ata-analysis-with-python-project-2

January 28, 2023

[1]:
```
import pandas as pd
```

1 How to analyze Data Frame?

[26]:
```
car=pd.read_csv("car data.csv")
car.head()
```

[26]:

	Car_Name	Year	Selling_Price	Present_Price	Kms_Driven	Fuel_Type	Seller_Type	Transmission	Owner
0	ritz	2014	3.35	5.59	27000	Petrol	Dealer	Manual	0
1	sx4	2013	4.75	9.54	43000	Diesel	Dealer	Manual	0
2	ciaz	2017	7.25	9.85	6900	Petrol	Dealer	Manual	0
3	wagon r	2011	2.85	4.15	5200	Petrol	Dealer	Manual	0
4	swift	2014	4.60	6.87	42450	Diesel	Dealer	Manual	0

[77]:
```
car.info()
# provid the basic information about data
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 301 entries, 0 to 300
Data columns (total 9 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   Car_Name       301 non-null    object
 1   Year           301 non-null    int64
 2   Selling_Price  301 non-null    float64
 3   Present_Price  5 non-null      float64
 4   Kms_Driven     301 non-null    int64
 5   Fuel_Type      301 non-null    object
 6   Seller_Type    301 non-null    object
 7   Transmission   301 non-null    object
```

```
 8   Owner          301 non-null    int64
dtypes: float64(2), int64(3), object(4)
memory usage: 21.3+ KB
```

[78]:
```
# index
#this attribute provides the index of the data frame
car.index
```

[78]:
```
RangeIndex(start=0, stop=301, step=1)
```

[27]:
```
car.shape
```

[27]:
```
(301, 9)
```

[63]:
```
#nunique() it shows the total no of unique values each columns
car.nunique()
```

[63]:
```
Car_Name          98
Year              16
Selling_Price    156
Present_Price      5
Kms_Driven       206
Fuel_Type          3
Seller_Type        2
Transmission       2
Owner              3
dtype: int64
```

[64]:
```
car.count()
# it show the total number of non null values in each columns or whole data␣
 ↪frame
```

[64]:
```
Car_Name         301
Year             301
Selling_Price    301
Present_Price      5
Kms_Driven       301
Fuel_Type        301
Seller_Type      301
Transmission     301
Owner            301
dtype: int64
```

2 For data Cleaning = Find all null value in the data set is any null value in any column .Then fill with the mean of that column

[28]:
```
car.isnull().sum()
```

[28]:
```
Car_Name         0
Year             0
Selling_Price    0
Present_Price    0
Kms_Driven       0
Fuel_Type        0
Seller_Type      0
Transmission     0
Owner            0
dtype: int64
```

[29]:
```
car.columns
```

[29]:
```
Index(['Car_Name', 'Year', 'Selling_Price', 'Present_Price', 'Kms_Driven',
       'Fuel_Type', 'Seller_Type', 'Transmission', 'Owner'],
      dtype='object')
```

[ ]:
```
mean=car['Selling_Price        '].mean()
```

3 chek what are the different type of Carname are there in our data set and what is count.

[47]:
```
car["Car_Name"].value_counts().head(60)
```

[47]:
```
city                          26
corolla altis                 16
verna                         14
fortuner                      11
brio                          10
ciaz                           9
innova                         9
i20                            9
grand i10                      8
jazz                           7
amaze                          7
Royal Enfield Classic 350      7
ertiga                         6
eon                            6
sx4                            6
alto k10                       5
i10                            5
```

```
swift                          5
Bajaj Pulsar 150               4
Royal Enfield Thunder 350      4
ritz                           4
wagon r                        4
etios liva                     4
dzire                          4
xcent                          3
etios cross                    3
Royal Enfield Thunder 500      3
TVS Apache RTR 160             3
Yamaha FZ S V 2.0              3
creta                          3
Honda CB Hornet 160R           3
etios g                        3
Bajaj Avenger 220              3
Bajaj Pulsar NS 200            3
Bajaj Discover 125             2
Bajaj Pulsar 220 F             2
TVS Apache RTR 180             2
Yamaha FZ  v 2.0               2
Honda Activa 4G                2
Hero Extreme                   2
Honda CB twister               2
Hero Splender iSmart           2
Activa 3g                      2
Hero Passion Pro               2
Honda CB Shine                 2
Bajaj Avenger 220 dtsi         2
Honda CBR 150                  2
elantra                        2
Royal Enfield Classic 500      2
KTM RC200                      2
Honda Karizma                  2
Bajaj  ct 100                  1
Hero  Ignitor Disc             1
TVS Jupyter                    1
Hero Honda Passion Pro         1
Hero Splender Plus             1
vitara brezza                  1
Bajaj Discover 100             1
Suzuki Access 125              1
s cross                        1
Name: Car_Name, dtype: int64
```

[66]:
```
# find how many Honda CB Hornet 160R
car[car.Car_Name=="Honda CB Hornet 160R"]
```

[66]:
```
                Car_Name  Year  Selling_Price  Present_Price  Kms_Driven  \
128  Honda CB Hornet 160R  2017           0.80            NaN        3000
130  Honda CB Hornet 160R  2017           0.75            NaN       11000
138  Honda CB Hornet 160R  2016           0.60            NaN       15000

    Fuel_Type Seller_Type Transmission  Owner
128    Petrol  Individual       Manual      0
130    Petrol  Individual       Manual      0
138    Petrol  Individual       Manual      0
```

[70]:
```
#groupby
car.groupby("Car_Name").get_group("Honda CB Hornet 160R")
```

[70]:
```
                Car_Name  Year  Selling_Price  Present_Price  Kms_Driven  \
128  Honda CB Hornet 160R  2017           0.80            NaN        3000
130  Honda CB Hornet 160R  2017           0.75            NaN       11000
138  Honda CB Hornet 160R  2016           0.60            NaN       15000

    Fuel_Type Seller_Type Transmission  Owner
128    Petrol  Individual       Manual      0
130    Petrol  Individual       Manual      0
138    Petrol  Individual       Manual      0
```

4 Filttering method= show all the recored where vechicel name is Bajaj Pulsar 150 or Royal Enfield Thunder 350

[53]:
```
car[car["Car_Name"].isin(["Bajaj Pulsar 150","Royal Enfield Thunder 350"])]
```

[53]:
```
                     Car_Name  Year  Selling_Price  Present_Price  \
107  Royal Enfield Thunder 350  2013           1.25           1.50
108  Royal Enfield Thunder 350  2016           1.20           1.50
111  Royal Enfield Thunder 350  2016           1.15           1.50
121  Royal Enfield Thunder 350  2011           1.05           1.50
135           Bajaj Pulsar 150  2015           0.65           0.74
184           Bajaj Pulsar 150  2008           0.25           0.75
190           Bajaj Pulsar 150  2008           0.20           0.75
200           Bajaj Pulsar 150  2006           0.10           0.75

     Kms_Driven Fuel_Type Seller_Type Transmission  Owner
107       15000    Petrol  Individual       Manual      0
108       18000    Petrol  Individual       Manual      0
111        8700    Petrol  Individual       Manual      0
121        6900    Petrol  Individual       Manual      0
135        5000    Petrol  Individual       Manual      0
184       26000    Petrol  Individual       Manual      1
190       60000    Petrol  Individual       Manual      0
```

```
200       92233    Petrol  Individual       Manual      0
```

[54]:
```
car[car["Fuel_Type"].isin(["Petrol"])]
```

[54]:
```
     Car_Name  Year  Selling_Price  Present_Price  Kms_Driven Fuel_Type  \
0        ritz  2014           3.35           5.59       27000    Petrol
2        ciaz  2017           7.25           9.85        6900    Petrol
3     wagon r  2011           2.85           4.15        5200    Petrol
6        ciaz  2015           6.75           8.12       18796    Petrol
10   alto 800  2017           2.85           3.60        2135    Petrol
..        …   …              …              …          …       …
293      city  2010           3.25           9.90       38000    Petrol
294     amaze  2014           3.75           6.80       33019    Petrol
297      brio  2015           4.00           5.90       60000    Petrol
298      city  2009           3.35          11.00       87934    Petrol
300      brio  2016           5.30           5.90        5464    Petrol

    Seller_Type Transmission  Owner
0        Dealer       Manual      0
2        Dealer       Manual      0
3        Dealer       Manual      0
6        Dealer       Manual      0
10       Dealer       Manual      0
..          …          …    …
293      Dealer       Manual      0
294      Dealer       Manual      0
297      Dealer       Manual      0
298      Dealer       Manual      0
300      Dealer       Manual      0

[239 rows x 9 columns]
```

[ ]:
```
Remove Unwanted Recored
Remove  all the records (rows) where KMs Driven is above 27000
```

[55]:
```
car[car["Kms_Driven"]>27000]
```

[55]:
```
    Car_Name  Year  Selling_Price  Present_Price  Kms_Driven Fuel_Type  \
1        sx4  2013           4.75           9.54       43000    Diesel
4      swift  2014           4.60           6.87       42450    Diesel
7    s cross  2015           6.50           8.61       33429    Diesel
9       ciaz  2015           7.45           8.92       42367    Diesel
11      ciaz  2015           6.85          10.38       51000    Diesel
..       …   …              …              …          …       …
294    amaze  2014           3.75           6.80       33019    Petrol
295     city  2015           8.55          13.09       60076    Diesel
296     city  2016           9.50          11.60       33988    Diesel
```

```
297     brio  2015           4.00           5.90       60000    Petrol
298     city  2009           3.35          11.00       87934    Petrol

    Seller_Type Transmission  Owner
1        Dealer       Manual      0
4        Dealer       Manual      0
7        Dealer       Manual      0
9        Dealer       Manual      0
11       Dealer       Manual      0
..          …          …    …
294      Dealer       Manual      0
295      Dealer       Manual      0
296      Dealer       Manual      0
297      Dealer       Manual      0
298      Dealer       Manual      0

[167 rows x 9 columns]
```

[60]:
```
car[~(car["Kms_Driven"]>27000)].head(5)
```

[60]:
```
         Car_Name  Year  Selling_Price  Present_Price  Kms_Driven Fuel_Type  \
0            ritz  2014           3.35          15.59       27000    Petrol
2            ciaz  2017           7.25          19.85        6900    Petrol
3         wagon r  2011           2.85          14.15        5200    Petrol
5  vitara brezza  2018           9.25          19.83        2071    Diesel
6            ciaz  2015           6.75          18.12       18796    Petrol

  Seller_Type Transmission  Owner
0      Dealer       Manual      0
2      Dealer       Manual      0
3      Dealer       Manual      0
5      Dealer       Manual      0
6      Dealer       Manual      0
```

5 Increase all the value of Present_Price by 5

[62]:
```
car["Present_Price"]=car["Present_Price"].apply(lambda x:x+5)
car.head()
```

[62]:
```
  Car_Name  Year  Selling_Price  Present_Price  Kms_Driven Fuel_Type  \
0     ritz  2014           3.35          25.59       27000    Petrol
1      sx4  2013           4.75          29.54       43000    Diesel
2     ciaz  2017           7.25          29.85        6900    Petrol
3  wagon r  2011           2.85          24.15        5200    Petrol
4    swift  2014           4.60          26.87       42450    Diesel
```

```
  Seller_Type Transmission  Owner
0      Dealer       Manual      0
1      Dealer       Manual      0
2      Dealer       Manual      0
3      Dealer       Manual      0
4      Dealer       Manual      0
```

6 find all instance when Selling_Price is above 4.60 and Present_Price is 29.5

[76]:
```
car[(car["Selling_Price"]>4.60)&(car["Present_Price"]==29.54)]
```

[76]:
```
  Car_Name  Year  Selling_Price  Present_Price  Kms_Driven Fuel_Type  \
1      sx4  2013           4.75          29.54       43000    Diesel

  Seller_Type Transmission  Owner
1      Dealer       Manual      0
```

[ ]:
```

```