In [ ]:
```
NAME : SHINDE SHUBHAM DNYANDEV,     ROLL NO. : EN23107121,     BATCH : C
```

In [264…
```
import pandas as pd
```

In [233…
```
df = pd.read_csv("/home/admin1/Downloads/data.csv")
df
```

Out[233…

	age	workclass	fnlwgt	education	educational-num	marital-status	occupation	relation
0	25	Private	226802	11th	7	Never-married	Machine-op-inspct	Own-
1	38	Private	89814	HS-grad	9	Married-civ-spouse	Farming-fishing	Husb
2	28	Local-gov	336951	Assoc-acdm	12	Married-civ-spouse	Protective-serv	Husb
3	44	Private	160323	Some-college	10	Married-civ-spouse	Machine-op-inspct	Husb
4	18	?	103497	Some-college	10	Never-married	?	Own-
...	...	...	...	...	...	...	...	
48837	27	Private	257302	Assoc-acdm	12	Married-civ-spouse	Tech-support	
48838	40	Private	154374	HS-grad	9	Married-civ-spouse	Machine-op-inspct	Husb
48839	58	Private	151910	HS-grad	9	Widowed	Adm-clerical	Unma
48840	22	Private	201490	HS-grad	9	Never-married	Adm-clerical	Own-
48841	52	Self-emp-inc	287927	HS-grad	9	Married-civ-spouse	Exec-managerial	

48842 rows × 15 columns

In [235…
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 48842 entries, 0 to 48841
Data columns (total 15 columns):
 #   Column           Non-Null Count  Dtype 
---  ------           --------------  ----- 
 0   age              48842 non-null  int64 
 1   workclass        48842 non-null  object
 2   fnlwgt           48842 non-null  int64 
 3   education        48842 non-null  object
 4   educational-num  48842 non-null  int64 
 5   marital-status   48842 non-null  object
 6   occupation       48842 non-null  object
 7   relationship     48842 non-null  object
 8   race             48842 non-null  object
 9   gender           48842 non-null  object
 10  capital-gain     48842 non-null  int64 
 11  capital-loss     48842 non-null  int64 
 12  hours-per-week   48842 non-null  int64 
 13  native-country   48842 non-null  object
 14  income           48842 non-null  object
dtypes: int64(6), object(9)
memory usage: 5.6+ MB
```

In [237…
```
df.describe()
```

Out[237…

	age	fnlwgt	educational-num	capital-gain	capital-loss	hours-w
count	48842.000000	4.884200e+04	48842.000000	48842.000000	48842.000000	48842.00
mean	38.643585	1.896641e+05	10.078089	1079.067626	87.502314	40.42
std	13.710510	1.056040e+05	2.570973	7452.019058	403.004552	12.39
min	17.000000	1.228500e+04	1.000000	0.000000	0.000000	1.00
25%	28.000000	1.175505e+05	9.000000	0.000000	0.000000	40.00
50%	37.000000	1.781445e+05	10.000000	0.000000	0.000000	40.00
75%	48.000000	2.376420e+05	12.000000	0.000000	0.000000	45.00
max	90.000000	1.490400e+06	16.000000	99999.000000	4356.000000	99.00

In [239…
```
df.dtypes
```

Out[239…
```
age                int64
workclass         object
fnlwgt             int64
education         object
educational-num    int64
marital-status    object
occupation        object
relationship      object
race              object
gender            object
capital-gain       int64
capital-loss       int64
hours-per-week     int64
native-country    object
income            object
dtype: object
```

In [241…
```
df['age'].median()
```

Out[241…
```
37.0
```

In [243…
```
df1 = df.select_dtypes('int64','float64')
```

```
df1
```

Out[243…

	age	fnlwgt	educational-num	capital-gain	capital-loss	hours-per-week
0	25	226802	7	0	0	40
1	38	89814	9	0	0	50
2	28	336951	12	0	0	40
3	44	160323	10	7688	0	40
4	18	103497	10	0	0	30
...	...	...	...	...	...	...
48837	27	257302	12	0	0	38
48838	40	154374	9	0	0	40
48839	58	151910	9	0	0	40
48840	22	201490	9	0	0	20
48841	52	287927	9	15024	0	40

48842 rows × 6 columns

In [245…

```
df1['income'] = df['income']
df1
```

Out[245…

	age	fnlwgt	educational-num	capital-gain	capital-loss	hours-per-week	income
0	25	226802	7	0	0	40	<=50K
1	38	89814	9	0	0	50	<=50K
2	28	336951	12	0	0	40	>50K
3	44	160323	10	7688	0	40	>50K
4	18	103497	10	0	0	30	<=50K
...	...	...	...	...	...	...	...
48837	27	257302	12	0	0	38	<=50K
48838	40	154374	9	0	0	40	>50K
48839	58	151910	9	0	0	40	<=50K
48840	22	201490	9	0	0	20	<=50K
48841	52	287927	9	15024	0	40	>50K

48842 rows × 7 columns

In [247…

```
combined = df1.groupby(['income','age'])
```

In [249…

```
combined.min()
```

Out[249…

income	age	fnlwgt	educational-num	capital-gain	capital-loss	hours-per-week
<=50K	**17**	19752	3	0	0	4
	18	20057	3	0	0	2
	19	20469	1	0	0	2
	20	19410	1	0	0	2
	21	20728	1	0	0	1
...	**...**	...	...	...	...	...
>50K	**83**	153183	6	0	0	50
	84	172907	10	0	0	35
	85	155981	13	0	0	40
	88	263569	7	6418	0	40
	90	46786	9	0	0	15

142 rows × 5 columns

In [251…

```
combined.max()
```

Out[251…

income	age	fnlwgt	educational-num	capital-gain	capital-loss	hours-per-week
<=50K	**17**	806316	10	34095	1721	50
	18	761006	14	34095	1721	72
	19	1047822	13	34095	2129	99
	20	745817	14	34095	2258	84
	21	811615	14	4865	2603	98
...	**...**	...	...	...	...	...
>50K	**83**	240150	13	20051	2392	55
	84	172907	10	0	0	35
	85	155981	13	0	0	40
	88	263569	7	6418	0	40
	90	313986	15	20051	1825	72

142 rows × 5 columns

In [253…

```
combined.std()
```

Out[253…

income	age	fnlwgt	educational-num	capital-gain	capital-loss	hours-per-week
<=50K	17	101243.296009	0.957188	1407.535432	232.467309	10.052279
	18	112046.448155	1.437882	1194.025643	245.870731	11.599559
	19	116351.515261	1.394263	1515.336478	269.259342	11.961223
	20	105724.045468	1.314072	1070.613636	250.178117	11.591010
	21	107995.818286	1.476579	477.993196	271.679717	11.649644
...	...	...	...	...	...	...
>50K	83	61494.955439	4.949747	14178.198070	1691.399421	3.535534
	84	NaN	NaN	NaN	NaN	NaN
	85	NaN	NaN	NaN	NaN	NaN
	88	NaN	NaN	NaN	NaN	NaN
	90	78487.937923	2.375084	8349.509253	506.163929	13.082088

142 rows × 5 columns

In [255…

```
combined.count()
```

Out[255…

income	age	fnlwgt	educational-num	capital-gain	capital-loss	hours-per-week
<=50K	17	595	595	595	595	595
	18	862	862	862	862	862
	19	1050	1050	1050	1050	1050
	20	1112	1112	1112	1112	1112
	21	1090	1090	1090	1090	1090
...	...	...	...	...	...	...
>50K	83	2	2	2	2	2
	84	1	1	1	1	1
	85	1	1	1	1	1
	88	1	1	1	1	1
	90	13	13	13	13	13

142 rows × 5 columns

In [ ]:

In [258…

```
df = pd.read_csv("/home/admin1/Downloads/Iris.csv")
df
```

Out[258…

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa
...	...	...	...	...	...	...
145	146	6.7	3.0	5.2	2.3	Iris-virginica
146	147	6.3	2.5	5.0	1.9	Iris-virginica
147	148	6.5	3.0	5.2	2.0	Iris-virginica
148	149	6.2	3.4	5.4	2.3	Iris-virginica
149	150	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 6 columns

In [260…

```
SETOSA = df[df['Species'] == 'Iris-setosa']
SETOSA
```

Out[260…

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa
5	6	5.4	3.9	1.7	0.4	Iris-setosa
6	7	4.6	3.4	1.4	0.3	Iris-setosa
7	8	5.0	3.4	1.5	0.2	Iris-setosa
8	9	4.4	2.9	1.4	0.2	Iris-setosa
9	10	4.9	3.1	1.5	0.1	Iris-setosa
10	11	5.4	3.7	1.5	0.2	Iris-setosa
11	12	4.8	3.4	1.6	0.2	Iris-setosa
12	13	4.8	3.0	1.4	0.1	Iris-setosa
13	14	4.3	3.0	1.1	0.1	Iris-setosa
14	15	5.8	4.0	1.2	0.2	Iris-setosa
15	16	5.7	4.4	1.5	0.4	Iris-setosa
16	17	5.4	3.9	1.3	0.4	Iris-setosa
17	18	5.1	3.5	1.4	0.3	Iris-setosa
18	19	5.7	3.8	1.7	0.3	Iris-setosa
19	20	5.1	3.8	1.5	0.3	Iris-setosa
20	21	5.4	3.4	1.7	0.2	Iris-setosa
21	22	5.1	3.7	1.5	0.4	Iris-setosa
22	23	4.6	3.6	1.0	0.2	Iris-setosa
23	24	5.1	3.3	1.7	0.5	Iris-setosa
24	25	4.8	3.4	1.9	0.2	Iris-setosa
25	26	5.0	3.0	1.6	0.2	Iris-setosa
26	27	5.0	3.4	1.6	0.4	Iris-setosa
27	28	5.2	3.5	1.5	0.2	Iris-setosa
28	29	5.2	3.4	1.4	0.2	Iris-setosa
29	30	4.7	3.2	1.6	0.2	Iris-setosa
30	31	4.8	3.1	1.6	0.2	Iris-setosa
31	32	5.4	3.4	1.5	0.4	Iris-setosa
32	33	5.2	4.1	1.5	0.1	Iris-setosa
33	34	5.5	4.2	1.4	0.2	Iris-setosa
34	35	4.9	3.1	1.5	0.1	Iris-setosa
35	36	5.0	3.2	1.2	0.2	Iris-setosa
36	37	5.5	3.5	1.3	0.2	Iris-setosa
37	38	4.9	3.1	1.5	0.1	Iris-setosa
38	39	4.4	3.0	1.3	0.2	Iris-setosa
39	40	5.1	3.4	1.5	0.2	Iris-setosa

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
40	41	5.0	3.5	1.3	0.3	Iris-setosa
41	42	4.5	2.3	1.3	0.3	Iris-setosa
42	43	4.4	3.2	1.3	0.2	Iris-setosa
43	44	5.0	3.5	1.6	0.6	Iris-setosa
44	45	5.1	3.8	1.9	0.4	Iris-setosa
45	46	4.8	3.0	1.4	0.3	Iris-setosa
46	47	5.1	3.8	1.6	0.2	Iris-setosa
47	48	4.6	3.2	1.4	0.2	Iris-setosa
48	49	5.3	3.7	1.5	0.2	Iris-setosa
49	50	5.0	3.3	1.4	0.2	Iris-setosa

In [221…

```
VERSICOLOR = df[df['Species'] == 'Iris-versicolor']
VERSICOLOR
```

Out[221…

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
50	51	7.0	3.2	4.7	1.4	Iris-versicolor
51	52	6.4	3.2	4.5	1.5	Iris-versicolor
52	53	6.9	3.1	4.9	1.5	Iris-versicolor
53	54	5.5	2.3	4.0	1.3	Iris-versicolor
54	55	6.5	2.8	4.6	1.5	Iris-versicolor
55	56	5.7	2.8	4.5	1.3	Iris-versicolor
56	57	6.3	3.3	4.7	1.6	Iris-versicolor
57	58	4.9	2.4	3.3	1.0	Iris-versicolor
58	59	6.6	2.9	4.6	1.3	Iris-versicolor
59	60	5.2	2.7	3.9	1.4	Iris-versicolor
60	61	5.0	2.0	3.5	1.0	Iris-versicolor
61	62	5.9	3.0	4.2	1.5	Iris-versicolor
62	63	6.0	2.2	4.0	1.0	Iris-versicolor
63	64	6.1	2.9	4.7	1.4	Iris-versicolor
64	65	5.6	2.9	3.6	1.3	Iris-versicolor
65	66	6.7	3.1	4.4	1.4	Iris-versicolor
66	67	5.6	3.0	4.5	1.5	Iris-versicolor
67	68	5.8	2.7	4.1	1.0	Iris-versicolor
68	69	6.2	2.2	4.5	1.5	Iris-versicolor
69	70	5.6	2.5	3.9	1.1	Iris-versicolor
70	71	5.9	3.2	4.8	1.8	Iris-versicolor
71	72	6.1	2.8	4.0	1.3	Iris-versicolor
72	73	6.3	2.5	4.9	1.5	Iris-versicolor
73	74	6.1	2.8	4.7	1.2	Iris-versicolor
74	75	6.4	2.9	4.3	1.3	Iris-versicolor
75	76	6.6	3.0	4.4	1.4	Iris-versicolor

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
76	77	6.8	2.8	4.8	1.4	Iris-versicolor
77	78	6.7	3.0	5.0	1.7	Iris-versicolor
78	79	6.0	2.9	4.5	1.5	Iris-versicolor
79	80	5.7	2.6	3.5	1.0	Iris-versicolor
80	81	5.5	2.4	3.8	1.1	Iris-versicolor
81	82	5.5	2.4	3.7	1.0	Iris-versicolor
82	83	5.8	2.7	3.9	1.2	Iris-versicolor
83	84	6.0	2.7	5.1	1.6	Iris-versicolor
84	85	5.4	3.0	4.5	1.5	Iris-versicolor
85	86	6.0	3.4	4.5	1.6	Iris-versicolor
86	87	6.7	3.1	4.7	1.5	Iris-versicolor
87	88	6.3	2.3	4.4	1.3	Iris-versicolor
88	89	5.6	3.0	4.1	1.3	Iris-versicolor
89	90	5.5	2.5	4.0	1.3	Iris-versicolor
90	91	5.5	2.6	4.4	1.2	Iris-versicolor
91	92	6.1	3.0	4.6	1.4	Iris-versicolor
92	93	5.8	2.6	4.0	1.2	Iris-versicolor
93	94	5.0	2.3	3.3	1.0	Iris-versicolor
94	95	5.6	2.7	4.2	1.3	Iris-versicolor
95	96	5.7	3.0	4.2	1.2	Iris-versicolor
96	97	5.7	2.9	4.2	1.3	Iris-versicolor
97	98	6.2	2.9	4.3	1.3	Iris-versicolor
98	99	5.1	2.5	3.0	1.1	Iris-versicolor
99	100	5.7	2.8	4.1	1.3	Iris-versicolor

In [223…

```
VIRGINICA = df[df['Species'] == 'Iris-virginica']
VIRGINICA
```

Out[223…

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
100	101	6.3	3.3	6.0	2.5	Iris-virginica
101	102	5.8	2.7	5.1	1.9	Iris-virginica
102	103	7.1	3.0	5.9	2.1	Iris-virginica
103	104	6.3	2.9	5.6	1.8	Iris-virginica
104	105	6.5	3.0	5.8	2.2	Iris-virginica
105	106	7.6	3.0	6.6	2.1	Iris-virginica
106	107	4.9	2.5	4.5	1.7	Iris-virginica
107	108	7.3	2.9	6.3	1.8	Iris-virginica
108	109	6.7	2.5	5.8	1.8	Iris-virginica
109	110	7.2	3.6	6.1	2.5	Iris-virginica
110	111	6.5	3.2	5.1	2.0	Iris-virginica
111	112	6.4	2.7	5.3	1.9	Iris-virginica
112	113	6.8	3.0	5.5	2.1	Iris-virginica
113	114	5.7	2.5	5.0	2.0	Iris-virginica
114	115	5.8	2.8	5.1	2.4	Iris-virginica
115	116	6.4	3.2	5.3	2.3	Iris-virginica
116	117	6.5	3.0	5.5	1.8	Iris-virginica
117	118	7.7	3.8	6.7	2.2	Iris-virginica
118	119	7.7	2.6	6.9	2.3	Iris-virginica
119	120	6.0	2.2	5.0	1.5	Iris-virginica
120	121	6.9	3.2	5.7	2.3	Iris-virginica
121	122	5.6	2.8	4.9	2.0	Iris-virginica
122	123	7.7	2.8	6.7	2.0	Iris-virginica
123	124	6.3	2.7	4.9	1.8	Iris-virginica
124	125	6.7	3.3	5.7	2.1	Iris-virginica
125	126	7.2	3.2	6.0	1.8	Iris-virginica

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
126	127	6.2	2.8	4.8	1.8	Iris-virginica
127	128	6.1	3.0	4.9	1.8	Iris-virginica
128	129	6.4	2.8	5.6	2.1	Iris-virginica
129	130	7.2	3.0	5.8	1.6	Iris-virginica
130	131	7.4	2.8	6.1	1.9	Iris-virginica
131	132	7.9	3.8	6.4	2.0	Iris-virginica
132	133	6.4	2.8	5.6	2.2	Iris-virginica
133	134	6.3	2.8	5.1	1.5	Iris-virginica
134	135	6.1	2.6	5.6	1.4	Iris-virginica
135	136	7.7	3.0	6.1	2.3	Iris-virginica
136	137	6.3	3.4	5.6	2.4	Iris-virginica
137	138	6.4	3.1	5.5	1.8	Iris-virginica
138	139	6.0	3.0	4.8	1.8	Iris-virginica
139	140	6.9	3.1	5.4	2.1	Iris-virginica
140	141	6.7	3.1	5.6	2.4	Iris-virginica
141	142	6.9	3.1	5.1	2.3	Iris-virginica
142	143	5.8	2.7	5.1	1.9	Iris-virginica
143	144	6.8	3.2	5.9	2.3	Iris-virginica
144	145	6.7	3.3	5.7	2.5	Iris-virginica
145	146	6.7	3.0	5.2	2.3	Iris-virginica
146	147	6.3	2.5	5.0	1.9	Iris-virginica
147	148	6.5	3.0	5.2	2.0	Iris-virginica
148	149	6.2	3.4	5.4	2.3	Iris-virginica
149	150	5.9	3.0	5.1	1.8	Iris-virginica

In [294…

```
SETOSA.describe()
```

Out[294…

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
count	50.00000	50.00000	50.000000	50.000000	50.00000
mean	25.50000	5.00600	3.418000	1.464000	0.24400
std	14.57738	0.35249	0.381024	0.173511	0.10721
min	1.00000	4.30000	2.300000	1.000000	0.10000
25%	13.25000	4.80000	3.125000	1.400000	0.20000
50%	25.50000	5.00000	3.400000	1.500000	0.20000
75%	37.75000	5.20000	3.675000	1.575000	0.30000
max	50.00000	5.80000	4.400000	1.900000	0.60000

In [298…

```
VERSICOLOR.describe()
```

Out[298…

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
count	50.00000	50.000000	50.000000	50.000000	50.000000
mean	75.50000	5.936000	2.770000	4.260000	1.326000
std	14.57738	0.516171	0.313798	0.469911	0.197753
min	51.00000	4.900000	2.000000	3.000000	1.000000
25%	63.25000	5.600000	2.525000	4.000000	1.200000
50%	75.50000	5.900000	2.800000	4.350000	1.300000
75%	87.75000	6.300000	3.000000	4.600000	1.500000
max	100.00000	7.000000	3.400000	5.100000	1.800000

In [296…

```
VIRGINICA.describe()
```

Out[296…

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
count	50.00000	50.00000	50.000000	50.000000	50.00000
mean	125.50000	6.58800	2.974000	5.552000	2.02600
std	14.57738	0.63588	0.322497	0.551895	0.27465
min	101.00000	4.90000	2.200000	4.500000	1.40000
25%	113.25000	6.22500	2.800000	5.100000	1.80000
50%	125.50000	6.50000	3.000000	5.550000	2.00000
75%	137.75000	6.90000	3.175000	5.875000	2.30000
max	150.00000	7.90000	3.800000	6.900000	2.50000

In [ ]: