

Examen Final – Tablero Analítico Repuestos Web

Instrucciones

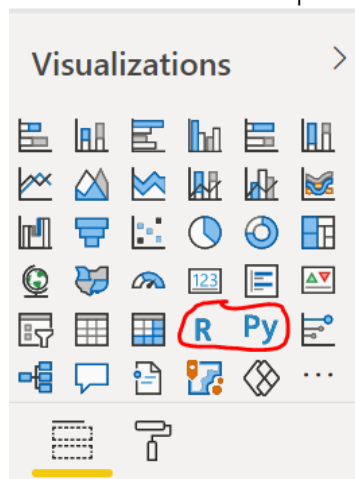
El siguiente es el enunciado del examen final de Análisis de Datos para la creación de un Tablero Analítico de Repuestos Web hecho en Power BI, el cual está basado en los modelos estadísticos creados en los lenguajes de R y Python vistos a lo largo del curso

El tablero deberá poseer los siguientes elementos basados en las hojas de trabajo realizadas en R y Python correspondientes a cada una de las técnicas de “Análisis Avanzado” vistas en clase:

- Pestaña Clustering → La pestaña de clustering mostrará a los usuarios como las diferentes “categorías de partes” se pueden clasificar en función los diferentes métricas de partes vendidas, cotizadas, etc..
- Pestaña de Clasificación → La pestaña de clasificación mostrara la eficiencia de los métodos de clasificación vistos en clase, en función de métricas tales como curva ROCR y valor del AUC, así como también la matriz de confusión, accuracy y plots
- Pestaña de Regresión → El motivo principal de la pestaña de regresión es mostrar plots y correlación de variables independientes vs. La variable dependiente, así como también el grado de “significancia” de cada una. Esto con el objetivo de mostrar cuales de las diferentes variables (región geográfica, categoría, vehiculo, etc) es la que mas impacta en métricas como “Total de Orden”...
- Pestaña de Series de Tiempo → La pestaña de series de tiempo mostrará pronósticos de estadísticas de Partes en función de por lo menos dos diferentes modelos de series de tiempo (los más precisos en función de RMSE y MAE) así como también las estadísticas de “accuracy” del mismo

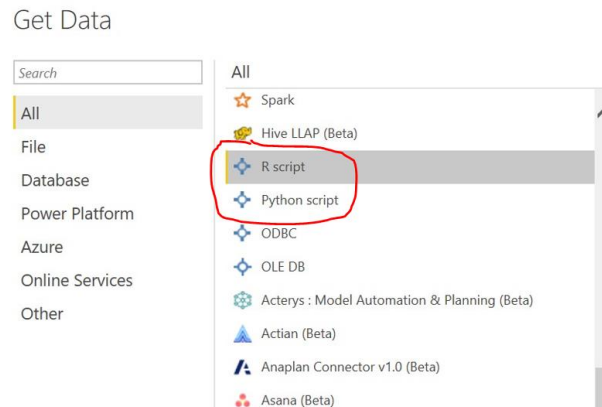
Todas y cada una de las visualizaciones de dichas pestañas debe de realizarse por medio de la ejecución de scripts de R y Python en Power BI, la cual se puede realizar por medio de las dos siguientes formas:

- La visualización “R script visual” ubicada en las visualizaciones:



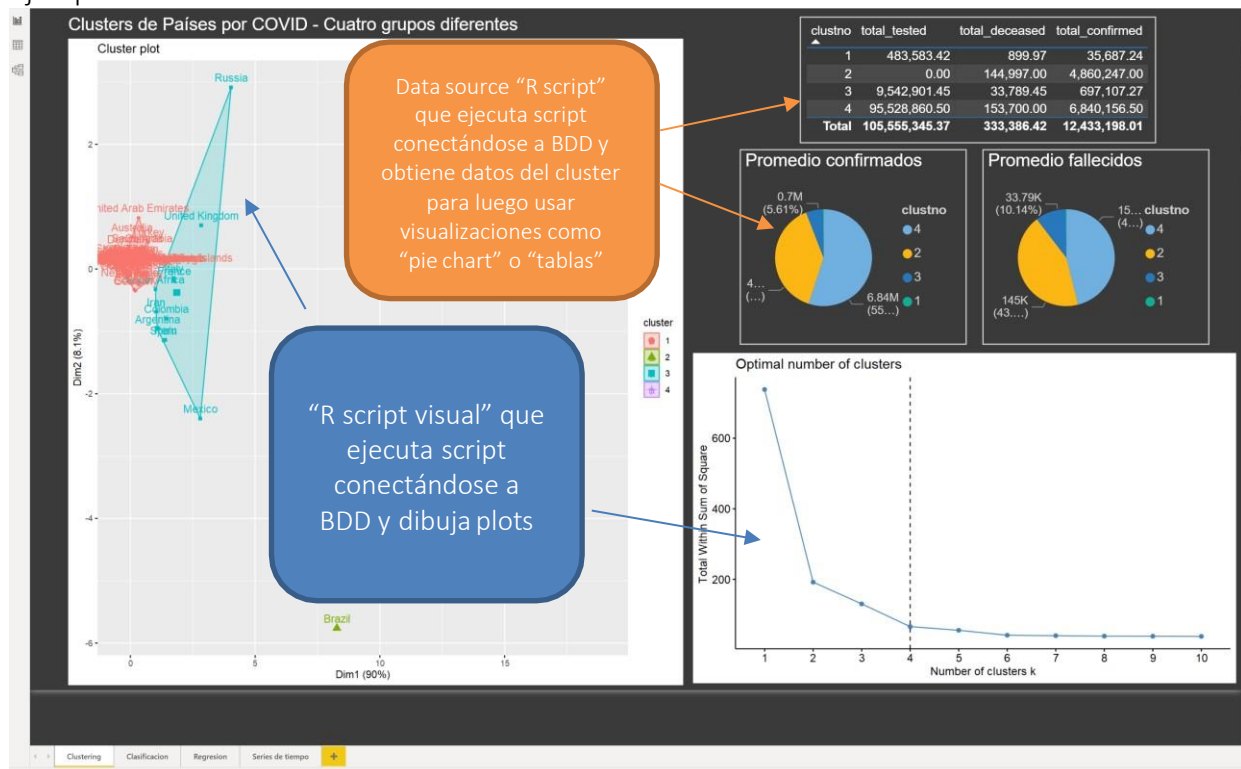
Esta visualización será la que utilizará para dibujar o llamar las funciones de “plot” o cualquier función equivalente que dibuja un plot (como por ejemplo la función “fviz_cluster” usada para graficar un cluster)

- Utilizando el “data source” de tipo “R script” y “Python script” que está en el listado de fuentes de datos disponibles:



Una vez se ejecuten scripts de R/Python para crear fuentes de datos o visualizaciones, utilice las mismas para crear “plots” que se ejecutan desde código u otras visualizaciones de Power BI que se conecten a la fuentes de datos de scripts, como se muestra en el siguiente ejemplo.

Ejemplo



Si bien es cierto el diseño de cada pestaña descrita anteriormente queda libre a discreción de que visualizaciones y plots quieren mostrarse, se le exhorta a realizar y utilizar visualizaciones atractivas y elegir los plots mas importantes en función de la interpretación que cada grupo dará a cada análisis hecho.

Sin embargo, se le recuerda lo siguiente a cada grupo:

- Recuerde que durante cada hoja de trabajo se realizaron plots y análisis que deben de estar presentes, sobre todo los más importantes
- Los scripts de R/Python deben de seguir las mismas instrucciones, como por ejemplo el separar la data en “train” (80% de los datos) y “test” (20% de los datos), ya que dichas instrucciones se validarán y tendrán ponderación.
- Para modelos tales como Clasificación o Series de tiempo en los cuales se vieron métricas para evaluar eficiencia y comprara modelos debe de incluir un resultado de dichas métricas en la pestaña (como por ejemplo mostrar MAE y RMSE en series de tiempo)
- Cada integrante de cada grupo deberá exponer una parte del tablero, por lo que deben de elegir bien que mostrar (explicarlo bien) y justificar por qué.

Entregables

Los entregables esperados son:

- Listado de reportes Power BI publicados en servidor
- Archivos de Power BI (.pbix o en su defecto .pbit)
- Todos y cada uno de los scripts de R/Python que fueron usados para generar visualizaciones odatos usados en cada pestaña.
- Scripts de base de datos, por ejemplo, si crearon una vista o tabla adjuntar dicho script
- Manual técnico – explicando detalles técnicos de como funcionan los diferentes modelos estadísticos de R/Pyton y como el tablero interactúa con dichos modelos por medio de cada script de R/Python
- Manual de usuario – explicando a un usuario no técnico como interactuar con los los reportes, pero también (y muy importante) explicando que modelos estadísticos son mas precisos o los hallazgos de cada modelo, por ejemplo:
 - En clustering descubrimos que los países “X” y “Y” se paracen en esto o aquello
 - En regresión descubrimos que lo que más impacta el “Total de ordenes” es el tipo de vehículo ya que modelos europeos tienden a cotizar partes mas costosas

Fecha de entrega

Miércoles 23 de Noviembre del 2021 a las 23:55 horas, para dicha fecha deberá de proveerlo siguiente:

- Subir al portal URL por medio del link llamado “Entregables Exámen Final” un archivo comprimido con:
 - Carpeta Archivos Power BI
 - Carpeta Scripts R/Python
 - Carpeta Scripts SQL
 - Manual técnico
 - Manual de usuario