

**Name: Muhammad Mobeen**

**Registration No: 200901097**

**BSCS01-(B)**

**Presented to: Mam Reeda Saeed**

# **Artificial Intelligence Lab**

**Lab #6**



# Pandas Plotting

## Code Documentation:-

This code reads a CSV file named "drinks.csv" using pandas and performs several operations on the resulting DataFrame. Here is a breakdown of each step

1. Import pandas and matplotlib.pyplot.
2. Read the "drinks.csv" file into a DataFrame called "drinks\_df" using the read\_csv() function from pandas.
3. Print the first five rows of the DataFrame using head().
4. Print the last five rows of the DataFrame using tail().
5. Print the default index of the DataFrame using index.
6. Print the data types of the columns in the DataFrame using dtypes.
7. Print the shape of the DataFrame using shape.
8. Print the "beer\_servings" column of the DataFrame using ['beer\_servings'].
9. Print the mean of the "beer\_servings" column of the DataFrame using mean().
10. Print the frequency of each continent in the "continent" column of the DataFrame using value\_counts().
11. Create a new DataFrame called "europe" that contains only the rows where the "continent" column is "Europe" and the "wine\_servings" column is more significant than 300.
12. Create a bar plot that shows the total litres of pure alcohol consumption for each continent using groupby() and plot().
13. Create a scatter plot that shows the relationship between "beer\_servings" and "spirit\_servings" using scatter() and plot().
14. Create a bar plot that shows the count of countries in each continent using value\_counts() and plot().

## Task 1

```
#how to read csv in pandas
import pandas as pd
drinks_df = pd.read_csv('drinks.csv')
```

These lines import the necessary libraries: pandas for data analysis and matplotlib for data visualization. This line reads in the CSV file named "drinks.csv" using pandas' read\_csv() function and creates a DataFrame called drinks\_df.

```
# Print the first 5 rows of the DataFrame
print(drinks_df.head())
# Print the last 5 rows of the DataFrame
print(drinks_df.tail())
```

This line prints the first 5 rows of the DataFrame using the head() method. This line prints the last 5 rows of the DataFrame using the tail() method.

```
# Print the default index of the DataFrame
print("Default Index: {}".format(drinks_df.index))
print("-----")

# Print the data types of the columns in the DataFrame
print("Data Types:-\n{}".format(drinks_df.dtypes))
print("-----")

# Print the shape of the DataFrame
print("Shape: {}".format(drinks_df.shape))
```

Print the default index of the DataFrame using index. Print the data types of the columns in the DataFrame using dtypes. Print the shape of the DataFrame using shape.

```
# Print the "beer_servings" column of the DataFrame
print(drinks_df['beer_servings'])
```

This line prints the "beer\_servings" column of the DataFrame using square bracket notation.

```
# Print the mean of the "beer_servings" column of the DataFrame
print("Mean of Beer Servings: {}".format(drinks_df['beer_servings'].mean()))
```

This line prints the mean of the "beer\_servings" column of the DataFrame using the mean() method.

```
# Print the frequency of each continent in the "continent" column of the DataFrame
print("Continents Frequency:-\n{}".format(drinks_df['continent'].value_counts()))
```

This line prints the frequency of each continent in the "continent" column of the DataFrame using the value\_counts() method.

```
# Create a new DataFrame called "europe" that contains only the rows where the "continent"
column is "Europe" and the "wine_servings" column is greater than 300
europe = drinks_df[(drinks_df['continent'] == 'Europe') & (drinks_df['wine_servings'] >
300)]
print(europe)
```

These lines create a new DataFrame called "europe" that contains only the rows where the "continent" column is "Europe" and the "wine\_servings" column is greater than 300, and then print the resulting DataFrame.

## Task 3

```
# Create a bar plot that shows the total litres of pure alcohol consumption for each continent
plt.figure(figsize=(8,6))
drinks_df.groupby('continent')['total_litres_of_pure_alcohol'].sum().plot(kind='bar')
plt.title('Total Litres of Pure Alcohol consumption by Continent')
plt.xlabel('Continent')
plt.ylabel('Total Litres of Pure Alcohol')
plt.show()
```

These lines create a bar plot that shows the total liters of pure alcohol consumption for each continent. The `groupby()` method groups the DataFrame by continent, the `sum()` method calculates the total liters of pure alcohol consumption for each continent, and the `plot()` method creates the bar plot.

```
# Create a scatter plot that shows the relationship between "beer_servings" and "spirit_servings"
plt.figure(figsize=(8,6))
plt.scatter(x=drinks_df['beer_servings'], y=drinks_df['spirit_servings'])
plt.title('Beer Servings vs Spirit Servings')
plt.xlabel('Beer Servings')
plt.ylabel('Spirit Servings')
plt.show()
```

These lines create a scatter plot that shows the relationship between "beer\_servings" and "spirit\_servings". The `scatter()` method creates the scatter plot with "beer\_servings" on the x-axis and "spirit\_servings" on the y-axis.

```
# Create a bar plot that shows the count of countries in each continent
plt.figure(figsize=(8,6))
drinks_df['continent'].value_counts().plot(kind='bar')
plt.title('Count of Countries by Continent')
plt.xlabel('Continent')
plt.ylabel('Count')
plt.show()
```

These lines create a bar plot showing the countries' count in each continent. The `value_counts()` method counts the number of countries in each continent, and the `plot()` method creates the bar plot.

Overall, the code reads in a CSV file using pandas, creates a DataFrame, and performs various analyses and visualizations on the data using pandas and matplotlib. The code demonstrates the capabilities of pandas and matplotlib for data analysis and visualization.