Bahria University

Karachi Campus

COURSE: CSL-460

DATA MINING LAB

TERM: SPRING 2024, CLASS: BSE- 6(B)

SUBMITTED BY:

**Mutayyab Imran**

**79315**

**(Name) (Enrollment No.)**

SUBMITTED TO:

Engr. HAMZA/Engr. HINA SHAKIR

SIGNED REMARKS: SCORE:

INDEX

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| SNO | DATE | LAB NO | LAB BJECTIVE | REMARKS/SIGN |
| 01 | 14-02-24 | 01 | GUI IN PYTHON USING GOOGLE COLAB & Data Mining Libraries |  |
| 02 | 21-02-24 | 02 | Exploratory data analysis using python best EDA libraries |  |
| 03 | 28-02-24 | 03 | Exploratory data analysis using python best EDA libraries |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Bahria University,

Karachi Campus

A logo with text on it

Description automatically generated

LAB EXPERIMENT NO.

03

LIST OF TASKS

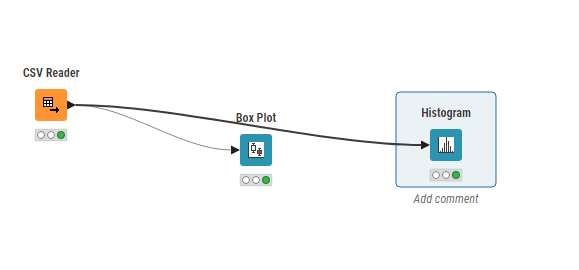
|  |  |
| --- | --- |
| TASK NO | OBJECTIVE |
| **01** | As part of a data analysis project, the team is investigating the Titanic dataset to uncover insights into passenger survival rates. One of the analysis tasks involves Implement the box plot and histogram for Titanic Data Set. |
| **02** | In a workflow group create an empty workflow “ Titanic Analysis”  A) read CSV file  B) remove irrelevant columns  C) enter the name and description for all nodes  D) performed basic statistics |
| **03** | Using KNIME, the team will analyze the Titanic dataset to identify the total number of male passengers who survived the disaster. This analysis will provide valuable insights into the survival patterns among male passengers aboard the Titanic, contributing to a better understanding of the factors influencing survival rates during the tragic event. |
| 04 | Calculate the total number of medals and appearances by male athletes in swimming disciplines in the Olympic Games between 1920 and 2020 by athlete and overall. |

Submitted On:

25-02-2024

(Date: DD/MM/YYYY)

**Task No. 01**: As part of a data analysis project, the team is investigating the Titanic dataset to uncover insights into passenger survival rates. One of the analysis tasks involves Implement the box plot and histogram for Titanic Data Set.

**Solution:**

**Output:**

A graph of a graph

Description automatically generatedA screenshot of a graph

Description automatically generated

**Task No. 02**: In a workflow group create an empty workflow “ Titanic Analysis”

A) read CSV file

B) remove irrelevant columns

C) enter the name and description for all nodes

D) performed basic statistics

E) performed the column and row filter and finally write the data to output file

F) save and execute workflow.

**A diagram of a missing value

Description automatically generatedSolution:**

**Output:**

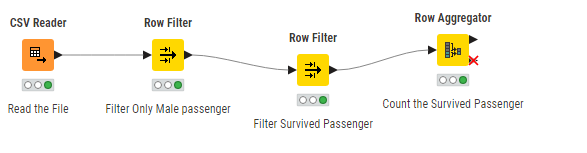
**A screenshot of a data report

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**Task No. 03**: Using KNIME, the team will analyze the Titanic dataset to identify the total number of male passengers who survived the disaster. This analysis will provide valuable insights into the survival patterns among male passengers aboard the Titanic, contributing to a better understanding of the factors influencing survival rates during the tragic event.

**Solution:**

**A screenshot of a computer

Description automatically generatedOutput:**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**Task No. 04**: Calculate the total number of medals and appearances by male athletes in swimming disciplines in the Olympic Games between 1920 and 2020 by athlete and overall.

A diagram of a computer

Description automatically generated**Solution:**

**Output:**

A pie chart with numbers and a green and blue circle

Description automatically generatedA pie chart with numbers and a number of different colors

Description automatically generated

A pie chart with numbers and text

Description automatically generated

A pie chart with numbers and a number of different colors

Description automatically generated