

Reporte de trabajo en equipo

Organización y distribución de tareas

Para este primer taller iniciamos con la asignación de grupos sugerida en el curso:

1. Análisis de negocio + Tablero de datos. (Catalina Gómez)
2. Ingeniería de datos + Ciencia de datos. (María Viviana Lara)
3. Análisis de datos + Despliegue. (Juan Rios)

Cada uno desarrollo las tareas propuestas para cada rol. Se mantuvo comunicación constante, mediante slack, whatsapp y durante las sesiones en el salón de clases para realizar un producto que cumpliera con las necesidades y pudiera dar respuesta de manera eficiente a las preguntas propuestas en el análisis de negocio.

Descripción de tareas

Catalina Gómez

Análisis de negocio

Se seleccionó el área de operaciones como cliente principal debido a que la información contenida en la base de datos presenta una relación directa con sus actividades y necesidades. A partir de esto, se definió la demanda de bicicletas como la variable de respuesta, mientras que las demás variables fueron analizadas para formular preguntas clave que permitieran abordar los retos del área de operaciones. Este enfoque busca generar información relevante y soluciones que apoyen en la toma de decisiones relacionadas con la planificación de recursos, el manejo de personal, la logística de transporte de bicicletas y el mantenimiento de los equipos.

Tablero de datos

Se desarrollaron diversas visualizaciones para analizar cómo las variables climáticas y temporales influyen en la demanda de bicicletas compartidas. Además, se ajustaron varias configuraciones predeterminadas de Plotly, incluyendo tamaños, colores y estilos, con el fin de generar gráficos más claros y añadir herramientas interactivas. Finalmente, todas las visualizaciones se integraron en un único tablero diseñado para ser amigable y accesible para el usuario.

Viviana Lara

Ingeniería de datos

Después de encontrar y cargar los datos, se realiza una limpieza cuidadosa, identificando y dándole un manejo adecuado a datos faltantes o atípicos. De igual forma, se alistan los datos en un formato que permita su posterior análisis.

Ciencia de datos.

Con base en los resultados de la exploración y limpieza de datos, se desarrollan modelos de regresión lineal con StatsModels teniendo en cuenta primer lugar, todas las variables resultantes de la limpieza y luego con una menos. Adicionalmente, se considera el desarrollo de modelos de predicción como el Arima para predecir a partir de los datos históricos.

Juan Rios

Análisis de datos

Para el análisis de datos se buscó entender un poco las tendencias y correlaciones entre las variables, con el objetivo que estas tendencias nos dieran un indicio de como podriamos modelar el problema y utilizar los datos de una manera adecuada en las demas treas del proyecto. El análisis de datos y la ingenieria de datos fueron trabajos que se hicieron de la mano y con constante comunicación para que el trabajo de ciencia de datos integrara estos hallazgos y transformaciones.

Despliegue

Para el despliegue se utilizó una maquina identica de los talleres del curso, donde pudimos poner a prueba nuestros conocimientos de autenticar el ingreso a la máquina virtual, subir archivos externos y desplegar correctamente el tablero al público usando la IP pública, es importante tener en cuenta que el tablero funcionara desde que la maquina se encuentre activa y el proceso de python3 este ejecutando el tablero.