

Exercises: Pandas

WE'LL COVER THE FOLLOWING



- Time to Test Your Skills!
 - Q1. Create a DataFrame from the given dictionary data and index labels and store it in the variable called “df”.
 - Q2. a) Select the column labelled “Listeners” and store it in the variable called “col”. b) Select the first row and store it in the variable called “row”.
 - Q3. Select all the rows where the Genre is ‘Pop’ and store the result in the variable “pop_artists”.
 - Q4. Select the artists who have more than 2,000,000 listeners and whose Genre is ‘Pop’ and save the output in the variable called “top_pop”.
 - Q5. Perform a grouping by Genre using sum() as the aggregation function and store the results in the variable called “grouped”.

Time to Test Your Skills!

Note: We are going to create a DataFrame called “*df*” in the first exercise, and we will keep referring to the same DataFrame as our input in the rest of the exercises.

Q1. Create a DataFrame from the given dictionary data and index labels and store it in the variable called “df”. #



```
import pandas as pd

# Input
data = {'Artist': ['Ariana Grande', 'Taylor Swift', 'Ed Sheeran', 'Justin Bieber', 'Lady Gaga'],
        'Genre': ['Jazz', 'Rock', 'Jazz', 'Pop', 'Pop', 'Rock'],
        'Listeners': [1300000, 2700000, 5000000, 2000000, 3000000, 1100000]}
```



```
labels = ['AG', 'TS', 'ED', 'JB', 'LG', 'BM']
```

```
# Your solution goes here
```

```
# Uncomment the print statement once done
```

```
# print(df)
```



Q2. a) Select the column labelled “Listeners” and store it in the variable called “col”.
b) Select the first row and store it in the variable called “row”. #

(Remember, we are still using the DataFrame called *df*.)



```
# Your solution goes here
```



```
# Uncomment the print statement once done
```

```
# print("Row:", row)
```

```
# print("Col:", col)
```



Q3. Select all the rows where the Genre is ‘Pop’ and store the result in the variable “pop_artists”. #



```
# Your solution goes here
```



```
# Uncomment the print statement once done
```

```
# print(pop_artists)
```



Q4. Select the artists who have more than 2,000,000 listeners and whose Genre is ‘Pop’ and save the output in the variable called “top_pop”. #



```
# Your solution goes here
```



```
# Uncomment the print statement once done
```

```
# print(top_pop)
```



Q5. Perform a grouping by Genre using sum() as the aggregation function and store the results in the variable called "grouped". #



Your solution goes here



```
grouped = df.groupby('Artist').sum()
```

Uncomment the print statement once done

```
# print(grouped)
```

