

```
1  """
2  10. us work
3      A. Row count
4      B. Each age group count
5      C. Each profession count [count desc]
6      D. Civil engineer dept and age above 30
7
8  """
9  import pandas as pd
10 data = pd.read_csv("C:/Users/anuru/OneDrive/Desktop/Luminar DSML/Data Science/Pandas1/Assignmen
11 df = pd.DataFrame(data)
12 df.columns = ["ID", "F_NAME", "L_NAME", "AGE", "PROF", "LOC"]
13 df_us = df.loc[df['LOC'] == 'us']
14 print(len(df_us))
15 print("="*100)
16
17 print(df_us.groupby('AGE')['AGE'].count())
18 print("="*100)
19
20 print(df_us.groupby('PROF')['PROF'].count().sort_values(ascending=False))
21 print("="*100)
22
23 print(df_us.loc[(df['PROF'] == 'Civil engineer') & (df['AGE'] > 30)])
```

```
2 9. India work
3     A. Row count
4     B. Each profession count [count desc order]
5     C. Age mxm 3 employees fname,lname,age,prof
6     D. Age minimum 3 employees fname,lname,age,prof
7     E. age above 40 full data
8     F. age range 30 to 40 [profession count]
9
10 import pandas as pd
11
12 data = pd.read_csv("C:/Users/anuru/OneDrive/Desktop/Luminar DSML/Data Science/Pandas1/Assignments/customer1.txt",header=None)
13 df = pd.DataFrame(data)
14 df.columns = ["ID","F_NAME","L_NAME","AGE","PROF","LOC"]
15 df_india = df.loc[df['LOC'] == 'india']
16 print("Row Count = ",len(df_india))
17 print("="*100)
18 prof_count = df_india.groupby('PROF') ['PROF'].count().sort_values(ascending=False)
19 print("Profession count\n",prof_count)
20 print("="*100)
21 print("Age mxm 3 employees fname,lname,age,prof\n",df_india.sort_values('AGE',ascending=False)[df.columns[1:5]].head(3))
22 print("="*100)
23 print("Age minim 3 employees fname,lname,age,prof\n",df_india.sort_values('AGE')[df.columns[1:5]].head(3))
24 print("="*100)
25 print("Age Above 40",df.loc[df['AGE'] > 40][df.columns[1:]])
26 print("="*100)
27 print("age range 30 to 40 [profession count]")
28 print(df.loc[(df['AGE']<=40)&(df['AGE']>=30)].groupby('PROF')['PROF'].count())
29 print("="*100)
30
```

 Q8.py ×

```
1 #8. Each profession count [count desc order]
2
3 import pandas as pd
4
5 data = pd.read_csv("C:/Users/anuru/OneDrive/Desktop/Luminar DSML/Data Science/Pa
6
7 df = pd.DataFrame(data)
8
9 df.columns = ["ID", "F_NAME", "L_NAME", "AGE", "PROF", "LOC"]
10
11 prof_count = df.groupby('PROF') ['PROF'].count().sort_values(ascending=False)
12
13 print(prof_count)
14
```


```
1 #7. Each age group count [age desc order]
2
3 import pandas as pd
4
5 data = pd.read_csv("C:/Users/anuru/OneDrive/Desktop/Luminar D
6
7 df = pd.DataFrame(data)
8
9 df.columns = ["ID", "F_NAME", "L_NAME", "AGE", "PROF", "LOC"]
10
11
12 age_count = df.groupby('AGE') ['AGE'].count()
13
14 print(age_count.sort_values(ascending=False))
15
16
```

```
1 #6. Full data
2
3 import pandas as pd
4
5 data = pd.read_csv("C:/Users/anuru/OneDrive/Desktop/Luminar DSM
6
7 df = pd.DataFrame(data)
8
9 df.columns = ["ID", "F_NAME", "L_NAME", "AGE", "PROF", "LOC"]
10
11 unique = df.groupby('LOC') ['LOC'].count()
12
13 print(unique.sort_values(ascending=False))
14 print("="*100)
15
16 print(df.loc[df['LOC'] == 'australia'])
17
18
```




```
1 #5. Each location count [count desc order]
2
3 import pandas as pd
4
5 data = pd.read_csv("C:/Users/anuru/OneDrive/Desktop/Luminar DSML
6
7 df = pd.DataFrame(data)
8
9 df.columns = ["ID", "F_NAME", "L_NAME", "AGE", "PROF", "LOC"]
10
11 unique = df.groupby('LOC') ['LOC'].count()
12
13 print(unique.sort_values(ascending=False))
14
15
```

 Q4.py ×

```
1 #4. Age minimum 5 employees fname,lname,prof,loc
2
3 import pandas as pd
4
5 data = pd.read_csv("C:/Users/anuru/OneDrive/Desktop/Luminar DSML/Data Scienc
6
7 df = pd.DataFrame(data)
8
9 df.columns = ["ID", "F_NAME", "L_NAME", "AGE", "PROF", "LOC"]
10 
11 print(df.sort_values(by='AGE',).head(5)[["F_NAME", "L_NAME", "PROF", "LOC"]])
```

 Q3.py ×

```
1 #3. Age maximum 10 fname,lname,prof,loc
2
3 import pandas as pd
4
5 data = pd.read_csv("C:/Users/anuru/OneDrive/Desktop/Luminar DSML/Data Scienc
6
7 df = pd.DataFrame(data)
8
9 df.columns = ["ID", "F_NAME", "L_NAME", "AGE", "PROF", "LOC"]
10 
11 print(df.sort_values(by='|AGE', ascending=False).head(10)[df.columns[1:]])
```



```
1 #2. Remove duplicates rows and find total row count
```

```
2  
3 import pandas as pd
```

```
4  
5 data = pd.read_csv("C:/Users/anuru/OneDrive/Desktop/Luminar DSML/
```

```
6  
7 df = pd.DataFrame(data)
```

```
8  
9 df.columns = ["ID", "F_NAME", "L_NAME", "AGE", "PROF", "LOC"]
```

```
10  
11 print(df.shape[0])
```

```
12 print("="*100)
```

```
13  
14 df1 = df.drop_duplicates()
```

```
15  
16 print(df1.shape[0])
```

```
17
```

Q1.py ×

```
1 #1. Find Row count
2 import pandas as pd
3
4 data = pd.read_csv("C:/Users/anuru/OneDrive/Desktop/Luminar DSML/Data S
5
6 df = pd.DataFrame(data)
7
8 df.columns = ["ID", "F_NAME", "L_NAME", "AGE", "PROF", "LOC"]
9 print(df.head())
10
11 print("="*100)
12
13 print("Number Of rows = ", df.shape[0])
14
15
```