DATA VISUALIZATION USING TABLEAU (TABLEAU DASHBOARD)

A REPORT

SUBMITTED FOR TABLEAU DASHBOARD ASSIGNMENT

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

By

DHANUSHREE C – 1GA22AI013
RACHANA P – 1GA22AI038

GLOBAL ACADEMY OF TECHNOLOGY

An Autonomous Institute, Affiliated to VTU, Belagavi, Approved by AICTE, New Delhi, Accredited NAAC, 'A' Grade
Ideal homes township, Rajarajeshwari Nagar, Bengaluru-560098

Problem Statement:

Objective:
Develop an interactive and informative healthcare dashboard in Tableau that analyzes and visualizes key data related to heart failure to enhance understanding, support decision-making, and promote proactive healthcare measures. The goal is to provide an interactive and comprehensive tool for analyzing and visualizing heart failure data.

Overview:

1. Prepare Your Data
1. Gather Data:
   - ➢ Sources: Collect heart failure data from sources such as hospital databases, clinical studies, or public datasets.
   - ➢ Fields: Ensure you have key fields such as patient ID, age, gender, heart failure stage, survival rates, and other relevant health metrics.
2. Clean Data:
   - ➢ Identify Missing Values: Use tools like Excel or pandas in Python to find and address missing or incomplete data.
   - ➢ Correct Errors: Verify and correct any inaccuracies or inconsistencies in the data.
   - ➢ Format Data: Ensure uniformity in date formats, numerical precision, and categorical values.

2.Setting Up Tableau:
   - ➢ Load the prepared dataset into Tableau.
   - ➢ Structure the data in Tableau by setting up the necessary dimensions (e.g., death event, sex, smoking) and measures (e.g., age, creatinine phosphokinase).

3.Creating Visualizations:
   - ➢ **Donut Charts:** Use donut charts to visualize categorical data, such as the distribution of heart failure cases across different age groups or genders.

- ➢ **Histograms:** Display continuous variables like age or blood pressure using histograms to show the distribution.
- ➢ **Box Plots:** Use box plots to analyze the distribution of key factors and identify outliers.
- ➢ **Scatter Plots:** Explore correlations between continuous variables (e.g., age vs. blood pressure) using scatter plots.

4.Dashboard Assembly:
- ➢ Combine the individual visualizations into a cohesive dashboard.
- ➢ Add filters to enable users to interact with the data by selecting specific dates, routes, or distances.

5.Enhancing User Experience:
- ➢ Implement a Show/Hide button for filters to keep the dashboard clean and user-friendly.
- ➢ Ensure the dashboard is intuitive and easy to navigate, allowing users to derive insights quickly.

6.Finalization and Review:
- ➢ Test the dashboard to ensure all functionalities work as expected.
- ➢ Review the dashboard for accuracy, clarity, and visual appeal.
- ➢ Make any necessary adjustments based on stakeholder feedback before final deployment.

Dataset Used:
- ➢ **Dataset Name:** heart_failure_clinical_records_dataset.csv
- ➢ **Source:** Kaggle
- ➢ **Number of columns:** 13
- ➢ **Number of rows:** 300

Cardiovascular diseases (CVDs) are the number 1 cause of death globally, taking an estimated 17.9 million lives each year, which accounts for 31% of all deaths worlwide.

Heart failure is a common event caused by CVDs and this dataset contains 12 features that can be used to predict mortality by heart failure.Most cardiovascular diseases can be prevented by addressing behavioural risk factors such as tobacco use, unhealthy diet and obesity, physical inactivity and harmful use of alcohol using population-wide strategies.

People with cardiovascular disease or who are at high cardiovascular risk (due to the presence of one or more risk factors such as hypertension, diabetes, hyperlipidaemia or already established disease) need early detection and management wherein a machine learning model can be of great help.

Instructions :

Step 1: Connect to Data
- Open Tableau and create a new workbook.
- Connect to your data source (heart_failure_clinical_records_dataset.csv).
- Check the data in Tableau's Data Source tab to ensure it's loaded correctly.

Step 2: Measures to Dimensions
- Convert death event, high blood pressure, sex, smoking, diabetes into dimensions.

Step 3: Aliases
- Right click on event then click on aliases set 0 as survival and 1 as death.
- Right click on diabetes then click on aliases set 0 as negative and 1 as positive. Follow the same for high blood pressure, anemia.
- Right click on sex select aliases set 0 as female and 1 as male.
- Right click on smoking select aliases set 0 as non-smoker and 1 as smoker.

Step 4: Create Visualizations
Donut Charts:
- **Purpose:** To show the positive and negative survival

status.
- **Steps:**
  - ➢ Drag and drop death event into color and anemia into detail.
  - ➢ Convert into pie and then anemia into color.
  - ➢ Hold ctrl key and drag and drop anemia into angle.
  - ➢ Right click on anemia(angle) select measure➔ count.
  - ➢ Hold ctrl key and drag and drop anemia, death event and CNT into label.
  - ➢ Right click on CNT and select Quick Table Calculation.
  - ➢ Double click on rows and write 0➔ press ctrl key and duplicate the value.
  - ➢ Go the SUM(0)(2) and remove everything.
  - ➢ In rows right click on SUM(0)(2) and select dual axis.
  - ➢ Name the title as **"ANEMIA SURVIVAL STATUS".**

Adjust the size and formatting to create the donut effect.
- **Repeat the above steps for High Blood Pressure, Sex, Diabetes, Smoking Survival Status.**

Histograms:
- **Purpose:** To visualize the continuous variables such as distribution of patient ages, ejection fraction and so on.
- **Steps:**
  - ➢ Click on new worksheet.
  - ➢ Double click on age and click on show me➔ select histograms.
  - ➢ Drag and drop death event into color, red for death and green for survival.
  - ➢ Name the title as "**Age Survival Status**".
- **Repeat the above steps for Serum Creatinine, Serum Sodium, Time, Ejection Fraction, Creatinine Phosphokinase, Platelets Survival Status.**

Scatter Plots:
- **Purpose:** To visualize Age -Time Survival Status.
- **Steps:**

➢ Drag time into rows and age into columns.
➢ Convert to dimension: Right click on age and time and select dimensions.
➢ Drag and drop death event into color.
➢ Click on shape and select the shape .
➢ Click on Show me and select Scatter Plot.
➢ Name the title as **"Age-Time Survival Status"**.

Box Plots:
- **Purpose:** To visualize Age - Sex Survival Status.
- **Steps:**
  ➢ Drag and drop age into rows and sex, death event into columns.
  ➢ Right click on age and select measure➔ Average.
  ➢ Click on Show me and select Box Plot.
  ➢ Drag and drop sex into columns.
  ➢ Click on Analysis and select aggregate measures.
  ➢ Select the Entire view.
  ➢ Name the title as **"Age-Sex Survival Status"**.

Step 5: Create Text Sheets
➢ Click on worksheet.
➢ Select Create Calculated Field.
  - Total individuals➔"Total individuals"
  - Total Death➔ "Total Death"
  - Total males➔ "Total males"
  - Total females➔ "Total females"
  - Average age➔ "Average age"

Click Apply and then OK.

➢ Drag and drop Total individuals and death event into text.
➢ Right click on death event select measures➔Count.
- **Repeat the above steps for Total death, Total males, Total females, Average age.**

Step 6: Assemble the Dashboard
1. **Create a New Dashboard:**

   ➢ Drag the sheets (" Sheet 1, Sheet 2,..........Sheet 15 ") onto the dashboard.

   ➢ Arrange them according to your design.

   ➢ Add a title to the dashboard (e.g., "HealthCare – Heart Failure Dashboard").

➢ Resize and format the elements for a balanced layout(1900*1050).

Step 8: Final Adjustments

1. **Test the Dashboard:**

   ➢ Test the filters and interactions to ensure everything functions correctly.

   ➢ Adjust the layout and formatting as needed.

OUTPUT: