|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Data visualization and analysis (CS-232 )** | | | | | |
|  | |  | **Namal University Mianwali**  **Faculty of Computer Science** | | |
| Lab Instructor | Sir Bilal | | | Dated | 5-03-2025 |
| Name | Samreen Fatima Kazmi | | | Roll No | NUM-BSCS-2022-52 |
| Max Marks |  | | | Obtained Marks |  |
| **Assignment** | | | | | |
| **Instructions:**  Using the files provided in the privision\_office\_dataset folder to complete the following tasks: | | | | | |
| **Task 1** | **Load and Explore the Data:**  • Read the "pirvision\_office\_dataset1.csv" file into a Pandas DataFrame.  • Display the total number of rows and columns, including their data types.  • Show the first 100 and last 100 rows of the dataset. | | | | |
| **Code** | import pandas as pd  all\_files\_data=pd.read\_csv(r'pirvision\_office\_dataset1.csv')  df = pd.DataFrame(all\_files\_data)  print(df.head(100))  print(df.tail(100))  print(df.dtypes)  print("rows columns = " ,df.shape) | | | | |
| **Output** |  | | | | |
| **Task 2** | **Identify Missing**  • Check how many missing (NaN) or empty values exist in each column. | | | | |
| **Code** | import pandas as pd  all\_files\_data=pd.read\_csv(r'pirvision\_office\_dataset1.csv')  print(all\_files\_data.isna().sum()) | | | | |
| **Output** |  | | | | |
| **Task 3** | **Handle Missing Values:**  • Replace missing (NaN) values with a suitable value (e.g., column mean or another appropriate  method). | | | | |
| **Code** | There will be no missing value column. | | | | |
| **Task 4** | **Extracting each colmun Information:**  • Finding mean of each feacture. | | | | |
| **Code** | import pandas as pd  # Load the dataset  file\_path = "pirvision\_office\_dataset1.csv"  df = pd.read\_csv(file\_path)  # Compute the mean for each numeric column  mean\_values = df.select\_dtypes(include=['number']).mean()  # Display the means  print(mean\_values) | | | | |
| **Output** |  | | | | |
| **Task 5** | **Extracting each colmun Information:**  • Finding mode of each feacture. | | | | |
| **Code** | import pandas as pd  # Load the dataset  file\_path = "pirvision\_office\_dataset1.csv"  df = pd.read\_csv(file\_path)  # Compute the mode for each column  mode\_values = df.mode().iloc[0]  # Display the modes  print(mode\_values) | | | | |
| **Output** |  | | | | |
| **Task 6** | **Extracting each colmun Information:**  • Finding median of each feacture. | | | | | **Analyze Worker Count Per Quarter:**  • Determine the number of workers who worked in each quarter.  • Sort the results in descending order based on worker count. |
| **Code** | import pandas as pd  # Load the dataset  file\_path = "pirvision\_office\_dataset1.csv"  df = pd.read\_csv(file\_path)  # Compute the median only for numeric columns  median\_values = df.select\_dtypes(include=['number']).median()  # Display the medians  print(median\_values) | | | | |
| **Output** |  | | | | |
| **Task 7** | **Extracting each colmun Information:**  • Finding variance of each feacture. | | | | |
| **Code** | import pandas as pd  # Load the dataset  file\_path = "pirvision\_office\_dataset1.csv"  df = pd.read\_csv(file\_path)  # Compute variance only for numeric columns  variance\_values = df.select\_dtypes(include=['number']).var()  # Display the variances  print(variance\_values) | | | | |
| **Output** | **Extracting each colmun Information:**  • Finding standard deviation of each feacture. | | | | |
|  | import pandas as pd  # Load the dataset  file\_path = "pirvision\_office\_dataset1.csv"  df = pd.read\_csv(file\_path)  # Compute standard deviation only for numeric columns  std\_dev\_values = df.select\_dtypes(include=['number']).std()  # Display the standard deviations  print(std\_dev\_values) | | | | |
|  |  | | | | |
| **THE END** | | | | | |