

Universidad Rafael Landívar
Facultad de Ingeniería
Análisis de Datos, sec. 01
Catedrático: Ing. Dan Stanly Bolaños



Manual de usuario

Andres Gálvez [1024718]
Alexander Villatoro [1182118]
Luis Chutá [1320016]
Sergio Lara [1044418]

Ciudad de Guatemala, 23 de noviembre de 2021

Contenido

Descripción de objetos y terminología	3
Base de datos	3
Dimensión	¡Error! Marcador no definido.
Hecho.....	¡Error! Marcador no definido.
Modelo estrella.....	¡Error! Marcador no definido.
Cubo	¡Error! Marcador no definido.
IDE	¡Error! Marcador no definido.
SQL.....	3
Visual Studio	¡Error! Marcador no definido.
SQL Server Management Studio.....	3
Power BI	4
Interactuar con los objetos	5
Desde SQL Server Management Studio.....	¡Error! Marcador no definido.
Desde Visual Studio	¡Error! Marcador no definido.
Desde Power BI	5

Descripción de objetos y terminología

Árbol de decisión

Se representa como un diagrama que determina un flujo con múltiples ramificaciones donde cada ramificación representa el camino de una condición y los nodos finales representan una etiqueta de clasificación y su probabilidad. Son empleados en múltiples campos para poder predecir probabilidades asociadas a uno o más eventos.

Base de datos

Conjunto de datos pertenecientes a un mismo contexto y almacenados sistemáticamente para su posterior uso.

Clasificador Bayesiano

Es un clasificador probabilístico que se fundamenta en el Teorema de Bayes, el cual evalúa la probabilidad de un evento dado otro. Es comúnmente conocido como Naive Bayes o clasificador bayesiano ingenuo, dado que asume ciertas hipótesis llamadas simplificaciones, ya que asume que la presencia o ausencia de una característica no está relacionada con la otra.

Curva ROC

Una curva Receiver Operating Characteristic es una representación gráfica que ilustra la relación entre la sensibilidad y la especificidad de un sistema clasificador para diferentes puntos.

Matriz de confusión

Es una herramienta de que permite la visualización del desempeño de un algoritmo que se emplea en aprendizaje supervisado. Contiene cuatro cuadrantes donde se detalla la cantidad de aciertos y desaciertos del modelo (falsos positivos / falsos negativos).

SQL

Structured Query Language, es un lenguaje de dominio específico, diseñado para administrar y recuperar información de sistemas de gestión de bases de datos relacionales.

SQL Server Management Studio

Es una aplicación de software que se utiliza para configurar y administrar todos los componentes dentro de SQL Server.

Power BI

Es un servicio de análisis de datos de Microsoft orientado a proporcionar visualizaciones interactivas y capacidades de inteligencia empresarial con una interfaz simple que permite la creación de informes y paneles.

Python

Es un lenguaje de programación interpretado cuya filosofía hace hincapié a la legibilidad del código. Es un lenguaje multiparadigma y multiplataforma.

R

R es un entorno y lenguaje de programación con un enfoque al análisis estadístico.

R Studio

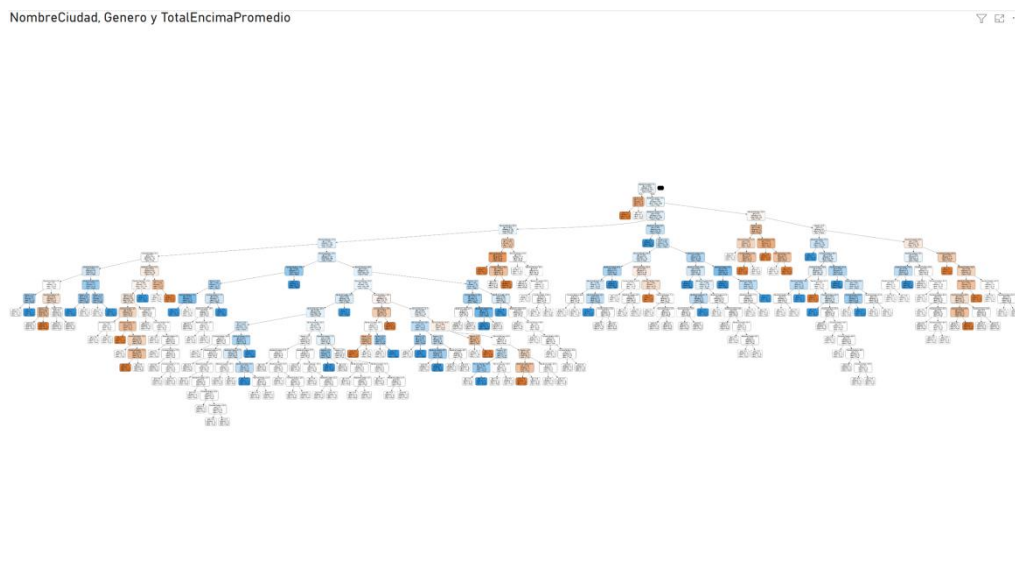
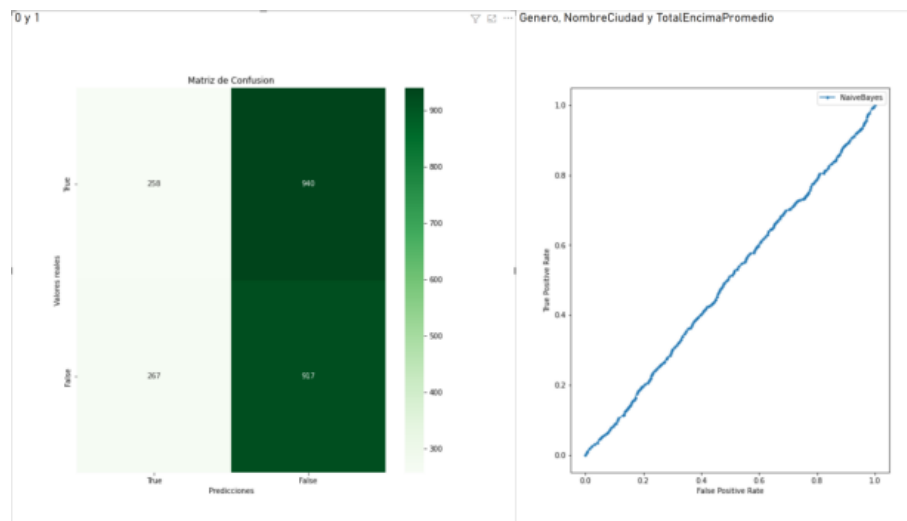
Es un entorno de desarrollo integrado para el lenguaje de programación R, dedicado a la computación estadística y gráficos.

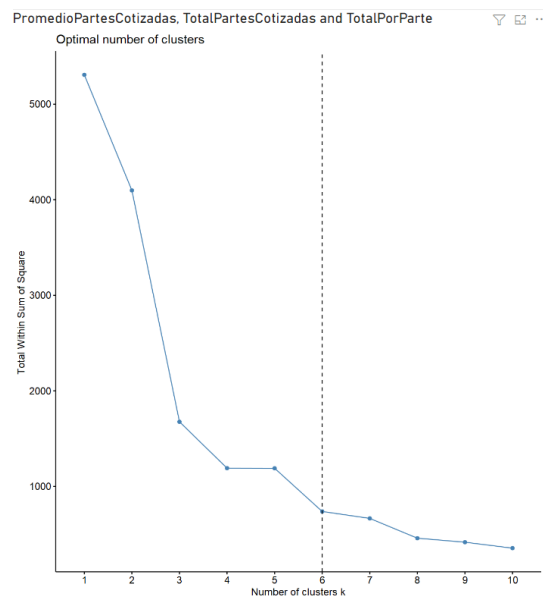
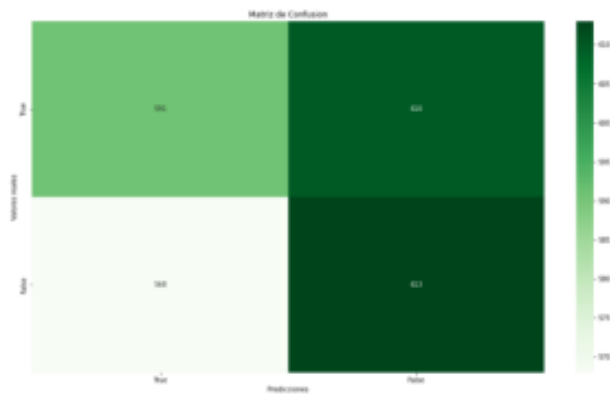
Interactuar con los objetos

Desde Power BI

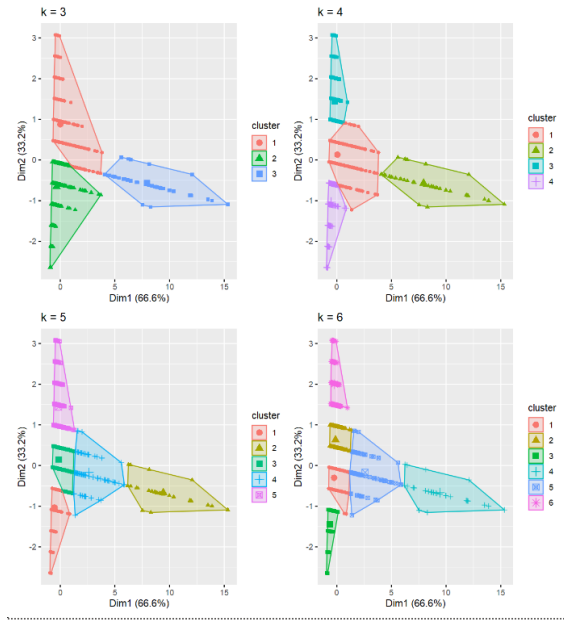
En esta herramienta se emplearon los scripts de R y Python para el manejo de los datos y la generación de gráficos.

Desde acá, es posible facilitar la visualización de distintos gráficos y métricas de manera amigable.

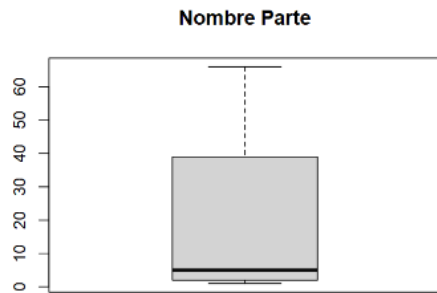




TotalPartesCotizadas, TotalPorParte and PromedioPartesCotizadas

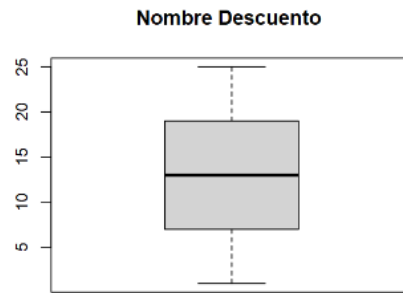


Anio, Cantidad, Marca, Modelo, NombreDescuento and NombreParte



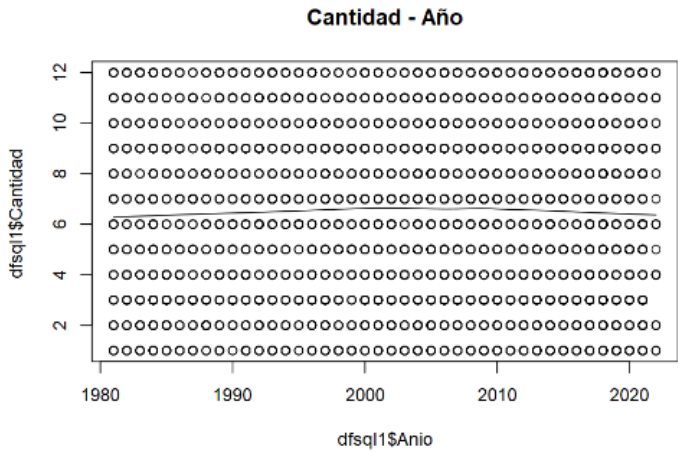
Outlier rows:

Anio, Cantidad, Marca, Modelo, NombreDescuento and NombreParte

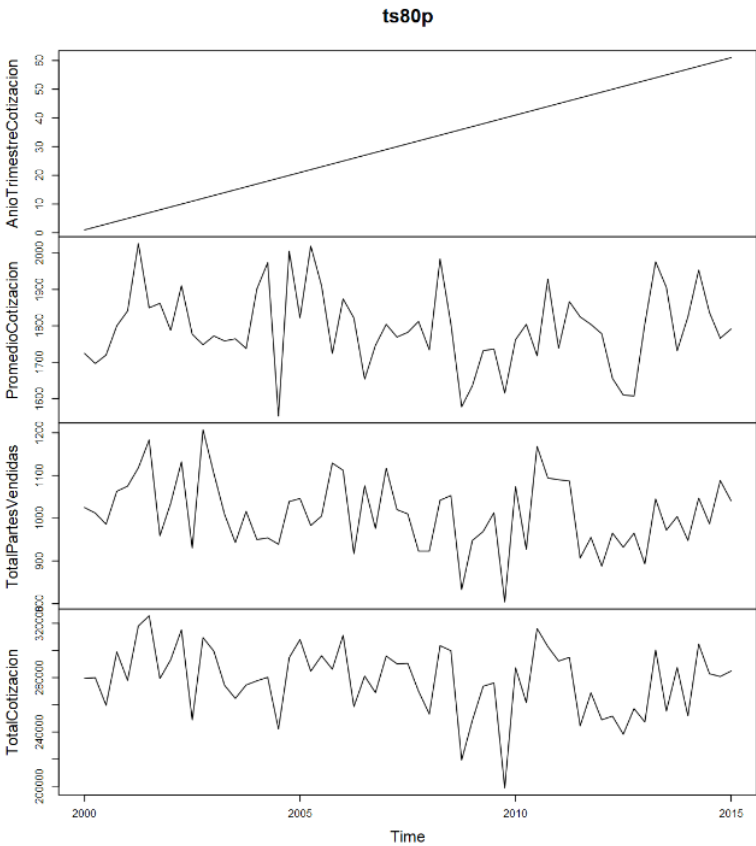


Outlier rows:

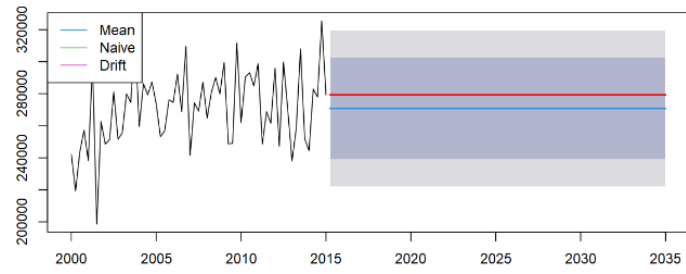
Anio, Cantidad, Modelo, Marca, NombreDescuento and NombreParte



AnioTrimestreCotizacion, PromedioCotizacion, TotalPartesVendidas an TotalCotizacion



PromedioCotizacion, AnioTrimestreCotizacion, TotalCotizacion and TotalPartesVendidas



AnioTrimestreCotizacion, PromedioCotizacion, TotalCotizacion and TotalPartesVendidas

