# Working on Real Project with Python

## (A part of Big Data Analysis)

## Cars Dataset

Here, The data of different cars is given with their specifications.

This data is available as a CSV file. We are going to analyze this data set using the Pandas DataFrame.

im	import pandas as pd										
	<pre>car = pd.read_csv(r"C:\Users\91999\Desktop\IT SPARKS\8. Project\7. Cars Data.csv")</pre>										
ca	r.head()										
МС	Make RP ∖	Model	Type 0	rigin	DriveTrai	n Unname	d: 5				
0	Acura	MDX	SUV	Asia	Al	ι	NaN	\$36,945			
1	Acura RS	X Type S 2dr	Sedan	Asia	Fron	t	NaN	\$23,820			
2	Acura	TSX 4dr	Sedan	Asia	Fron	t	NaN	\$26,990			
3	Acura	TL 4dr	Sedan	Asia	Fron	t	NaN	\$33,195			
4	Acura	3.5 RL 4dr	Sedan	Asia	Fron	t	NaN	\$43,755			
\	Invoice	EngineSize	Cylinder	s Hor	sepower I	MPG_City	MPG_	Highway			
ò	\$33,337	3.5	6.	0	265.0	17.0		23.0			
1	\$21,761	2.0	4.	0	200.0	24.0		31.0			
2	\$24,647	2.4	4.	0	200.0	22.0		29.0			
3	\$30,299	3.2	6.	0	270.0	20.0		28.0			
4	\$39,014	3.5	6.	0	225.0	18.0		24.0			

```
Wheelbase Length
   Unnamed: 13 Weight
0
           NaN
                4451.0
                             106.0
                                     189.0
1
           NaN
                2778.0
                             101.0
                                     172.0
2
                3230.0
                             105.0
                                     183.0
           NaN
3
                3575.0
                             108.0
                                     186.0
           NaN
4
                            115.0
                                     197.0
           NaN 3880.0
car.shape
(432, 17)
```

#### 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.

```
car.isnull().sum()
Make
                  4
Model
                 4
Type
                 4
Origin
DriveTrain
Unnamed: 5
               432
MSRP
                 4
Invoice
                 4
EngineSize
                 4
Cylinders
                 6
Horsepower
                 4
MPG_City
                 4
                 4
MPG Highway
Unnamed: 13
               432
Weight
Wheelbase
                 4
Length
dtype: int64
car['Cylinders'].fillna(car['Cylinders'].mean(), inplace = True)
```

#### 2) Question (Based on Value Counts)

• 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?

```
car.head(2)
```

\	Make			Мо	del	Туре	Origi	n Driv	/eTra	ain	M	ISRP	Invoice	е
0	Acura				MDX	SUV	Asi	а	Α	III	\$36,9	45	\$33,337	
1	Acura	RSX	Туре	S	2dr	Sedan	Asi	а	Fro	nt	\$23,8	320	\$21,761	
0 1	EngineS	Size 3.5 2.0	Cyli		ers 6.0 4.0	Horse	265 200	MPG_0	17 24	MP	G_High	1way 23 31	Weight 4451 2778	\
0		ase 106 101	Lengt 18 17	39										
ca	r['Make'	'].va	alue_c	ou	nts(	)								
Che Fo Me Fo BM A Hoi Che Ni C	di nda ssan rysler lkswager tsubishi dge guar lvo undai a ntiac baru zda xus rcury ick ncoln turn zuki dillac C finiti rsche	ı İ	28 27 26 23 20 19 17 17 15 13 13 12 12 12 11 11 11 11 11 11 11 11 11 11											

Land Rover 3
Scion 2
Isuzu 2
MINI 2
Hummer 1
Name: Make, dtype: int64

## 3) Instruction (Filtering)

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

ca	r.head(	2)									
	Make			Model	Туре	0rigi	n Driv	eTrain	MSRP	Invoic	e
0	Acura			MDX	SUV	Asi	La	All	\$36,945	\$33,337	
1	Acura	RSX	Туре	S 2dr	Sedan	Asi	La	Front	\$23,820	\$21,761	
	Engino	C i 70	Cv1 i	ndore	Horco	2011015	MDC C	i+v MD(	G Highway	Woight	\
0 1	Engine	3.5	Cyli	6.0 4.0	Horse	265 200	MPG_C	17 24	3_nigliway 23 31	Weight 4451 2778	\
0		ase 106 101	Lengt 18 17	9							
ca	r[car[ˈ	Origi	in'].i	sin([	'Asia',	'Euro	pe'])]				
MS	Mak RP ∖	е			Mo	odel	Type	Origin	DriveTrain	1	
0	Acur	a				MDX	SUV	Asia	All	\$36,9	45
1	Acur	a		RSX	Type S	2dr	Sedan	Asia	Front	\$23,8	20
2	Acur	a			TSX	4dr	Sedan	Asia	Front	\$26,9	90
3	Acur	a			TL	4dr	Sedan	Asia	Front	\$33,1	95
4	Acur	a			3.5 RL	4dr	Sedan	Asia	Front	\$43,7	55
		•									
42	3 Volv	o C7	70 LPT	conve	ertible	2dr	Sedan	Europe	Front	\$40,5	65
42	4 Volv	o C	70 HPT	conve	ertible	2dr	Sedan	Europe	Front	\$42,5	65
42	5 Volv	0			S80 T6	4dr	Sedan	Europe	Front	\$45,2	10

426	Volvo		V40	Wagon	Europe	Front	\$26,135
427	Volvo		XC70	Wagon	Europe	All	\$35,145
MDC		EngineSize	Cylinders	Horse	power MF	PG_City	
0 _	Highway \ \$33,337	3.5	6.0		265	17	
23 1	\$21,761	2.0	4.0		200	24	
31							
2 29	\$24,647	2.4	4.0		200	22	
3 28	\$30,299	3.2	6.0		270	20	
4	\$39,014	3.5	6.0		225	18	
24							
-	¢20, 202						
423 28	\$38,203	2.4	5.0		197	21	
424 26	\$40,083	2.3	5.0		242	20	
425	\$42,573	2.9	6.0		268	19	
26 426	\$24,641	1.9	4.0		170	22	
29 427					208	20	
27	\$33,112	2.5	5.0		200	20	
	Weight Whe	eelbase Ler	ngth				
0	4451	106	189				
1 2	2778 3230	101 105	172 183				
3 4	3575 3880	108 115	186 197				
423 424	3450 3450	105 105	186 186				
425	3653	110	190				
426 427	2822 3823	101 109	180 186				
[281	rows x 15 (	columnsl					
_	'Origin'].va	_	( )				
		a cuc_counts	( )				
Asia USA	158 147						

```
Europe 123
Name: Origin, dtype: int64

158 + 123

281
```

### 4) Instruction ( Removing unwanted records )

• Remove all the records (rows) where Weight is above 4000.

ca	r.head(	2)		·	•	J					
	Make		ŀ	1odel	Туре	0rigi	n Driv	eTrain	MSRP	Invoice	е
0	Acura			MDX	SUV	Asi	la	All	\$36,945	\$33,337	
1	Acura	RSX	Type S	3 2dr	Sedan	Asi	La	Front	\$23,820	\$21,761	
0 1	Engine	Size 3.5 2.0	Cylir	6.0 4.0	Horsep	265 200	MPG_C	ity MP0 17 24	G_Highway 23 31	Weight 4451 2778	\
0		ase 106 101	Length 189 172	)							
ca	r[~(car	['Wei	lght']	> 400							
Uni	Mak named:				Mo	odel	Type	Origin	DriveTrain	า	
1 Nal	Acur	-		RSX	Type S	2dr	Sedan	Asia	Front	t	
2	Acur	a			TSX	4dr	Sedan	Asia	Front	t	
Nal	N Acur	а			TI	4dr	Sedan	Asia	Front	Н	
Nal	V										
4 Nai	Acur N	a			3.5 RL	4dr	Sedan	Asia	Front	t	
5 Nal	Acur	a 3.	5 RL v	v/Navi	gation	4dr	Sedan	Asia	Fron	t	
421 Nal	7 Volv	o C7	0 LPT	conve	ertible	2dr	Sedan	Europe	Fron	t	
428 Nal	3 Volv	o C7	0 HPT	conve	ertible	2dr	Sedan	Europe	Front	t	
429		0			S80 T6	4dr	Sedan	Europe	Fron	t	

NaN 430	Volvo			V40	Wagon	Europe	e Front	
NaN					_			
431 NaN	Volvo			XC70	Wagon	Europe	e All	
	MSRP	Invoice	Engin	eSize	Cylind	lers Ho	orsepower	
MPG_ 1	City \ \$23,820	\$21,761		2.0		4.0	200.0	24.0
2	\$26,990	\$24,647		2.4		4.0	200.0	22.0
3	\$33,195	\$30,299		3.2		6.0	270.0	20.0
4	\$43,755	\$39,014		3.5		6.0	225.0	18.0
5	\$46,100	\$41,100		3.5		6.0	225.0	18.0
427	\$40,565	\$38,203		2.4		5.0	197.0	21.0
428	\$42,565	\$40,083		2.3		5.0	242.0	20.0
429	\$45,210	\$42,573		2.9		6.0	268.0	19.0
430	\$26,135	\$24,641		1.9		4.0	170.0	22.0
431	\$35,145	\$33,112		2.5		5.0	208.0	20.0
1	MPG_Highwa 31	•	d: 13 NaN	Weigh <sup>-</sup> 2778.0		lbase 101.0	Length 172.0	
2	29 28		NaN NaN	3230.0 3575.0		105.0 108.0	183.0 186.0	
4	24	.0	NaN	3880.0	9	115.0	197.0	
5	24	.⊍	NaN 	3893.0		115.0	197.0	
427 428	28 26		NaN NaN	3450.0 3450.0		105.0 105.0	186.0 186.0	
429	26	.0	NaN	3653.0	9	110.0	190.0	
430 431	29 27		NaN NaN	2822.0 3823.0		101.0 109.0	180.0 186.0	
[329	rows x 17	columns]						
428	- 103							
325								

## 5) Instruction ( Applying function on a column )

• Increase all the values of 'MPG\_City' column by 3.

	HICIEC	JSE all L	iic valt	103 01 1	0_0	-icy	cotun	iii by 3.			
ca	r.head(2	2)									
	Make		Ŋ	lodel	Ту	pe	Orig	in Dri	veTrain	Unnamed: 5	
	RP \ Acura			MDX	S	UV	As	ia	All	NaN	\$36,945
1	Acura	RSX T	vne (	S 2dr	Sed	an	Δς	ia	Front	NaN	\$23,820
_	Acura	113/1	урс	201	Sca	an	7.5	<b>1</b> 0	TTOTIC	Nan	Ψ25,020
	Invoid	ce En	gineS	Size	Cyli	nde	rs	Horsep	ower MP	G_City MPG	_Highway
0	\$33,337	7		3.5		6	.0	20	65.0	17.0	23.0
1	\$21,761			2.0			.0		00.0	24.0	31.0
_	\$21,701	<u>L</u>		2.0		4	.0	21	00.0	24.0	31.0
0 1	Unnamed	d: 13 NaN NaN	445]		1		0	ength 189.0 172.0			
са	r['MPG_(	City']	= (	car['N	1PG_C	ity	'].a	pply(l	ambda x:	x+ <mark>3</mark> )	
ca	r										
	Make	9				Мо	del	Tvpe	Origin	DriveTrain	
_	named: 5	5 \					MDX	SUV	-		
0 Nal											
1 Nal		a .		RSX	Type	S	2dr	Sedan	Asia	Front	
2	Acura	Э			T	SX	4dr	Sedan	Asia	Front	
Nal 3	N Acura	a				TL	4dr	Sedan	Asia	Front	
Nal 4	N Acura	a			3.5	RI	4dr	Sedan	Asia	Front	
Nal					3.5			Scaan	71514	110110	
		•									
42 <sup>°</sup> Nal		C70	LPT	conve	ertib	le	2dr	Sedan	Europe	Front	
42	8 Volvo	C70	HPT	conve	ertib	le	2dr	Sedan	Europe	Front	
Nal 42		)			S80 <sup>-</sup>	T6	4dr	Sedan	Europe	Front	
Nal 43	N								•		
/I - ( )	0 Volvo	)					V40	Wagon	Europe	Front	
Nal	N										
	1 Volvo	)				X	C70	Wagon	Europe	All	

MPG	MS ∖ City		nvoice	Engin	eSize	Cylinde	rs H	orsepowe	er	
0 _	\$36,94		3,337		3.5	6	.0	265	. 0	20.0
1	\$23,82	0 \$2	1,761		2.0	4	.0	200	. 0	27.0
2	\$26,99	0 \$2	4,647		2.4	4	. 0	200	. 0	25.0
3	\$33,19	5 \$3	0,299		3.2	6	. 0	270	. 0	23.0
4	\$43,75	5 \$3	9,014		3.5	6	.0	225	. 0	21.0
427	\$40,56	5 \$3	8,203		2.4	5	. 0	197	. 0	24.0
428	\$42,56	5 \$4	0,083		2.3	5	. 0	242	. 0	23.0
429	\$45,21	.0 \$4	2,573		2.9	6	. 0	268	. 0	22.0
430	\$26,13	5 \$2	4,641		1.9	4	. 0	170	. 0	25.0
431	\$35,14	5 \$3	3,112		2.5	5	. 0	208	. 0	23.0
0 1 2 3 4	MPG_Hi	ghway 23.0 31.0 29.0 28.0 24.0	Unnamed	d: 13 NaN NaN NaN NaN NaN	Weight 4451.0 2778.0 3230.0 3575.0 3880.0	) 10 ) 10 ) 10	base 06.0 01.0 05.0 08.0	Length 189.0 172.0 183.0 186.0 197.0		
427 428 429 430 431		28.0 26.0 26.0 29.0 27.0		NaN NaN NaN NaN NaN	3450.0 3450.0 3653.0 2822.0 3823.0	) 10 ) 10 ) 1	05.0 05.0 10.0 01.0	186.0 186.0 190.0 180.0		
[432	rows x	17 co	lumns]							