lab-3-1

April 19, 2024

[1]:
```
import pandas as pd
```

[6]:
```
df= pd.read_csv("Customer Segmentation.csv")
```

[7]:
```
df.head()
```

[7]:

	CustomerID	Gender	Age	Annual Income (k$)	Spending Score (1-100)
0	1	Male	19	15	39
1	2	Male	21	15	81
2	3	Female	20	16	6
3	4	Female	23	16	77
4	5	Female	31	17	40

[8]:
```
df.tail()
```

[8]:

	CustomerID	Gender	Age	Annual Income (k$)	Spending Score (1-100)
195	196	Female	35	120	79
196	197	Female	45	126	28
197	198	Male	32	126	74
198	199	Male	32	137	18
199	200	Male	30	137	83

[9]:
```
df.info
```

[9]:
```
<bound method DataFrame.info of      CustomerID  Gender  Age  Annual Income (k$)
Spending Score (1-100)
0             1    Male   19                  15                      39
1             2    Male   21                  15                      81
2             3  Female   20                  16                       6
3             4  Female   23                  16                      77
4             5  Female   31                  17                      40
..          …     …  …                 …                       …
195         196  Female   35                 120                      79
196         197  Female   45                 126                      28
197         198    Male   32                 126                      74
198         199    Male   32                 137                      18
199         200    Male   30                 137                      83
```

```
[200 rows x 5 columns]>
```

[10]:
```
df.shape
```

[10]:
```
(200, 5)
```

[11]:
```
df.describe
```

[11]:
```
<bound method NDFrame.describe of      CustomerID  Gender  Age  Annual Income
(k$)  Spending Score (1-100)
0             1    Male   19                  15                      39
1             2    Male   21                  15                      81
2             3  Female   20                  16                       6
3             4  Female   23                  16                      77
4             5  Female   31                  17                      40
..          …     …  …                  …                     …
195         196  Female   35                 120                      79
196         197  Female   45                 126                      28
197         198    Male   32                 126                      74
198         199    Male   32                 137                      18
199         200    Male   30                 137                      83

[200 rows x 5 columns]>
```

[12]:
```
df.dtypes
```

[12]:
```
CustomerID                int64
Gender                   object
Age                       int64
Annual Income (k$)        int64
Spending Score (1-100)    int64
dtype: object
```

[21]:
```
summary = df.groupby('Spending Score (1-100)')['Age'].
  ↪agg(['mean','median','min','max'])

print(summary)
```

```
                         mean  median  min  max
Spending Score (1-100)
1                        35.5    35.5   34   37
3                        64.0    64.0   64   64
4                        56.5    56.5   53   60
5                        35.5    33.0   19   57
6                        27.5    27.5   20   35
…                       …      …  …  …
```

```
94                    23.0    23.0   23   23
95                    36.5    36.5   33   40
97                    29.0    29.0   28   30
98                    35.0    35.0   35   35
99                    35.0    35.0   35   35

[84 rows x 4 columns]
```

[ ]: