

```
reading.py data_analysis.py long_rivers.py vw_cars_data_analysis.py X vw_extension.py
session_4 > vw_cars_data_analysis.py > ...
1 import csv
2 from collections import namedtuple
3
4 VW_Car = namedtuple("Car", "model year price transmission mileage fuelType tax mpg engineSize")
5
6 vw_cars = []
7
8 with open("vw.csv", "r", encoding="utf-8") as csvfile:
9     print(csvfile)
10
11     reader = csv.reader(csvfile, skipinitialspace=True)
12     next(reader)
13     for row in reader:
14         new_car = VW_Car(*row)
15         vw_cars.append(new_car)
16 print(vw_cars)
17
18 # What is the most expensive vw car listed?
19
20 most_expensive_vw_car = []
21 expensive_vw_car = vw_cars[0]
22
23 for car in vw_cars:
24     if int(car.price) > int(expensive_vw_car.price):
25         expensive_vw_car = car
26 most_expensive_vw_car.append(expensive_vw_car)
27 print("\nTask 1:")
28 print("The most expensive vw car:")
29 print(most_expensive_vw_car)
30
Ln 57, Col 16 Spaces: 4 UTF-8 CRLF Python 3.11.5 64-bit (microsoft sto
```

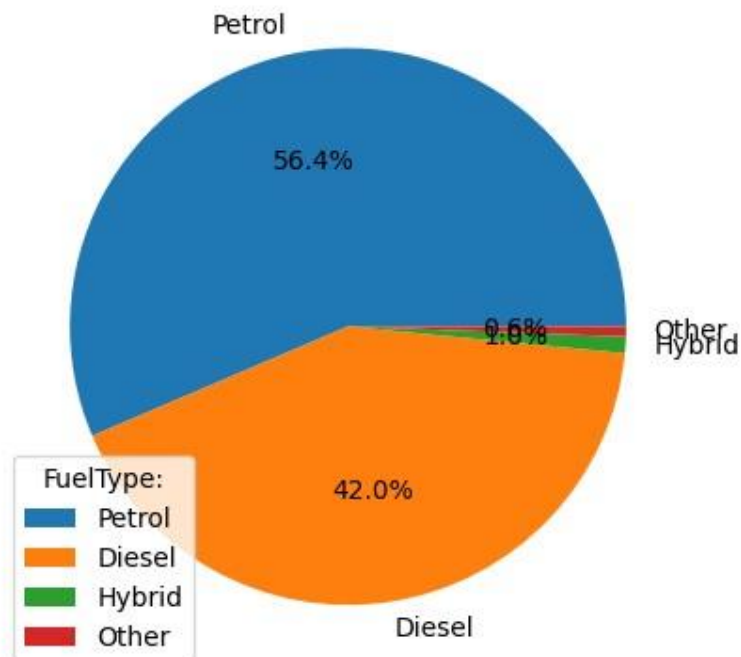
```
reading.py data_analysis.py long_rivers.py vw_cars_data_analysis.py X vw_extension.py
session_4 > vw_cars_data_analysis.py > ...
31 # Find all the VW Golf models. What is their average price?
32 sum_of_vw_golf_models = 0
33 number_of_vw_cars = 0
34
35 for car in vw_cars:
36     if car.model == "Golf":
37         # check the sum total and number of VW Golf cars
38         sum_of_vw_golf_models += int(car.price)
39         number_of_vw_cars += 1
40 # Calculate the average price of the Golf cars
41 avg_price_of_vw_golf = sum_of_vw_golf_models / number_of_vw_cars
42 print("\nTask 2:")
43 print(f"The average price of VW Golf cars is {round(avg_price_of_vw_golf)}")
44
45 # What is the average mileage for VW Polo models registered in 2020?
46
47 sum_of_2020_Polo_models = 0
48 number_of_2020_Polo_models = 0
49
50 for car in vw_cars:
51     if car.model == "Polo" and int(car.year) == 2020:
52         sum_of_2020_Polo_models += int(car.mileage)
53         number_of_2020_Polo_models += 1
54 print(sum_of_2020_Polo_models)
55 print(number_of_2020_Polo_models)
56 avg_mileage = sum_of_2020_Polo_models / number_of_2020_Polo_models
57 print("\nTask 3:")
58 print(avg_mileage)
59 print(f"The average mileage of VW Polo models registered in 2020 is {round(avg_mileage)}")
60
Ln 57, Col 16 Spaces: 4 UTF-8 CRLF Python 3.11.5 64-bit (n
```

```
reading.py  data_analysis.py  long_rivers.py  vw_cars_data_analysis.py  vw_extension.py
session_4 > vw_cars_data_analysis.py > ...
57 print("\nTask 3:")
58 print(avg_mileage)
59 print(f"The average mileage of VW Polo models registered in 2020 is {round(avg_mileage)}")
60
61 used_vw_cars = []
62
63 for car in vw_cars:
64     cars = VW_Car(
65         car.model,
66         car.year,
67         car.price,
68         car.transmission,
69         car.mileage,
70         car.fuelType,
71         car.tax,
72         car.mpg,
73         car.engineSize
74     )
75     used_vw_cars.append(cars)
76 print(used_vw_cars)
77
78 with open("amend_vw.csv", "w") as new_csvfile:
79     writer = csv.writer(new_csvfile, quoting=csv.QUOTE_ALL)
80     writer.writerow(VW_Car\
81         ("Model", "Year", "Price", "Transmission", "Mileage", "FuelType", "Tax", "Mpg", "EngineSize"))
82     for car in used_vw_cars:
83         writer.writerow(car)
84
```

```
reading.py  data_analysis.py  long_rivers.py  vw_cars_data_analysis.py  vw_extension.py
session_4 > vw_extension.py > ...
1  import pandas as pd
2  import matplotlib.pyplot as plt
3
4  # 1. A pie chart showing the distribution between fuel types
5  used_vw_cars = pd.read_csv("amend_vw.csv")
6  #used_vw_cars = used_vw_cars.drop_duplicates(subset="Model", keep="last")
7
8  num_of_vw_cars_by_FuelType = used_vw_cars.groupby('FuelType')\
9  [['Model']].count().sort_values("Model", ascending=False).reset_index()
10 print("\nExtension 1:")
11 print("The FuelType are:")
12 print(num_of_vw_cars_by_FuelType)
13
14
15 plt.pie(
16     num_of_vw_cars_by_FuelType.Model,
17     labels=num_of_vw_cars_by_FuelType.FuelType,
18     autopct="%1.1f%%")
19
20 plt.title("VW Cars Fuel Distrition")
21 plt.legend(title='FuelType:')
22 plt.show()
23
24 # 2. A bar chart showing the average mileage for each model.
25
26 model_type = used_vw_cars.groupby('Model')\
27     [['Mileage']].mean().sort_values\
28     ("Mileage", ascending=False).head(10).reset_index()
29
```

Ln 31, Col 30 Spaces: 4 UTF-8 CRLF Python 3.11.5 64-bit

VW Cars Fuel Distrition



```

session_4 > vw_extension.py > ...
21 plt.legend(title='FuelType:')
22 plt.show()
23
24 # 2. A bar chart showing the average mileage for each model.
25
26 model_type = used_vw_cars.groupby('Model')\
27     [['Mileage']].mean().sort_values\
28     ("Mileage", ascending=False).head(10).reset_index()
29
30 print("\nExtension 2:")
31 print("The Model Types are:")
32 print(round(model_type))
33
34 plt.bar(
35     model_type.Model,
36     model_type.Mileage,
37     color = "Orange",
38     width=0.5
39 )
40
41 plt.xlabel("VW Model")
42 plt.ylabel("Average Mileage")
43 plt.title("Distribution of Average Model against VW Model")
44 plt.xticks(rotation=45) # Rotate x-axis
45 plt.tight_layout()
46 plt.show()

```

