

|  |
| --- |
| Experiment No. 11 |
| Program to demonstrate manipulation using Pandas |
| Date of Performace: 10/04/2024 |
| Date of Submission: 17/04/2024 |



**Experiment No. 11**



**Code:**import pandas as pd

import matplotlib.pyplot as plt

# Sample DataFrame creation

data = {'name': ["Anjali", "Teena", "Smart", "Yami", "Anjali", "Teena", "Smart", "Yami"],

        'age': [25, 30, 35, 40, 22, 28, 36, 42]}

df = pd.DataFrame(data)

# Mapping of old names to new names

name\_mapping = {"Anjali": "Yash Chavan", "Teena": "Krisha Chikka", "Smart": "Kamal Chhotaray", "Yami": "Sarth Choudhary"}

# Replace names in the DataFrame

df['name'].replace(name\_mapping, inplace=True)

# Display DataFrame type and content

print("Type:", type(df))

print("DataFrame:")

print(df)

# Plot histogram of name frequencies

df['name'].value\_counts().plot(kind='bar', color='skyblue')

plt.title('Name Frequencies')

plt.xlabel('Names')

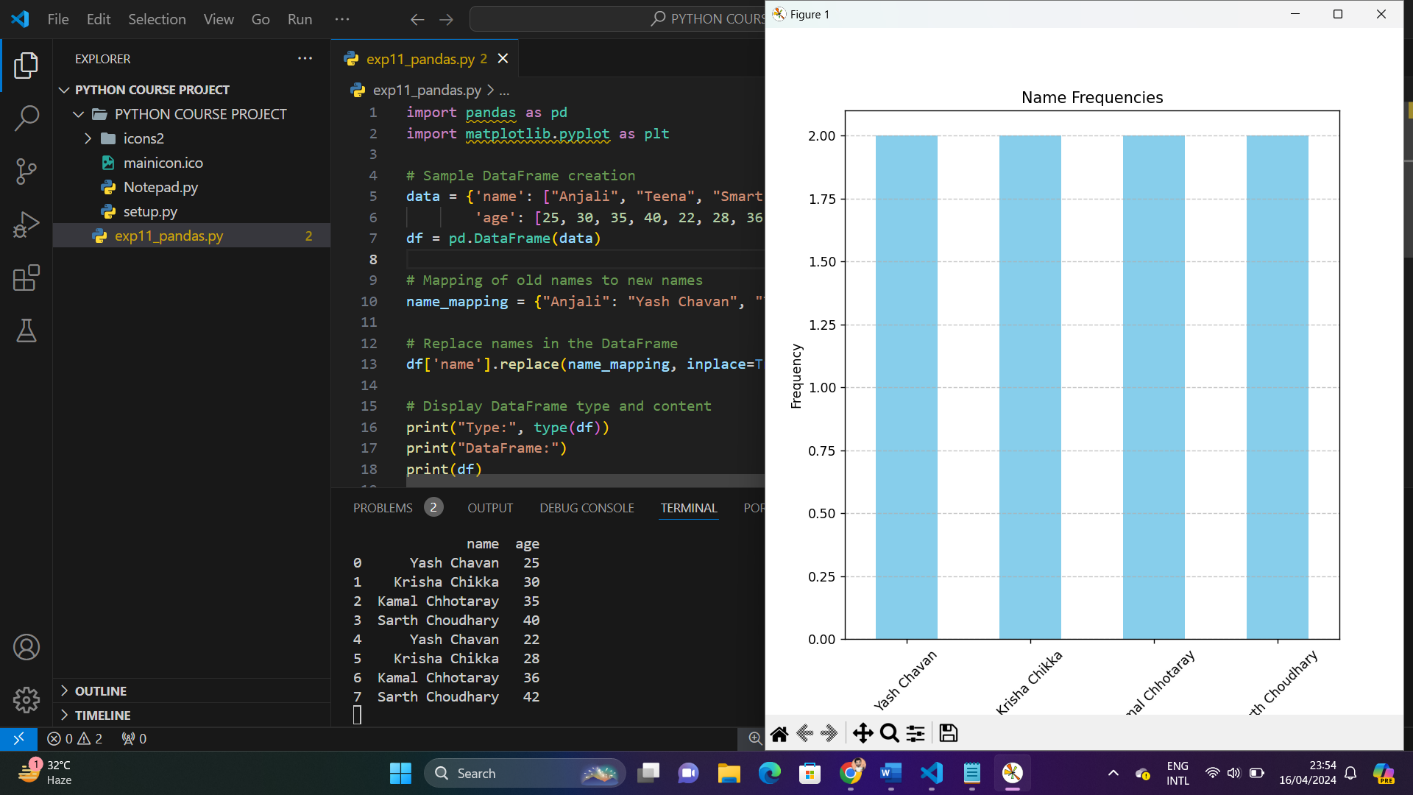
plt.ylabel('Frequency')

plt.xticks(rotation=45)

plt.grid(axis='y', linestyle='--', alpha=0.7)

plt.show()

**Output:**





**Conclusion:**

Pandas is an open-source data manipulation and analysis library built on top of the Python programming language. Provides powerful data structures: primarily Series (1D) and DataFrame (2D), which are built on top of NumPy arrays. Offers tools for reading and writing data in various formats, including CSV, Excel, SQL databases, and more. Facilitates data manipulation tasks like filtering, sorting, grouping, merging, and reshaping datasets. Supports handling missing data through methods like dropna() and fillna(). Provides time series functionality for working with time-indexed data. Enables easy data visualization using integration with Matplotlib and Seaborn.



