

Desafío: NYC Taxis Verdes

Este desafío tiene como objetivo evaluar tus habilidades e intuición sobre datos del mundo real. Utilizaremos datos recopilados por la Comisión de Taxis y Limusinas de la Ciudad de Nueva York sobre los "Taxis Verdes". Los Taxis Verdes (a diferencia de los amarillos) no tienen permitido recoger pasajeros dentro de las áreas densamente pobladas de Manhattan, descargados de la página web de registros de viajes de taxis y limusinas de NYC: http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml

Si haces referencia a materiales, bibliotecas o código encontrados en línea, recuerda citarlos adecuadamente. *Si tienes alguna pregunta o inquietud, no dudes en consultar.*

Q1:

- Informa cuántas filas y columnas tiene el archivo.

Q2:

- Grafica un histograma de la distancia del viaje ("Trip Distance").
- Informa cualquier estructura que encuentres y cualquier hipótesis que tengas sobre esa estructura.

Q3:

- Informa la distancia promedio y la mediana de los viajes agrupados por hora del día.
- Nos interesa identificar viajes que tengan origen o destino en uno de los aeropuertos del área de Nueva York. ¿Puedes proporcionar el número de transacciones que cumplen con este criterio, el costo promedio del viaje y cualquier otra característica interesante de estos viajes?

Q4:

- Crea una variable binaria que indique si el % de la propina sobre el total pagado ($\text{tip_amount} / \text{total_amount}$) es menor al 5%.
- Tienes que ayudar con tips a un taxista que pretende recibir no menos del 5% de propina. ¿Podrías darle algunas recomendaciones?

El resultado del análisis se puede preparar en el formato que mejor te sientas: notebook, excel o cualquier herramienta de visualización de datos.

A medida que avances en el desarrollo de las preguntas, vas a tener más oportunidades para demostrar tus conocimientos. Por favor, incluí todos tus pensamientos, suposiciones y conclusiones para que podamos conocer más sobre tus habilidades.

¡Buena suerte!