****

**Analisis de ingreso de mercado al transporte de pasajeros en base a vehículos pequeños**

**Analisis EDA**

Glosario:

* Stack Tecnologico
* Datasets NYC-TLC
* Dataset Air Quality
* Dataset Modelos aprovados
* Dataset Economia Autos
* Dataset Zones
* Dataset Puente Color
* Dataset Puente Estado
* Dataset Puente Ciudad

**Stack Tecnologico**

Para la realización de este análisis se usaron las bibliotecas:

1. Pandas
2. Matplotlib
3. Seaborn
4. Visual Studio Code

**Datasets NYC-TLC**

Se hizo un análisis exhaustivo de los datasets, en primer lugar se usaron los siguientes archivos para generar un dataset concatenado

nyc\_green\_tripdata(01 al 12 y de 2018 a 2023).

Ej: nyc\_green\_tripdata\_01\_2018.parquet

Resultando en los siguientes concatenados

1. nyc\_yellow\_2018.parquet
2. nyc\_green\_2018.parquet
3. nyc\_yellow\_2019.parquet
4. nyc\_green\_2019.parquet
5. nyc\_yellow\_2020.parquet
6. nyc\_green\_2020.parquet
7. nyc\_yellow\_2021.parquet
8. nyc\_green\_2021.parquet
9. nyc\_yellow\_2022.parquet
10. nyc\_green\_2022.parquet
11. nyc\_yellow\_2023.parquet
12. nyc\_green\_2023.parquet

Se hicieron de esta manera para poder optimizar recursos y para hacer análisis más rápidamente ya que solo se llama a un archivo.

**Análisis de nulos**

Se detectaron en todos los datasets de colores verdes las mismas columnas con datos nulos,

* store\_and\_fwd\_flag
* RatecodeID
* passenger\_count
* Payment\_type
* Trip\_type
* congestion\_surcharge

Por el lado de los datasets de colores Amarillos las mismas columnas con datos nulos,

* store\_and\_fwd\_flag
* RatecodeID
* passenger\_count
* Payment\_type
* Trip\_type
* congestion\_surcharge

Estos nulos se deben en gran parte a la falta o la mala carga de los viajes por parte del taxista o por problemas de conexión con la base central.

Un caso especial es el de los taxis verdes que el congestion\_surcharge Generalmente se debe a que ingresa poco a la zona de congestión Ya que solo puede dejar pasajeros y no recogerlos en la calle, solo si es pedido por teléfono con anterioridad al viaje, rige la misma restricción en los aeropuertos del Tri state (Manhattan, Queens y Brooklyn).

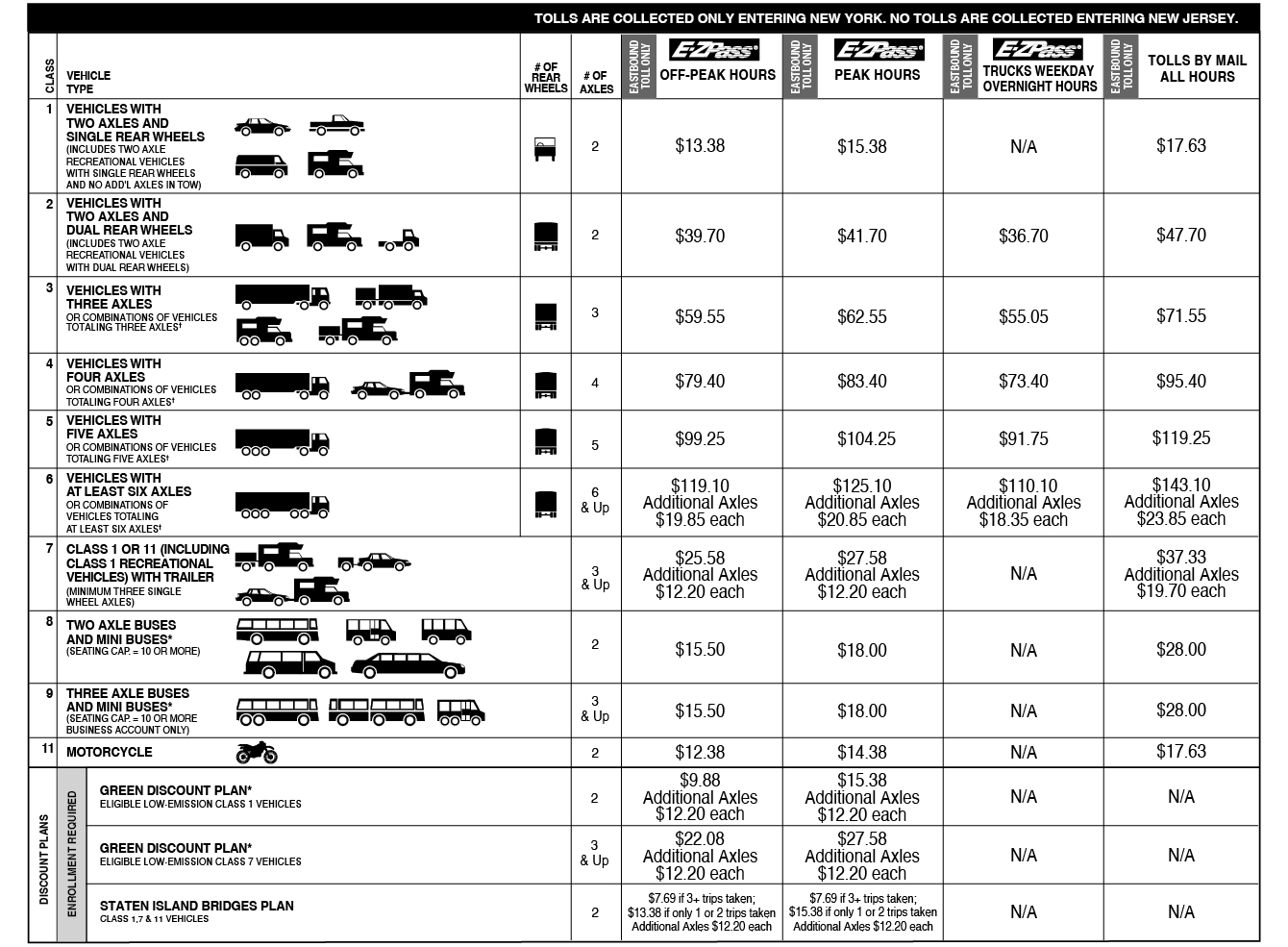
La zona de congestión es la siguiente



Esta zona, tiene puntos de cobro, en los ingresos a Manhattan, desde la calle 60 (borde de central park) hasta la punta del Downtown, los ingresos son:

* Queensboro Bridge
* Manhattan Bridge
* Brooklyn Bridge.
* Queens Tunnel.
* Abraham Lincoln Tunnel
* Holland Tunnel.

A estos cargos se le suma el peaje que tiene estas escalas



Los peajes solo aplican a los túneles, no así a los puentes de acceso a la ciudad dentro de la zona de congestión, hay otros puentes con peaje pero no se toman en consideración de momento, no se cobra peaje en sentido a New Jersey.

A estos valores se le suman los cargos de congestión, en caso de vehículos de transporte (Taxis, Uber, Lyft), se les cobra ambos, ya que vehículos particulares tienen un descuento de 5 dólares si paga el peaje Ej:

Un vehículo particular entra en la zona de congestión viniendo por uno de los túneles estaría pagando $13.38 - $5 + $ 15 = $23.38

Un vehículo particular hibrido o electrico que entra en la zona de congestión viniendo por uno de los túneles estaría pagando $9.88 - $5 + $ 15 = $19,88

Esto nos muestra la ventaja del uso de vehículos ecológicos por sobre vehículos de combustible tradicional

Tomando estos datos en cuenta, nos sirve para poder tomar en cuenta la propuesta del cliente de querer insertarse al mercado de transporte, estos solo son uno de los tantos impuestos que se deben pagar para el ingreso.

Datasets trip\_data 2018 a 2023