DSBDAL 02

Aditya T. Salagare

Roll no: 13320 (C1 Batch)

```
In [45]: import pandas as pd
  import numpy as np
  import matplotlib.pyplot as plt
```

In [46]: df=pd.read_csv("C:/Users/Welcome/Downloads/DSBDA02.csv")
df

Out[46]:		Math_Score	Reading_Score	Writing_Score	Placement_Score	Club_Join_Date	Placement_Offer_
	0	72.0	86	64.0	75.0	2020.0	
	1	65.0	91	67.0	111.0	2020.0	
	2	66.0	81	88.0	95.0	NaN	
	3	59.0	77	64.0	90.0	2000.0	
	4	66.0	89	76.0	81.0	2018.0	
	5	61.0	92	76.0	82.0	2019.0	
	6	NaN	93	72.0	66.0	2018.0	
	7	85.0	85	59.0	93.0	2019.0	
	8	73.0	95	68.0	NaN	NaN	
	9	65.0	81	67.0	86.0	2023.0	
	10	79.0	91	NaN	100.0	2018.0	
	11	75.0	80	77.0	89.0	2018.0	
	12	78.0	81	73.0	69.0	2002.0	
	13	65.0	93	76.0	NaN	2019.0	
	14	81.0	84	17.0	76.0	2021.0	
	15	62.0	80	64.0	81.0	NaN	
	16	70.0	77	68.0	86.0	2021.0	
	17	75.0	86	70.0	76.0	2020.0	
	18	71.0	95	55.0	93.0	2011.0	
	19	NaN	79	70.0	NaN	2021.0	
	20	70.0	86	71.0	94.0	2018.0	
	21	78.0	92	67.0	78.0	NaN	
	22	66.0	82	74.0	75.0	2019.0	
	23	68.0	77	77.0	100.0	2022.0	
	24	65.0	75	NaN	101.0	2021.0	
	25	55.0	89	73.0	91.0	2018.0	
	26	73.0	88	79.0	77.0	2020.0	
	27	80.0	80	68.0	83.0	2019.0	
	28	74.0	92	60.0	88.0	2019.0	
	29	75.0	78	66.0	97.0	2021.0	
4							•
In [47]:	df.c	columns					
Out[47]:	Inde		oin_Date', 'P		'Writing_Score r_Count'],	e', 'Placement	t_Score',

In [48]: df.isnull()

Out[48]: Math_Score Reading_Score Writing_Score Placement_Score Club_Join_Date Placement_Offer_ 0 False False False False False 1 False False False False False 2 False False False False True 3 False False False False False 4 False False False False False 5 False False False False False 6 True False False False False 7 False False False False False 8 False False False True True 9 False False False False **False** 10 False False True False False 11 False False **False False False** 12 False False False False False 13 False False False True False 14 False False False False False 15 False False False False True 16 False False False False False 17 False False False False False 18 False False False False False 19 True False False True False 20 False False False False False 21 False False False False True 22 False False False False False 23 False False False False False 24 False False True False False 25 False **False** False False False 26 False False False False False 27 False False False False False 28 False False False False False 29 False False False False False series = pd.isnull(df['Math_Score ']) In [49]:

df[series]

Out[49]:		Math_Score	Reading_Score	Writing_Score	Placement_Score	Club_Join_Date	Placement_Offer_
	6	NaN	93	72.0	66.0	2018.0	
	19	NaN	79	70.0	NaN	2021.0	
4)
In [50]:	df.	notnull()					

DSBDA 02 13/02/2024, 15:15

Out[50]:		Math_Score	Reading_Score	Writing_Score	Placement_Score	Club_Join_Date	Placement_Offer_
	0	True	True	True	True	True	
	1	True	True	True	True	True	
	2	True	True	True	True	False	
	3	True	True	True	True	True	
	4	True	True	True	True	True	
	5	True	True	True	True	True	
	6	False	True	True	True	True	
	7	True	True	True	True	True	
	8	True	True	True	False	False	
	9	True	True	True	True	True	
	10	True	True	False	True	True	
	11	True	True	True	True	True	
	12	True	True	True	True	True	
	13	True	True	True	False	True	
	14	True	True	True	True	True	
	15	True	True	True	True	False	
	16	True	True	True	True	True	
	17	True	True	True	True	True	
	18	True	True	True	True	True	
	19	False	True	True	False	True	
	20	True	True	True	True	True	
	21	True	True	True	True	False	
	22	True	True	True	True	True	
	23	True	True	True	True	True	
	24	True	True	False	True	True	
	25	True	True	True	True	True	
	26	True	True	True	True	True	
	27	True	True	True	True	True	
	28	True	True	True	True	True	
	29	True	True	True	True	True	
4							+
In [51]:		ries1 = pd.r series1]	notnull(df['Ma	th_Score ']			

df[series1]

Out[51]:		Math_Score	Reading Score	Writing Score	Discoment Score	Chale Join Date	
			J	wiiding_score	riaceillelit_3core	Club_Join_Date	Placement_Offer_
	0	72.0	86	64.0	75.0	2020.0	
	1	65.0	91	67.0	111.0	2020.0	
	2	66.0	81	88.0	95.0	NaN	
	3	59.0	77	64.0	90.0	2000.0	
	4	66.0	89	76.0	81.0	2018.0	
	5	61.0	92	76.0	82.0	2019.0	
	7	85.0	85	59.0	93.0	2019.0	
	8	73.0	95	68.0	NaN	NaN	
	9	65.0	81	67.0	86.0	2023.0	
	10	79.0	91	NaN	100.0	2018.0	
	11	75.0	80	77.0	89.0	2018.0	
	12	78.0	81	73.0	69.0	2002.0	
	13	65.0	93	76.0	NaN	2019.0	
	14	81.0	84	17.0	76.0	2021.0	
	15	62.0	80	64.0	81.0	NaN	
	16	70.0	77	68.0	86.0	2021.0	
	17	75.0	86	70.0	76.0	2020.0	
	18	71.0	95	55.0	93.0	2011.0	
	20	70.0	86	71.0	94.0	2018.0	
	21	78.0	92	67.0	78.0	NaN	
	22	66.0	82	74.0	75.0	2019.0	
	23	68.0	77	77.0	100.0	2022.0	
	24	65.0	75	NaN	101.0	2021.0	
	25	55.0	89	73.0	91.0	2018.0	
	26	73.0	88	79.0	77.0	2020.0	
	27	80.0	80	68.0	83.0	2019.0	
	28	74.0	92	60.0	88.0	2019.0	
	29	75.0	78	66.0	97.0	2021.0	
4)

In [53]: missing_values = ["Na", "na"]
 df = pd.read_csv("C:/Users/Welcome/Downloads/DSBDA02.csv", na_values =missing_value
 df

72021, 10.10							
Out[53]:		Math_Score	Reading_Score	Writing_Score	Placement_Score	Club_Join_Date	Placement_Offer_
	0	72.0	86	64.0	75.0	2020.0	
	1	65.0	91	67.0	111.0	2020.0	
	2	66.0	81	88.0	95.0	NaN	
	3	59.0	77	64.0	90.0	2000.0	
	4	66.0	89	76.0	81.0	2018.0	
	5	61.0	92	76.0	82.0	2019.0	
	6	NaN	93	72.0	66.0	2018.0	
	7	85.0	85	59.0	93.0	2019.0	
	8	73.0	95	68.0	NaN	NaN	
	9	65.0	81	67.0	86.0	2023.0	
	10	79.0	91	NaN	100.0	2018.0	
	11	75.0	80	77.0	89.0	2018.0	
	12	78.0	81	73.0	69.0	2002.0	
	13	65.0	93	76.0	NaN	2019.0	
	14	81.0	84	17.0	76.0	2021.0	
	15	62.0	80	64.0	81.0	NaN	
	16	70.0	77	68.0	86.0	2021.0	
	17	75.0	86	70.0	76.0	2020.0	
	18	71.0	95	55.0	93.0	2011.0	
	19	NaN	79	70.0	NaN	2021.0	
	20	70.0	86	71.0	94.0	2018.0	
	21	78.0	92	67.0	78.0	NaN	
	22	66.0	82	74.0	75.0	2019.0	
	23	68.0	77	77.0	100.0	2022.0	
	24	65.0	75	NaN	101.0	2021.0	
	25	55.0	89	73.0	91.0	2018.0	
	26	73.0	88	79.0	77.0	2020.0	
	27	80.0	80	68.0	83.0	2019.0	
	28	74.0	92	60.0	88.0	2019.0	
	29	75.0	78	66.0	97.0	2021.0	
4							>
In [54]:	d - F -	df.head(11)					
In [55]:		=df -fillna(0)					

localhost:8888/nbconvert/html/TE 13320/DSBDA 02.ipynb?download=false

	ivia	tn_Score	Reading_Score	Writing_Score	Placement_Score	Club_Join_Date	Placellielit_O
	0	72.0	86	64.0	75.0	2020.0	
	1	65.0	91	67.0	111.0	2020.0	
	2	66.0	81	88.0	95.0	0.0	
	3	59.0	77	64.0	90.0	2000.0	
	4	66.0	89	76.0	81.0	2018.0	
	5	61.0	92	76.0	82.0	2019.0	
	6	0.0	93	72.0	66.0	2018.0	
	7	85.0	85	59.0	93.0	2019.0	
	8	73.0	95	68.0	0.0	0.0	
	9	65.0	81	67.0	86.0	2023.0	
	10	79.0	91	0.0	100.0	2018.0	
i [57]:	df	th_Score rs\Welcom				8491271.pv:2:	SettingWith
[5/]:	C:\Use pyWarn: A value See the e/user	rs\Welcom ing: e is tryi e caveats _guide/in	e\AppData\Lo ng to be set in the docu dexing.html#	cal\Temp\ipyk on a copy of	ernel_10212\234 a slice from a ttps://pandas.py	a DataFrame vdata.org/pand	-
	C:\Use pyWarn: A value See the e/user_ df['I	rs\Welcom ing: e is tryi e caveats _guide/in Math_Scor	e\AppData\Lo ng to be set in the docu dexing.html# e '].fillna	cal\Temp\ipyk on a copy of mentation: ht returning-a-v (value=m_v, i	ernel_10212\234 a slice from a ttps://pandas.py	n DataFrame vdata.org/pand opy	as-docs/stal
	C:\Use pyWarn: A value See the e/user_ df['I	rs\Welcom ing: e is tryi e caveats _guide/in Math_Scor	e\AppData\Lo ng to be set in the docu dexing.html# e '].fillna	cal\Temp\ipyk on a copy of mentation: ht returning-a-v (value=m_v, i	ernel_10212\234 a slice from a tps://pandas.py iew-versus-a-co nplace=True)	n DataFrame vdata.org/pand opy	as-docs/stal
	C:\Use pyWarn: A value See the e/user df['I	rs\Welcom ing: e is tryi e caveats _guide/in Math_Score	e\AppData\Lo ng to be set in the docu dexing.html# e '].fillna Reading_Score	cal\Temp\ipyk on a copy of mentation: ht returning-a-v (value=m_v, i Writing_Score	ernel_10212\234 a slice from a ttps://pandas.py iew-versus-a-co nplace=True) Placement_Score	DataFrame vdata.org/pand ppy Club_Join_Date	as-docs/stal
	C:\Use pyWarn: A value See the e/user df['I Ma	rs\Welcom ing: e is tryi e caveats _guide/in Math_Scor th_Score	e\AppData\Lo ng to be set in the docu dexing.html# e '].fillna Reading_Score	cal\Temp\ipyk on a copy of mentation: ht returning-a-v (value=m_v, i Writing_Score	ernel_10212\234 a slice from a ttps://pandas.py iew-versus-a-co nplace=True) Placement_Score	DataFrame vdata.org/pand opy Club_Join_Date 2020.0	as-docs/stal
t[57]:	C:\Use pyWarn: A value See the e/user df['I Ma	rs\Welcom ing: e is tryi e caveats _guide/in Math_Score 72.0 65.0	e\AppData\Lo ng to be set in the docu dexing.html# e '].fillna Reading_Score 86 91	cal\Temp\ipyk on a copy of mentation: ht returning-a-v (value=m_v, i Writing_Score 64.0 67.0	ternel_10212\234 a slice from a stps://pandas.pyriew-versus-a-conplace=True) Placement_Score 75.0	DataFrame vdata.org/pand opy Club_Join_Date 2020.0 2020.0	as-docs/stal
	C:\Use pyWarn: A value See the e/user df['I Ma	rs\Welcom ing: e is tryi e caveats _guide/in Math_Score 72.0 65.0 66.0	e\AppData\Lo ng to be set in the docu dexing.html# e '].fillna Reading_Score 86 91 81	cal\Temp\ipyk on a copy of mentation: ht returning-a-v (value=m_v, i Writing_Score 64.0 67.0 88.0	ternel_10212\234 a slice from a stps://pandas.pyiew-versus-a-conplace=True) Placement_Score 75.0 111.0 95.0	DataFrame vdata.org/pand opy Club_Join_Date 2020.0 2020.0 NaN	as-docs/stal
	C:\Use pyWarn: A value See the e/user df['I Ma 0 1 2 3	rs\Welcom ing: e is tryi e caveats _guide/in Math_Score 72.0 65.0 66.0 59.0	e\AppData\Lo ng to be set in the docu dexing.html# e '].fillna Reading_Score 86 91 81 77	cal\Temp\ipyk on a copy of mentation: ht returning-a-v (value=m_v, i Writing_Score 64.0 67.0 88.0 64.0	ternel_10212\234 a slice from a stps://pandas.pyiew-versus-a-conplace=True) Placement_Score 75.0 111.0 95.0 90.0	Club_Join_Date 2020.0 NaN 2000.0	as-docs/stal
	C:\Use pyWarn: A value See the e/user df['I Ma 0 1 2 3	rs\Welcom ing: e is tryi e caveats _guide/in Math_Score 72.0 65.0 66.0 59.0 66.0	e\AppData\Lo ng to be set in the docu dexing.html# e '].fillna Reading_Score 86 91 81 77 89	cal\Temp\ipyk on a copy of mentation: ht returning-a-v (value=m_v, i Writing_Score 64.0 67.0 88.0 64.0 76.0	ternel_10212\234 a slice from a stps://pandas.py riew-versus-a-co nplace=True) Placement_Score 75.0 111.0 95.0 90.0 81.0	DataFrame data.org/pand opy Club_Join_Date 2020.0 2020.0 NaN 2000.0 2018.0	as-docs/stal
	C:\Use pyWarn: A value See the e/user df['I Ma 0 1 2 3 4 5	rs\Welcom ing: e is tryi e caveats _guide/in Math_Score 72.0 65.0 66.0 59.0 66.0 61.0	e\AppData\Lo ng to be set in the docu dexing.html# e '].fillna Reading_Score 86 91 81 77 89 92	cal\Temp\ipyk on a copy of mentation: ht returning-a-v (value=m_v, i Writing_Score 64.0 67.0 88.0 64.0 76.0	ternel_10212\234 a slice from a stps://pandas.py riew-versus-a-co nplace=True) Placement_Score 75.0 111.0 95.0 90.0 81.0 82.0	Club_Join_Date 2020.0 2020.0 NaN 2000.0 2018.0 2019.0	as-docs/stal
	C:\Use pyWarn: A value See the e/user df['I Ma 0 1 2 3 4 5	rs\Welcom ing: e is tryi e caveats guide/in Math_Score 72.0 65.0 66.0 59.0 66.0 61.0 69.1	e\AppData\Lo ng to be set in the docu dexing.html# e '].fillna Reading_Score 86 91 81 77 89 92 93	cal\Temp\ipyk on a copy of mentation: ht returning-a-v (value=m_v, i Writing_Score 64.0 67.0 88.0 64.0 76.0 76.0 72.0	rernel_10212\234 a slice from a stps://pandas.py riew-versus-a-co nplace=True) Placement_Score 75.0 111.0 95.0 90.0 81.0 82.0 66.0	Club_Join_Date 2020.0 2020.0 NaN 2000.0 2018.0 2018.0	as-docs/stal
	C:\Use pyWarn: A value See the e/user df['I Ma 0 1 2 3 4 5 6 7	rs\Welcom ing: e is tryi e caveats guide/in Math_Score 72.0 65.0 66.0 59.0 66.0 61.0 69.1 85.0	e\AppData\Lo ng to be set in the docu dexing.html# e '].fillna Reading_Score 86 91 81 77 89 92 93 85	cal\Temp\ipyk on a copy of mentation: ht returning-a-v (value=m_v, i Writing_Score 64.0 67.0 88.0 64.0 76.0 76.0 72.0	rernel_10212\234 a slice from a stps://pandas.py riew-versus-a-conplace=True) Placement_Score 75.0 111.0 95.0 90.0 81.0 82.0 66.0 93.0	Club_Join_Date 2020.0 2020.0 NaN 2000.0 2018.0 2019.0	as-docs/stal
	C:\Use pyWarn: A value See the e/user df['I' Ma 0 1 2 3 4 5 6 7 8	rs\Welcom ing: e is tryi e caveats guide/in Math_Score 72.0 65.0 66.0 59.0 66.0 61.0 69.1 85.0 73.0	e\AppData\Lo ng to be set in the docu dexing.html# e '].fillna Reading_Score 86 91 81 77 89 92 93 85 95	cal\Temp\ipyk on a copy of mentation: ht returning-a-v (value=m_v, i Writing_Score 64.0 67.0 88.0 64.0 76.0 76.0 75.0 68.0	rernel_10212\234 Fa slice from a stps://pandas.py riew-versus-a-complace=True) Placement_Score 75.0 111.0 95.0 90.0 81.0 82.0 66.0 93.0 NaN	Club_Join_Date 2020.0 2020.0 NaN 2000.0 2018.0 2019.0 2018.0 2019.0 NaN	as-docs/stal

Out[58]:							
our[30]:		Math_Score	Reading_Score	Writing_Score	Placement_Score	Club_Join_Date	Placement_Offer
	0	72.0	86	64.0	75.0	2020.0	
	1	65.0	91	67.0	111.0	2020.0	
	2	66.0	81	88.0	95.0	-99.0	
	3	59.0	77	64.0	90.0	2000.0	
	4	66.0	89	76.0	81.0	2018.0	
	5	61.0	92	76.0	82.0	2019.0	
	6	69.1	93	72.0	66.0	2018.0	
	7	85.0	85	59.0	93.0	2019.0	
	8	73.0	95	68.0	-99.0	-99.0	
	9	65.0	81	67.0	86.0	2023.0	
	10	79.0	91	-99.0	100.0	2018.0	
							•
In [59]:	ndf	dropna()					
111 1771:		· uropha()					
							DI
		Math_Score			Placement_Score		Placement_Offer_
	0	Math_Score	86	64.0	75.0	2020.0	Placement_Offer_
	0	Math_Score 72.0 65.0	86 91	64.0 67.0	75.0 111.0	2020.0	Placement_Offer_
Out[59]:	0 1 3	72.0 65.0 59.0	86 91 77	64.0 67.0 64.0	75.0 111.0 90.0	2020.0 2020.0 2000.0	Placement_Offer_
	0 1 3 4	72.0 65.0 59.0 66.0	86 91 77 89	64.0 67.0 64.0 76.0	75.0 111.0 90.0 81.0	2020.0 2020.0 2000.0 2018.0	Placement_Offer_
	0 1 3	72.0 65.0 59.0 66.0 61.0	86 91 77 89 92	64.0 67.0 64.0 76.0	75.0 111.0 90.0 81.0 82.0	2020.0 2020.0 2000.0 2018.0 2019.0	Placement_Offer_
	0 1 3 4 5	72.0 65.0 59.0 66.0 61.0	86 91 77 89 92 93	64.0 67.0 64.0 76.0 76.0	75.0 111.0 90.0 81.0 82.0 66.0	2020.0 2020.0 2000.0 2018.0 2019.0 2018.0	Placement_Offer_o
	0 1 3 4 5	72.0 65.0 59.0 66.0 61.0	86 91 77 89 92	64.0 67.0 64.0 76.0	75.0 111.0 90.0 81.0 82.0	2020.0 2020.0 2000.0 2018.0 2019.0	Placement_Offer_0
	0 1 3 4 5	72.0 65.0 59.0 66.0 61.0	86 91 77 89 92 93	64.0 67.0 64.0 76.0 76.0	75.0 111.0 90.0 81.0 82.0 66.0	2020.0 2020.0 2000.0 2018.0 2019.0 2018.0	Placement_Offer_

localhost:8888/nbconvert/html/TE 13320/DSBDA 02.ipynb?download=false

N				DSBDA 02		
	/lath_Score	Reading_Score	Writing_Score	Placement_Score	Club_Join_Date	Placement_Offer_
0	72.0	86	64.0	75.0	2020.0	
1	65.0	91	67.0	111.0	2020.0	
2	66.0	81	88.0	95.0	NaN	
3	59.0	77	64.0	90.0	2000.0	
4	66.0	89	76.0	81.0	2018.0	
5	61.0	92	76.0	82.0	2019.0	
6	69.1	93	72.0	66.0	2018.0	
7	85.0	85	59.0	93.0	2019.0	
8	73.0	95	68.0	NaN	NaN	
9	65.0	81	67.0	86.0	2023.0	
10	79.0	91	NaN	100.0	2018.0	
)
		nns with at Le	ast 1 null va	ıl ue		
	dropna(axi			icuc.		
		.s = 1)				
N	/lath_Score	Reading_Score		er_Count		
0	Math_Score	Reading_Score		er_Count 2		
0 1	72.0 65.0	Reading_Score 86 91		2 3		
0 1 2	72.0 65.0 66.0	Reading_Score 86 91 81		2 3 3		
0 1 2 3	72.0 65.0 66.0 59.0	Reading_Score 86 91 81 77		2 3 3 3		
0 1 2 3 4	72.0 65.0 66.0 59.0 66.0	Reading_Score 86 91 81 77 89		2 3 3 3 2		
0 1 2 3 4 5	72.0 65.0 66.0 59.0 66.0 61.0	Reading_Score 86 91 81 77 89 92		2 3 3 3 3 2 2 2		
0 1 2 3 4 5	72.0 65.0 66.0 59.0 66.0 61.0	Reading_Score 86 91 81 77 89 92 93		2 3 3 3 2 2 2 1		
	4 5 6 7 8 9 10	4 66.0 5 61.0 6 69.1 7 85.0 8 73.0 9 65.0 10 79.0	4 66.0 89 5 61.0 92 6 69.1 93 7 85.0 85 8 73.0 95 9 65.0 81 10 79.0 91	4 66.0 89 76.0 5 61.0 92 76.0 6 69.1 93 72.0 7 85.0 85 59.0 8 73.0 95 68.0 9 65.0 81 67.0	4 66.0 89 76.0 81.0 5 61.0 92 76.0 82.0 6 69.1 93 72.0 66.0 7 85.0 85 59.0 93.0 8 73.0 95 68.0 NaN 9 65.0 81 67.0 86.0 10 79.0 91 NaN 100.0	4 66.0 89 76.0 81.0 2018.0 5 61.0 92 76.0 82.0 2019.0 6 69.1 93 72.0 66.0 2018.0 7 85.0 85 59.0 93.0 2019.0 8 73.0 95 68.0 NaN NaN 9 65.0 81 67.0 86.0 2023.0 10 79.0 91 NaN 100.0 2018.0

In [62]: new_data = ndf.dropna(axis = 0, how ='any')
 new_data

3

91

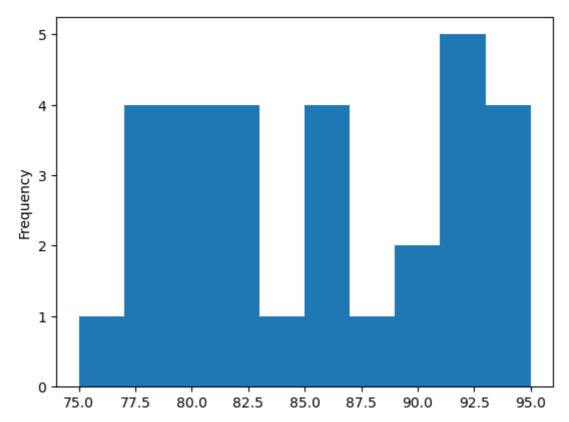
10

79.0

Out[62]:	Mat	h_Score	Reading_Score	Writing_Score	Placement_Score	Club_Join_Date	Placement_Offer_C
	0	72.0	86	64.0	75.0	2020.0	
	1	65.0	91	67.0	111.0	2020.0	
	3	59.0	77	64.0	90.0	2000.0	
	4	66.0	89	76.0	81.0	2018.0	
	5	61.0	92	76.0	82.0	2019.0	
	6	69.1	93	72.0	66.0	2018.0	
	7	85.0	85	59.0	93.0	2019.0	
	9	65.0	81	67.0	86.0	2023.0	
4							•
In []:							
T: [100	d£_nd	noad co	v/"C+/Usons/h	lol como /Downle	oads/DSBDA02.csv	,"\	
In [108	ит=ри.	reau_cs	v(C:/osers/w	iercome/pownic	Jaus/DSBDA02.CSV	/)	
In [109		<pre>['Math_ plot(co</pre>		Reading_Score'	, 'Writing_Sco	re ', 'Placeme	nt_Score']
	plt.sh		-/				
	[
						T	_
	100 -						
	80 -						
							_
						_	-
	60 -			•			
	40 -						
	20 -				φ		
		Math	_Score P	teading_Score	e Writing_Sco	re Placemen	t_Score
In [110	print(np.wher	e(df['Math_Sc	ore ']>90))			
	(array	([], dt	ype=int64),)				
In [111	print(np.wher	e(df[' Readin	ig_Score']<25)))		
			ype=int64),)				
In [112	fig, a	x = plt	.subplots(fig	gsize = (18,10	9))		

```
ax.scatter(df['Placement_Score'], df['Placement_Offer_Count'])
In [113...
           plt.show()
           3.00
           2.75
           2.50
           2.25
           1.75
           1.50
           1.25
           1.00
           print(np.where((df['Placement_Score']<50) & (df['Placement_Offer_Count']>1)))
In [114...
           (array([], dtype=int64),)
           print(np.where((df['Placement_Score']>85) & (df['Placement_Offer_Count']<3)))</pre>
In [115...
           (array([], dtype=int64),)
           from scipy import stats
In [116...
           z = np.abs(stats.zscore(df[' Reading_Score']))
In [117...
           print(z)
```

```
0
                 0.188365
          1
                 0.861099
          2
                 0.484368
          3
                 3.309850
          4
                 0.592006
          5
                 0.995646
          6
                 1.130193
          7
                 0.053819
          8
                 1.399286
          9
                 0.484368
                 0.861099
          10
          11
                 0.618915
          12
                 0.484368
          13
                 1.130193
          14
                 0.080728
                 0.618915
          15
          16
                 1.022555
          17
                 0.188365
          18
                 1.399286
          19
                 0.753462
          20
                 0.188365
                 0.995646
          21
                 0.349822
          22
          23
                 1.022555
          24
                 1.291649
          25
                 0.592006
          26
                 0.457459
          27
                 0.618915
          28
                 0.995646
          29
                 0.888008
          Name: Reading_Score, dtype: float64
In [118...
           sorted_rscore= sorted(df[' Reading_Score'])
           print(sorted_rscore)
          [60, 75, 77, 77, 78, 79, 80, 80, 80, 81, 81, 81, 82, 84, 85, 86, 86, 86, 88, 89, 8
          9, 91, 91, 92, 92, 93, 93, 95, 95]
           q1 = np.percentile(sorted_rscore, 25)
In [119...
           q3 = np.percentile(sorted_rscore, 75)
           print(q1,q3)
           IQR = q3-q1
           IQR
          80.0 91.0
          11.0
Out[119]:
           lwr bound = q1-(1*IQR)
In [120...
           upr bound = q3+(1*IQR)
           print(lwr_bound, upr_bound)
          69.0 102.0
In [121...
           r_outliers = []
           for i in sorted rscore:
               if (i<lwr_bound or i>upr_bound):
                   r_outliers.append(i)
           print(r_outliers)
          [60]
          df[' Reading_Score'].plot(kind = 'hist')
 In [26]:
           plt.show()
```



```
In [ ]: df['log_math'] = np.log10(df['math score'])
    df['log_math'].plot(kind = 'hist')
```