveda, M. A. (2012). Introducción al análisis factorial exploratorio. *Revista Colombiana de Psiquiatría*, 41(1):197–207. Accedido el 15 de mayo de 2024.
- Nunnally, J. C. and Bernstein, I. H. (1994). *Psychometric Theory*. McGraw-Hill.
- Ochoa Muñoz, A. F. (2018). Análisis de correspondencias múltiples en presencia de datos faltantes: el principio de datos disponibles del algoritmo nipals (acmpdd). Trabajo de investigación de maestría, Universidad del Valle, Santiago de Cali, Colombia. Facultad de Ingeniería, Escuela de Estadística. Director: Ph.D. Victor Manuel Gonzalez Rojas.
- Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, 44(4):443–460.
- Olsson, U., Drasgow, F., and Dorans, N. J. (1982). The polyserial correlation coefficient. *Psychometrika*, 47(3):337–347.

- Organización Mundial del Turismo (2005). *Indicators of Sustainable Development for Tourism Destinations: A Guidebook*. Organización Mundial del Turismo.
- Organización Mundial del Turismo (2016). *Informe Anual 2016*. Organización Mundial del Turismo (OMT).
- Organización Mundial del Turismo (2018). *Informe Anual 2018*. Organización Mundial del Turismo (OMT).
- Organización Mundial del Turismo (2023). El turismo. <https://www.unwto.org/es/turismo>. Consultado el 24 de abril de 2024.
- Otero, J., Pérez, A., and García, L. (2017). El papel de la satisfacción en la imagen y competitividad de los destinos turísticos. *Journal of Tourism Research*, 12:34–48.
- Otero Gómez, M., Giraldo, W., and Duarte, L. (2016). Satisfacción con un destino turístico: Una investigación en Villavicencio (Colombia). *In Vestigium Ire*, 10:97–132.
- Pagès, J. (2004). Multiple factor analysis: Main features and application to sensory data. *Revista de Investigación Operacional*, 25(1):37–55.
- Perron, B. E. and Gillespie, D. F. (2015). *Latent Variables*. Pocket Guides to Social Work Research Methods. Oxford Academic, New York, online edition edition. Accessed: 2024-09-04.
- Pérez Porto, J. and Gardey, A. (2022). Definición de turismo - qué es, definición, historia y tipos. <https://definicion.de/turismo/>. Recuperado el 24 de mayo de 2024.
- Rojas, E. N. (2008). Escalas o índices para la medición de constructos: El dilema del analista de datos. *Avances en medición*, 6:15–24.
- Sistema de Información Turística Regional del Valle del Cauca (2024). Qué es el sistema de información turística en el Valle del Cauca - situr. <https://siturvalle.com/estadisticas>. Consultado el 24 de mayo de 2024.
- Spreng, R. A. and Mackoy, R. D. (1996). An empirical examination of a model of perceived service quality and satisfaction. *Journal of Retailing*, 72(2):201–214.
- Tabachnick, B. G. and Fidell, L. S. (2013). *Using Multivariate Statistics*. Pearson Education, 6th edition.
- Taborga, C. E. V. (2016). Comparación de los modelos formativo, reflexivo y de antecedentes de evaluación estudiantil del servicio de docencia. *Revista de Métodos Cuantitativos para la Economía y la Empresa*, 16:95–120.

- Taylor, S. A. and Baker, T. L. (1994). An assessment of the relationship between service quality and customer satisfaction in the formation of consumers' purchase intentions. *Journal of Retailing*, 70(2):163–178.
- Thurstone, L. L. (1944). *Multiple-Factor Analysis*. University of Chicago Press, Chicago.
- Tourism Notes (2024). Turismo cultural: Definiciones, tipos, ventajas y desventajas. Accedido: 2024-09-04.
- World Tourism Organization (2024). Tourism's importance for growth highlighted in world economic outlook report. Accedido: 2024-09-04.