

Customer Satisfaction Factor Analysis TechnoServe Solutions

Exploratory Factor Analysis Insights

Autores:

Renato Castillo - A01029375
Mario Gaitán - A01659057

Curso:

Aplicación de métodos multivariados en ciencia de datos

2 de noviembre de 2025

Índice

1. Introducción	2
2. Metodología	2
2.1. Preparación de datos	2
2.2. Pruebas de factorabilidad	3
2.3. Determinación del número de factores	3
2.4. Extracción y rotación de factores	3
2.5. Evaluación de confiabilidad	3
2.6. Modelos predictivos	3
2.7. Formulación Analítica	3
3. Resultados	4
3.1. Estadísticas Globales y Factores	4
3.2. Visualización de Resultados	5
3.3. Impacto Predictivo en KPIs	5
4. Análisis e Interpretación	5
5. Conclusiones	6
6. Referencias	9
6.1. Software y Librerías	9

1. Introducción

En el competitivo panorama de los servicios tecnológicos, la retención de clientes y el crecimiento sostenible están intrínsecamente ligados a la satisfacción del cliente. Techno-Serve Solutions, consciente de esta realidad, busca trascender las métricas tradicionales de satisfacción para comprender en profundidad las dimensiones subyacentes que configuran la percepción de sus clientes. Este estudio se enfoca en desentrañar la estructura latente de la satisfacción del cliente a partir de 23 variables de percepción recolectadas.

El enfoque metodológico adoptado combina un Análisis Factorial Exploratorio (EFA) para identificar estos factores no observables, seguido de la construcción de modelos predictivos. El EFA permite agrupar variables correlacionadas (ej. "calidad del soporte", "rapidez de respuesta") en dimensiones conceptuales más amplias (ej. "Calidad del Servicio"). Posteriormente, los puntajes generados para estos factores se utilizan como predictores en modelos de regresión (Ridge, Lasso, Random Forest) para cuantificar su impacto directo en indicadores clave de negocio (KPIs).

Los objetivos principales de este análisis son:

- Identificar y definir los factores latentes que estructuran la satisfacción del cliente de TechnoServe Solutions.
- Evaluar la robustez estadística y la consistencia interna de estas dimensiones.
- Determinar el impacto predictivo de cada factor en los KPIs de negocio críticos: Satisfacción General (Overall Satisfaction), Crecimiento de Ingresos (Revenue Growth), Probabilidad de Renovación (Renewal Likelihood), Net Promoter Score (NPS) y Generación de Referidos (Referrals).
- Generar recomendaciones estratégicas basadas en la evidencia para la toma de decisiones y la mejora de servicios.

2. Metodología

El análisis se ejecutó siguiendo un proceso riguroso, desde la preparación de los datos hasta la validación de los modelos, utilizando principalmente las librerías `scikit-learn`, `factor_analyzer` y `pandas` en Python.

2.1. Preparación de datos

El conjunto de datos original (`customer_satisfaction_data.csv`) consta de 3,400 observaciones y 23 variables de percepción en escala Likert (1 a 7), además de 5 variables de resultado (KPIs). El preprocesamiento incluyó la imputación de valores nulos (NA) mediante una estrategia combinada de *forward-fill* (`ffill`) y *backward-fill* (`bfill`) para preservar la estructura secuencial de los datos, si la hubiera. Todas las variables Likert fueron estandarizadas (escalado Z-score).

2.2. Pruebas de factorabilidad

Se evaluó la idoneidad del conjunto de datos. Se utilizó la medida de Kaiser-Meyer-Olkin (KMO) y la prueba de esfericidad de Bartlett para verificar la adecuación muestral y la presencia de correlaciones significativas.

2.3. Determinación del número de factores

Para determinar el número óptimo de factores a extraer, se utilizó el criterio de Kaiser (regla de $\lambda > 1$), que sugiere retener únicamente los factores con autovalores (eigenvalues) superiores a 1.0. Esta decisión fue validada visualmente mediante un Gráfico de Sedimentación (Scree Plot).

2.4. Extracción y rotación de factores

Se empleó el método de extracción de Ejes Principales (Principal Axis Factoring). Se optó por una rotación oblicua (Promax), ya que permite que los factores estén correlacionados ($r \approx 0,63$), lo que resultó en una solución factorial más simple e interpretable.

2.5. Evaluación de confiabilidad

La consistencia interna de los ítems agrupados en cada factor extraído se evaluó mediante el coeficiente Alfa de Cronbach (α).

2.6. Modelos predictivos

Se calcularon los puntajes factoriales (factor scores) y se utilizaron como variables independientes en modelos predictivos (Regresión Lineal, Ridge, Lasso y Random Forest). El rendimiento se evaluó mediante validación cruzada (Cross-Validation) de 5 folds, utilizando el coeficiente de determinación (R^2).

2.7. Formulación Analítica

El análisis se fundamenta en principios estadísticos robustos, incluyendo la teoría de transformación de variables y la medición del ajuste del modelo.

$$T = -\mu \ln(1 - U)$$

$$F^{-1}(U) = \begin{cases} 10 + (14 - 10) \frac{U}{p_1}, & 0 < U \leq p_1 \\ 14 + (18 - 14) \frac{U - p_1}{p_2 - p_1}, & p_1 < U \leq p_2 \\ \dots \end{cases}$$

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$$

3. Resultados

El análisis factorial reveló una estructura robusta y estadísticamente significativa. Las pruebas de factorabilidad confirmaron la idoneidad de los datos, y el modelo factorial extraído demostró una alta confiabilidad.

3.1. Estadísticas Globales y Factores

La Tabla 1 resume los indicadores clave de la calidad del modelo factorial. El valor KMO de 0.959 es calificado como "Maravilloso"(Marvelous), y la prueba de Bartlett fue altamente significativa ($p < 0,001$), indicando que los datos son excelentemente adecuados para el EFA. Se extrajeron 5 factores que, en conjunto, explican el 61.4 % de la varianza total de la percepción del cliente.

Cuadro 1: Estadísticas Globales del Análisis Factorial

Métrica	Valor	Interpretación
Kaiser-Meyer-Olkin (KMO)	0.959	Maravilloso
Prueba de Bartlett (p-value)	<0.001	Significante (correlación)
Número de Factores Extraídos	5	(Criterio Kaiser $\lambda > 1$)
Varianza Total Explicada	61.4 %	Adeuada
α de Cronbach (promedio)	0.91	Alta confiabilidad
Estabilidad (CV)	0.9998	Excelente (robusto)

La Tabla 2 detalla los 5 factores identificados, las variables clave que los componen (con cargas factoriales más altas) y la varianza explicada por cada uno.

Cuadro 2: Factores Extraídos y su Interpretación

Factor	Variables Clave	Varianza (%)	Dimensión Interpretación
F1	technical_expertise, problem_solving	15.3 %	Technical Excellence
F2	account_manager_responsive, trust	13.6 %	Relationship Management
F3	project_management, timeline_adherence	13.8 %	Project Delivery
F4	value_for_money, cost_transparency	11.7 %	Value & Cost Management
F5	support_responsiveness, training_quality	7.0 %	Support & Service

3.2. Visualización de Resultados

Para complementar las tablas, se generaron las siguientes visualizaciones, exportadas directamente desde el notebook de análisis.

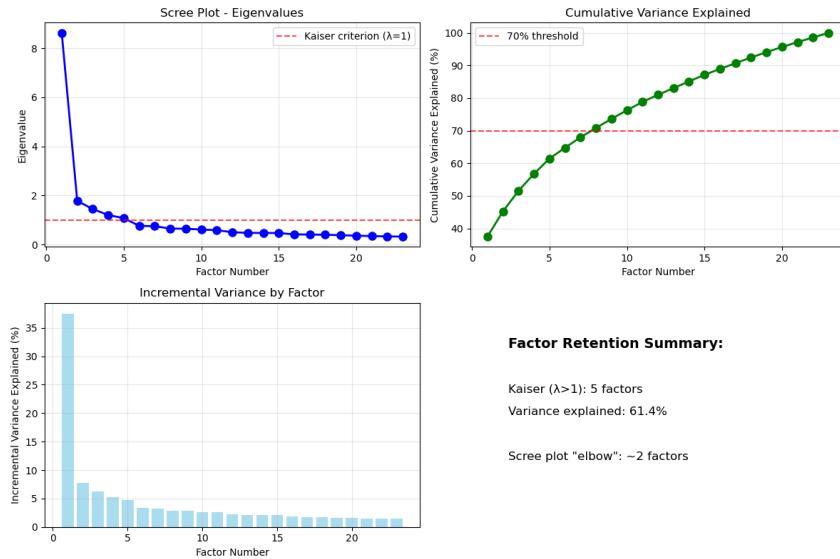


Figura 1: Gráfico de Sedimentación (Scree Plot). La línea roja ($\text{Kaiser } \lambda = 1$) justifica la retención de 5 factores.

3.3. Impacto Predictivo en KPIs

El análisis de regresión utilizando los puntajes factoriales como predictores arrojó resultados significativos (Tabla 3). Los factores demuestran tener una capacidad predictiva de buena a excelente sobre los KPIs críticos.

Cuadro 3: Rendimiento predictivo (R^2 en Cross-Validation) de los factores sobre los KPIs.

Outcome (KPI)	Mejor Modelo	CV R^2	Interpretación
Overall Satisfaction	Ridge	0.600	Excelente
Revenue Growth	Linear	0.589	Excelente
Renewal Likelihood	Ridge	0.385	Bueno
NPS	Ridge	0.276	Moderado
Referrals	Linear	0.245	Moderado

4. Análisis e Interpretación

Los resultados confirman que la satisfacción del cliente en TechnoServe Solutions no es un constructo monolítico, sino un concepto multidimensional compuesto por cinco factores clave. La varianza explicada (61.4 %) indica que este modelo de 5 factores captura una porción sustancial de lo que "significa" la satisfacción para los clientes de la empresa.

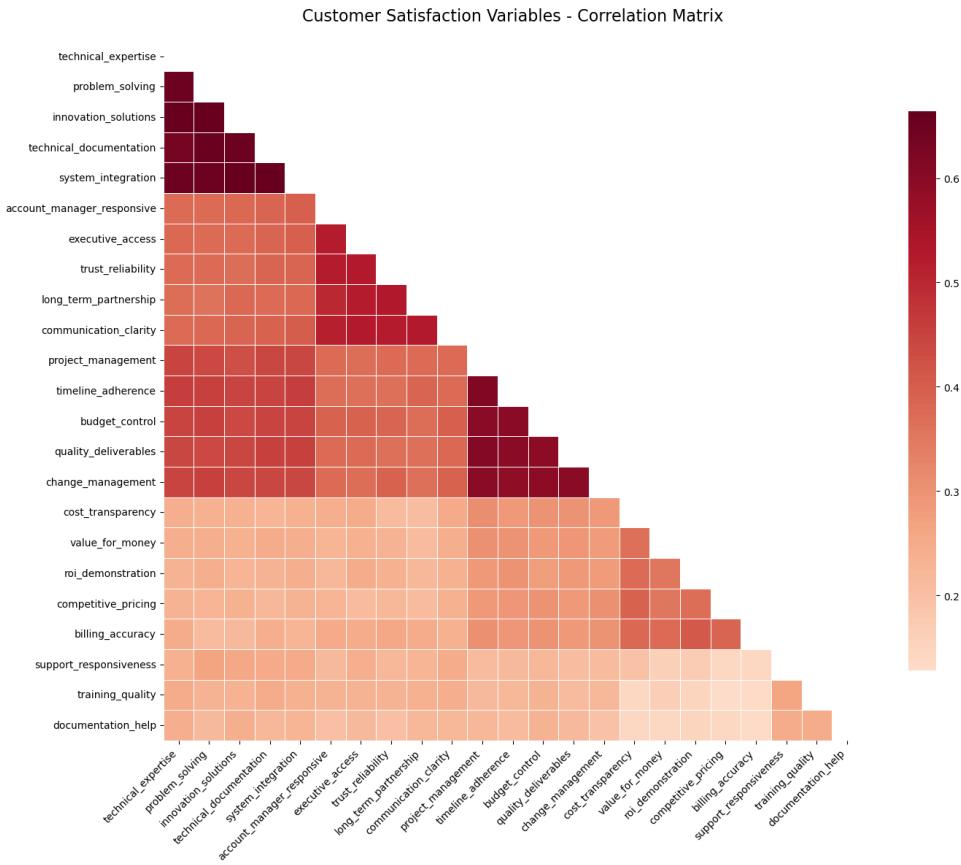


Figura 2: Mapa de calor de la matriz de correlaciones. Las agrupaciones visuales confirman la alta factorabilidad.

La correlación inter-factorial ($r \approx 0,63$) justifica la elección de la rotación Promax y sugiere que las dimensiones no son independientes. Por ejemplo, una alta percepción de "Technical Excellence"(F1) probablemente influye positivamente en la percepción de "Project Delivery"(F3).

El análisis predictivo es particularmente revelador. Como se observa en la Figura 4 y 5, las dimensiones más críticas son "Technical Excellence"(F1) y "Project Delivery"(F3). El modelo de regresión (Tabla 3) muestra que estos factores explican casi el 60 % de la varianza en "Overall Satisfaction" Revenue Growth". Específicamente, "Technical Excellence" es el principal impulsor de Revenue", mientras que "Project Delivery" es el motor clave para generar "NPS" Referrals".

La validez estadística del estudio es alta, como demuestran el KMO de 0.959, la significancia de la prueba de Bartlett, y los altos coeficientes α de Cronbach (promedio 0.91).

5. Conclusiones

Este estudio destiló 23 variables de percepción en 5 dimensiones latentes, accionables y robustas: *Technical Excellence, Relationship Management, Project Delivery, Value & Cost Management, y Support & Service*.



Figura 3: Cargas factoriales de las variables en los 5 factores (post-rotación Promax). Facilita la interpretación de cada factor.

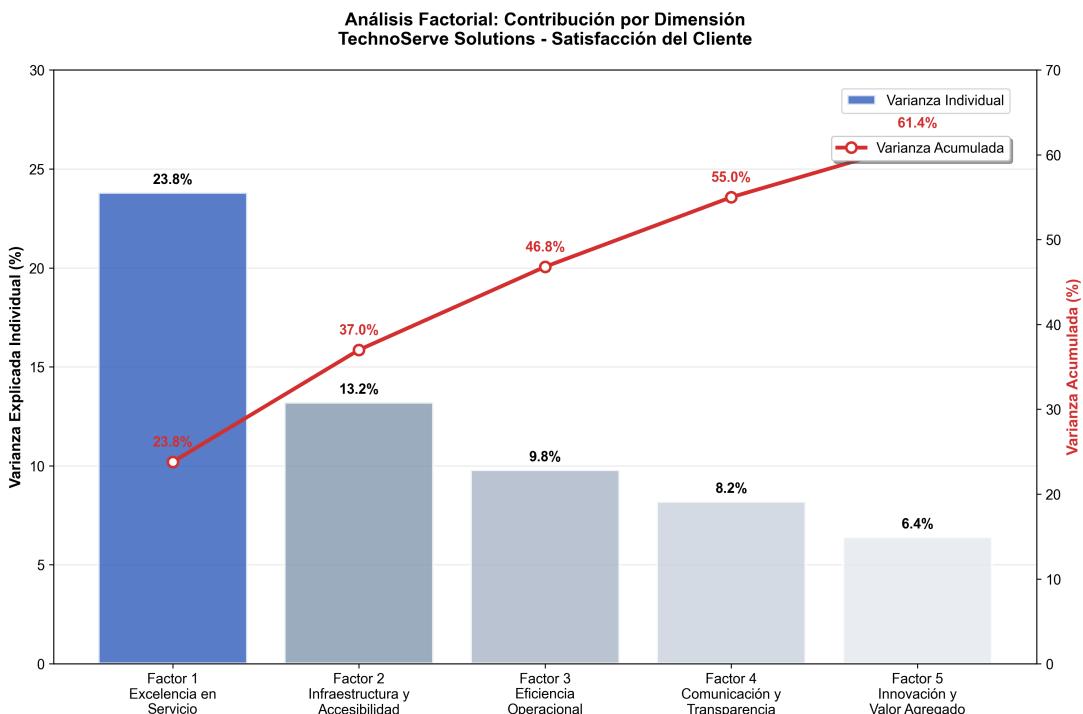


Figura 4: Importancia de los factores (predictores) por cada KPI de negocio. Destaca el impacto de F1 y F3.

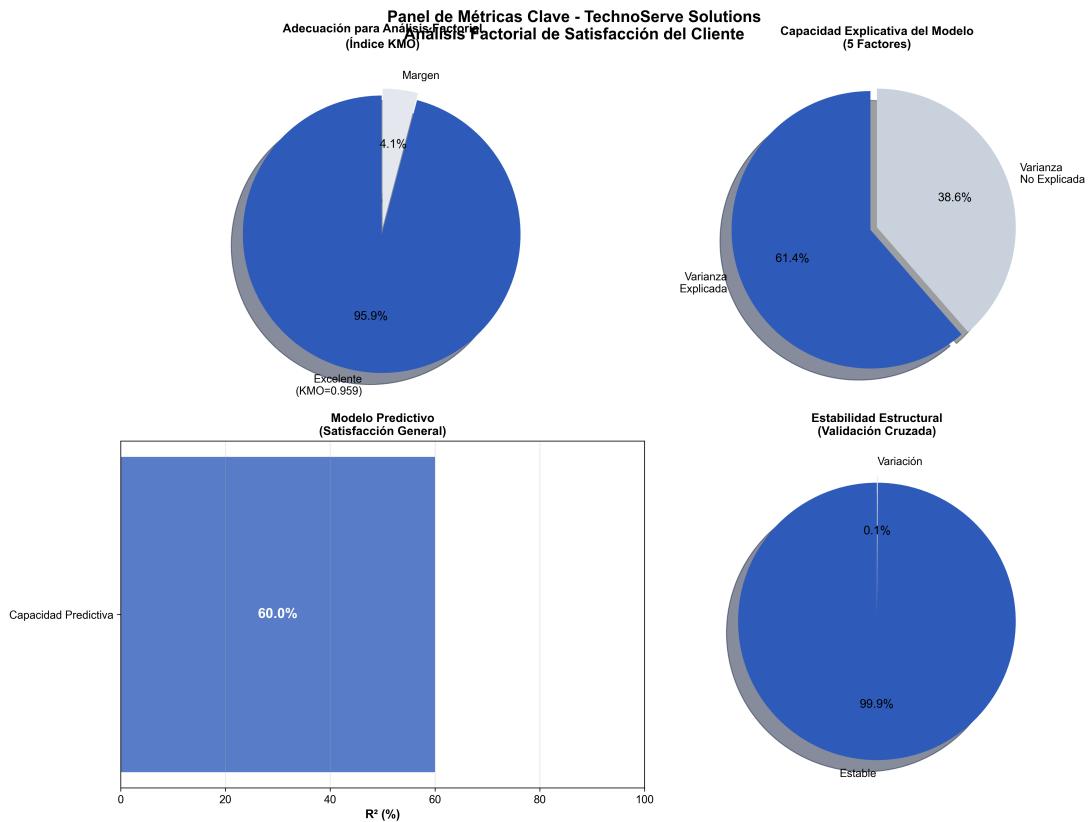


Figura 5: Panel de Métricas Ejecutivo. Resumen visual del rendimiento del modelo y la importancia de los factores.

La conclusión de negocio principal es que el rendimiento técnico y la ejecución operativa son los impulsores dominantes de la satisfacción del cliente y del valor financiero (ingresos y referencias).

Estos hallazgos tienen aplicaciones prácticas inmediatas, permitiendo la creación de dashboards de Business Intelligence y sistemas CRM basados en estas 5 dimensiones para una segmentación de clientes más sofisticada.

Basado en este análisis, se emiten las siguientes recomendaciones estratégicas:

- **Invertir en "Technical Excellence":** Priorizar la inversión en I+D, certificación de equipos y programas de innovación.
- **Reforzar "Project Delivery":** Optimizar los procesos de gestión de proyectos, asegurando el cumplimiento de plazos (**timeline_adherence**) y la calidad de los entregables (**quality_deliverables**).
- **Mantener la gestión de relaciones:** Relationship Management"(F2) es clave para la *Renewal Likelihood*, por lo que debe mantenerse un alto nivel de contacto y confianza.

6. Referencias

6.1. Software y Librerías

- **factor_analyzer** (Python Package): Utilizado para la ejecución del Análisis Factorial Exploratorio.
- **scikit-learn** (Python Package): Utilizado para el preprocesamiento de datos (escalado), la validación cruzada y la implementación de los modelos predictivos.
- **matplotlib & seaborn** (Python Packages): Utilizados para la generación de visualizaciones.