URL:https://raw.githubusercontent.com/ameenmanna8824/DATASETS/main/Mall_Customers.csv

```
import pandas as pd
df=pd.read_csv("https://raw.githubusercontent.com/ameenmanna8824/DATASETS/main/Mall_Customers.csv")
df
```

	CustomerID	Genre	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 × 5 columns

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200 entries, 0 to 199
Data columns (total 5 columns):
 #   Column                  Non-Null Count  Dtype 
---  ------                  --------------  ----- 
 0   CustomerID              200 non-null    int64 
 1   Genre                   200 non-null    object
 2   Age                     200 non-null    int64 
 3   Annual Income (k$)      200 non-null    int64 
 4   Spending Score (1-100)  200 non-null    int64 
dtypes: int64(4), object(1)
memory usage: 7.9+ KB
```

method - .info()

attribute - .shape,.size

```
df.size
```

```
1000
```

```
df.shape
```

```
(200, 5)
```

```
#df.iloc[row slicing,column slicing]
df.iloc[5:10,2:10]
```

	Age	Annual Income (k$)	Spending Score (1-100)
5	22	17	76
6	35	18	6
7	23	18	94
8	64	19	3
9	30	19	72

```
df[10:20]
```

	CustomerID	Genre	Age	Annual Income (k$)	Spending Score (1-100)
10	11	Male	67	19	14
11	12	Female	35	19	99
12	13	Female	58	20	15
13	14	Female	24	20	77
14	15	Male	37	20	13
15	16	Male	22	20	79
16	17	Female	35	21	35
17	18	Male	20	21	66

```
id=df.CustomerID#variable id to store CustomerID
id
```

```
0        1
1        2
2        3
3        4
4        5
       ...
195    196
196    197
197    198
198    199
199    200
Name: CustomerID, Length: 200, dtype: int64
```

```
m=id[0:64]
m
```

```
0      1
1      2
2      3
3      4
4      5
       ..
59    60
60    61
61    62
62    63
63    64
Name: CustomerID, Length: 64, dtype: int64
```

```
k=df.Genre.unique()#shows the unique value
#nunique is to count unique values
k
```

```
array(['Male', 'Female'], dtype=object)
```

```
len(k)
```

```
2
```

```
l=df.groupby('Annual Income (k$)',sort = False).size()
l
#groupby sets my values in Alphabetical order
```

```
Annual Income (k$)
15     2
16     2
17     2
18     2
19     4
       ..
103    4
113    2
120    2
126    2
137    2
Length: 64, dtype: int64
```

```
len(l)
```

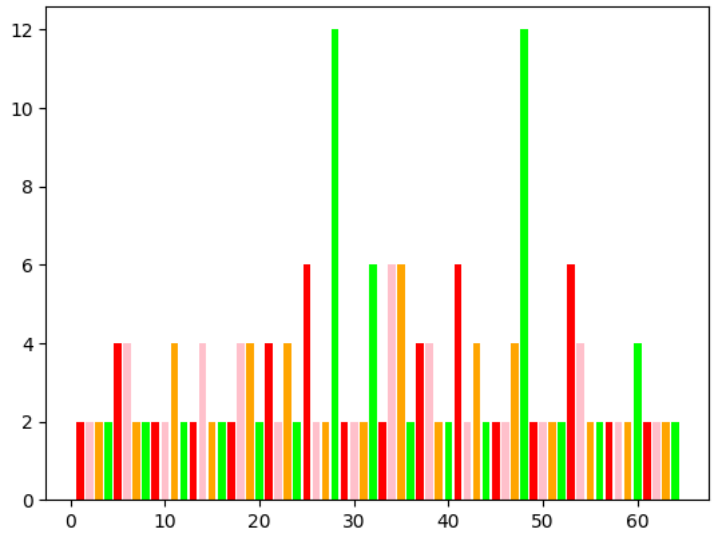
```
64
```

```
len(m)
```

```
64
```

```
import matplotlib.pyplot as plt
plt.bar(m,l,color = ['red','pink','orange','lime'])
```

```
<BarContainer object of 64 artists>
```




```
df.isnull().sum()
```

```
CustomerID              0
Genre                   0
Age                     0
Annual Income (k$)      0
Spending Score (1-100)  0
dtype: int64
```

```
df['Spending Score (1-100)'].value_counts()
```

```
42    8
55    7
46    6
73    6
35    5
     ..
31    1
44    1
53    1
65    1
18    1
Name: Spending Score (1-100), Length: 84, dtype: int64
```

```
import numpy as np
u = np.sum((df['Spending Score (1-100)']>=0)&(df['Spending Score (1-100)']<20))
v = np.sum((df['Spending Score (1-100)']>=20)&(df['Spending Score (1-100)']<40))
r = np.sum((df['Spending Score (1-100)']>=40)&(df['Spending Score (1-100)']<60))
s = np.sum((df['Spending Score (1-100)']>=60))
print(u)
print(v)
print(r)
print(s)
```

```
34
25
76
65
```

```
np.max(df['Spending Score (1-100)'])
```

```
99
```

```
np.min(df['Spending Score (1-100)'])
```

```
1
```

```
df.dropna()
```

	CustomerID	Genre	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

```
df['Genre'] = df['Genre'].str.replace('sex','')
df['Genre']
```

```
0        Male
1        Male
2      Female
3      Female
4      Female
        ...
195    Female
196    Female
197      Male
198      Male
199      Male
Name: Genre, Length: 200, dtype: object
```

```
df['Age'] = df['Age'].astype('float64')
df['Age']
```

```
0      19.0
1      21.0
2      20.0
3      23.0
4      31.0
       ...
195    35.0
196    45.0
197    32.0
198    32.0
199    30.0
Name: Age, Length: 200, dtype: float64
```