

CAR SALES ANALYSIS

Dataset:

auto_purchases.csv: auto purchase data

- purchaser_id: hashed identifier from auto purchase dataset
- transactiontype: C=cash, D=loan, L=lease
- newusedflag: N=new, U=used
- modelyear: model year of vehicle purchased
- vehprice: price of vehicle purchased
- segment: type of vehicle purchased

onboarder_matches.csv: purchaser_id to device_id matches

- process_date: date when onboarder sent that match
- purchaser_id: hashed identifier from auto purchase dataset
- device_id: hashed identifier from third party dataset

attribute_data.csv: third party attributes data

- display_code: numeric code for third party attribute
- device_id: hashed identifier from third party dataset

attribute_lookup.csv: mapping of third party attribute code to attribute name

- display_code: numeric code for third party attribute
- attribute_name: name of third party attribute

Technologies used:

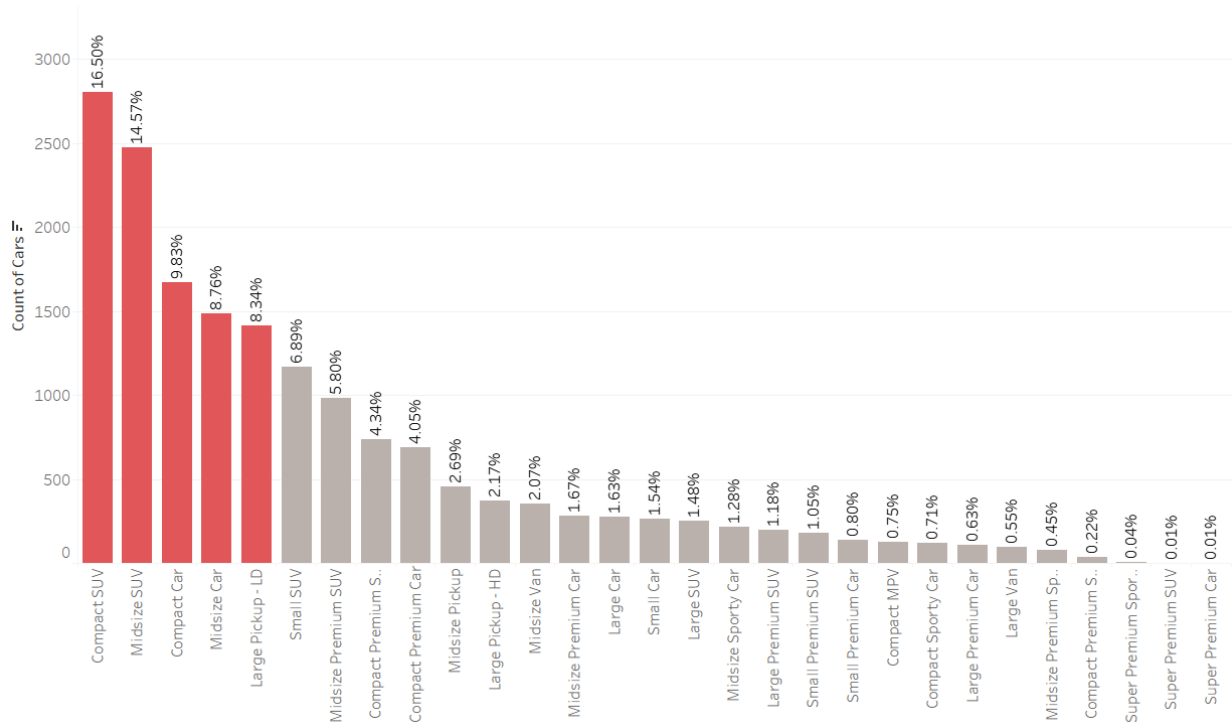
SQL, Tableau, Python

Libraries used:

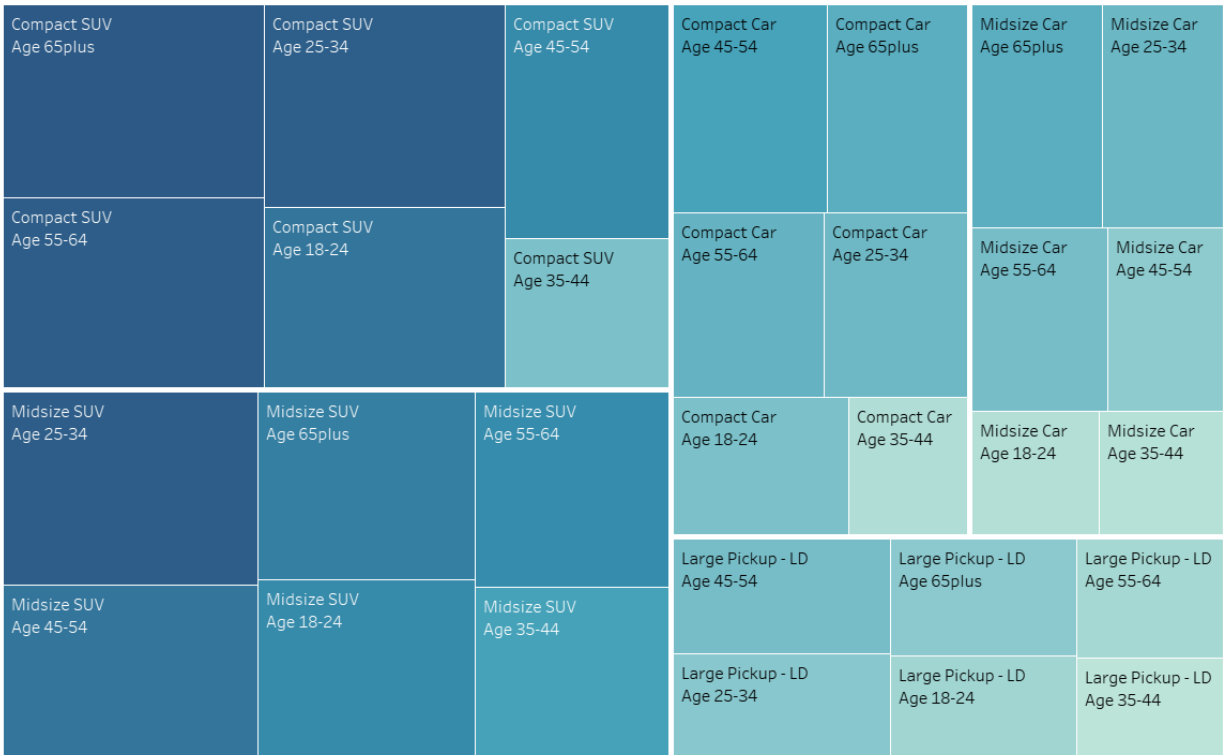
Pandas, Numpy, Sklearn, Seaborn, Yellowbrick, Matplotlib

Analysis

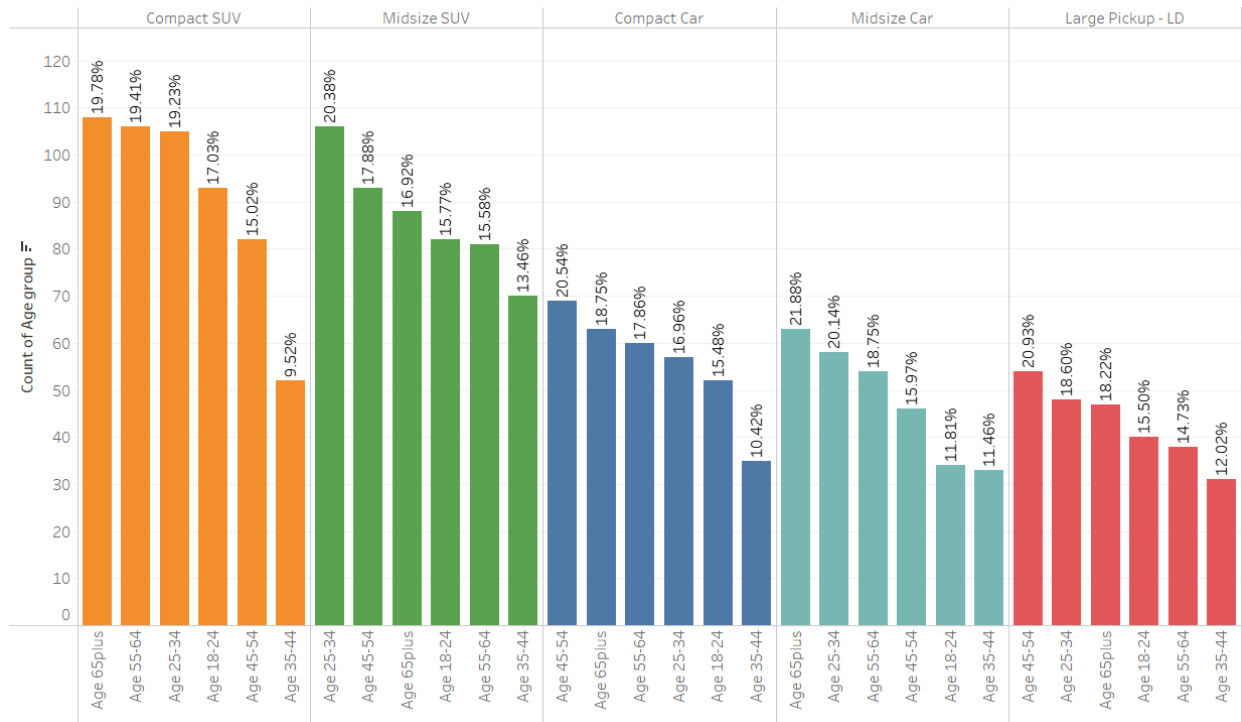
Most bought car type



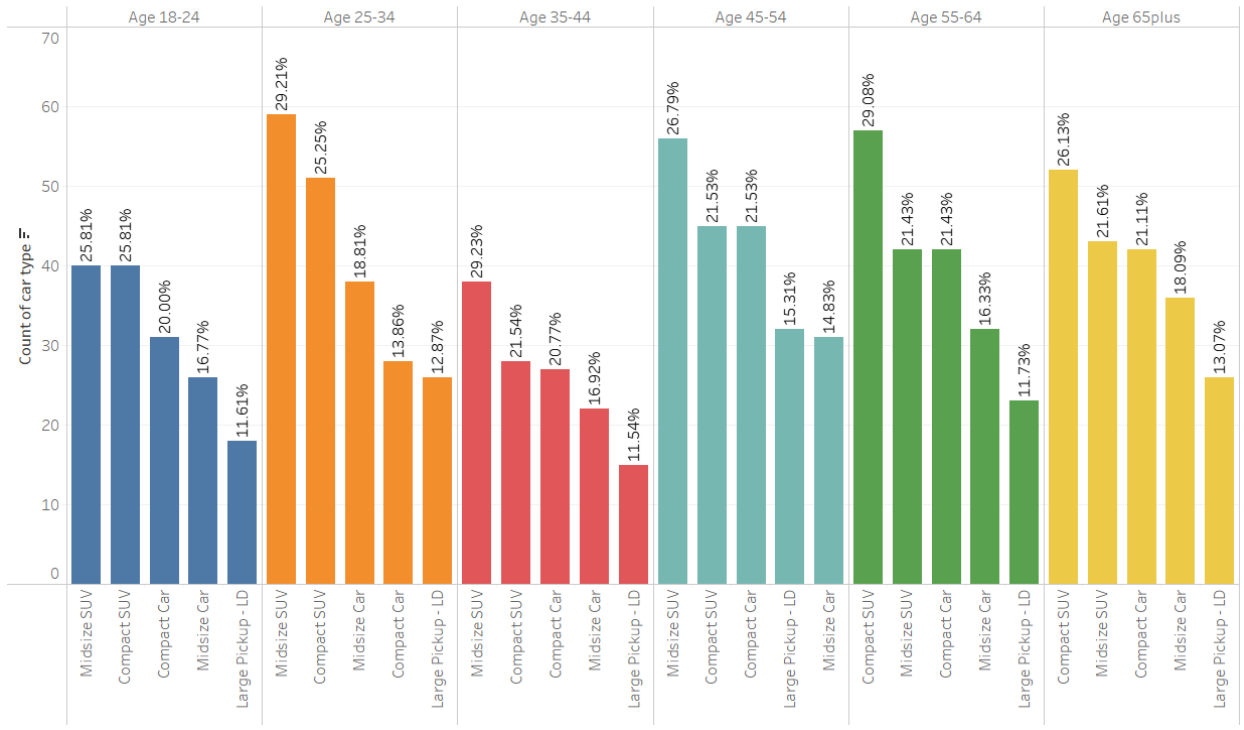
Car type and Age relation



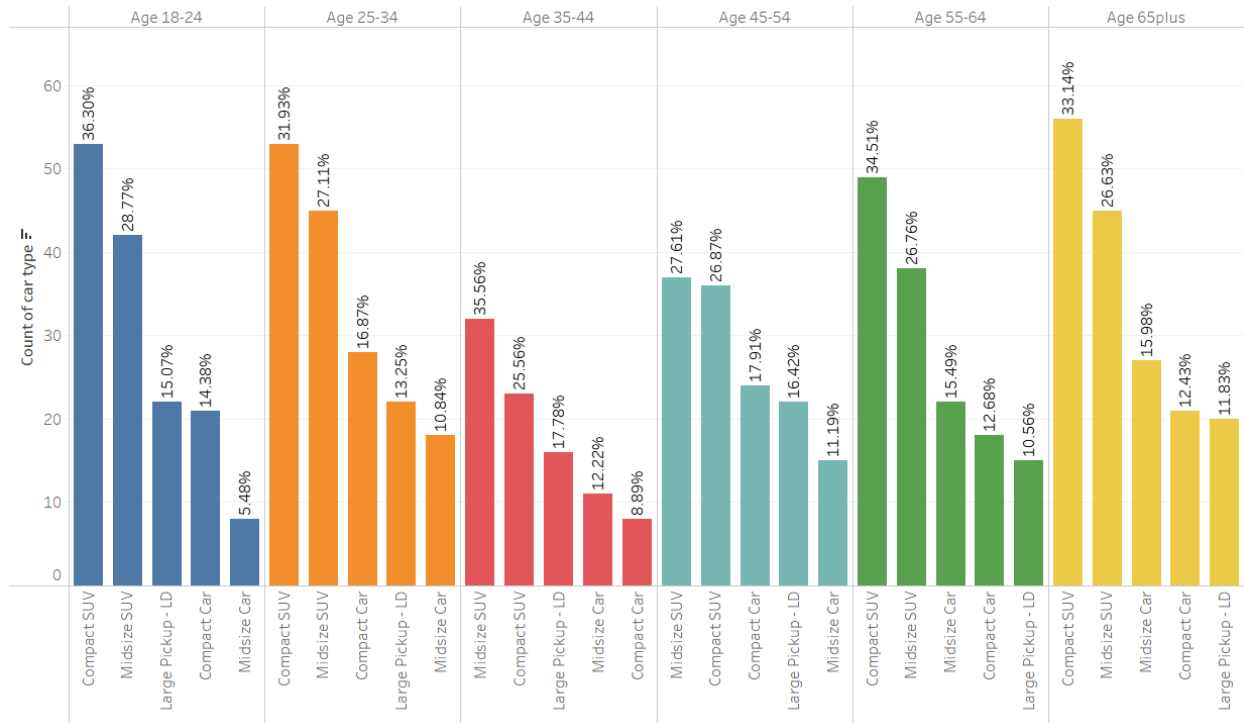
Car type and Age relation



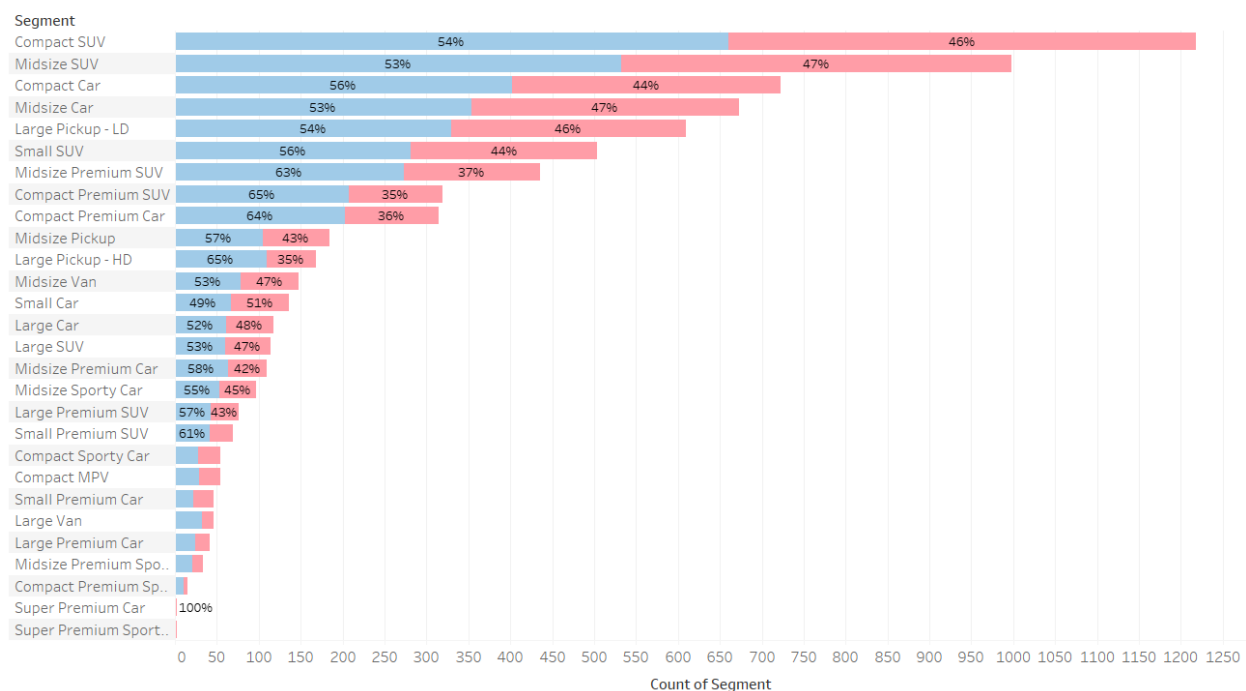
Car type % within age groups - used cars



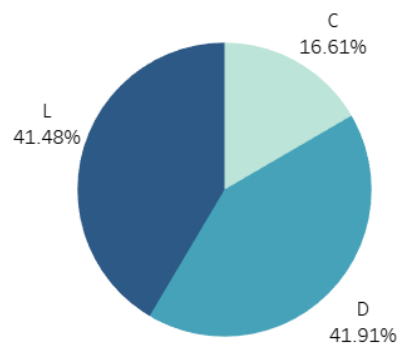
Car type % within age groups - new cars



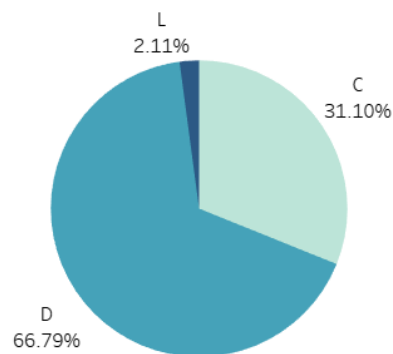
Car type and gender relation



New cars transaction type %

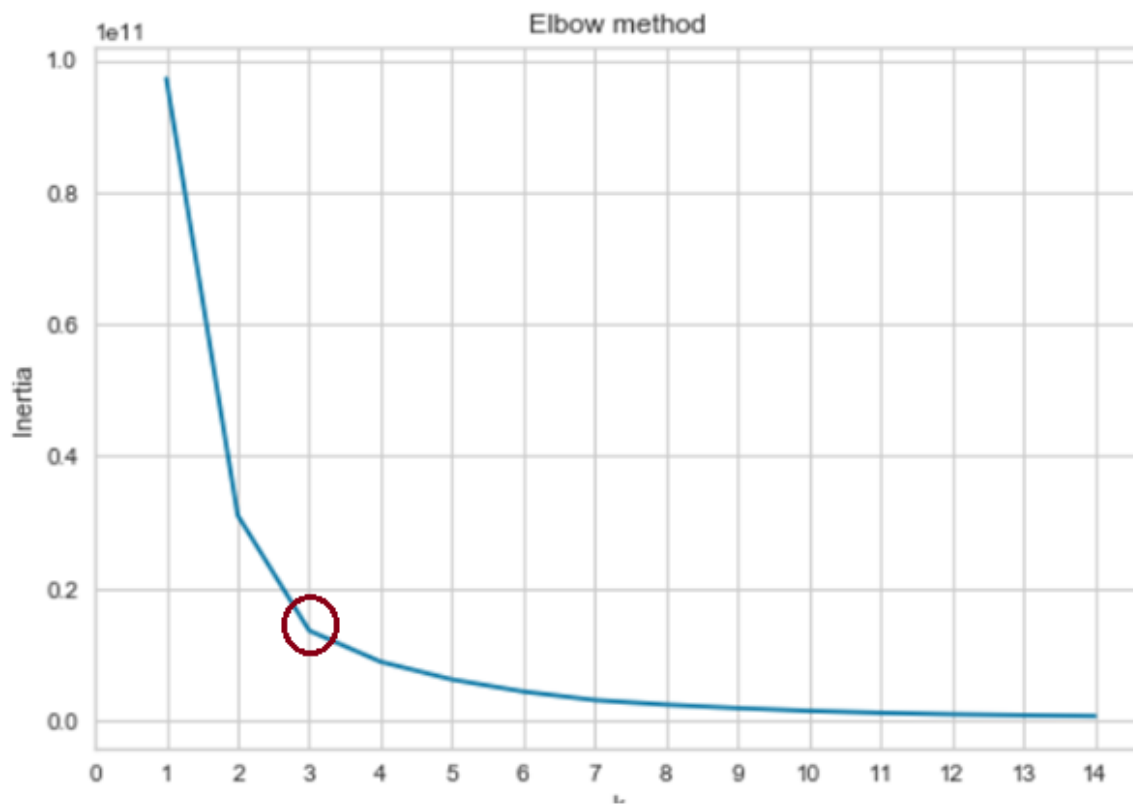


Used cars transaction type %

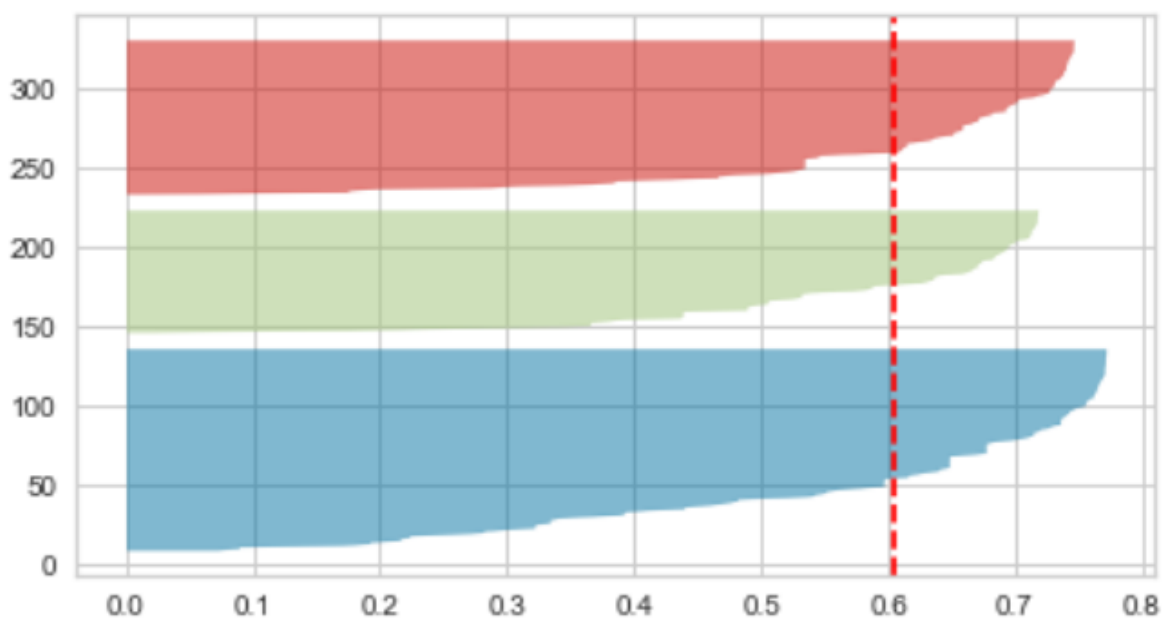


L - Lease
D - Loan
C - Cash

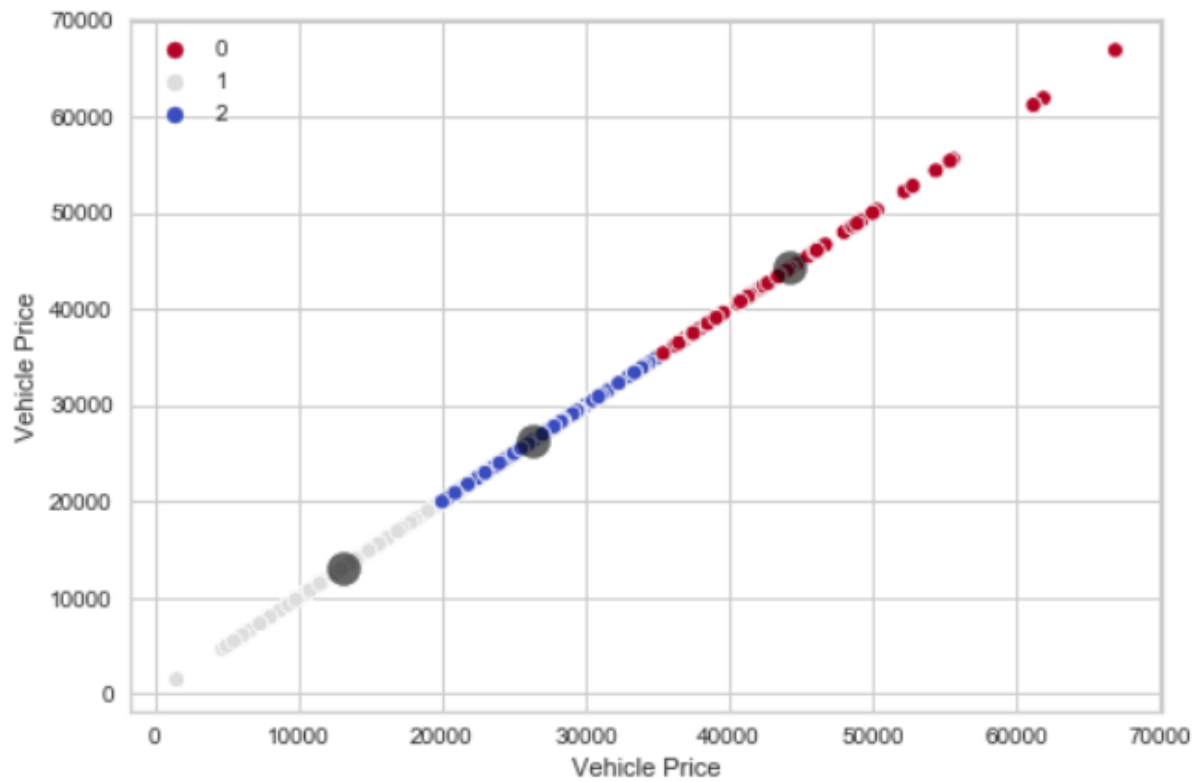
Vehicle price clustering for age group 18-24: Elbow method suggesting 3 clusters



Further verified number of clusters using Silhouette Visualizer



Clusters shown in scatter plot:



```
2    127
1     97
0     77
Name: cluster
```

Most popular travel destination for auto purchasers with travel intent: Caribbean

proportion
0.31

Following is the analysis and graphs for each group (New cars, Used cars, Luxury cars)

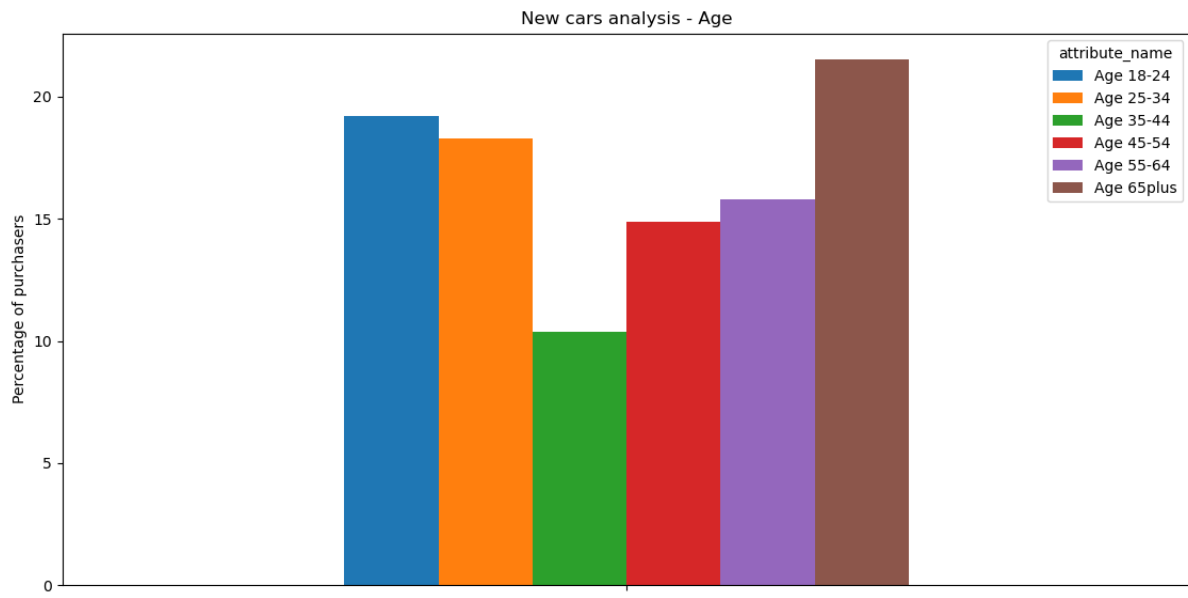
Used Python to analyze and generate graphs

Purchasers of new cars

1) Age:

Dataframe generated for plotting

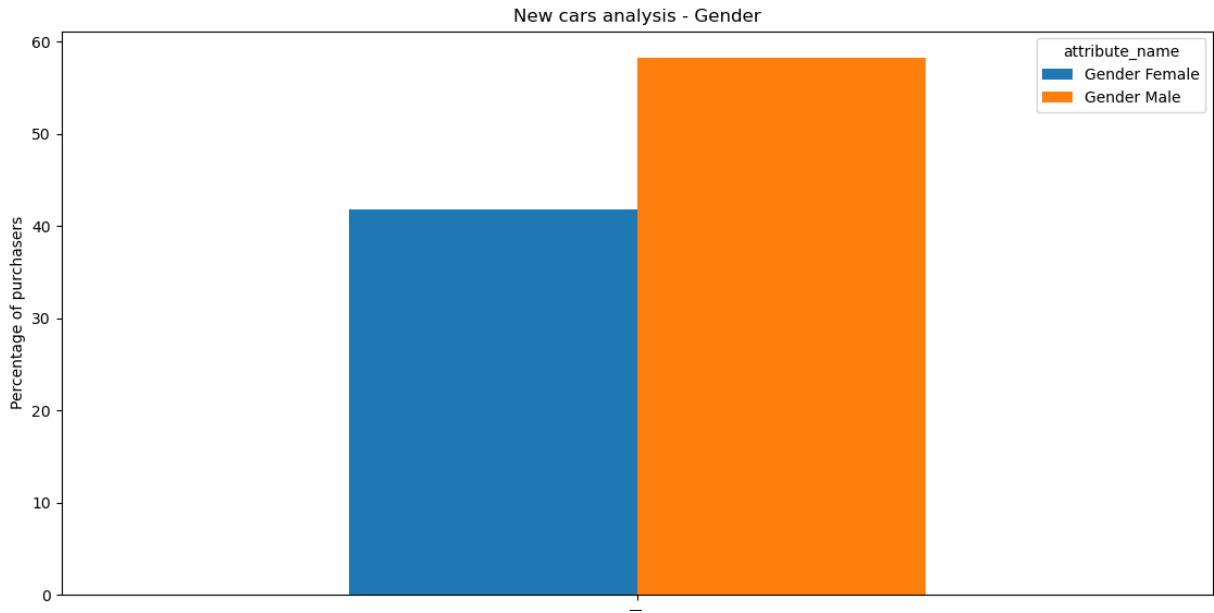
attribute_name	Age 18-24	Age 25-34	Age 35-44	Age 45-54	Age 55-64	Age 65plus
purchaser_id%	19.19	18.26	10.37	14.86	15.78	21.51



It seems strange at first looking at the high number of new car buyer are of age 65plus but it actually makes sense since many of them might have access to their 401k funds and avoid maintenance hustle of used cars

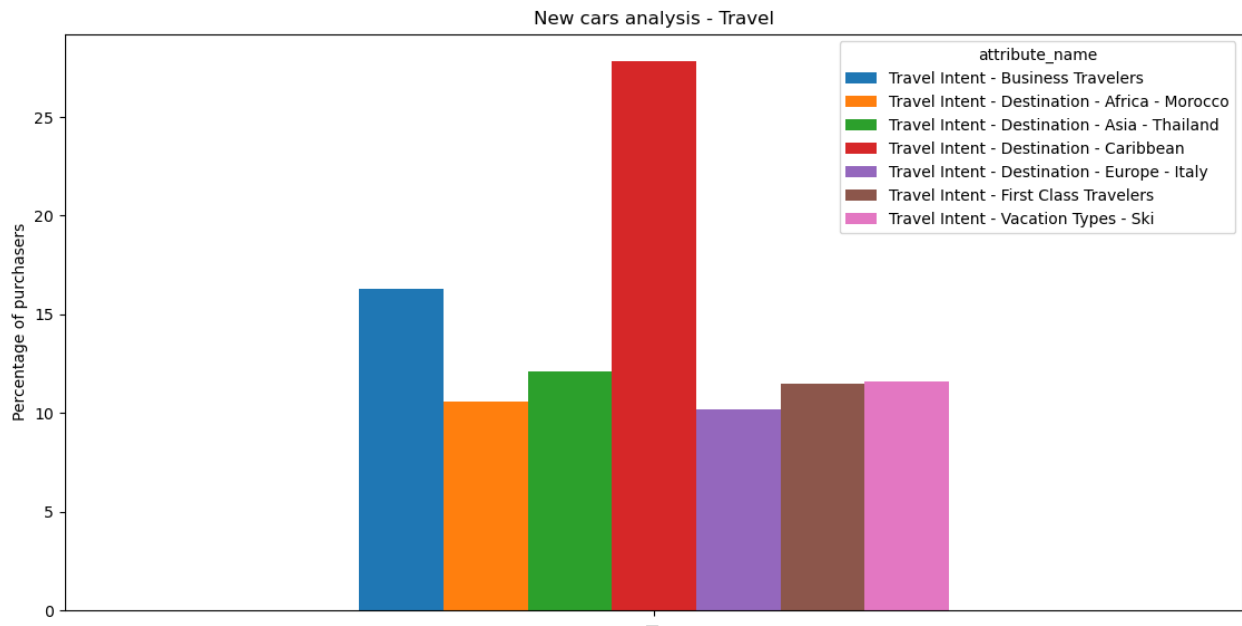
2) Gender:

attribute_name	Gender Female	Gender Male
purchaser_id%	41.76	58.23



3) Travel:

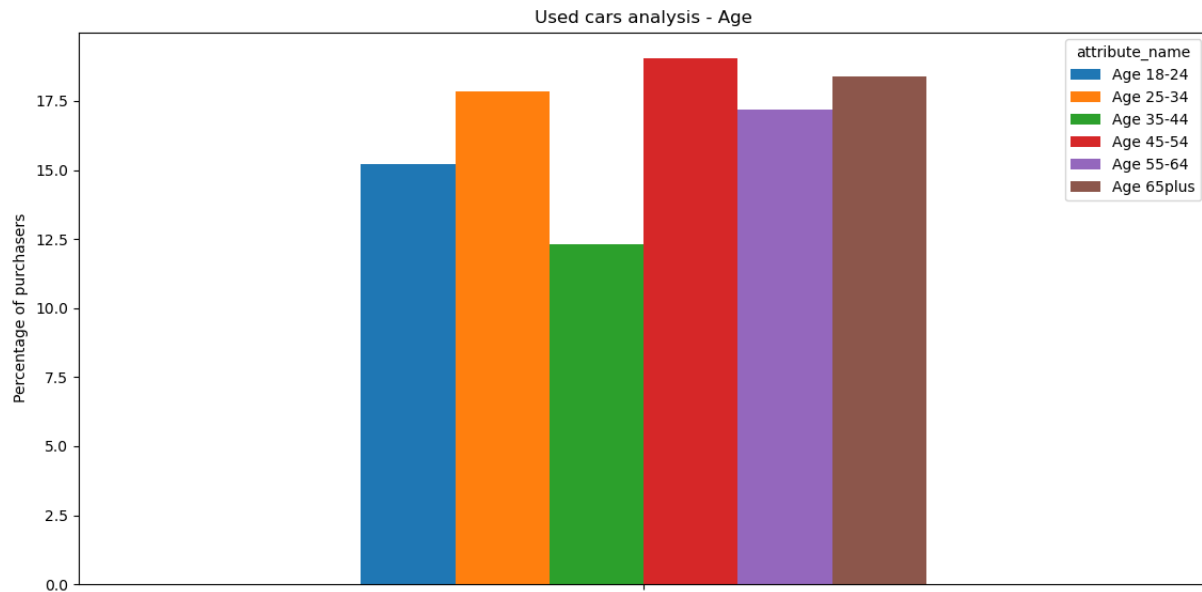
attribute_name	Travel Intent - Business Travelers	Travel Intent - Destination - Africa - Morocco	Travel Intent - Destination - Asia - Thailand	Travel Intent - Destination - Caribbean	Travel Intent - Destination - Europe - Italy	Travel Intent - First Class Travelers	Travel Intent - Vacation Types - Ski
purchaser_id%	16.29	10.57	12.09	27.82	10.16	11.48	11.57



Purchasers of used cars

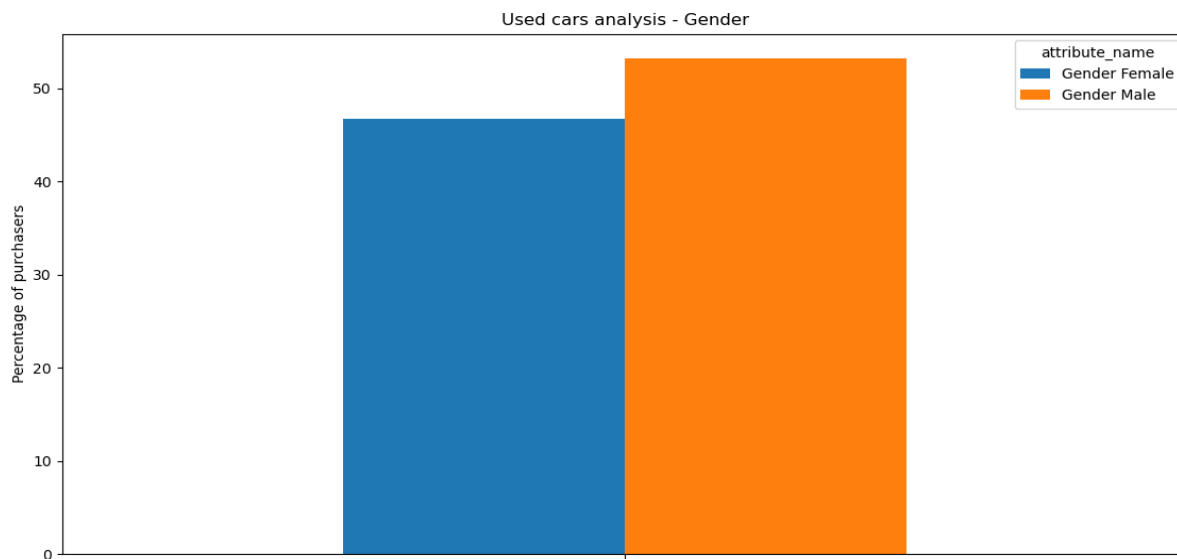
1) Age:

attribute_name	Age 18-24	Age 25-34	Age 35-44	Age 45-54	Age 55-64	Age 65plus
purchaser_id%	15.20	17.82	12.33	19.04	17.18	18.40



2) Gender:

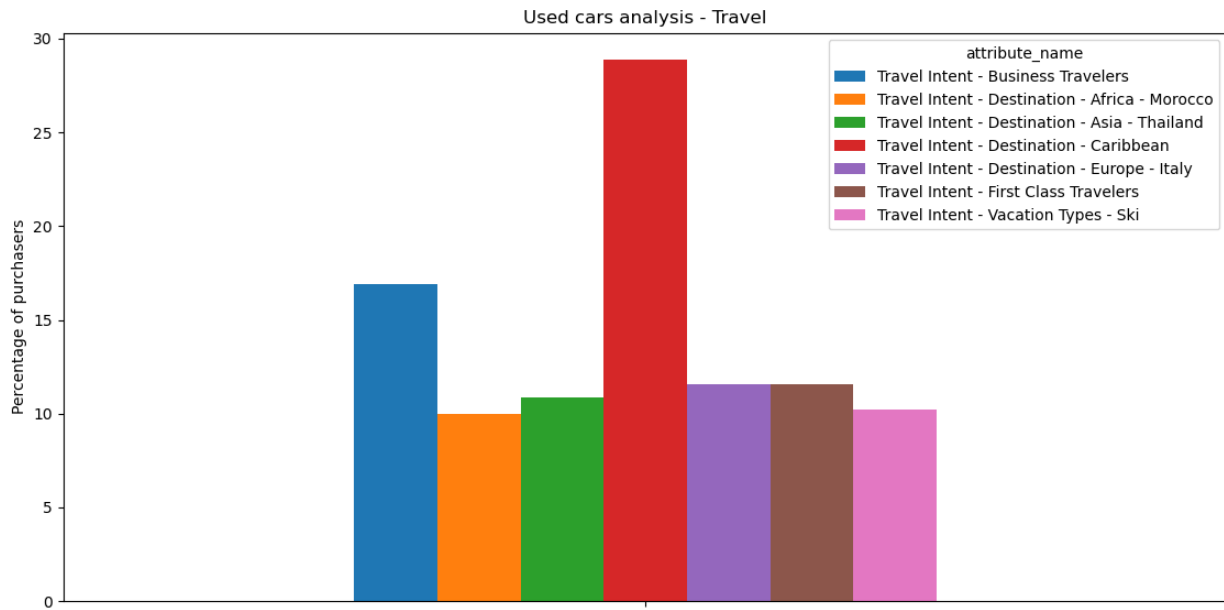
attribute_name	Gender Female	Gender Male
purchaser_id%	46.78	53.21



Females prefer to buy used cars than men. Men prefer to buy new cars instead.

3) Travel:

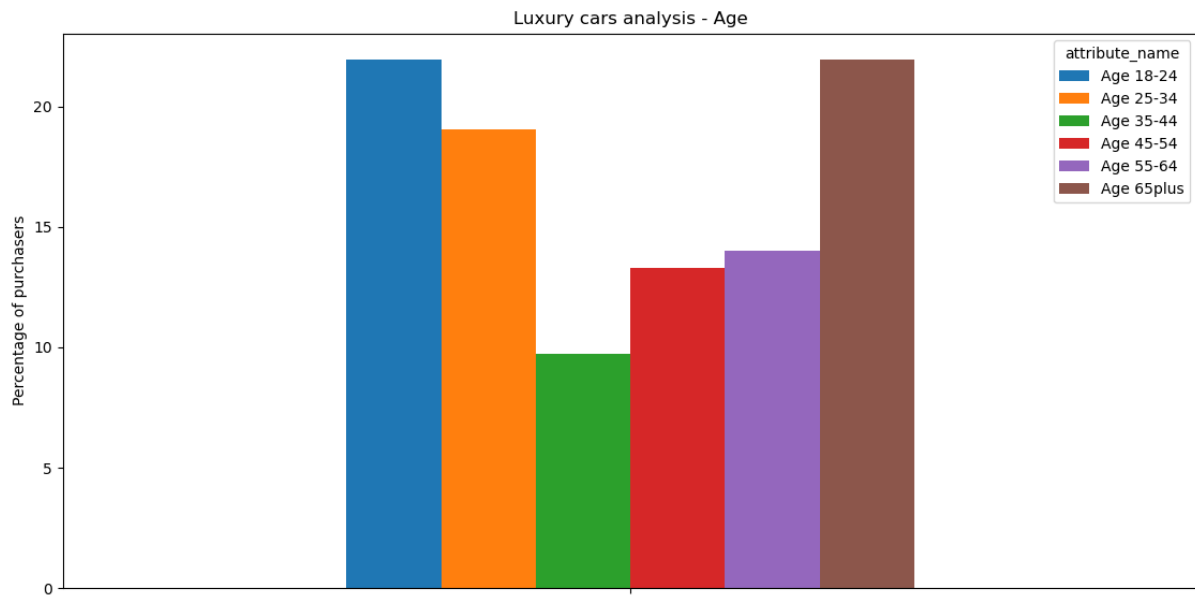
attribute_name	Travel Intent - Business Travelers	Travel Intent - Destination - Africa - Morocco	Travel Intent - Destination - Asia - Thailand	Travel Intent - Destination - Caribbean	Travel Intent - Destination - Europe - Italy	Travel Intent - First Class Travelers	Travel Intent - Vacation Types - Ski
purchaser_id%	16.92	9.97	10.84	28.87	11.56	11.59	10.21



Purchasers of luxury vehicles (price >\$50k)

1) Age:

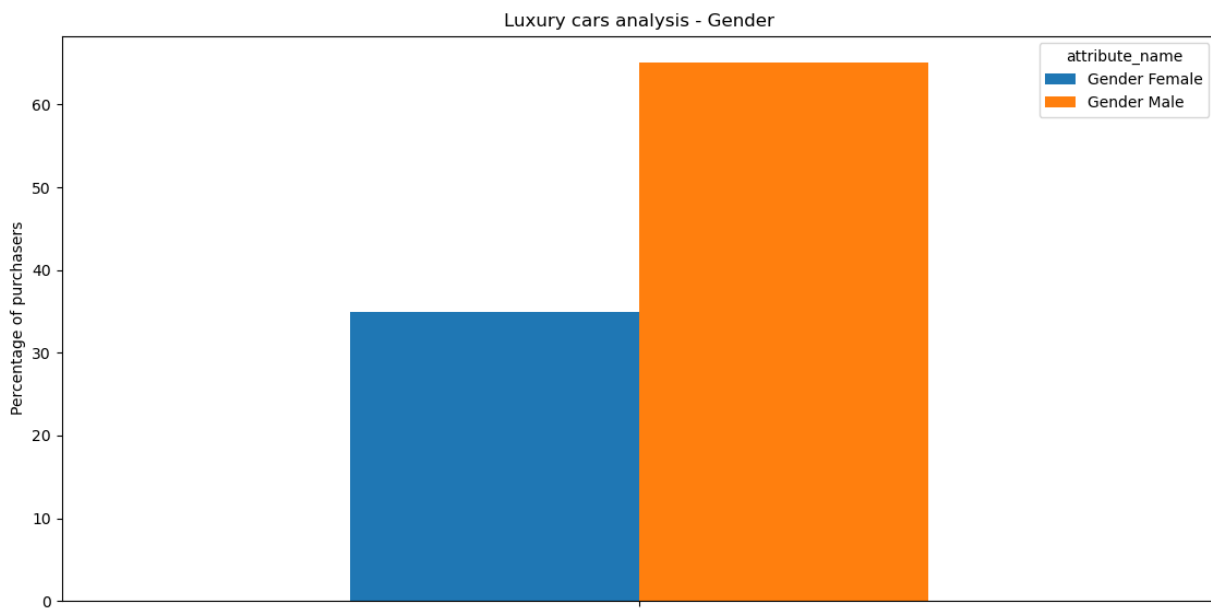
attribute_name	Age 18-24	Age 25-34	Age 35-44	Age 45-54	Age 55-64	Age 65plus
purchaser_id%	21.94	19.06	9.71	13.30	14.02	21.94



I see Age 25-34 as opportunistic here and if handled appropriately there are chances of them buying luxury cars in more number

2) Gender:

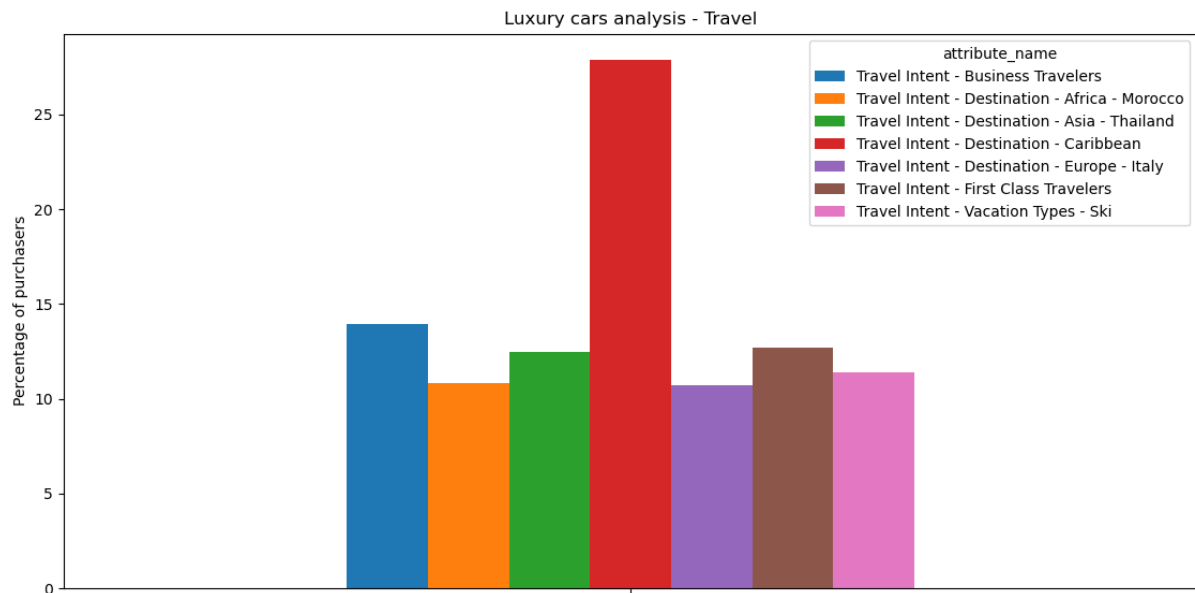
attribute_name	Gender Female	Gender Male
purchaser_id%	34.93	65.06



Since the buy difference is higher between Male and Female luxury car buyers. We also see that age group 35-44 have bought less no. of luxury cars. This gives us an indication that if client is Female and in age group 35-44 then she is less likely to buy luxury cars and hence should be presented with other options.

3) Travel:

attribute_name	Travel Intent - Business Travelers	Travel Intent - Destination - Africa - Morocco	Travel Intent - Destination - Asia - Thailand	Travel Intent - Destination - Caribbean	Travel Intent - Destination - Europe - Italy	Travel Intent - First Class Travelers	Travel Intent - Vacation Types - Ski
purchaser_id%	13.93	10.84	12.50	27.87	10.73	12.72	11.39



Insights/Findings:

- Most preferred car types across all age groups are Compact SUV and Midsize SUV.
- Age group between '35-44' have purchased least no. of cars in New, Used and Luxury category
- Compact SUV are most preferred by elderly people irrespective of used or new cars. While youngsters prefer Midsize SUV when buying used cars and Compact SUV when buying used cars.
- Cluster analysis on age group 18-24 shows that the most people in this age group buy mid ranged vehicle. The low range car buyers can be encouraged to buy med ranged vehicle.
- Age group 35-44 and 45-54 always prefers midsize SUV.
- Small cars are bought more by women and premium cars are more bought by men.
- Age groups 25-34 are seen as opportunistic here and if handled appropriately there are chances of them buying luxury cars in more number

- h. It seems strange at first looking at the high number of new car buyer are of age 65plus but it actually makes sense since many of them might have access to their 401k funds and avoid maintenance hustle of used cars
- i. Buy difference is higher between Male and Female luxury car buyers. We also see that age group 35-44 has bought less no. of luxury cars. This gives an indication that if client is Female and in age group 35-44 then she is less likely to buy luxury cars and hence should be presented with other options
- j. New luxury cars are mostly either loaned or leased and used luxury cars are mostly loaned.

N	C	280
N	D	539
N	L	594
U	C	66
U	D	116
U	L	38