Chirag Rao KV

240962180

PC 66

Week 9 demo

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

```python
fruit_datafruit = pd.Series(['apple','orange','apple','apple']*2)
fruit_datafruit.value_counts()
```

```
apple     6
orange    2
Name: count, dtype: int64
```

```python
fruit_datafruit.unique()
```

```
array(['apple', 'orange'], dtype=object)
```

```python
prod_ser = pd.Series(["a","b","c","a","b","c","a"])
prod_cat= prod_ser.astype("category")
prod_cat
```

```
0    a
1    b
2    c
3    a
4    b
5    c
6    a
dtype: category
Categories (3, object): ['a', 'b', 'c']
```

```python
prod_typ = prod_cat.array
prod_typ.categories
prod_typ.codes
```

```
array([0, 1, 2, 0, 1, 2, 0], dtype=int8)
```

```python
data = pd.Series(["low","medium","high","low","high"])
Income_cat = pd.Categorical(data)
print(Income_cat)
```

```
['low', 'medium', 'high', 'low', 'high']
Categories (3, object): ['high', 'low', 'medium']
```

```
data_cat = pd.Categorical(data,categories =
["low","medium","high"],ordered = True)
print(data_cat)
```

```
['low', 'medium', 'high', 'low', 'high']
Categories (3, object): ['low' < 'medium' < 'high']
```

```
num_series = pd.Series([10,25,50,75,45,105,13,45])
bins= [0,30,60,100]
labels = ["low","medium", "high"]
num_cat = pd.cut(num_series, bins = bins , labels = labels)
print("Count in each bin")
num_cat.value_counts()
```

```
Count in each bin
```

```
low       3
medium    3
high      1
Name: count, dtype: int64
```

```
data = pd.read_csv("mtcars.csv", index_col=0)
mpg = data["mpg"]
bins = [0,18,25,40]
labels = ["low", "medium","high"]
mpg_cat=pd.cut(mpg, bins = bins, labels = labels)
data['mpgCat']= mpg_cat
data
```

```
                      mpg  cyl   disp   hp  drat     wt   qsec  vs  am
gear  \
model

Mazda RX4            21.0    6  160.0  110  3.90  2.620  16.46   0   1
4
Mazda RX4 Wag        21.0    6  160.0  110  3.90  2.875  17.02   0   1
4
Datsun 710           22.8    4  108.0   93  3.85  2.320  18.61   1   1
4
Hornet 4 Drive       21.4    6  258.0  110  3.08  3.215  19.44   1   0
3
Hornet Sportabout    18.7    8  360.0  175  3.15  3.440  17.02   0   0
3
Valiant              18.1    6  225.0  105  2.76  3.460  20.22   1   0
3
Duster 360           14.3    8  360.0  245  3.21  3.570  15.84   0   0
3
Merc 240D            24.4    4  146.7   62  3.69  3.190  20.00   1   0
```

```
4
Merc 230             22.8    4  140.8   95  3.92  3.150  22.90    1    0
4
Merc 280             19.2    6  167.6  123  3.92  3.440  18.30    1    0
4
Merc 280C            17.8    6  167.6  123  3.92  3.440  18.90    1    0
4
Merc 450SE           16.4    8  275.8  180  3.07  4.070  17.40    0    0
3
Merc 450SL           17.3    8  275.8  180  3.07  3.730  17.60    0    0
3
Merc 450SLC          15.2    8  275.8  180  3.07  3.780  18.00    0    0
3
Cadillac Fleetwood   10.4    8  472.0  205  2.93  5.250  17.98    0    0
3
Lincoln Continental  10.4    8  460.0  215  3.00  5.424  17.82    0    0
3
Chrysler Imperial    14.7    8  440.0  230  3.23  5.345  17.42    0    0
3
Fiat 128             32.4    4   78.7   66  4.08  2.200  19.47    1    1
4
Honda Civic          30.4    4   75.7   52  4.93  1.615  18.52    1    1
4
Toyota Corolla       33.9    4   71.1   65  4.22  1.835  19.90    1    1
4
Toyota Corona        21.5    4  120.1   97  3.70  2.465  20.01    1    0
3
Dodge Challenger     15.5    8  318.0  150  2.76  3.520  16.87    0    0
3
AMC Javelin          15.2    8  304.0  150  3.15  3.435  17.30    0    0
3
Camaro Z28           13.3    8  350.0  245  3.73  3.840  15.41    0    0
3
Pontiac Firebird     19.2    8  400.0  175  3.08  3.845  17.05    0    0
3
Fiat X1-9            27.3    4   79.0   66  4.08  1.935  18.90    1    1
4
Porsche 914-2        26.0    4  120.3   91  4.43  2.140  16.70    0    1
5
Lotus Europa         30.4    4   95.1  113  3.77  1.513  16.90    1    1
5
Ford Pantera L       15.8    8  351.0  264  4.22  3.170  14.50    0    1
5
Ferrari Dino         19.7    6  145.0  175  3.62  2.770  15.50    0    1
5
Maserati Bora        15.0    8  301.0  335  3.54  3.570  14.60    0    1
5
Volvo 142E           21.4    4  121.0  109  4.11  2.780  18.60    1    1
4
```

```
                     carb  mpgCat
model
Mazda RX4               4  medium
Mazda RX4 Wag           4  medium
Datsun 710              1  medium
Hornet 4 Drive          1  medium
Hornet Sportabout       2  medium
Valiant                 1  medium
Duster 360              4     low
Merc 240D               2  medium
Merc 230                2  medium
Merc 280                4  medium
Merc 280C               4     low
Merc 450SE              3     low
Merc 450SL              3     low
Merc 450SLC             3     low
Cadillac Fleetwood      4     low
Lincoln Continental     4     low
Chrysler Imperial       4     low
Fiat 128                1    high
Honda Civic             2    high
Toyota Corolla          1    high
Toyota Corona           1  medium
Dodge Challenger        2     low
AMC Javelin             2     low
Camaro Z28              4     low
Pontiac Firebird        2  medium
Fiat X1-9               1    high
Porsche 914-2           2    high
Lotus Europa            2    high
Ford Pantera L          4     low
Ferrari Dino            6  medium
Maserati Bora           8     low
Volvo 142E              2  medium
```

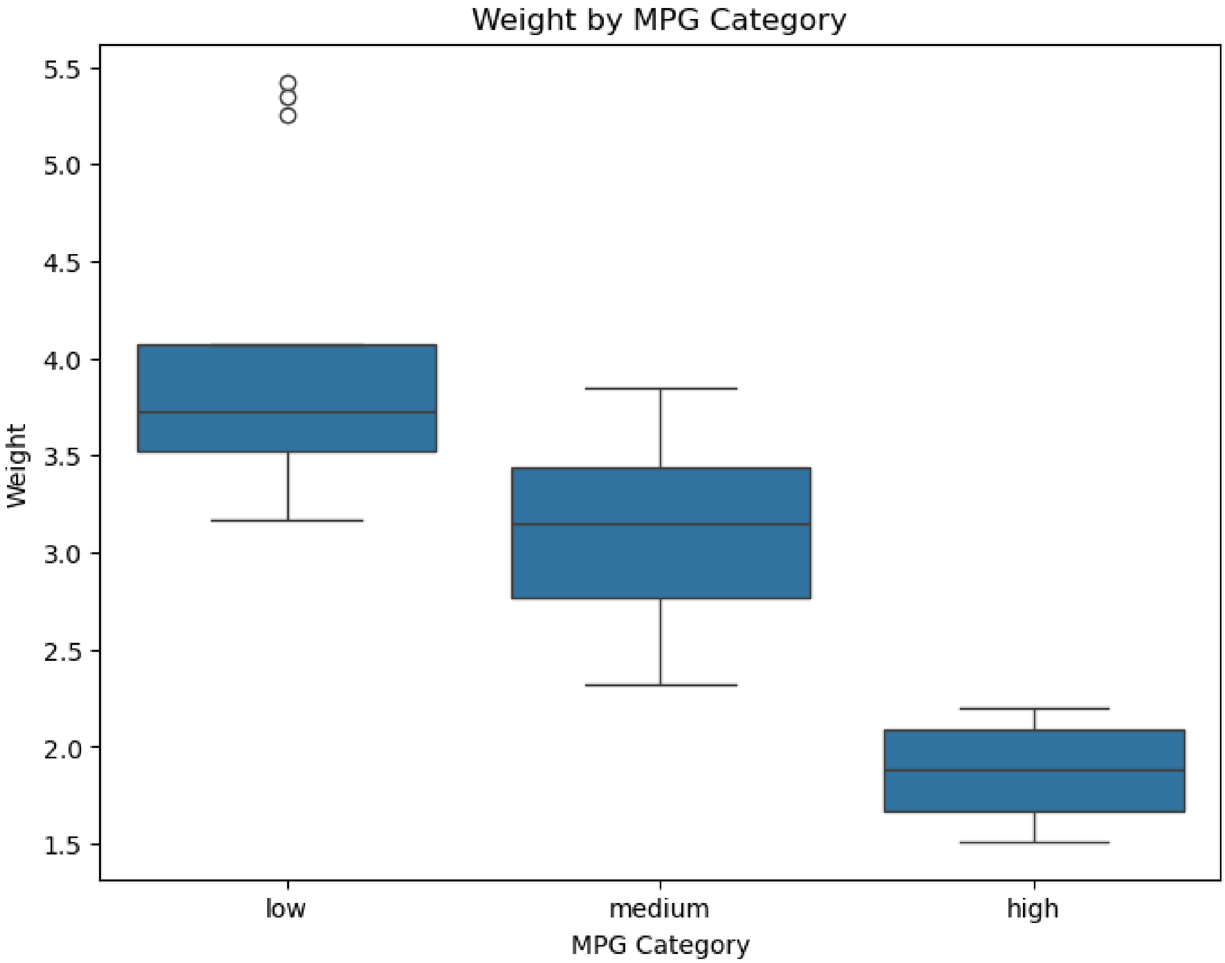
```
tab3 = data.groupby('mpgCat')['wt'].mean()
print("Table displaying categories of MPG and their average weight")
print(tab3)
```

```
Table displaying categories of MPG and their average weight
mpgCat
low       4.011077
medium    3.043846
high      1.873000
Name: wt, dtype: float64
```

```
/tmp/ipykernel_110065/1719140303.py:1: FutureWarning: The default of
observed=False is deprecated and will be changed to True in a future
version of pandas. Pass observed=False to retain current behavior or
```

```
observed=True to adopt the future default and silence this warning.
  tab3 = data.groupby('mpgCat')['wt'].mean()
```

```
df1 = data[['mpgCat','cyl', 'wt', 'mpg']]
plt.figure(figsize=(8, 6))
sns.boxplot(x='mpgCat', y='wt', data=df1)
plt.title('Weight by MPG Category')
plt.xlabel('MPG Category')
plt.ylabel('Weight')
plt.show()
```



```
str1= "  Data visualization is very interesting. I love Python
Programming.    "
print('String length',len(str1))
str1=str1.strip()
print('String 1:',str1)
str2=str1.split('.')
print('String 2:',str2)
```

```
str3=str1.lower()
print('String 3:', str3)
```

```
String length 71
String 1: Data visualization is very interesting. I love Python
Programming.
String 2: ['Data visualization is very interesting', ' I love Python
Programming', '']
String 3: data visualization is very interesting. i love python
programming.
```

```
"python" in str3
print(str1.index('Python') )
str2=str1.replace('interesting','amazing')
print(str2)
str2.count('.')
```

```
47
Data visualization is very amazing. I love Python Programming.
```

```
2
```