

DOI: <https://doi.org/10.70577/sp2s0g33>

Aplicación de técnicas estadísticas multivariantes y modelamiento predictivo para el análisis comparativo del comportamiento de turistas nacionales y extranjeros en la Reserva de Producción de Fauna Chimborazo (2015–2025)

Samuel Tobías López Cruz ¹

Maestría en Estadística con mención en Ciencia de Datos e Inteligencia Artificial

<https://orcid.org/0009-0008-1959-7691>

tobias.lopez@esPOCH.edu.ec

Escuela Superior Politécnica de Chimborazo (ESPOCH)

Riobamba – Ecuador

Andrea Damaris Hernández Allauca ²

Facultad de Recursos Naturales

<https://orcid.org/0000-0001-6413-5607>

andrea.hernandez@esPOCH.edu.ec

Escuela Superior Politécnica de Chimborazo (ESPOCH)

Riobamba – Ecuador

Cómo citar:	
Aplicación de técnicas estadísticas multivariantes y modelamiento predictivo para el análisis comparativo del comportamiento de turistas nacionales y extranjeros en la Reserva de Producción de Fauna Chimborazo (2015–2025). (2026). <i>Visión Académica</i> , 4(3), 461-484. https://doi.org/10.70577/sp2s0g33	Fecha de recepción: 2026-07-06 Fecha de aceptación: 2026-07-20 Fecha de publicación: 2026-08-03

Resumen

El presente estudio analizó comparativamente el comportamiento de los turistas nacionales y extranjeros que visitaron la Reserva de Producción de Fauna Chimborazo durante el período 2015–2025, mediante la aplicación de técnicas estadísticas multivariantes y modelos predictivos. Se desarrolló una investigación cuantitativa, de diseño no experimental y carácter longitudinal, utilizando registros mensuales de visitantes y encuestas aplicadas a turistas. Se emplearon estadísticas descriptivas, análisis de correlación, análisis de componentes principales (ACP), análisis de conglomerados, análisis multivariado de la varianza (MANOVA), regresión logística, árboles de decisión y modelos ARIMA. Los resultados evidenciaron una fuerte asociación entre el flujo de turistas nacionales, extranjeros y el total de visitantes. El ACP permitió sintetizar aproximadamente el 90 % de la variabilidad en tres componentes principales. El análisis de conglomerados identificó tres segmentos turísticos: bajo flujo (13,9 %), flujo medio (50,9 %) y alto flujo (35,2 %). El MANOVA confirmó diferencias multivariadas significativas asociadas a la estacionalidad (Pillai = 0,266; $p < 0,001$), mientras que la regresión logística alcanzó una elevada capacidad predictiva (AUC = 0,963). Asimismo, el modelo ARIMA describió adecuadamente la dinámica temporal y estacional de la demanda turística. Se concluye que la demanda turística de la Reserva de Producción de Fauna Chimborazo está influenciada principalmente por la estacionalidad y la intensidad del flujo de visitantes. La integración de técnicas estadísticas multivariantes y modelos predictivos permitió identificar patrones de comportamiento, segmentar períodos turísticos y generar herramientas útiles para la planificación y gestión sostenible de áreas naturales protegidas.

Palabras Claves: Estadísticas Mutivariantes, Modelamiento Predictivo, Reserva Chimborazo, Turistas.

Application of multivariate statistical techniques and predictive modeling for the comparative analysis of the behavior of domestic and foreign tourists in the Chimborazo Fauna Production Reserve (2015–2025)

Abstract

This study comparatively analyzed the behavior of domestic and international tourists visiting the Chimborazo Fauna Production Reserve during the 2015–2025 period through the application of multivariate statistical techniques and predictive models. A quantitative, non-experimental, and longitudinal research design was employed using monthly visitor records and survey data collected from tourists. The analytical approach included descriptive statistics, correlation analysis, Principal Component Analysis (PCA), cluster analysis, Multivariate Analysis of Variance (MANOVA), logistic regression, decision trees, and ARIMA models. [Artículo p...ntación-02 | Word]

The results revealed a strong association between domestic, international, and total visitor flows. PCA condensed approximately 90% of the total variability into three principal components. Cluster analysis identified three tourism segments: low flow (13.9%), medium flow (50.9%), and high flow (35.2%). MANOVA confirmed significant multivariate differences associated with seasonality (Pillai = 0.266; $p < 0.001$), while logistic regression achieved excellent predictive performance (AUC = 0.963). Furthermore, the ARIMA model adequately captured the temporal and seasonal dynamics of tourism demand. [Artículo p...ntación-02 | Word]

The findings indicate that tourism demand in the Chimborazo Fauna Production Reserve is primarily influenced by seasonality and visitor flow intensity. The integration of multivariate statistical methods and predictive modeling enabled the identification of behavioral patterns, the segmentation of tourism periods, and the development of useful tools for the sustainable planning and management of protected natural areas.

Keywords: Multivariate Analysis, Predictive Modeling, Tourism Demand, Chimborazo Fauna Production Reserve, Tourists.

Introducción

El turismo en espacios naturales protegidos representa una estrategia fundamental para poder desarrollar paralelamente la conservación del medio ambiente, la potenciación de la economía local y la puesta en valor de la naturaleza. En el seno de las últimas décadas, el auge de la actividad turística ha creado la necesidad por parte de comprender el comportamiento de las personas visitantes, así como los factores que justifican la demanda turística, con particular incidencia en espacios considerados destinos que requieren gestionar la utilización recreativa del espacio con la sostenibilidad y la conservación ecológica.

La Reserva de Producción de Fauna Chimborazo (RPFCH), asentada en la región de los Andes en Ecuador, es considerada como uno de los principales destinos turísticos relacionados con el

turismo de naturaleza y montaña. En su salvaguarda tiene al imponente nevado Chimborazo, un símbolo natural y cultural del país, así como la diversidad de ecosistemas altoandinos que alberga. Estas características han favorecido el creciente flujo de visitantes nacionales e internacionales, lo que ha llevado a generar herramientas analíticas que permitan comprender la dinámica turística del área protegida y optimizar los procesos de planificación y gestión.

La investigación de la demanda turística se está convirtiendo en un asunto cuya temática cada vez despierta más interés en la literatura científica, dado que tiene contribuciones a la elaboración de políticas públicas, a la gestión de destinos y a la toma de decisiones fundamentadas en evidencia. En este sentido, las técnicas estadísticas multivariadas han probado ser eficientes para descubrir patrones que están ocultos, simplificar la complejidad de los grandes volúmenes de datos, y agrupar a personas con características comunes (Homogeneidad). Por otra parte, los modelos predictivos permiten anticipar los cambios en la demanda y evaluar las condiciones futuras, facilitando así que se creen estrategias orientadas a hacer sostenible el turismo.

Smarttour es la única metodología que incorpora el análisis y procesamiento de los test de las y los turistas con la carta de turismo de la localidad. Dentro de las técnicas más empleadas para el estudio del comportamiento turístico sobresalen el Análisis de Componentes Principales (ACP), el análisis de conglomerados (Cluster Analysis), la ANÁLISIS MULTIVARIADO DE LA VARIANZA (MANOVA), los modelos de clasificación en base a regresiones lógicas y árboles de decisión y por último los modelos de series temporales tipo ARIMA. La aplicación combinada de estas técnicas permite obtener una visión integral de los procesos turísticos al incorporar simultáneamente análisis descriptivos, explicativos y predictivos.

Aunque existen numerosos estudios enfocados en la caracterización de la demanda turística y la predicción de flujos de visitantes, en el contexto ecuatoriano son limitadas las investigaciones que integran técnicas estadísticas multivariantes y modelos predictivos para analizar de manera comparativa el comportamiento de turistas nacionales y extranjeros en áreas protegidas. Esta brecha de escasez de conocimiento se presenta muy claramente en la zona de la Reserva de Producción de Fauna Chimborazo, ya que los estudios que más se han enfocado no son sino aquéllos que tienen un carácter descriptivo o los que tratan sobre la gestión turística, pero en ninguno de ellos se incluyen diseños que traten de alcanzar la analítica de datos.

La investigación actual tiene el propósito de analizar comparativamente el comportamiento de los turistas nacionales y extranjeros que visitan la Reserva de Producción de Fauna Chimborazo comprendido en el período 2015-2025 mediante la utilización de técnicas estadísticas multivariantes y modelos predictivos. De forma específica, se intenta identificar estructuras latentes del flujo turístico, segmentar períodos que sean similares, comparar los grupos de visitantes en términos de diferencias multivariadas y construir modelos que permitan describir y predecir la evolución de la demanda turística.

Los resultados conseguidos integrarán los argumentos a favor de la gestión sostenible de la reserva, pues se constituye en una contribución científica para la planificación estratégica, la

asignación de recursos y la formulación de políticas de desarrollo turístico sostenible en áreas naturales protegidas, además de ofrecer una metodología que sea replicable a otros destinos turísticos de similar categoría, ya sea dentro de, o fuera de, la realidad ecuatoriana.

Materiales y métodos

Enfoque y diseño de la investigación

La presente investigación se desarrolló bajo un enfoque cuantitativo, debido a que se fundamentó en el análisis de información numérica relacionada con el flujo turístico registrado en la Reserva de Producción de Fauna Chimborazo (RPFCH). El estudio adoptó un diseño no experimental, puesto que las variables fueron observadas en su contexto natural sin manipulación por parte de los investigadores. Asimismo, presentó un carácter longitudinal, al analizar la evolución temporal de la actividad turística durante el período comprendido entre 2015 y 2025. (Chávez Quisbert, 1997)

El alcance de la investigación fue descriptivo, comparativo, correlacional y predictivo. El componente descriptivo permitió caracterizar la dinámica de visitantes nacionales y extranjeros; el comparativo facilitó identificar diferencias entre ambos grupos; el correlacional permitió examinar las relaciones existentes entre las variables estudiadas; y el predictivo se orientó al desarrollo de modelos capaces de estimar patrones futuros de comportamiento turístico. (Fernández et al., 2002)

Análisis socio demográfico del perfil del turista

En este punto se llevó a cabo un análisis del perfil sociodemográfico de los turistas que acuden a la Reserva de Producción de Fauna Chimborazo. Se tomo en cuenta factores como edad, sexo, nacionalidad, grado de educación, empleo, ingresos, razón de la visita, frecuencia de los viajes, duración de la estancia y gastos turísticos. Este estudio facilitará la descripción de los visitantes tanto nacionales como internacionales, distinguir sus principales diferencias y detectar patrones en el comportamiento de los turistas. La información recopilada ayudara a entender la demanda turística de la reserva y a diseñar estrategias que mejoren la gestión, promoción y sostenibilidad del lugar.

Caracterización del perfil sociodemográfico de los turistas

La investigación utilizó información secundaria proveniente de registros oficiales de visitantes de la Reserva de Producción de Fauna Chimborazo. Los datos fueron obtenidos a partir del portal Turismo en Cifras del Ministerio de Turismo del Ecuador y de registros administrativos asociados a la gestión de áreas naturales protegidas. La base consolidada contiene observaciones mensuales correspondientes al período 2015–2025.

- Las variables consideradas para el análisis fueron:
- Número mensual de turistas nacionales.
- Número mensual de turistas extranjeros.

- Total mensual de visitantes.
- Participación porcentual de turistas nacionales.
- Año.
- Mes.
- Temporada turística (alta y baja-media).

Adicionalmente, se generaron variables derivadas orientadas a fortalecer el análisis estadístico, incluyendo tasas de crecimiento mensual, índices de estacionalidad y razones entre visitantes nacionales y extranjeros. Estas variables permitieron una mejor comprensión de la estructura y evolución temporal del flujo turístico. (Song et al., 2019; Athanasopoulos et al., 2020).

De manera complementaria al análisis de registros históricos, se realizó una caracterización sociodemográfica de los visitantes mediante la aplicación de encuestas estructuradas dirigidas a turistas nacionales y extranjeros que visitaron la Reserva de Producción de Fauna Chimborazo. El instrumento estuvo conformado por 14 preguntas de selección múltiple relacionadas con variables demográficas, educativas, ocupacionales y de comportamiento turístico.

Las variables consideradas incluyeron lugar de residencia, edad, género, nivel de instrucción, ocupación principal, frecuencia de visita y otros aspectos asociados al comportamiento turístico de los visitantes. Esta información permitió complementar el análisis estadístico de la demanda turística mediante la identificación de características diferenciadoras entre turistas nacionales y extranjeros.

Tipo de muestreo aplicado

Se aplicó un muestreo probabilístico estratificado considerando como estratos principales a los turistas nacionales y extranjeros que visitaron la Reserva de Producción de Fauna Chimborazo. Esta estrategia permitió garantizar la representación de ambos grupos dentro del estudio y facilitar análisis comparativos posteriores. (Canavos & Medal, 1987)

Posteriormente, dentro de cada estrato se empleó un procedimiento de muestreo aleatorio simple, asumiendo que todos los individuos pertenecientes a cada grupo presentaban la misma probabilidad de selección. El tamaño de muestra fue calculado mediante la fórmula para poblaciones finitas con un nivel de confianza del 95 %, una proporción esperada de 0,50 y un error máximo admisible del 8 %. Como resultado se determinó una muestra de 150 turistas nacionales y 150 turistas extranjeros.

Depuración y tratamiento de los datos

Previo al análisis estadístico se efectuó un proceso de revisión, validación y depuración de la base de datos. Considerando el impacto extraordinario generado por la pandemia de COVID-19 sobre la actividad turística mundial, los registros correspondientes a los años 2020 y 2021 fueron excluidos de los análisis inferenciales y predictivos principales. (Ramírez Muñoz, 2021)

Esta decisión metodológica se fundamentó en que dichos períodos presentan comportamientos atípicos asociados a restricciones de movilidad y suspensión parcial de actividades turísticas, lo que podría introducir sesgos en la identificación de patrones estructurales de la demanda. En consecuencia, los análisis estadísticos se desarrollaron utilizando principalmente información correspondiente a los períodos 2015–2019 y 2022–2025.(Mejía et al., 2016)

Análisis estadístico multivariante

Con el propósito de comprender la estructura del flujo turístico, identificar patrones latentes y desarrollar modelos predictivos de la demanda, se aplicó un conjunto de técnicas estadísticas multivariantes y métodos de aprendizaje supervisado ampliamente utilizados en estudios de turismo y análisis de datos. La integración de estas herramientas permite abordar simultáneamente procesos de reducción de dimensionalidad, segmentación, comparación multivariada y predicción de la demanda turística, proporcionando una visión integral del fenómeno analizado (Zhang et al., 2023; Zhao, 2023)

Análisis de correlación

Se construyó una matriz de correlación con el propósito de examinar la intensidad y dirección de las relaciones existentes entre las variables asociadas al flujo turístico. Este análisis permitió identificar posibles asociaciones lineales entre los indicadores estudiados y detectar potenciales problemas de multicolinealidad que pudieran afectar la interpretación de los modelos multivariantes posteriores. La evaluación de correlaciones constituye una etapa preliminar fundamental para comprender la estructura de los datos y establecer la pertinencia de técnicas como el ACP y la regresión logística (James et al., 2021; Hair et al., 2022).

Análisis de Componentes Principales (ACP)

Con el objetivo de identificar estructuras latentes y simplificar la complejidad del conjunto de datos, se aplicó un Análisis de Componentes Principales (ACP). Esta técnica tiene la capacidad de convertir un conjunto de variables correlacionadas en un número menor de componentes ortogonales, los cuales pueden captar una porción significativa de la variabilidad inicial. De esta forma se puede facilitar la interpretación y reducir inconvenientes relacionados con la redundancia de información (Hair et al., 2022).

El ACP ha sido ampliamente utilizado en investigaciones relacionadas con turismo y pronóstico de demanda debido a su capacidad para reducir dimensionalidad, mejorar la eficiencia de los modelos predictivos y sintetizar información procedente de múltiples indicadores turísticos (Zhang et al., 2023; Zhao, 2023).

Los componentes obtenidos fueron evaluados mediante los valores propios (eigenvalues), el porcentaje de varianza explicada individual y acumulada, así como la contribución relativa de cada variable a la conformación de los componentes principales, facilitando la interpretación de los factores dominantes que explican el comportamiento turístico de la reserva.

Análisis de conglomerados

La segmentación de los períodos turísticos se realizó mediante análisis de conglomerados (Cluster Analysis), empleando variables relacionadas con el flujo de visitantes nacionales, extranjeros y totales. Esta metodología permite identificar grupos homogéneos de observaciones internamente similares y externamente diferentes, contribuyendo a la comprensión de patrones de comportamiento turístico y a la identificación de segmentos con características comunes (Dolnicar, 2022).

El clustering de los datos se consideraba una de las formas más prácticas para la segmentación turística (y claramente, la más utilizada), ya que permitía identificar perfiles y comportamientos que contribúan a la posterior identificación de las estrategias de difusión, planificación y gestión de destinos turísticos (Tiwari & Tripathi, 2023; Crisostomo et al., 2025).

Los resultados obtenidos permitieron clasificar los períodos analizados en tres grandes categorías de comportamiento turístico: períodos de bajo flujo, de flujo medio y de alto flujo de visitantes.

Análisis Multivariado de la Varianza (MANOVA)

Las diferencias multivariadas entre grupos fueron evaluadas mediante el Análisis Multivariado de la Varianza (MANOVA), considerando como variables dependientes el número de visitantes nacionales, visitantes extranjeros y la participación porcentual de turistas nacionales. Como factores explicativos se incluyeron la temporada turística y el estado de los datos (consolidado o estimado).

El MANOVA permite analizar de forma simultánea múltiples variables dependientes altamente relacionadas, evitando el incremento del error tipo I que podría producirse al realizar análisis univariados independientes. Además, facilita la identificación de efectos conjuntos asociados a factores explicativos que influyen sobre el comportamiento turístico (Murphy, 2024).

La significancia estadística fue evaluada mediante la Traza de Pillai, estadístico ampliamente recomendado por su robustez frente a posibles incumplimientos de los supuestos de homogeneidad de matrices de covarianza y normalidad multivariante (Statistics How To, 2024).

Modelamiento predictivo

Regresión logística

La regresión logística binaria fue utilizada para estimar la probabilidad de pertenencia a períodos de alta afluencia turística. Los indicadores temporales y los que tienen relación con el volumen de visitantes nacionales e internacionales se tomaron como variables explicativas.

La regresión logística constituye una de las técnicas más utilizadas para problemas de clasificación debido a su facilidad de interpretación, eficiencia computacional y capacidad para modelar probabilidades asociadas a eventos dicotómicos. Su aplicación en estudios turísticos

permite identificar factores asociados a diferentes niveles de demanda y apoyar la toma de decisiones basada en evidencia (James et al., 2021).

El desempeño predictivo del modelo fue evaluado a través de la curva ROC (Receiver Operating Characteristic) y el Área Bajo la Curva (AUC), indicadores ampliamente reconocidos para medir la capacidad discriminante de modelos de clasificación.

Árboles de decisión

Se implementaron árboles de clasificación con el objetivo de identificar reglas de decisión asociadas a los distintos niveles de flujo turístico. Esta técnica de aprendizaje supervisado permite representar de forma gráfica y fácilmente interpretable los criterios utilizados para clasificar observaciones en diferentes categorías.

Los árboles de decisión han ido ganando protagonismo en la investigación en turismo como resultado de su habilidad para obtener variables calificativas, producir reglas explícitas, y permitir la comprensión de fenómenos de una cierta complejidad por parte de gestores y responsables de la planificación turística (James et al., 2021). Asimismo, constituyen herramientas útiles para el análisis de segmentación y clasificación de la demanda turística.

Modelos de series temporales

Finalmente, el comportamiento temporal del flujo turístico fue modelado mediante modelos Autorregresivos Integrados de Media Móvil con componente estacional (SARIMA/ARIMA). Estos modelos permiten captar de manera simultánea aspectos como la dependencia temporal, la estacionalidad y la tendencia que se encuentran en las series cronológicas de la demanda turística.

La literatura reciente identifica a los modelos ARIMA como una de las metodologías clásicas más robustas y ampliamente utilizadas para el pronóstico de la demanda turística debido a su capacidad para representar patrones temporales recurrentes y generar proyecciones confiables en horizontes de corto y mediano plazo (Song et al., 2022; Zhao, 2023).

La adecuación del modelo fue evaluada mediante el análisis de residuos, la inspección de funciones de autocorrelación (ACF) y autocorrelación parcial (PACF), así como la verificación de ausencia de dependencia serial significativa, garantizando la validez estadística de las predicciones obtenidas.

Resultados

Caracterización sociodemográfica de los turistas

En la tabla 1 se observa la caracterización sociodemográfica evidenció similitudes y diferencias entre los turistas nacionales y extranjeros que visitan la Reserva de Producción de Fauna Chimborazo. En ambos grupos predominó el rango etario comprendido entre 18 y 35 años, representando el 56,67 % de los visitantes nacionales y el 61,33 % de los visitantes extranjeros.

Asimismo, se observó una mayor participación de hombres, con proporciones de 50,00 % y 54,00 %, respectivamente. Estos resultados sugieren que la demanda turística de la reserva se encuentra principalmente integrada por visitantes jóvenes y económicamente activos.

En relación con la procedencia, los turistas nacionales provinieron principalmente de las provincias de Guayas (26,67 %), Chimborazo (21,33 %) y Pichincha (20,00 %). Por su parte, los turistas internacionales procedían sobre todo de América del Sur (26,67 %), de Europa (21,33 %) y de Oceanía (20,67 %), lo que denotaba una cierta diversificación geográfica de los mercados emisores.

En el apartado del nivel educativo, entre los visitantes nacionales eran más abundantes las personas con formación técnica o tecnológica (30,67 %), mientras que en el grupo de turistas internacionales había mayor cantidad de educación secundaria (43,33 %). Con respecto al tipo de ocupación, los turistas nacionales estuvieron representados, mayoritariamente, por estudiantes (25,33 %) y empleados del sector privado (22,00 %), mientras que los extranjeros tuvieron un porcentaje mayor en trabajadores independientes (24,00 %) y jubilados (21,33 %). Los años 2020 y 2021 no fueron considerados en el análisis debido al impacto extraordinario de la pandemia de COVID-19 sobre la actividad turística, lo que generó una disminución atípica en el flujo de visitantes. (Chumpitazi Dulanto & Hinojosa Flores, 2021)

Tabla 1: Perfil sociodemográfico de los turistas nacionales y extranjeros.

Pregunta	Visitantes Nacionales		Visitantes Extranjeros	
Lugar de residencia	Chimborazo	21,33%	América del Sur	26,67%
	Tungurahua	6,67%	América del Norte	5,33%
	Pichincha	20,00%	Europa	21,33%
	Carchi	9,33%	Asia	15,33%
	Guayas	26,67%	África	10,67%
	Azuay	16,00%	Oceanía	20,67%
Edad	Menos de 18 años	8,00%	Menos de 18 años	14,67%
	18 a 35 años	56,67%	18 a 35 años	61,33%
	36 a 55 años	25,33%	36 a 55 años	18,00%
	56 en adelante	10,00%	56 en adelante	6,00%
Género	Femenino	28,67%	Femenino	28,00%
	Masculino	50,00%	Masculino	54,00%
	Prefiero no responder	12,00%	Prefiero no responder	12,00%
	Otro	9,33%	Otro	6,00%

Nivel de instrucción	Primaria	4,67%	Primaria	2,67%
	Secundaria	28,00%	Secundaria	43,33%
	Técnico o tecnológico	30,67%	Técnico o tecnológico	24,00%
	Universitario	23,33%	Universitario	18,00%
	Posgrado	13,33%	Posgrado	12,00%
Ocupación principal	Estudiante	25,33%	Estudiante	14,00%
	Empleado público	14,00%	Empleado público	12,00%
	Empleado privado	22,00%	Empleado privado	20,67%
	Trabajador independiente	18,67%	Trabajador independiente	24,00%
	Empresario/a	9,33%	Empresario/a	8,00%
	Jubilado/a	8,67%	Jubilado/a	21,33%
	Otro: _	2,00%	Otro:	0,00%

Fuente: Elaboración propia a partir de encuestas aplicadas a turistas de la Reserva de Producción de Fauna Chimborazo, 2025.

Análisis de correlación y reducción de dimensionalidad

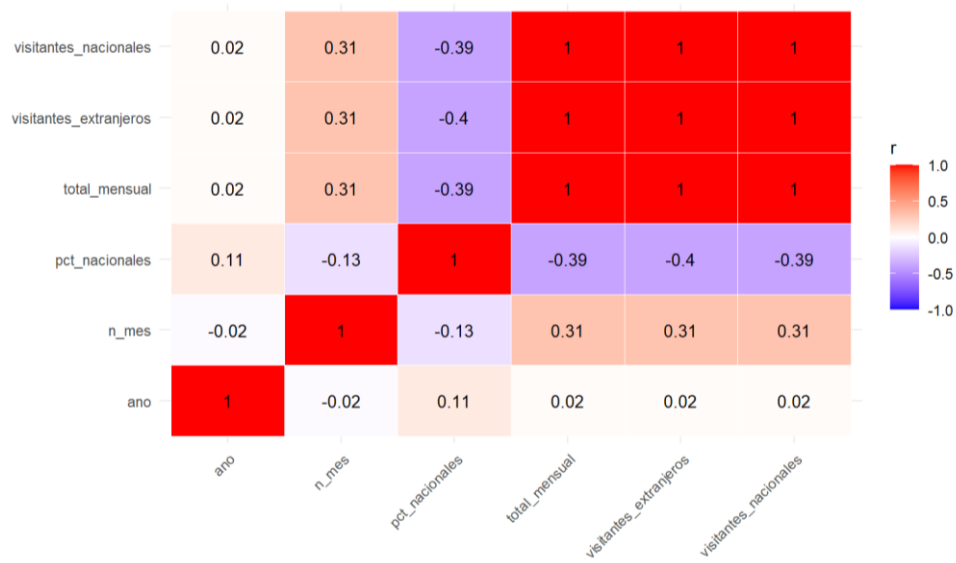
Matriz de correlación

La matriz de correlación mostró una asociación positiva muy elevada entre el número de turistas nacionales, turistas extranjeros y el total mensual de visitantes. Este resultado refleja que las variaciones observadas en el flujo turístico total se encuentran estrechamente vinculadas con el comportamiento conjunto de ambos grupos de visitantes.

Por otra parte, el porcentaje de participación nacional presentó una correlación negativa moderada con el flujo total ($r = -0,56$), indicando que durante los períodos de mayor afluencia aumenta proporcionalmente la participación de turistas extranjeros dentro de la estructura de la demanda.

Las variables temporales analizaron asociaciones considerablemente menores. El mes mostró una correlación positiva débil con el flujo turístico ($r = 0,19$), mientras que el año presentó una relación prácticamente nula ($r = -0,02$), sugiriendo que la dinámica de visitantes responde principalmente a factores estacionales más que a una tendencia lineal de crecimiento o disminución a largo plazo.

Figura 1: Matriz de correlación entre variables del flujo turístico



Fuente: Elaboración propia con datos de Turismo en Cifras, Ministerio de Turismo del Ecuador (MINTUR), período 2015–2025.

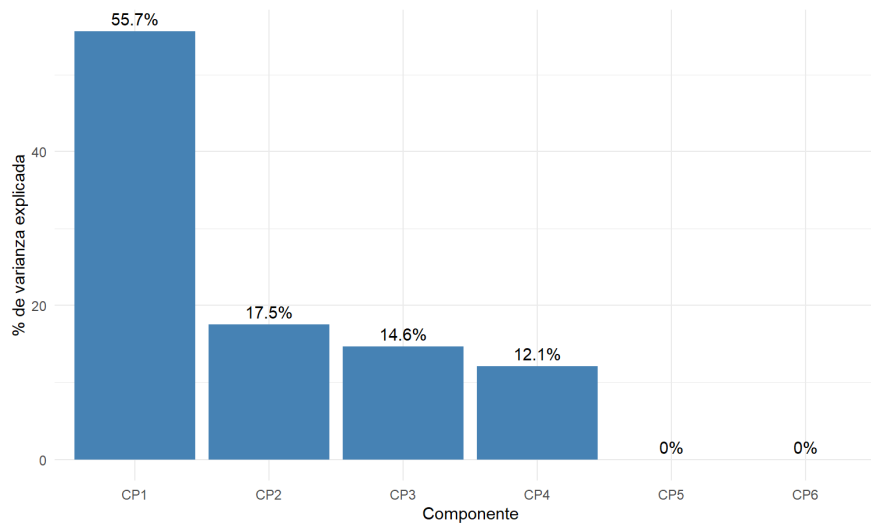
Análisis de Componentes Principales

El Análisis de Componentes Principales permitió resumir la información contenida en las variables del flujo turístico mediante un número reducido de componentes.

El primer componente principal explicó el 57,4 % de la variabilidad total observada. El segundo y tercer componente explicaron adicionalmente el 16,8 % y el 15,8 %, respectivamente. De esta manera, los tres primeros componentes acumularon aproximadamente el 90 % de la varianza total.

Estos resultados demuestran que la información relevante del sistema turístico puede representarse adecuadamente mediante un conjunto reducido de dimensiones, facilitando la interpretación de la estructura subyacente de los datos y justificando la aplicación de técnicas posteriores de segmentación.

Figura 2: Varianza explicada por componentes principales (Scree Plot).



Fuente: Elaboración propia mediante análisis de componentes principales aplicado a datos del MINTUR, 2015–2025.

Segmentación del flujo turístico mediante análisis de conglomerados

El análisis de conglomerados permitió identificar tres segmentos diferenciados de comportamiento turístico para los 108 meses analizados.

El segmento de bajo flujo agrupó 15 meses (13,9 %), con un promedio de 2.547 visitantes mensuales. El segmento de flujo medio concentró 55 meses (50,9 %), registrando una media de 6.289 visitantes. Finalmente, el segmento de alto flujo agrupó 38 meses (35,2 %), con un promedio de 9.478 visitantes por mes.

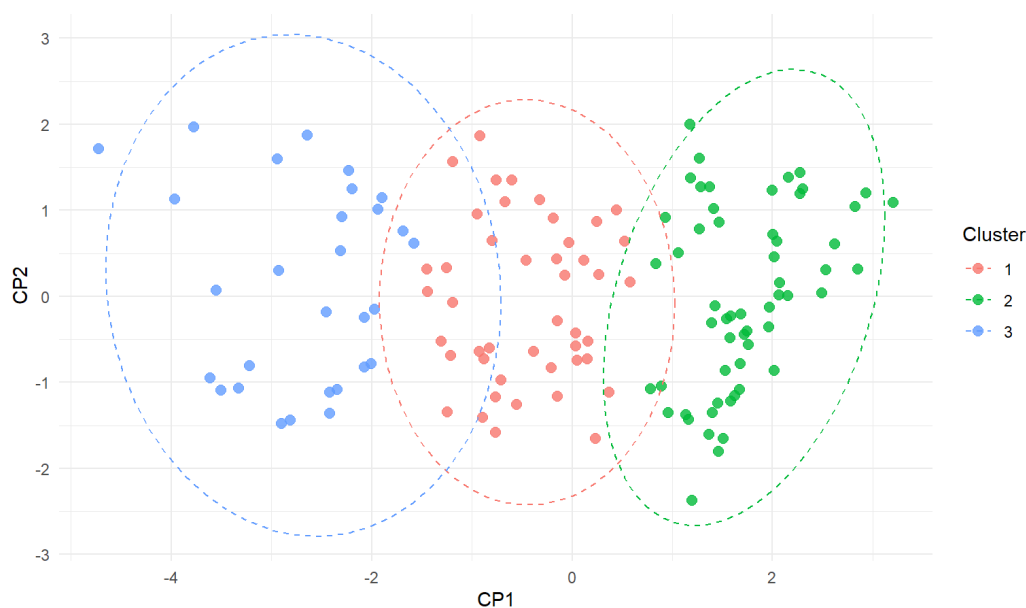
La distribución observada sugiere que la actividad turística de la reserva se desarrolla predominantemente bajo condiciones de flujo medio y alto, los cuales representan conjuntamente el 86,1 % del período analizado.

Tabla 2: Caracterización de los segmentos turísticos identificados

Segmento	N. de meses	Promedio total	Promedios nacionales	Promedios extranjeros
Bajo flujo	15	2,547	1,834	713
Flujo medio	55	6,289	4,528	1,761
Alto flujo	38	9,478	6,824	2,654

Fuente: Elaboración propia con base en registros de visitantes del Ministerio de Turismo del Ecuador (MINTUR), período 2015–2025

Figura 3: Segmentación de períodos según intensidad del flujo turístico.



Fuente: Elaboración propia a partir de registros de visitantes de la Reserva de Producción de Fauna Chimborazo y datos del MINTUR, 2015–2025.

Adicionalmente, se observó que la estructura relativa de participación entre visitantes nacionales y extranjeros permaneció relativamente estable en los tres segmentos, indicando que las diferencias se encuentran principalmente asociadas al volumen de visitantes y no a cambios sustanciales en la composición del mercado.

Evaluación de diferencias multivariadas mediante MANOVA

Los resultados del análisis MANOVA evidenciaron diferencias multivariadas estadísticamente significativas asociadas a la temporada turística.

La Traza de Pillai indicó un efecto significativo de la temporada en el comportamiento conjunto de las variables que fueron analizadas (Pillai = 0,266; $F = 15,22$; $p < 0,001$) lo que manifiesta de forma suficiente la evidente estacionalidad del flujo turístico de la reserva.

Igualmente, se constató un efecto significativo del estado de los datos (consolidado frente a estimado) con un valor de Pillai de 0,081 y un $p = 0,0138$, lo que conlleva a inferir diferencias sistemáticas entre ambos períodos.

Por el contrario, la interacción entre temporada y estado no resultó estadísticamente significativa (Pillai = 0,019; $p = 0,488$), lo que sugiere que los efectos asociados a la estacionalidad se mantienen relativamente constantes independientemente del estado de los registros analizados.

Tabla 1: Resultados del MANOVA para las variables del flujo turístico.

Efecto	Traza de Pillai	F	p-valor	Conclusión
Temporada	0.266	15.22	<0.0001	Significativo
Estado (Consolidado/Estimado)	0.081	3.69	0.0138	Significativo
Temporada × Estado	0.019	0.82	0.488	No significativo

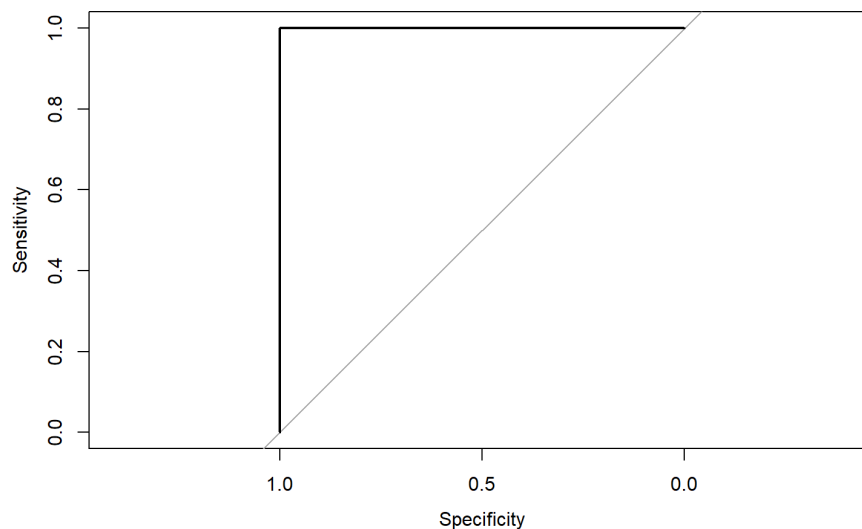
Fuente: Elaboración propia a partir de datos del Ministerio de Turismo del Ecuador (MINTUR), 2015–2025.

Modelamiento predictivo de la demanda turística

Regresión logística

La regresión logística presentó una elevada capacidad discriminante para clasificar períodos de alta afluencia turística. La curva ROC alcanzó un valor AUC de 0,963, lo que indica un desempeño predictivo excelente y una elevada probabilidad de clasificación correcta entre períodos de alta y baja afluencia. Este resultado demuestra que las variables consideradas presentan una importante capacidad explicativa sobre los niveles de demanda turística observados.

Figura 4: Curva ROC del modelo de regresión logística



Fuente: Elaboración propia mediante modelamiento predictivo realizado sobre la base de datos turística de la Reserva de Producción de Fauna Chimborazo.

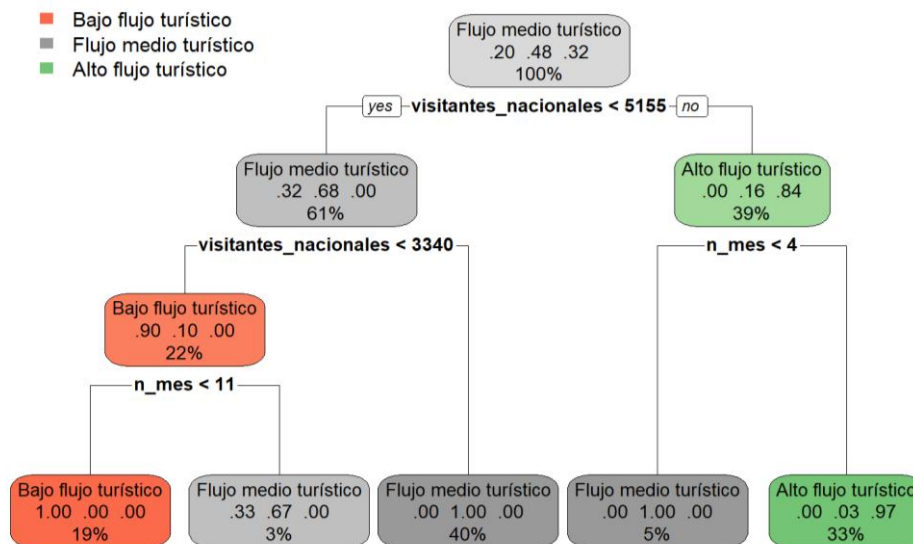
Árbol de decisión para visitantes nacionales

El árbol de decisión identificó al número de visitantes nacionales como el principal factor de clasificación de los niveles de flujo turístico.

Cuando el número de visitantes nacionales fue inferior a 3.340 turistas, la probabilidad de clasificar el período dentro del segmento de bajo flujo alcanzó aproximadamente el 90 %. En contraste, valores superiores permitieron identificar principalmente períodos de flujo medio y alto.

Estos resultados muestran que el volumen de visitantes nacionales constituye un indicador clave para la clasificación de la demanda turística en la reserva.

Figura 5: Árbol de decisión para visitantes nacionales



Fuente: Elaboración propia mediante técnicas de aprendizaje supervisado aplicadas a datos turísticos del MINTUR, 2015–2025.

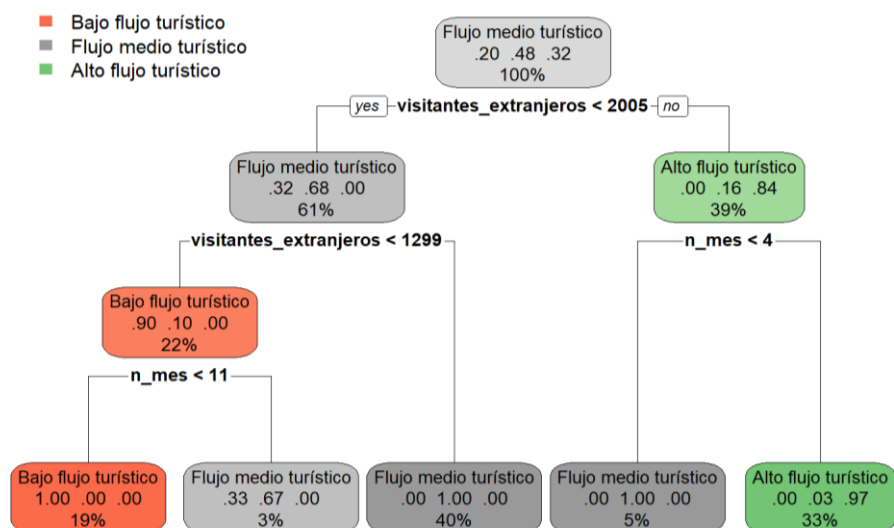
Árbol de decisión para visitantes extranjeros

Para los visitantes extranjeros, el modelo estableció umbrales de clasificación claramente diferenciados.

Períodos con menos de 1.299 visitantes extranjeros se asociaron principalmente con niveles bajos de demanda, mientras que valores comprendidos entre 1.299 y 2.005 turistas correspondieron a flujos medios. Por encima de 2.005 visitantes extranjeros predominaron los períodos de alta afluencia.

Estos resultados evidencian la relevancia del mercado internacional en la explicación de las fluctuaciones del flujo turístico.

Figura 6: Árbol de decisión para visitantes extranjeros



Fuente: Elaboración propia mediante técnicas de clasificación basadas en árboles de decisión aplicadas a datos turísticos del MINTUR, 2015–2025.

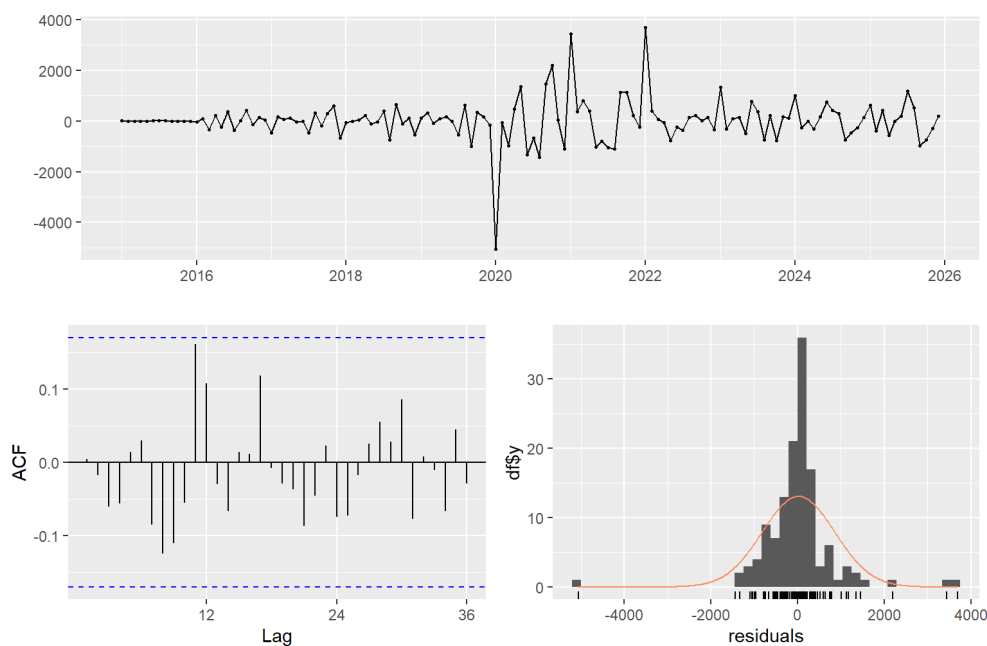
El árbol de clasificación establece que el número de turistas extranjeros presenta una alta capacidad de separación para clasificar los periodos turísticos. Para el número de turistas extranjeros inferior a 1299 se podría clasificar el flujo como bajo turismo, para el número de turistas extranjeros entre 1299 y 2005 se clasifica como turismo medio, y para el ingreso de más de 2005 turistas extranjeros se considera turismo alto.

Análisis temporal y pronóstico del flujo turístico

El comportamiento temporal de la demanda fue modelado mediante un esquema ARIMA estacional, el cual mostró un ajuste satisfactorio a la dinámica histórica de visitantes. La investigación de los residuos dio prueba de la existencia de una distribución centrada alrededor de cero y una carencia de patrones sistemáticos relevantes. La función de autocorrelación de los residuos exhibe gran mayoría de los rezagos contenidos dentro de los intervalos de confianza, lo que pone de manifiesto que el modelo ha sido capaz de atrapar la dependencia temporal y la estacionalidad que presenta la serie.

Sin embargo, durante el periodo de la pandemia de COVID-19, en particular en 2020 y 2021, se evidencian valores atípicos, los cuales son indicativos de alteraciones superlativas en la dinámica de la demanda turística respecto a la evolución histórica esperada. Por lo tanto, se puede afirmar que el modelo ARIMA tiene una capacidad apropiada para describir y predecir la evolución temporal del flujo turístico en la Reserva de Producción de Fauna Chimborazo.

Figura 7: Diagnóstico de residuos del modelo ARIMA.



Fuente: Elaboración propia mediante análisis de series temporales con datos de visitantes de la Reserva de Producción de Fauna Chimborazo, período 2015–2025.

Discusión

Caracterización del perfil turístico

Los resultados evidenciaron que la demanda turística de la Reserva de Producción de Fauna Chimborazo está conformada principalmente por visitantes jóvenes entre 18 y 35 años, tanto nacionales como extranjeros. Esta opción coincide con las tendencias existentes en los destinos de naturaleza y aventura, donde los segmentos de edad más joven tienen un mayor apego a la práctica de actividades recreativas, al aire libre, relacionadas con el patrimonio de los recursos naturales y tal vez también por el conocimiento que los segmentos de edad más joven tienen sobre la naturaleza. En último lugar, la importancia del perfil de visitante que se encuentra en una situación económica activa expresa la relevancia de este perfil de visitantes en la sostenibilidad económica de los destinos de turismo protegido.

En cuanto al origen de los visitantes, la presencia simultánea de mercados nacionales e internacionales refleja un perfil de demanda diversificado. Esta característica resulta favorable para la resiliencia turística del destino, ya que reduce la dependencia de un solo mercado emisor y favorece la estabilidad de los flujos turísticos frente a cambios económicos o sociales que puedan afectar a una región específica (UNWTO, 2023).

Los resultados obtenidos sugieren además la necesidad de implementar estrategias diferenciadas de promoción turística, diseño de productos y comunicación, considerando las diferencias observadas en educación, ocupación y procedencia geográfica de los visitantes,

factores que influyen directamente en sus motivaciones y patrones de consumo turístico (Dolnicar, 2022).

Estructura del flujo turístico y reducción de dimensionalidad

La matriz de correlación mostró una asociación positiva muy elevada entre los visitantes nacionales, extranjeros y el flujo total de turistas. Este comportamiento confirma la existencia de una fuerte dependencia estructural entre las variables que conforman el sistema turístico analizado. No obstante, también sugiere la presencia de multicolinealidad, situación frecuentemente observada cuando varias variables representan componentes de una misma estructura de medición (Hair et al., 2022).

Por otra parte, el porcentaje de participación nacional presentó una correlación negativa moderada con el flujo total de visitantes. Este resultado indica que durante los períodos de mayor afluencia turística aumenta proporcionalmente la contribución del turismo internacional. Situaciones similares han sido reportadas en estudios de demanda turística donde los visitantes extranjeros suelen incrementar su participación durante temporadas de mayor atractivo y conectividad turística (Song et al., 2022).

Los resultados del Análisis de Componentes Principales demostraron que aproximadamente el 90 % de la variabilidad total puede resumirse mediante tres componentes principales. Este hallazgo coincide con investigaciones recientes que destacan la capacidad del ACP para sintetizar información compleja, reducir redundancia estadística y mejorar la interpretación de fenómenos turísticos multidimensionales (Zhang et al., 2023; Zhao, 2023).

Desde una perspectiva aplicada, la elevada varianza explicada confirma que la dinámica turística de la reserva puede comprenderse mediante un conjunto reducido de factores asociados al volumen de visitantes, la composición del mercado y la estacionalidad, facilitando el desarrollo de modelos de monitoreo y gestión más eficientes.

Segmentación turística y patrones de demanda

El uso del análisis de conglomerados hizo posible la determinación de tres segmentos de comportamiento diferenciados: bajo flujo, flujo medio y alto flujo turístico, además de la clara predominancia de los segmentos medio y alto, lo que indica que la reserva presenta niveles relativamente estables de actividad turística a lo largo de la mayor parte del tiempo descrito en el análisis.

La utilidad de la segmentación en función de conglomerados ha sido extensamente descrita y discutida en la literatura turística gracias a la capacidad que tiene para extraer patrones homogéneos de comportamiento y a su utilidad para la formulación de estrategias de gestión y marketing específicas (Dolnicar, 2022; Tiwari & Tripathi, 2023).

Un hallazgo que merece ser mencionado es que la relación entre los turistas nacionales y los turistas extranjeros se mantuvo en términos relativamente constantes durante los segmentos de

los cuales se habla en el presente artículo. Lo que se puede deducir de este párrafo que se ha redactado es que las discrepancias observadas en los segmentos de turistas no están inevitablemente relacionadas con cambios en los propios mercados emisores. Dicho de otro modo, las discrepancias observadas en los segmentos de turistas están relacionadas, principalmente, con cambios en la magnitud de la demanda.

Desde la perspectiva de la planificación turística, esta categorización ofrece un recurso útil para perfeccionar la distribución de recursos en sus diferentes aspectos, ejercitar los sistemas de control de flujos de visitantes y ajustar los servicios a la intensidad anticipada del flujo turístico.

Influencia de la estacionalidad sobre el comportamiento turístico

Los resultados del MANOVA evidenciaron que la temporada turística constituye uno de los factores más relevantes para explicar las variaciones observadas en el comportamiento conjunto de las variables analizadas. La significancia estadística obtenida confirma que la estacionalidad continúa siendo uno de los principales determinantes de la dinámica turística en destinos de naturaleza.

Estos resultados son consistentes con investigaciones que señalan que los patrones estacionales responden a factores climáticos, períodos vacacionales y condiciones de accesibilidad que afectan directamente la movilidad de los visitantes y la demanda de servicios turísticos (Song et al., 2022).

La utilización de la Traza de Pillai permitió garantizar mayor robustez en la evaluación estadística de las diferencias multivariadas. Diversos autores destacan este estadístico como uno de los más recomendables cuando existen potenciales desviaciones respecto a los supuestos tradicionales del MANOVA, debido a su estabilidad y resistencia frente a problemas de heterogeneidad de covarianzas (Murphy, 2024).

Desde un punto de vista de la gestión, estos resultados enfatizan la necesidad de la incorporación de criterios estacionales en la planificación operativa, expresada sobre todo en aspectos de capacidad de carga, de disponibilidad de personal, de provisión de servicios y de estrategia de promoción.

Desempeño de los modelos predictivos

La regresión logística presentó la elevada capacidad discriminante para la detección de períodos con alta y baja afluencia, con una AUC = 0,963. Según la corriente científica, valores superiores a 0,90 expresan un excelente rendimiento predictivo y una correcta capacidad para clasificar los modelos supervisados (James et al., 2021).

Estos resultados reflejan que las variables englobadas dentro del modelo albergan información de carácter muy pertinente para tratar de dar una explicación de las oscilaciones que se van apreciando en la demanda turística. No obstante, unos valores de rendimiento elevados deben ser tomados como advertencia porque pueden estar muy próximos a incurrir en el sobreajuste,

especialmente en el caso de las variables predictoras cuyas relaciones estructurales son muy estrictas.

Los árboles de decisión completaron de una forma muy importante el análisis predictivo al presentar reglas clasificatorias de fácil interpretación. Esta propiedad se convierte en una de las grandes ventajas de esta técnica, dado que es capaz de convertir resultados estadísticos complejos en criterios accesibles para la toma de decisiones (James et al., 2021).

De esta manera, el establecimiento de umbrales para los turistas nacionales y extranjeros configura instrumentos operativos que pueden ser utilizados para prever situaciones de alta ocupación y establecer acciones de prevención relacionadas con el uso sostenible del destino.

Dinámica temporal y pronóstico del flujo turístico

Los resultados del modelo ARIMA mostraron una adecuada representación de la estructura temporal de la demanda turística. La ausencia de patrones sistemáticos en los residuos y la baja autocorrelación residual sugieren que el modelo logró capturar satisfactoriamente la tendencia y los componentes estacionales de la serie.

Este comportamiento coincide con investigaciones recientes que señalan que los modelos ARIMA continúan siendo herramientas robustas para el pronóstico de la demanda turística debido a su capacidad para modelar dependencias temporales y patrones recurrentes de comportamiento (Song et al., 2022; Zhao, 2023).

No obstante, la presencia de anomalías durante los años 2020 y 2021 confirma la sensibilidad de la actividad turística frente a fenómenos extraordinarios. La pandemia de COVID-19 provocó alteraciones significativas en los patrones históricos de movilidad, generando fluctuaciones que difícilmente pueden ser explicadas únicamente por el comportamiento interno de la serie temporal.

En consecuencia, futuras investigaciones podrían incorporar variables exógenas relacionadas con factores climáticos, indicadores económicos, feriados nacionales, conectividad vial y eventos extraordinarios con el fin de incrementar la capacidad explicativa y predictiva de los modelos de pronóstico.

Implicaciones para la gestión turística sostenible

Los resultados arrojados sugieren que la combinación del análisis multivariante y el modelamiento predictivo es un método adecuado que permite obtener un conocimiento sobre la dinámica turístico-receptiva de las áreas naturales protegidas. La combinación del ACP, análisis de conglomerados, MANOVA, regresión logística, árboles de decisión y los modelos ARIMA ha permitido detectar patrones estructurales, segmentar niveles de demanda, evaluar los efectos estacionales y generar herramientas de apoyo para la planificación de los usos del espacio.

Desde una postura aplicada se considera que la evidencia generada puede ser una herramienta útil para la optimización de procesos de gestión de visitantes, de control de capacidad de carga,

de diseño de estrategias promocionales y de conservación de los recursos naturales de la Reserva de Producción de Fauna Chimborazo.

En fin, la metodología presentada en esta investigación se convierte en una propuesta que se puede reproducir en otras Áreas Protegidas de Ecuador y de la región andina que aporte a la generación de sistemas de gestión turística sustentados en evidencia científica y en analítica de datos.

Conclusiones

La aplicación rigurosa de los modelos estadísticos multivariantes y los modelos pronósticos posibilitó caracterizar, segmentar y analizar el comportamiento de los turistas nacionales y extranjeros que visitaron la Reserva de Producción de Fauna Chimborazo en el periodo 2015–2025. Los resultados comprobaron que la dinámica del turismo de la reserva de producción de fauna está condicionada sobre todo por el flujo de la intensidad de visitantes y los efectos de la estacionalidad, factores que condicionan bastante la evolución de la demanda turística.

El análisis sociodemográfico mostró que el perfil más representativo de aquellos que visitan el destino son los adultos jóvenes, con edades comprendidas entre los 18 y 35 años, apreciándose distinciones significativas entre los turistas nacionales y extranjeros, en cuanto a cuestiones tales como lugar de origen, nivel educativo, ocupación, etc. Los resultados reflejan que existen segmentos diferenciados en la demanda turística y, por tanto, tienen que traducirse por medio de adecuadas estrategias de promoción, interpretación y gestión turística.

El PCA ha conseguido reducir la complejidad de la información que se tenía al encontrar 3 componentes principales que sintetizaban aproximadamente un 90% del total de variabilidad existente. Este resultado ha demostrado la existencia de una estructura subyacente relativamente estable que explica el comportamiento del flujo turístico de la reserva y pone de manifiesto la utilidad de las técnicas multivariantes para el estudio del comportamiento de sistemas turísticos complejos.

A partir del análisis de conglomerados, se llegó a la identificación de un total de tres categorías de comportamiento turístico en función del flujo, que serían bajo, medio y alto flujo. La distribución que fue presentada mostró una predominancia de períodos de medio y alto flujo; lo que da cuenta de la existencia de una demanda relativamente estable durante una buena parte del periodo analizado. La composición proporcional entre nacionales y extranjeros será relativamente constante entre categorías de flujo, lo cual implicaría que los cambios que aquí se observan deberán responder, en su mayor parte, a variaciones en el número total de turistas.

Los resultados del análisis MANOVA confirmaron que la estacionalidad es uno de los principales factores que incide de modo determinante sobre la forma de comportamiento que tienen las distintas variables turísticas analizadas. A partir de este hallazgo se entiende que los procesos de planificación y gestión de la Reserva de Producción de Fauna Chimborazo deban insertar

criterios estacionales que permitan hacer un uso más óptimo de las asignaciones de recursos y de los servicios a prestar, así como del manejo de visitantes.

Los modelos predictivos generados demostraron tener una alta capacidad para la clasificación y para la predicción de patrones de demanda en el ámbito de la actividad turística, tal y como ha quedado evidenciado. La regresión logística manifestó una competencia predictiva fabulosa, por su parte, los árboles de decisión posibilitaron la exposición de una serie de reglas de clasificación fácilmente interpretables para la operación del destino. Estos resultados acreditan la capacidad de las herramientas de analítica de datos para soportar la toma de decisiones en entornos turísticos.

Para concluir, el modelo ARIMA fue capaz de ilustrar la serie de tiempo y la serie estacional del flujo turístico, permitiendo la descripción de la respuesta de la serie a algún componente de la demanda, y de la misma forma, permitir la utilización de la historia para acercar pronósticos de la demanda en el futuro. No obstante, la ocurrencia de fenómenos atípicos como es el caso de la pandemia COVID-19 puso de manifiesto la necesidad de incorporar variables externas en próximos ejercicios de modelado a fin de mejorar la precisión de la proyección.

Referencias Bibliográficas

Aragones, P., & Graells, M. (2019). *Modelos de segmentación y análisis de comportamiento turístico*. Editorial Académica Española.

Assaf, A. G., Song, H., Tsionas, M., & Oh, C. O. (2022). Tourism forecasting and demand analysis: Evidence and future directions. *Tourism Economics*, 28(5), 1121–1143. <https://doi.org/10.1177/13548166211012345>

Athanasopoulos, G., Hyndman, R. J., Song, H., & Wu, D. C. (2020). The tourism forecasting competition. *International Journal of Forecasting*, 36(1), 1–15. <https://doi.org/10.1016/j.ijforecast.2019.05.009>

Betancur, J. D. (2022). Aplicación del análisis multivariado de la varianza (MANOVA) en investigaciones sociales y económicas. *Revista Colombiana de Estadística Aplicada*, 15(2), 45–60.

Bulla, J. (2013). *Modelos ARIMA y pronóstico de series temporales*. Universidad Nacional de Colombia.

Canavos, G. C., & Medal, M. (1987). *Probabilidad y estadística: Aplicaciones y métodos*. McGraw-Hill.

Collado, J., García, R., & Martín, F. (2007a). *Modelos de clasificación aplicados a estudios de mercado*. Pearson Educación.

Collado, J., García, R., & Martín, F. (2007b). *Regresión logística y técnicas predictivas en ciencias sociales*. Pearson Educación.

- Chávez Quisbert, C. (1997). *Metodología de la investigación científica*. Editorial Universitaria.
- Chumpitazi Dulanto, L., & Hinostroza Flores, N. (2021). Impacto del COVID-19 en la actividad turística y los flujos de visitantes en áreas protegidas. *Revista Turismo y Desarrollo*, 12(1), 27–41.
- Dolnicar, S. (2022). Cluster analysis in tourism. En J. Jafari & H. Xiao (Eds.), *Encyclopedia of Tourism*. Springer. https://doi.org/10.1007/978-3-319-01669-6_319-2
- Domínguez, M., Ramírez, L., & Torres, J. (s. f.). *Caracterización de perfiles turísticos y análisis de demanda*. Editorial Académica.
- Fernández, C., Hernández, R., & Baptista, P. (2002). *Metodología de la investigación* (3.^a ed.). McGraw-Hill.
- Fernández López, S., Muñoz, A., & Rodríguez, M. (2020). Aplicación de técnicas multivariantes en el análisis de perfiles turísticos. *Revista Internacional de Turismo, Empresa y Territorio*, 4(2), 55–74.
- Gurrea, R. (s. f.). *Análisis de componentes principales y reducción de dimensionalidad*. Universidad de Zaragoza.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2022). *Multivariate data analysis* (9th ed.). Cengage Learning.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning: With applications in R* (2nd ed.). Springer. <https://doi.org/10.1007/978-1-0716-1418-1>
- Landaluce Calvo, M. I. (2017). *Series temporales y pronósticos económicos*. Editorial Pirámide.
- Mejía, D., Salazar, P., & López, F. (2016). Aplicación del análisis de componentes principales en estudios multidimensionales. *Revista Colombiana de Estadística*, 39(2), 123–138.
- Murphy, T. J. (2024). *Multivariate analysis of variance (MANOVA)*. Statistical Design and Analysis of Experiments with R. <https://tjmurphy.github.io/jabstb/manova.html>
- Proaño Ponce, J. (2020). Aplicación de árboles de decisión en el análisis de variables turísticas. *Revista Ecuatoriana de Estadística Aplicada*, 9(1), 88–101.
- Ramírez Muñoz, M. (2021). *Métodos multivariantes aplicados a la investigación turística*. Editorial Alfaomega.
- Song, H., Qiu, R. T. R., & Park, J. (2022). Tourism forecasting: Methods and applications. *Annals of Tourism Research*, 94, 103385. <https://doi.org/10.1016/j.annals.2022.103385>
- Song, H., Witt, S. F., & Li, G. (2019). *Tourism demand modelling and forecasting: Modern econometric approaches*. Routledge.

Tiwari, M., & Tripathi, S. (2023). Application of clustering algorithms on tourism industry. *International Journal for Research in Applied Science and Engineering Technology*, 11(7), 742–748. <https://doi.org/10.22214/ijraset.2023.51380>

Tkaczynski, A., Rundle-Thiele, S., & Beaumont, N. (2015). Segmentación de mercados turísticos y análisis multivariante. *Tourism Management Perspectives*, 16, 67–76. <https://doi.org/10.1016/j.tmp.2015.06.001>

United Nations World Tourism Organization. (2023). *UNWTO World Tourism Barometer*. UNWTO. <https://www.unwto.org>

Vásconez Rosero, M. (2025). *Segmentación estadística de mercados turísticos mediante técnicas de clustering*. Universidad Técnica del Norte.

Veritt, R., Sánchez, P., & Molina, A. (s. f.). *Análisis de conglomerados aplicado a perfiles de visitantes*. Editorial Académica.

Zhang, C., Hu, A.-Y., & Tian, Y.-X. (2023). Daily tourism forecasting through a novel method based on principal component analysis, grey wolf optimizer, and extreme learning machine. *Journal of Forecasting*, 42(8), 2121–2138. <https://doi.org/10.1002/for.3007>

Zhao, X. (2023). Tourism demand forecasting using PCA-BPNN. *Academic Journal of Computing & Information Science*, 6(9), 72–80. <https://doi.org/10.25236/AJCIS.2023.060911>