VEHICLE CO2 EMISSIONS
Data Analytics – Python

Background

You volunteer for a public policy advocacy organization in Canada, and your colleague asked you to help her draft recommendations for guidelines on CO2 emissions rules.

After researching emissions data for a wide range of Canadian vehicles, she would like you to investigate which vehicles produce lower emissions.

Data Validation

```python
# Import the pandas and numpy packages
import pandas as pd
import numpy as np

# Load the data
cars = pd.read_csv('data/co2_emissions_canada.csv')

# create numpy arrays
cars_makes = cars['Make'].to_numpy()
cars_models = cars['Model'].to_numpy()
cars_classes = cars['Vehicle Class'].to_numpy()
cars_engine_sizes = cars['Engine Size(L)'].to_numpy()
cars_cylinders = cars['Cylinders'].to_numpy()
cars_transmissions = cars['Transmission'].to_numpy()
cars_fuel_types = cars['Fuel Type'].to_numpy()
cars_fuel_consumption = cars['Fuel Consumption Comb (L/100 km)'].to_numpy()
cars_co2_emissions = cars['CO2 Emissions(g/km)'].to_numpy()

# Preview the dataframe
cars
```

	Make ⌄	Model ⌄	Vehicle Class ⌄	Engine Size(L) ⌄	Cylinders ⌄	Transmission ⌄	Fuel Type ⌄	Fuel Consumption
0	ACURA	ILX	COMPACT	2	4	AS5	Z	
1	ACURA	ILX	COMPACT	2.4	4	M6	Z	
2	ACURA	ILX HYBRID	COMPACT	1.5	4	AV7	Z	
3	ACURA	MDX 4WD	SUV - SMALL	3.5	6	AS6	Z	
4	ACURA	RDX AWD	SUV - SMALL	3.5	6	AS6	Z	
5	ACURA	RLX	MID-SIZE	3.5	6	AS6	Z	
6	ACURA	TL	MID-SIZE	3.5	6	AS6	Z	
7	ACURA	TL AWD	MID-SIZE	3.7	6	AS6	Z	
8	ACURA	TL AWD	MID-SIZE	3.7	6	M6	Z	
9	ACURA	TSX	COMPACT	2.4	4	AS5	Z	

Table Chart

« ‹ 1 of 223 › » Rows per page 10 ⌄ 7385 rows (truncated to 2222) ↓

```
# Look at the first ten items in the CO2 emissions array
cars_co2_emissions[:10]
```
[3] ●

```
array([196, 221, 136, 255, 244, 230, 232, 255, 267, 212])
```

1. What is the median engine size in liters?
The median engine size is 3.0 liters.

```
print(cars['Engine Size(L)'].median())
```
[4] ●

```
3.0
```

2. What is the average fuel consumption for regular gasoline (Fuel Type = X), premium gasoline (Z), ethanol (E), and diesel (D)?

Regular gasoline (X) = 10.084575
Premium gasoline (Z) = 11.422767
Ethanol (E) = 16.861351
Diesel (D) = 8.835429

```
fuel_consumption = cars.groupby("Fuel Type")["Fuel Consumption Comb (L/100 km)"].mean()
print(fuel_consumption)
```
[5] ●

```
Fuel Type
D      8.835429
E     16.861351
N     12.700000
X     10.084575
Z     11.422767
Name: Fuel Consumption Comb (L/100 km), dtype: float64
```

4. Which vehicle class has lower average CO2 emissions, 'SUV - SMALL' or 'MID-SIZE'?
Mid Size

```
CO2 = cars.groupby("Vehicle Class")["CO2 Emissions(g/km)"].mean()    [6] ●
print(CO2[["MID-SIZE","SUV - SMALL"]])

Vehicle Class
MID-SIZE        222.455428
SUV - SMALL     236.292523
Name: CO2 Emissions(g/km), dtype: float64
```

5. What are the average CO2 emissions for all vehicles? For vehicles with an engine size of 2.0 liters or smaller?

All vehicles: 250.58469871360867
Vehicles with <= 2.0 liter engines: 198.26783530370975

```
all = cars["CO2 Emissions(g/km)"].mean()                             [7] ●
two_liters = cars[cars["Engine Size(L)"] <= 2.0]["CO2 Emissions(g/km)"].mean()
print(str("All vehicles: ") + str(all))
print(str("Vehicles with <= 2.0 liter engines: ") + str(two_liters))

All vehicles: 250.58469871360867
Vehicles with <= 2.0 liter engines: 198.26783530370975
```

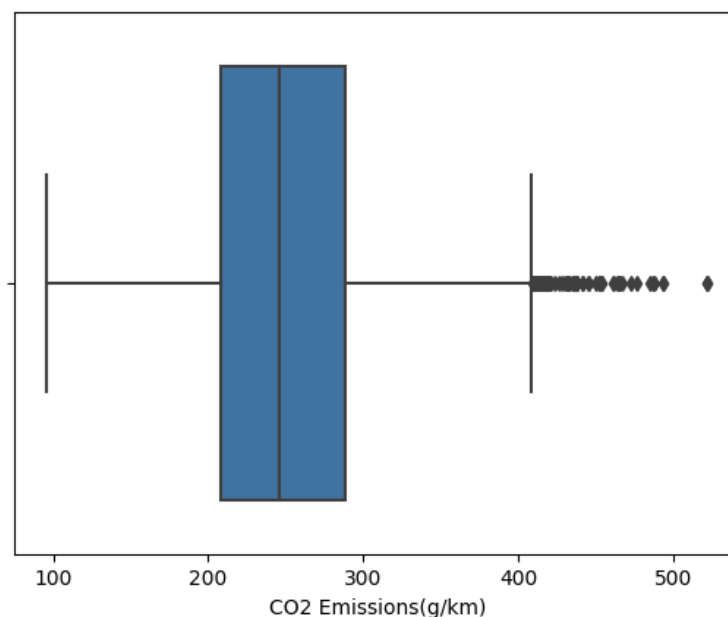**6. Any other insights you found during your analysis?**
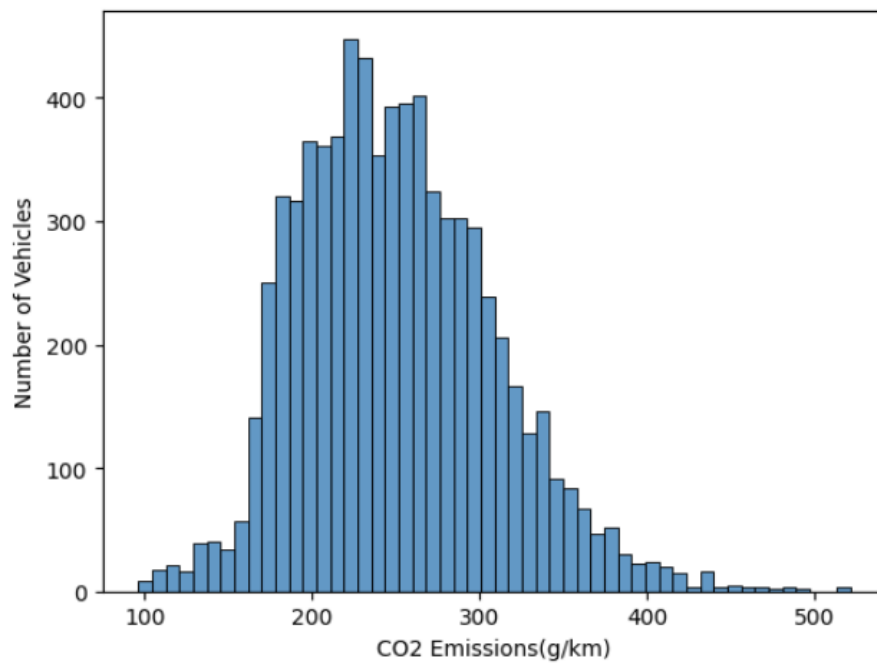
Most vehicles emit 200-300 g/km CO2.

```
import matplotlib.pyplot as plt                                      [8] ●
import seaborn as sns
sns.boxplot(data=cars, x="CO2 Emissions(g/km)")
plt.show()

sns.histplot(data=cars, x="CO2 Emissions(g/km)")
plt.ylabel("Number of Vehicles")
plt.show()
```

As one might expect, the level of CO2 emissions increases with the number of engine cylinders. However, all fuel types emit similar levels of CO2. Overall, natural gas is the cleanest, and ethanol is the worst, with premium gasoline a close second followed by diesel and regular gasoline in that order.

```
sns.lineplot(data=cars, x="Cylinders", y="CO2 Emissions(g/km)")
plt.title("CO2 Emissions by Cylinder Size")
plt.xlabel("Cylinder Size")
plt.show()
```
[9] ●
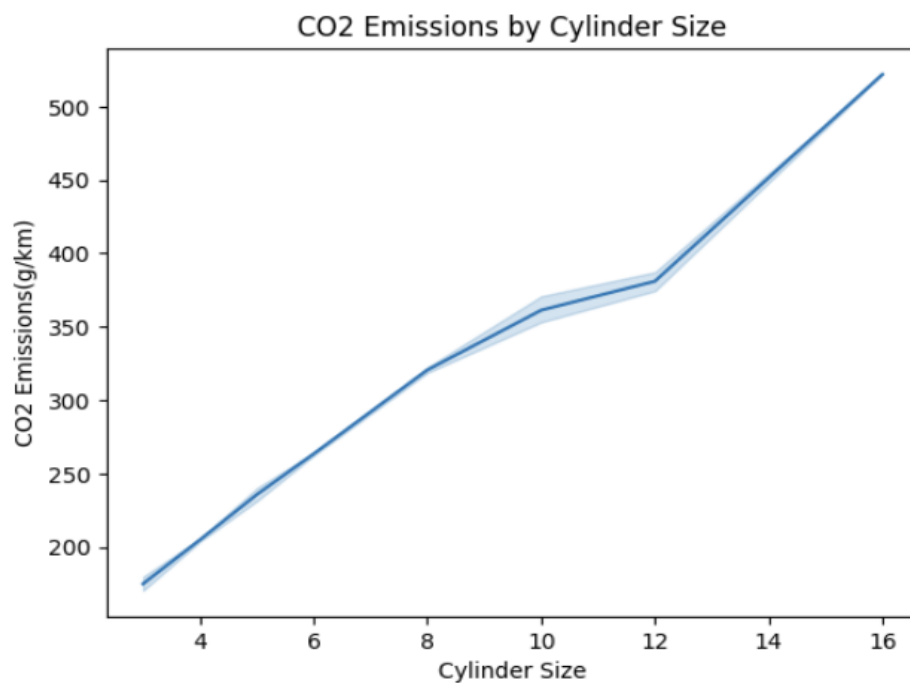

```
labels=["Z = Premium gasoline", "D = Diesel", "X = Regular gasoline", "E = Ethanol (E85)", [10] ●
natural gas"]
sns.barplot(data=cars, x="Fuel Type", y="CO2 Emissions(g/km)")
plt.title("CO2 Emissions By Fuel Type")
plt.legend(labels)
plt.show()
```

CO2 Emissions By Fuel Type

Legend:
- Z = Premium gasoline
- D = Diesel
- X = Regular gasoline
- E = Ethanol (E85)
- N = natural gas
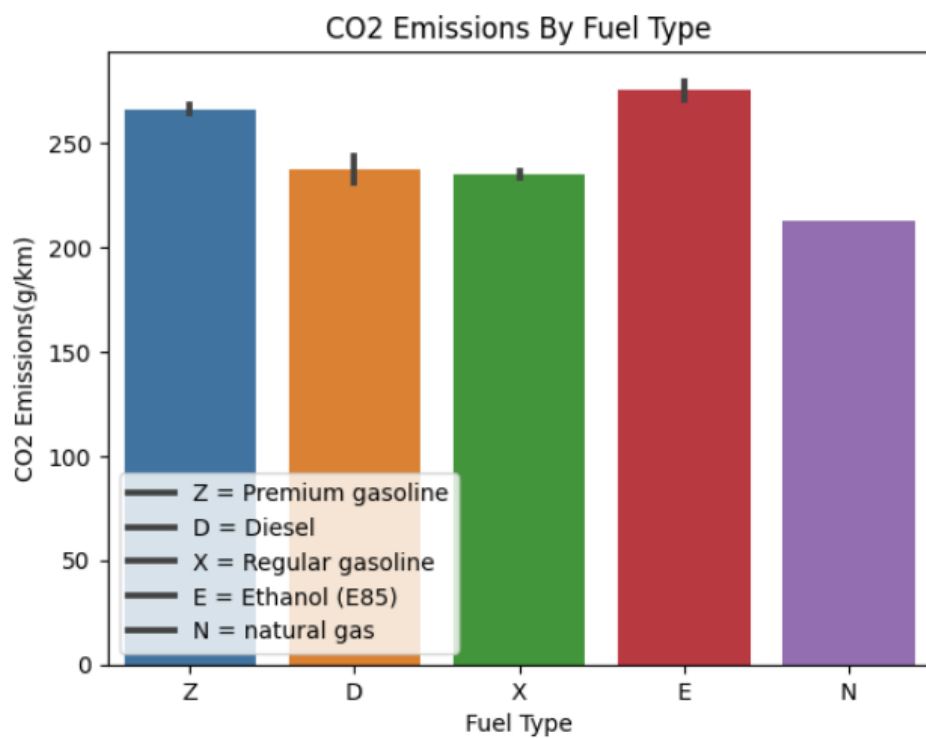
On average, the vehicle class "small station wagon" emits the least CO2 overall, with "passeneger vans" emitting the most.

Overall, the car makes which emit the most CO2 are all luxury, sports cars: Bugatti, Lamborghini, Rolls Royce, SRT, Aston Martin and Bentley. Smart emit the least CO2.

```
sns.catplot(data=cars, kind="box", x ="Vehicle Class", y="CO2 Emissions(g/km)", height=8.2[11] ●
aspect=11.7/8.27)
plt.xticks(rotation=90)
plt.title("CO2 Emissions by Vehicle Class")

plt.show()
```


CO2 Emissions by Vehicle Class