Cars EDA

September 20, 2023

0.1 ## Cars Data Analysis by Chidi Ikebude.

```
[1]: import pandas as pd
    import numpy as np
[2]: pip install pandoc
    Collecting pandoc
      Downloading pandoc-2.3.tar.gz (33 kB)
      Preparing metadata (setup.py): started
      Preparing metadata (setup.py): finished with status 'done'
    Collecting plumbum (from pandoc)
      Downloading plumbum-1.8.2-py3-none-any.whl (127 kB)
                                                0.0/127.0 kB ? eta -:--:--
                                                10.2/127.0 kB ? eta -:--:--
                                             41.0/127.0 kB 219.4 kB/s eta 0:00:01
                                             61.4/127.0 kB 273.8 kB/s eta 0:00:01
                                             61.4/127.0 kB 273.8 kB/s eta 0:00:01
         -----
         _____
                                             61.4/127.0 kB 273.8 kB/s eta 0:00:01
                                             61.4/127.0 kB 273.8 kB/s eta 0:00:01
         -----
                                             61.4/127.0 kB 273.8 kB/s eta 0:00:01
         _____
                                             61.4/127.0 kB 273.8 kB/s eta 0:00:01
         ______
                                             71.7/127.0 kB 106.3 kB/s eta 0:00:01
         _____
                                             71.7/127.0 kB 106.3 kB/s eta 0:00:01
         -----
                                             71.7/127.0 kB 106.3 kB/s eta 0:00:01
                                             71.7/127.0 kB 106.3 kB/s eta 0:00:01
                                             71.7/127.0 kB 106.3 kB/s eta 0:00:01
                                             71.7/127.0 kB 106.3 kB/s eta 0:00:01
                                     ----- 127.0/127.0 kB 141.1 kB/s eta 0:00:00
    Requirement already satisfied: ply in c:\users\user\anaconda3\lib\site-packages
    (from pandoc) (3.11)
```

Requirement already satisfied: pywin32 in c:\users\user\appdata\roaming\python\python311\site-packages (from plumbum->pandoc) (306) Building wheels for collected packages: pandoc Building wheel for pandoc (setup.py): started Building wheel for pandoc (setup.py): finished with status 'done' Created wheel for pandoc: filename=pandoc-2.3-py3-none-any.whl size=33290 Stored in directory: c:\users\user\appdata\local\pip\cache\wheels\1c\a9\c4\625 4542c4e8202d52fcd69798d2507aaad1f2a4bb60f2f0fea Successfully built pandoc Installing collected packages: plumbum, pandoc Successfully installed pandoc-2.3 plumbum-1.8.2 Note: you may need to restart the kernel to use updated packages. [4]: import plotly.io as pio pio.renderers.default = 'notebook' []: [12]: pip install nbconvert>=6.4.4 Note: you may need to restart the kernel to use updated packages. [13]: pip install jedi>=0.17.2 Note: you may need to restart the kernel to use updated packages. 0.2 Working on Real Project With Python The Cars Dataset [3]: cars = pd.read_csv(r"C:\Users\user\Projects\Data-Analysis-Projects\Cars Data_ ⇔Analysis\CarsDataset.csv") [4]: cars [4]: \ Make Model Type Origin DriveTrain **MSRP** 0 MDX SUV Asia All \$36,945 Acura 1 Acura RSX Type S 2dr Sedan Asia Front \$23,820 2 Acura TSX 4dr Sedan Asia Front \$26,990 3 Acura TL 4dr Sedan Front \$33,195 Asia Acura 3.5 RL 4dr Sedan Front \$43,755 Asia 423 Volvo C70 LPT convertible 2dr Sedan Front \$40,565 Europe 424 Volvo C70 HPT convertible 2dr Sedan Europe Front \$42,565 425 Volvo S80 T6 4dr Sedan Europe Front \$45,210

Wagon

XC70 Wagon Europe

Europe

V40

Front \$26,135

All \$35,145

426 Volvo

427 Volvo

	Invoice	e EngineSize	Cylinders	Horsepower	$\mathtt{MPG_City}$	MPG_Highway	\
0	\$33,337	3.5	6.0	265	17	23	
1	\$21,761	2.0	4.0	200	24	31	
2	\$24,647	2.4	4.0	200	22	29	
3	\$30,299	3.2	6.0	270	20	28	
4	\$39,014	3.5	6.0	225	18	24	
	•••	•••	•••		•••		
423	\$38,203	2.4	5.0	197	21	28	
424	\$40,083	2.3	5.0	242	20	26	
425	\$42,573	2.9	6.0	268	19	26	
426	\$24,641	1.9	4.0	170	22	29	
427	\$33,112	2.5	5.0	208	20	27	
	Weight	Wheelbase Le	ength				
0	4451	106	189				
1	2778	101	172				
2	3230	105	183				
3	3575	108	186				
4	3880	115	197				
	•••						
423	3450	105	186				
424	3450	105	186				
425	3653	110	190				
426	2822	101	180				

[428 rows x 15 columns]

0.3 How to Analyze DataFrames

Lets look at the description of the dataframe

5]: ca	rs.head(()							
5]:	Make		Model	Туре	Origin	n DriveTra	in MSRP	Invoice	. \
0	Acura		MDX	SUV	Asia	ı A	11 \$36,945	\$33,337	
1	Acura	RSX	Type S 2dr	Sedan	Asia	ı Fro	nt \$23,820	\$21,761	
2	Acura		TSX 4dr	Sedan	Asia	ı Fro	nt \$26,990	\$24,647	
3	Acura		TL 4dr	Sedan	Asia	ı Fro	nt \$33,195	\$30,299	
4	Acura		3.5 RL 4dr	Sedan	Asia	r Fro	nt \$43,755	\$39,014	
	EngineS	Size	Cylinders	Horse	power	MPG_City	MPG_Highway	Weight	\
0		3.5	6.0		265	17	23	4451	
1		2.0	4.0		200	24	31	2778	
2		2.4	4.0		200	22	29	3230	
3		3.2	6.0		270	20	28	3575	

4 3.5 6.0 225 18 24 3880 Wheelbase Length

0 106 189 1 101 172 2 105 183 3 108 186 4 115 197

Lets look at the total number of rows and Columns

[6]: cars.shape

[6]: (428, 15)

To look at all the codes information about the whole data feame we use .info

[7]: cars.info

Г т П .	41		. J. D. + . F.				M = 1= =			M - J - J	Т	
[7]:			od DataFi				Make			Model	Ty]	pe
	_	in Driv	errain	MSF	RP \	MDV	OIII.	۸	_ ^ 7	ነገ ውስራ (245	
	0	Acura				MDX	SUV	Asi				
	1	Acura		RSX 1	Type S		Sedan	Asi				
	2	Acura				4dr	Sedan	Asi				
	3	Acura				4dr	Sedan	Asi	a Fron	nt \$33,1	195	
	4	Acura		3	3.5 RL	4dr	Sedan	Asi	a Fron	nt \$43,7	755	
		•••			•••	•••	•••					
	423	Volvo	C70 LPT	conver	rtible	2dr	Sedan	Europ	e Fron	nt \$40,	565	
	424	Volvo	C70 HPT	conver	rtible	2dr	Sedan	Europ	e Fron	nt \$42,	565	
	425	Volvo		5	380 T6	4dr	Sedan	Europ	e Fron	nt \$45,2	210	
	426	Volvo				V40	Wagon	Europ	e Fron	nt \$26,3	135	
	427	Volvo			2	XC70	Wagon	Europ	e Al	ll \$35,	145	
		Invoi	ce Engi	neSize	Cyli	nders	Horse	power	MPG_City	MPG_High	nway	\
	0	\$33,33	7	3.5		6.0		265	17		23	
	1	\$21,76	1	2.0		4.0		200	24		31	
	2	\$24,64	7	2.4		4.0		200	22		29	
	3	\$30,29	9	3.2		6.0		270	20		28	
	4	\$39,01	4	3.5		6.0		225	18		24	
							•••	•••	•••			
	423	\$38,20	3	2.4		5.0		197	21		28	
	424	\$40,083		2.3		5.0		242	20		26	
	425	\$42,57		2.9		6.0		268	19		26	
	426	\$24,64		1.9		4.0		170	22		29	
	427	\$33,11		2.5		5.0		208	20		27	
	141	400,11	_	2.0		0.0		200	20		21	

Weight Wheelbase Length
0 4451 106 189

1	2778	101	172
2	3230	105	183
3	3575	108	186
4	3880	115	197
	•••		
423	3450	105	186
		-00	
424	3450	105	186
424 425	3450 3653		186 190
		105	

[428 rows x 15 columns]>

Now lets dive into analyzing the data in full

0.4 Lets run some Analysis to view somw certain things

- 1) Instruction (For Data Cleaning)
- Find all Null Values in the dataset. If there is any null value in any column, then fill it with the mean of that column.

```
[8]:
     cars.isnull().sum()
[8]: Make
                      0
     Model
                      0
     Type
                      0
                      0
     Origin
     DriveTrain
                      0
     MSRP
                      0
     Invoice
                      0
     EngineSize
                      0
                      2
     Cylinders
                      0
     Horsepower
     MPG_City
                      0
                      0
     MPG_Highway
                      0
     Weight
     Wheelbase
                      0
     Length
                      0
```

There are 2 null values in the Cylinders column, lets fix that up with the mean of the column Cylinders

```
[9]: cars['Cylinders'].fillna(cars['Cylinders'].mean(), inplace = True)
```

The inplace = TRUE makes the code be applicable to the column in the dataframe

2) Question (Based on Value Counts)

dtype: int64

• 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

[10]: cars.Make.value_counts()

[10]:	Toyota	28
	Chevrolet	27
	Mercedes-Benz	26
	Ford	23
	BMW	20
	Audi	19
	Honda	17
	Nissan	17
	Volkswagen	15
	Chrysler	15
	Dodge	13
	Mitsubishi	13
	Volvo	12
	Jaguar	12
	Hyundai	12
	Subaru	11
	Pontiac	11
	Mazda	11
	Lexus	11
	Kia	11
	Buick	9
	Mercury	9
	Lincoln	9
	Saturn	8
	Cadillac	8
	Suzuki	8
	Infiniti	8
	GMC	8
	Acura	7
	Porsche	7
	Saab	7
	Land Rover	3
	Oldsmobile	3
	Jeep	3
	Scion	2
	Isuzu	2
	MINI	2
	Hummer	1

Name: Make, dtype: int64

Toyota has the highest make with 27

3) Instruction (Filtering)

• Show all the records where Origin is Asia or Europe.

```
[11]: cars[(cars.Origin == 'Asia') | (cars.Origin == 'Europe')]
[11]:
             Make
                                        Model
                                                 Туре
                                                       Origin DriveTrain
                                                                                 MSRP
                                                                             $36,945
      0
                                          MDX
            Acura
                                                  SUV
                                                          Asia
                                                                       All
      1
                              RSX Type S 2dr
                                                                             $23,820
            Acura
                                                Sedan
                                                          Asia
                                                                     Front
      2
                                      TSX 4dr
                                                                     Front
                                                                             $26,990
            Acura
                                                Sedan
                                                          Asia
      3
                                       TL 4dr
                                                                             $33,195
            Acura
                                                Sedan
                                                          Asia
                                                                     Front
      4
                                  3.5 RL 4dr
                                                Sedan
                                                                     Front
                                                                             $43,755
            Acura
                                                          Asia
      423
           Volvo
                   C70 LPT convertible 2dr
                                                                     Front
                                                                             $40,565
                                                Sedan
                                                       Europe
      424
            Volvo
                   C70 HPT convertible 2dr
                                                Sedan
                                                       Europe
                                                                     Front
                                                                             $42,565
      425
           Volvo
                                  S80 T6 4dr
                                                       Europe
                                                                     Front
                                                                             $45,210
                                                Sedan
      426
           Volvo
                                          V40
                                                       Europe
                                                                     Front
                                                                             $26,135
                                                Wagon
      427
            Volvo
                                         XC70
                                                Wagon
                                                       Europe
                                                                       All
                                                                             $35,145
                                    Cylinders
                                                Horsepower
                                                              MPG_City
                                                                         MPG_Highway
             Invoice
                       EngineSize
      0
            $33,337
                               3.5
                                           6.0
                                                         265
                                                                     17
                                                                                    23
      1
            $21,761
                               2.0
                                           4.0
                                                         200
                                                                     24
                                                                                    31
      2
            $24,647
                               2.4
                                           4.0
                                                         200
                                                                     22
                                                                                    29
                                                         270
      3
                               3.2
                                           6.0
                                                                     20
                                                                                    28
            $30,299
      4
            $39,014
                               3.5
                                           6.0
                                                         225
                                                                     18
                                                                                    24
      . .
                 •••
                                                           •••
      423
            $38,203
                               2.4
                                           5.0
                                                         197
                                                                     21
                                                                                    28
      424
            $40,083
                               2.3
                                           5.0
                                                         242
                                                                     20
                                                                                    26
      425
            $42,573
                               2.9
                                           6.0
                                                         268
                                                                     19
                                                                                    26
      426
            $24,641
                                                                     22
                                                                                    29
                               1.9
                                           4.0
                                                         170
      427
            $33,112
                               2.5
                                           5.0
                                                         208
                                                                     20
                                                                                    27
            Weight
                     Wheelbase
                                 Length
      0
              4451
                            106
                                     189
      1
              2778
                            101
                                     172
      2
              3230
                            105
                                     183
      3
              3575
                            108
                                     186
      4
              3880
                                     197
                            115
               •••
      423
                            105
                                     186
              3450
      424
              3450
                            105
                                     186
      425
              3653
                                     190
                            110
      426
              2822
                            101
                                     180
      427
              3823
                            109
                                     186
      [281 rows x 15 columns]
      Total of 281 Rows. we could use another function
```

7

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

[12]:

[12]:		Make		Мо	odel	Туре	Origin	DriveTra	in MSRP	, /
	0	Acura			MDX	SUV	Asia	ı A.	11 \$36,945	
	1	Acura	RSX T	ype S	2dr	Sedan	Asia	ı Fron	nt \$23,820	
	2	Acura		TSX	4dr	Sedan	Asia	ı Fron	nt \$26,990	
	3	Acura		TL	4dr	Sedan	Asia	ı Fron	nt \$33,195	
	4	Acura	3	.5 RL	4dr	Sedan	Asia	ı Fron	nt \$43,755	
		•••		•••	•••	•••				
	423		O LPT conver			Sedan	Europe	e Fron	nt \$40,565	
	424	Volvo C7	O HPT conver	tible	2dr	Sedan	Europe	e Froi	nt \$42,565	
	425	Volvo	S	80 T6	4dr	Sedan	Europe	e Froi	nt \$45,210	
	426	Volvo			V40	Wagon	Europe	e Froi		
	427	Volvo		2	XC70	Wagon	Europe	e AI	11 \$35,145	
		Invoice	EngineSize	Cylin		Horse	-	•	MPG_Highway	
	0	\$33,337	3.5		6.0		265	17	23	
	1	\$21,761	2.0		4.0		200	24	31	
	2	\$24,647	2.4		4.0		200	22	29	
	3	\$30,299	3.2		6.0		270	20	28	;
	4	\$39,014	3.5		6.0		225	18	24	:
		•••	•••	•••		•••	•••	•••		
	423	\$38,203	2.4		5.0		197	21	28	
	424	\$40,083	2.3		5.0		242	20	26	j
	425	\$42,573	2.9		6.0		268	19	26	;
	426	\$24,641	1.9		4.0		170	22	29	l
	427	\$33,112	2.5		5.0		208	20	27	,
		0		ngth						
	0	4451	106	189						
	1	2778	101	172						
	2	3230	105	183						
	3	3575	108	186						
	4	3880	115	197						
		•••	•••							
	423	3450	105	186						
	424	3450	105	186						
	425	3653	110	190						
	426	2822	101	180						
	427	3823	109	186						

[281 rows x 15 columns]

They both bring the same data of 281 rows

- 4) Instruction (Removing unwanted records)
- Remove all the records (rows) where Weight is above 4000.

```
[13]: cars[~(cars['Weight'] > 4000)]
```

```
[13]:
            Make
                                      Model
                                               Type
                                                     Origin DriveTrain
                                                                              MSRP \
                             RSX Type S 2dr
      1
           Acura
                                              Sedan
                                                        Asia
                                                                   Front
                                                                          $23,820
      2
           Acura
                                    TSX 4dr
                                              Sedan
                                                                   Front
                                                                          $26,990
                                                        Asia
      3
           Acura
                                     TL 4dr
                                              Sedan
                                                        Asia
                                                                   Front
                                                                          $33,195
      4
                                 3.5 RL 4dr
                                                                          $43,755
           Acura
                                              Sedan
                                                        Asia
                                                                   Front
      5
                   3.5 RL w/Navigation 4dr
                                                                   Front
                                                                          $46,100
           Acura
                                              Sedan
                                                        Asia
      . .
                   C70 LPT convertible 2dr
      423
           Volvo
                                              Sedan
                                                      Europe
                                                                   Front
                                                                          $40,565
      424
           Volvo
                   C70 HPT convertible 2dr
                                                                          $42,565
                                              Sedan
                                                      Europe
                                                                   Front
      425
           Volvo
                                 S80 T6 4dr
                                              Sedan
                                                      Europe
                                                                   Front
                                                                          $45,210
      426 Volvo
                                                                   Front
                                         V40
                                              Wagon
                                                      Europe
                                                                          $26,135
      427 Volvo
                                        XC70
                                              Wagon
                                                      Europe
                                                                     All
                                                                          $35,145
            Invoice
                      EngineSize
                                   Cylinders
                                               Horsepower
                                                            MPG_City MPG_Highway
           $21,761
                                          4.0
      1
                              2.0
                                                       200
                                                                   24
                              2.4
                                          4.0
      2
           $24,647
                                                       200
                                                                   22
                                                                                 29
      3
           $30,299
                              3.2
                                          6.0
                                                       270
                                                                   20
                                                                                 28
      4
           $39,014
                              3.5
                                          6.0
                                                       225
                                                                   18
                                                                                 24
      5
           $41,100
                              3.5
                                          6.0
                                                       225
                                                                   18
                                                                                 24
      . .
                                                         •••
                                                                                 28
      423
           $38,203
                              2.4
                                          5.0
                                                       197
                                                                   21
      424
           $40,083
                              2.3
                                          5.0
                                                                   20
                                                                                 26
                                                       242
      425
           $42,573
                              2.9
                                          6.0
                                                       268
                                                                   19
                                                                                 26
      426
           $24,641
                              1.9
                                          4.0
                                                       170
                                                                   22
                                                                                 29
      427
           $33,112
                              2.5
                                          5.0
                                                       208
                                                                   20
                                                                                 27
           Weight
                    Wheelbase Length
                                   172
      1
              2778
                           101
      2
              3230
                           105
                                   183
      3
              3575
                           108
                                   186
      4
              3880
                           115
                                   197
      5
              3893
                           115
                                   197
               ...
      . .
      423
                           105
                                   186
              3450
      424
                           105
                                   186
              3450
      425
                                   190
              3653
                           110
      426
                           101
                                   180
              2822
      427
              3823
                           109
                                   186
```

[325 rows x 15 columns]

That could work. cars[cars.Weight <= 4000] could also work

[21]: cars[(cars['Weight'] <= 4000)]

[21]:	Make	Model	Туре	Origin	DriveTrain	MSRP	\
1	Acura	RSX Type S 2dr	Sedan	Asia	Front	\$23,820	
2	Acura	TSX 4dr	Sedan	Asia	Front	\$26,990	

```
3
     Acura
                                TL 4dr
                                         Sedan
                                                   Asia
                                                               Front
                                                                       $33,195
4
                            3.5 RL 4dr
                                                                       $43,755
     Acura
                                         Sedan
                                                   Asia
                                                               Front
5
     Acura
             3.5 RL w/Navigation 4dr
                                         Sedan
                                                   Asia
                                                               Front
                                                                       $46,100
. .
423
     Volvo
             C70 LPT convertible 2dr
                                                                       $40,565
                                         Sedan
                                                 Europe
                                                               Front
             C70 HPT convertible 2dr
424
     Volvo
                                         Sedan
                                                 Europe
                                                               Front
                                                                       $42,565
425
     Volvo
                            S80 T6 4dr
                                                                       $45,210
                                         Sedan
                                                 Europe
                                                               Front
426
     Volvo
                                    V40
                                         Wagon
                                                 Europe
                                                               Front
                                                                       $26,135
427
     Volvo
                                  XC70
                                         Wagon
                                                 Europe
                                                                 All
                                                                       $35,145
                                                        MPG_City
      Invoice
                EngineSize
                              Cylinders
                                          Horsepower
                                                                   MPG_Highway
1
     $21,761
                        2.0
                                     4.0
                                                  200
                                                               24
                        2.4
2
     $24,647
                                     4.0
                                                  200
                                                               22
                                                                             29
3
     $30,299
                        3.2
                                     6.0
                                                  270
                                                               20
                                                                             28
4
                        3.5
     $39,014
                                     6.0
                                                  225
                                                               18
                                                                             24
5
     $41,100
                        3.5
                                     6.0
                                                  225
                                                               18
                                                                             24
. .
                                                    •••
423
     $38,203
                        2.4
                                     5.0
                                                  197
                                                               21
                                                                             28
424
     $40,083
                        2.3
                                     5.0
                                                  242
                                                               20
                                                                             26
425
     $42,573
                        2.9
                                     6.0
                                                  268
                                                               19
                                                                             26
426
     $24,641
                                                               22
                        1.9
                                     4.0
                                                  170
                                                                             29
427
     $33,112
                        2.5
                                     5.0
                                                               20
                                                                             27
                                                  208
     Weight
              Wheelbase
                          Length
1
       2778
                     101
                              172
2
       3230
                     105
                              183
3
       3575
                     108
                              186
4
       3880
                              197
                     115
5
       3893
                     115
                              197
423
                              186
       3450
                     105
424
       3450
                     105
                              186
425
                              190
       3653
                     110
426
       2822
                     101
                              180
427
       3823
                     109
                              186
```

[325 rows x 15 columns]

- 5) Instruction (Applying function on a column)
- Increase all the values of 'MPG_City' column by 3.

```
cars['MPG_City'] = cars['MPG_City'].apply(lambda x:x+3)
[23]:
      cars
[23]:
            Make
                                      Model
                                                     Origin DriveTrain
                                                                             MSRP
                                              Туре
      0
           Acura
                                        MDX
                                               SUV
                                                       Asia
                                                                    All
                                                                         $36,945
```

1	Acura	RSX T	ype S 2dr	Sedan	Asia	a Fron	t \$23,820	
2	Acura		TSX 4dr	Sedan	Asia	a Fron	t \$26,990	
3	Acura		TL 4dr	Sedan	Asia	a Fron	t \$33,195	
4	Acura	3	.5 RL 4dr	Sedan	Asia	a Fron	t \$43,755	
	•••		•••			•••		
423	Volvo C	70 LPT conver	tible 2dr	Sedan	Europe	e Fron	t \$40,565	
424	Volvo C	70 HPT conver	tible 2dr	Sedan	Europe	e Fron	t \$42,565	
425	Volvo	S	80 T6 4dr	Sedan	Europe	e Fron	t \$45,210	
426	Volvo		V40	Wagon	Europe	e Fron	t \$26,135	
427	Volvo		XC70	Wagon	Europe	e Al	1 \$35,145	
	Invoice	EngineSize	Cylinders	Horse	epower	MPG_City	MPG_Highway	\
0	\$33,337	3.5	6.0		265	20	23	
1	\$21,761	2.0	4.0		200	27	31	
2	\$24,647	2.4	4.0		200	25	29	
3	\$30,299	3.2	6.0		270	23	28	
4	\$39,014	3.5	6.0		225	21	24	
	•••	***	•••	•••	•••	•••		
423	\$38,203	2.4	5.0		197	24	28	
424	\$40,083	2.3	5.0		242	23	26	
425	\$42,573	2.9	6.0		268	22	26	
426	\$24,641	1.9	4.0		170	25	29	
427	\$33,112	2.5	5.0		208	23	27	
	Weight	Wheelbase Le	ngth					
0	4451	106	189					
1	2778	101	172					
2	3230	105	183					
3	3575	108	186					
4	3880	115	197					
	•••							
423	3450	105	186					
424	3450	105	186					
425	3653	110	190					
426	2822	101	180					
427	3823	109	186					

[428 rows x 15 columns]