

```
1 # === binningTED.ipynb ===
2
3 #%%
4 # Import Dependencies
5 import pandas as pd
6
7
8 #%%
9 # Create a path to the csv and read it into a Pandas DataFrame
10 csv_path = "Resources/ted_talks.csv"
11 ted_df = pd.read_csv(csv_path)
12
13 ted_df.head()
14
15
16 #%%
17 # Figure out the minimum and maximum views for a TED Talk
18 print(ted_df["views"].max())
19 print(ted_df["views"].min())
20
21
22 #%%
23 # Create bins in which to place values based upon TED Talk views
24 bins = [0, 200000, 400000, 600000, 800000, 1000000,
25         2000000, 3000000, 4000000, 5000000, 50000000]
26
27 # Create labels for these bins
28 group_labels = ["0 to 200k", "200k to 400k", "400k to 600k", "600k to 800k", "800k to
... 1mil", "1mil to 2mil", "2mil to 3mil", "3mil to 4mil", "4mil to 5mil", "5mil to 50mil"]
29
30
31 #%%
32 # Slice the data and place it into bins
33 pd.cut(ted_df["views"], bins, labels=group_labels).head()
34
35
36 #%%
37 # Place the data series into a new column inside of the DataFrame
38 ted_df["View Group"] = pd.cut(ted_df["views"], bins, labels=group_labels)
39 ted_df.head()
40
41
42 #%%
43 # Create a GroupBy object based upon "View Group"
44 ted_group = ted_df.groupby("View Group")
45
46 # Find how many rows fall into each bin
47 print(ted_group["comments"].count())
48
49 # Get the average of each column within the GroupBy object
50 ted_group[["comments", "duration", "languages"]].mean()
51
52
53
```