Working on real project with python

RSX Type S 2dr

TSX 4dr

Sedan

Sedan

Asia

Asia

(A part of big data analysis)

Cars Dataset

1 Acura

2 Acura

4 Acura

car.isnull().sum()

car.isnull().sum()

4

4 4

0

4

2) instruction (filtering)

In [5]

In [9]:

Out[9]:

In [15]:

Out[15]:

In [16]:

Out[16]:

In [17]:

Out[17]: Toyota

BMW

Chrysler

Mitsubishi

Dodge

Volvo

Jaguar

Lexus

Buick

Mercury Lincoln Saturn Cadillac

Suzuki

GMC

Acura Porsche Saab

Jeep Scion Isuzu

MINI

In [18]:

Out[18]:

Out[20]:

Hummer

Make

Make

0 Acura

1 Acura

4 Acura

431 Volvo

car.head(2)

Make

Acura

0

18

20

419

405 Volkswagen

415 Volkswagen

416 Volkswagen

car.head(2)

Make

4 Acura

432 rows × 15 columns

In [26]:

Out[26]:

In [33]:

Out[33]:

Volkswagen

In [25]: car[car['Weight'] > 4000]

Make

Acura

Audi

Audi

Out[22]:

Out[25]:

281 rows × 15 columns

0 Acura

Infiniti

Land Rover

Oldsmobile

Kia

Make

Model

Type **Origin** DriveTrain

MSRP Invoice EngineSize Cylinders

Horsepower MPG_City

MPG_Highway Weight Wheelbase Length

dtype: int64

car.head(2)

car['Make'].value_counts()

28

20 19

15

13

13

12 12

11

11

9 9

8

8

8

8

3

3 3

2

2

1

3) instruction (filtering)

Model

MDX

1 Acura RSX Type S 2dr Sedan

Type

SUV

Model

MDX

car[car['Origin'].isin (['Asia', 'Europe'])]

Origin DriveTrain

Asia

Asia

Type

SUV

Sedan

Sedan

XC70 Wagon Europe

Type Origin DriveTrain

Asia

4)instruction (removing unwanted records)

Name: Make, dtype: int64

Here, the data of different cars is given their specifications. This data is available as a CSV file. we are going to analyze the data set using the pandas dataframe.

import pandas as pd

Front \$23,820 \$21,761

\$26,990

Front

Front \$43,755 \$39,014

find all null value in the dataset.if there is any value in any column, then fill it with the mean of that column.

car=pd.read_csv(r"C:\Users\hp\Downloads\2. Cars Data1.csv")

In [3]: car Model Make Origin DriveTrain MSRP Invoice EngineSize Cylinders Horsepower MPG_City MPG_Highway Weight Wheelbase Length Out[3]: Type **0** Acura MDX SUV \$36,945 \$33,337 3.5 265.0 17.0 4451.0 106.0 189.0 Asia ΑII 23.0

\$24,647

\$33,195 \$30,299 3.2 6.0 270.0 20.0 186.0 TL 4dr Sedan 28.0 3575.0 108.0 3 Acura Asia Front 3.5 RL 4dr Front \$43,755 \$39,014 3.5 6.0 225.0 18.0 24.0 3880.0 197.0 4 Acura Sedan Asia 115.0 C70 LPT convertible 2dr \$40,565 \$38,203 427 Volvo Sedan Europe Front 2.4 5.0 197.0 21.0 28.0 3450.0 105.0 186.0 Volvo C70 HPT convertible 2dr Front \$42,565 \$40,083 2.3 5.0 242.0 20.0 26.0 3450.0 186.0 428 105.0 Sedan Europe Volvo \$45,210 \$42,573 2.9 6.0 268.0 19.0 3653.0 110.0 190.0 429 S80 T6 4dr Sedan Europe Front \$26,135 \$24,641 1.9 4.0 170.0 22.0 29.0 2822.0 180.0 430 Volvo V40 Wagon Europe 101.0

2.0

2.4

4.0

4.0

225.0

18.0

200.0

200.0

24.0

22.0

31.0 2778.0

3880.0

24.0

3230.0

101.0

105.0

197.0

115.0

Weight Wheelbase Length

Weight Wheelbase Length

4451.0

3230.0

3575.0

3880.0

31.0 2778.0

27.0 3823.0

Weight Wheelbase Length

106.0

101.0

23.0

25.0

24.0

24.0

22.0

20.0

22.0

19.0

25.0

189.0

172.0

4451.0

4013.0

4024.0

4399.0

4024.0

5086.0

5194.0

5399.0

4067.0

189.0

172.0

115.0

189.0

197.0

106.0

101.0

106.0

105.0

109.0

121.0

109.0

112.0

118.0

118.0

106.0

113.0

189.0

180.0

193.0

204.0

191.0

187.0

204.0

204.0

184.0

189.0

106.0

101.0

189.0

172.0

Weight Wheelbase Length

106.0

101.0

105.0

108.0

115.0

105.0

105.0

110.0

101.0

109.0

189.0

172.0

183.0

186.0

197.0

186.0

186.0

190.0

180.0

186.0

106.0

101.0

189.0

172.0

4451.0

31.0 2778.0

17.0

24.0

172.0

183.0

431 Volvo All \$35,145 \$33,112 2.5 5.0 208.0 20.0 27.0 3823.0 109.0 186.0 XC70 Wagon Europe 432 rows × 15 columns

car.head()

Invoice EngineSize Cylinders Horsepower MPG_City MPG_Highway Make Model Origin DriveTrain MSRP Weight Wheelbase Length Out[4]: Type MDX SUV All \$36,945 \$33,337 3.5 265.0 17.0 4451.0 189.0 O Acura Asia 6.0 23.0 106.0 **1** Acura RSX Type S 2dr Sedan Asia Front \$23,820 \$21,761 2.0 4.0 200.0 24.0 31.0 2778.0 101.0 172.0 4.0 200.0 3230.0 183.0 2 Acura TSX 4dr Sedan Front \$26,990 \$24,647 2.4 22.0 29.0 105.0 Asia 3 Acura TL 4dr Sedan Front \$33,195 3.2 270.0 20.0 3575.0 108.0 186.0 Asia

3.5

car.shape Out[5]: (432, 15) 1)instruction (for data cleaning)

6.0

Make Model

3.5 RL 4dr Sedan

Asia

Type **Origin** DriveTrain **MSRP** Invoice EngineSize Cylinders 6 Horsepower 4 MPG_City 4 MPG_Highway 4 Weight Wheelbase Length dtype: int64 In [13]: car['Cylinders'].fillna(car['Cylinders'].mean(),inplace= True)

Type Origin DriveTrain Make **MSRP** Invoice EngineSize Cylinders Horsepower MPG_City MPG_Highway Model MDX All \$36,945 \$33,337 3.5 6.0 265.0 **0** Acura SUV Asia 200.0 1 Acura RSX Type S 2dr Sedan Front \$23,820 \$21,761 2.0 4.0 Asia

27 Chevrolet Mercedes-Benz 26 Ford 23

• check what are the different types of make are there in our dataset.and, what is the count (occurrence) of each make in the data?

Audi 17 Honda Nissan 17 Volkswagen 15

Hyundai 12 Subaru 11 Pontiac 11 11 Mazda

 show all the records where origin is asia or europe. car.head(2)

\$33,337

MSRP

\$36.945

\$26,990

Front \$23,820 \$21,761

Front \$43,755 \$39,014

All \$35,145 \$33,112

\$33,195 \$30,299

ΑII

Front

Front

Invoice EngineSize Cylinders Horsepower MPG_City MPG_Highway

265.0

200.0

6.0

4.0

4.0

6.0

6.0

5.0

265.0

200.0

3.5

3.0

4.2

4.2

4.2

3.2

4.2

6.0

4.0

Invoice EngineSize Cylinders Horsepower MPG_City MPG_Highway

17.0

24.0

265.0

200.0

200.0

270.0

225.0

208.0

17.0

24.0

6.0

6.0

8.0

8.0

6.0

8.0

12.0

8.0

Invoice EngineSize Cylinders Horsepower MPG_City MPG_Highway Weight Wheelbase Length

265.0

200.0

6.0

17.0

24.0

225.0

26.0

23.0

31.0

17.0

24.0

22.0

20.0

18.0

20.0

MPG_Highway

265.0

220.0

300.0

330.0

450.0

220.0

335.0

420.0

270.0

23.0

31.0

Invoice EngineSize Cylinders Horsepower MPG_City MPG_Highway Weight Wheelbase Length

17.0

18.0

17.0

17.0

15.0

15.0

16.0

12.0

18.0

23.0

4451.0

27.0

3880.0

31.0 2778.0

4451.0

2778.0

4451.0

2778.0

23.0

29.0

28.0

24.0

6.0

4.0

3.5

2.0

2.4

3.2

3.5

2.5

Invoice EngineSize Cylinders Horsepower MPG_City

6.0

3.5

MSRP

\$44,240

\$49,690

\$69,190

Front

\$84,600 \$76,417

\$35,515 \$32,243

Front \$65,000 \$59,912

Front \$75,000 \$69,130

\$33,337

\$40,075

\$44,936

\$64,740

\$36,956

All \$36,945

MSRP

Front \$23,820 \$21,761

All \$36,945

Origin DriveTrain

Asia

Asia

Asia

2 Acura TSX 4dr Sedan Asia 3 Acura TL 4dr Sedan Asia

3.5 RL 4dr

RSX Type S 2dr

C70 LPT convertible 2dr Sedan Europe **427** Volvo Front \$40,565 \$38,203 2.4 5.0 197.0 21.0 28.0 3450.0 428 Volvo C70 HPT convertible 2dr Sedan Europe \$42,565 \$40,083 2.3 5.0 242.0 20.0 26.0 3450.0 429 Volvo S80 T6 4dr Europe \$45,210 \$42,573 2.9 6.0 268.0 19.0 26.0 3653.0 Sedan Front 170.0 22.0 2822.0 430 Volvo V40 Wagon Europe Front \$26,135 \$24,641 1.9 4.0 29.0

3.5

2.0

\$33,337

\$24,647

1 Acura RSX Type S 2dr Sedan Front \$23,820 \$21,761 2.0 4.0 Asia

Type

SUV

Sedan Europe

Sports Europe

SUV Europe

Sedan Europe

Sedan Europe

Wagon Europe

Model

MDX

MSRP

\$36,945

\$33,337

Origin DriveTrain

Asia

15 Audi A4 3.0 Quattro convertible 2dr Sedan Europe 17 Audi A6 4.2 Quattro 4dr Sedan Europe

A8 L Quattro 4dr

RS 6 4dr

Touareg V6

Phaeton 4dr

Passat W8

Phaeton W12 4dr

• remove all the records (rows) where weight is above 4000.

SUV

Model

MDX

4638.0 420 XC90 T6 All \$41,250 \$38,851 2.9 6.0 268.0 15.0 Volvo SUV Europe 20.0 103 rows × 15 columns

\$40,235

MDX **0** Acura SUV Asia All \$36,945 \$33,337 3.5 6.0 1 Acura RSX Type S 2dr Sedan Front \$23,820 \$21,761 2.0 4.0 Asia

MSRP

\$43,755

Front

\$39,014

Origin DriveTrain

5)instruction (applying function on a column)

car

3.5 RL 4dr

car['MPG_Highway']= car['MPG_Highway'].apply(lambda x:x+3)

Sedan

Asia

• increase all the value of 'MPG_Highway' column by 3.

Type

Model

Make Model Type Origin DriveTrain MSRP Invoice EngineSize Cylinders Horsepower MPG_City MPG_Highway Weight Wheelbase Length \$33,337 0 Acura MDX SUV All \$36,945 3.5 6.0 265.0 25.0 26.0 4451.0 106.0 Asia 2.0 4.0 200.0 32.0 34.0 2778.0 1 Acura RSX Type S 2dr Sedan Asia Front \$23,820 \$21,761 101.0

172.0 TSX 4dr Front \$26,990 \$24,647 2.4 4.0 200.0 30.0 32.0 3230.0 105.0 183.0 2 Acura Sedan Asia 3 Acura TL 4dr Sedan Asia Front \$33,195 \$30,299 3.2 6.0 270.0 28.0 31.0 3575.0 108.0 186.0

Volvo C70 LPT convertible 2dr 197.0 3450.0 \$40,565 \$38,203 2.4 5.0 29.0 31.0 105.0 186.0 427 Sedan Europe Front \$42,565 \$40,083 2.3 5.0 242.0 28.0 3450.0 428 Volvo C70 HPT convertible 2dr Sedan Europe 29.0 105.0 186.0 Front 429 Volvo S80 T6 4dr Sedan Europe \$45,210 \$42,573 2.9 6.0 268.0 27.0 29.0 3653.0 110.0 190.0 430 Volvo V40 Wagon \$26,135 \$24,641 1.9 4.0 170.0 30.0 32.0 2822.0 101.0 180.0 Europe XC70 Wagon 2.5 208.0 30.0 3823.0 431 Volvo Europe All \$35,145 \$33,112 5.0 28.0 109.0 186.0

3.5