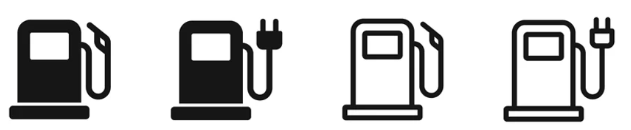
**Análisis preliminar y calidad de los datos**

Bajo la misma línea de pensamiento de Urban Green Solutions, sin perder de vista los objetivos planteados y con la finalidad de obtener un panorama real del estado y los retos a los que nos enfrentamos, se realiza un análisis preliminar de los datos:

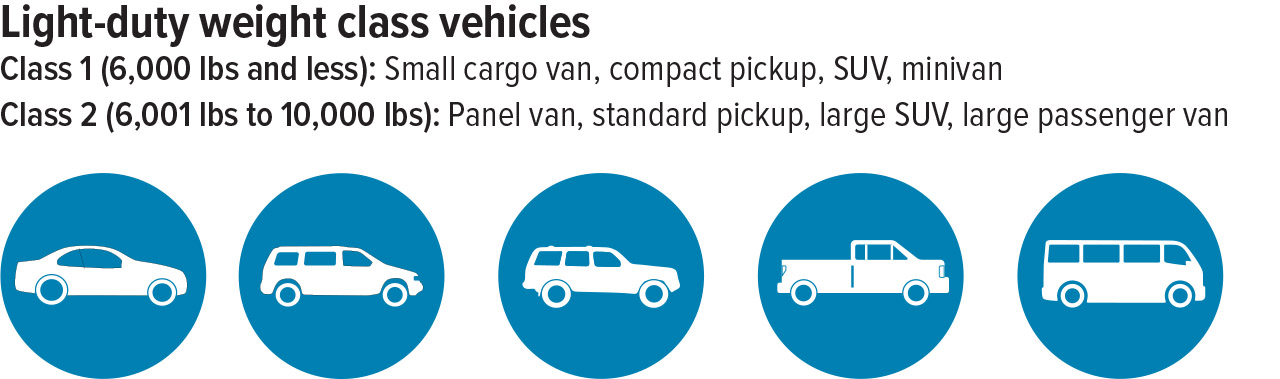
Para ello se extraen de las Agencias gubernamentales NYC Taxi and Limousine Comisión y Environment & Health de New York datos reales y actuales útiles para este proyecto a trabajar

1. **Electric and Alternative Fuel Charging Stations:** Este dataset contiene información sobre las estaciones de carga de combustibles alternativos para vehículos (electricidad, Gas Natural, Gas Licuado, Gas de biodigestores, Etanol e hidrógeno) en EE.UU. después de analizar los datos se obtuvieron los siguientes insights:
   1. Existen gran cantidad de nulos por columnas especificas de cada tipo de combustible, sin existir problema filtrando data.
   2. Casi 60k estaciones son destinadas para todo publico
   3. Mas de 45k estaciones abren las 24 horas
   4. Mas de 10k estaciones de carga ofrecen carga de vehículos eléctricos de forma gratuita y otras más solo cobran tiempo de estacionamiento
   5. En NY existe un total de 3409 estaciones de carga para vehículos eléctricos y mas de 600 ofrecen servicio gratuito. 17%



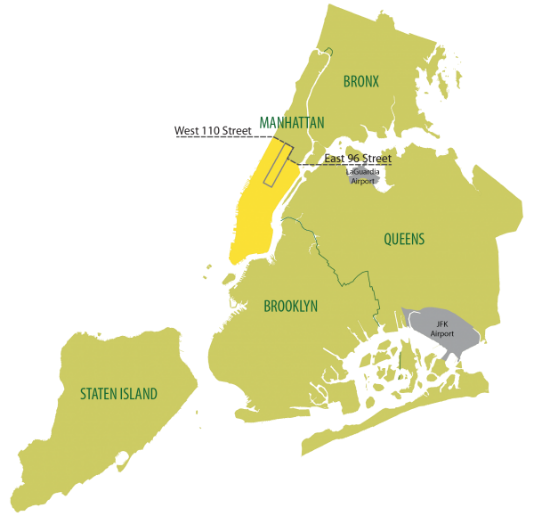
1. **Vehículos ligeros:** este dataset contiene información adicional de 1993 a 2022 sobre vehículos ligeros (autos a camionetas hasta 2.7 ton) en él se almacenan datos referentes a tipo de combustible) eficiencia de combustible alternativo (ciudad, carretera y combinado)
   1. La fábrica predominante es Ford
   2. La categoría de estos vehículos predominante es sedan/wagon y le sigue SUV
   3. Casi 15% son vehículos eléctricos.



1. **Electric Car Data:** Este set de datos contiene la información de los autos eléctricos disponibles en el mercado de EE.UU
   1. No existen datos nulos
   2. Existen información de marca, modelo, eficiencia, autonomía, Precio, entre otras
   3. Existen outliers en eficiencia y precio



1. **Taxi-Zones**. Este dataset contiene datos de las zonas de Taxis divididas en
   1. Las ciudades de New York
   2. Áreas o espacios de zonas de taxis
   3. Longitud de recorridos
   4. Zonas de aeropuertos



1. **Vehicle fuel Economy Data:** Este dataset contiene información acerca de la economía y emisiones de CO2 del combustible tradicional y alternativo en un periodo comprendido de 1984 a 2023. La relevancia de los datos es:
   1. 120 gr/milla de emisiones en promedio por auto de combustible tradicional contra 0 gr/milla de vehículos electricos
   2. Casi $3k dólares contra $15.36 el promedio de comparación de costos de combustibles
2. **AirQuality:** Calidad del aire: Dentro de este dataset se encuentran valores que miden la calidad del aire de acuerdo a una tabla con la cual se compara la concentración de contaminantes, medidos en **µg/m3,** así como registros de temporalidad por hora, día, mes y de requerir por año. Junto a este set de datos se analizan datos que contiene información sobre las ubicaciones de las mediciones:
   1. **No nulos**
   2. Valor máximo: 265.21
   3. Valor mínimo: 2.19
   4. Promedio: 9.88
   5. Asimismo, se pudo observar que los horarios del día en que existe menor calidad de aire son de 12 a 17 horas

**En conclusión**

1. Confiabilidad: Alta por venir de organismos públicos
2. Problemas detectados:
   1. Gran cantidad de datos tiene nulos, solo en algunos datasets por ser datos de vehículos específicos
   2. Buena variabilidad: