Healthcare PGP

Project 2

DESCRIPTION

Problem Statement

- NIDDK (National Institute of Diabetes and Digestive and Kidney Diseases) research creates knowledge about and treatments for the most chronic, costly, and consequential diseases.
- The dataset used in this project is originally from NIDDK. The objective is to predict whether or not a patient has diabetes, based on certain diagnostic measurements included in the dataset.
- Build a model to accurately predict whether the patients in the dataset have diabetes or not.

Dataset Description

The datasets consists of several medical predictor variables and one target variable (Outcome). Predictor variables includes the number of pregnancies the patient has had, their BMI, insulin level, age, and more.

Variables	Description
Pregnancies	Number of times pregnant
Glucose	Plasma glucose concentration in an oral glucose tolerance test
BloodPressure	Diastolic blood pressure (mm Hg)
SkinThickness	Triceps skinfold thickness (mm)
Insulin	Two hour serum insulin
BMI	Body Mass Index
DiabetesPedigreeFunction	Diabetes pedigree function
Age	Age in years
Outcome	Class variable (either 0 or 1). 268 of 768 values are 1, and the others are 0

Data Reporting:

Create a dashboard in tableau by choosing appropriate chart types and metrics useful for the business. The dashboard must entail the following:

Tableau Public Link :

https://public.tableau.com/profile/anand.jha#!/vizhome/Diabetesanalysis_16078457260250/DiabeticAnalysisHealthCareProjectVisualization

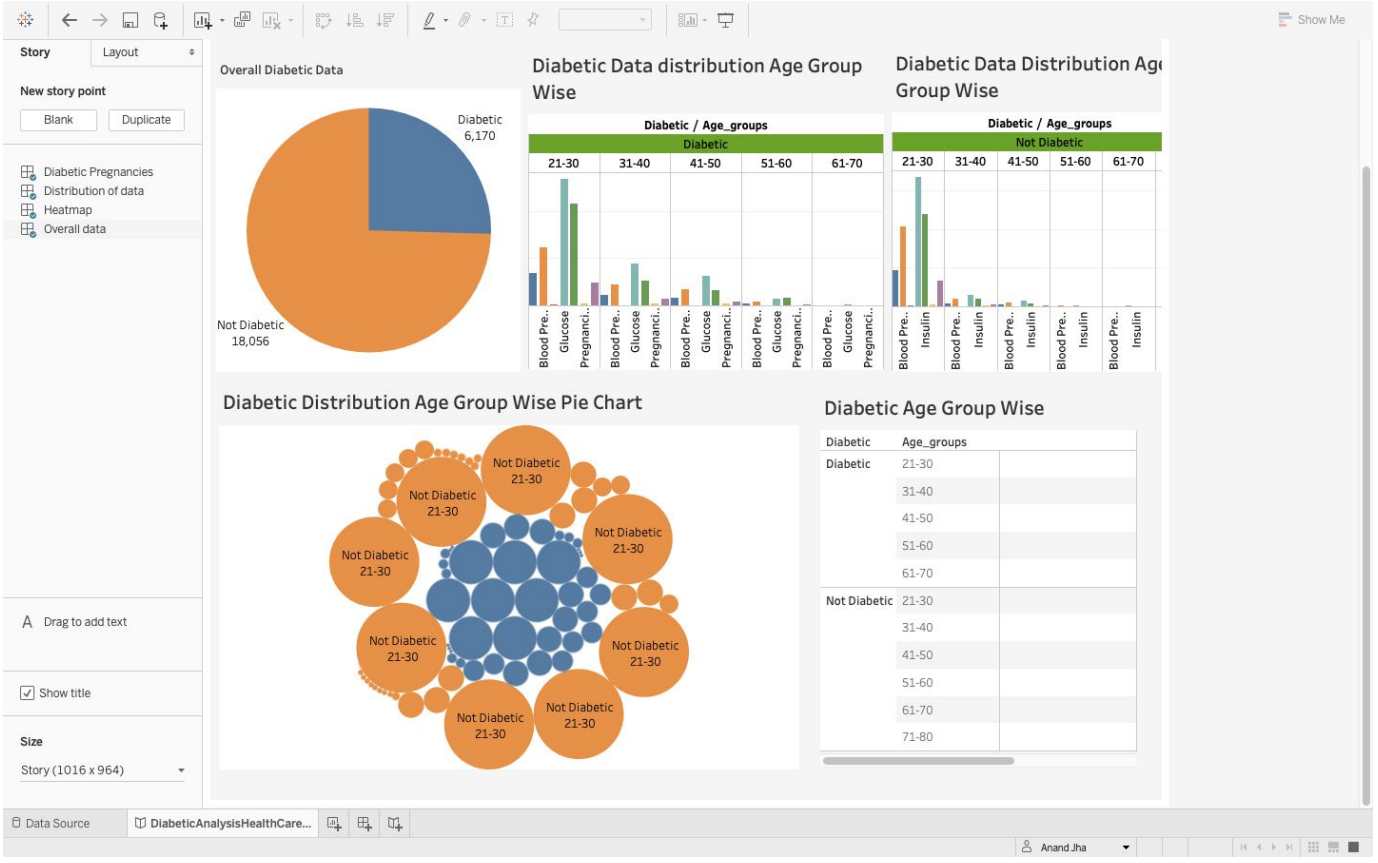
a. Pie chart to describe the diabetic or non-diabetic population

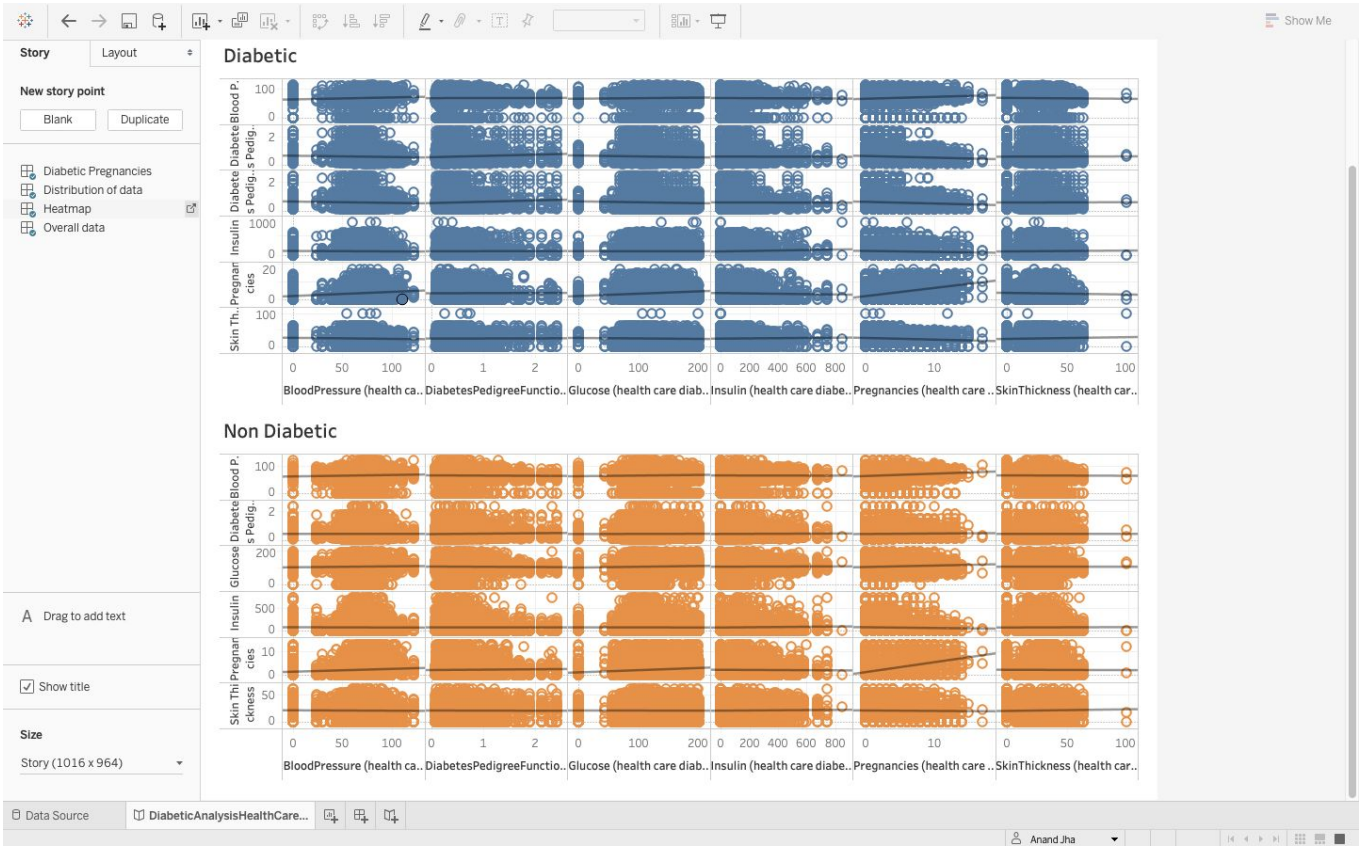
b. Scatter charts between relevant variables to analyze the relationships

c. Histogram or frequency charts to analyze the distribution of the data

d. Heatmap of correlation analysis among the relevant variables

e. Create bins of these age values: 20-25, 25-30, 30-35, etc. Analyze different variables for these age brackets using a bubble chart.




Story
Layout
New story point
Blank
Duplicate
Diabetic Pregnancies
Distribution of data
Heatmap
Overall data
Drag to add text
Show title
Size
Story (1016 x 964)
Diabetic
Non Diabetic
BloodPressure (health ca..
DiabetesPedigreeFunctio..
Glucose (health care diab..
Insulin (health care diabe..
Pregnancies (health care ..
SkinThickness (health car..
Data Source
DiabeticAnalysisHealthCare...
Show Me


Dashboard
Layout
Default
Phone
Device Preview
Size
Automatic
Sheets
DiabeticDataDistri...
NonDiabeticDataDi...
Objects
Horizontal
Blank
Vertical
Navigation
Text
Download
Image
Extension
Web Page
Tiled
Floating
Show dashboard title
Diabetic Data distribution Age Group Wise
Diabetic / Age_groups
Diabetic
21-30
31-40
41-50
51-60
61-70
Diabetic Data Distribution Age Group Wise
Not Diabetic
71-80
Measure Names
BMI
Blood Pressure
Diabetes Pedigree F..
Glucose
Insulin
Pregnancies
Skin Thickness
Data Source
DiabeticAnalysisHealthCareP...
Distribution of data
Show Me


Entire View
Show Me
Dashboard
Layout
Default
Phone
Device Preview
Size
Automatic
Sheets
DiabeticPregnancies
Objects
Horizontal
Blank
Vertical
Navigation
Text
Download
Image
Extension
Web Page
Tiled
Floating
Show dashboard title
iabetic Pregnancies
Diabetic
Pregnancies
Not Diabetic
Blood Pressure
Glucose
Insulin
Diabetes Pedigree Function
Skin Thickness
Data Source
DiabeticAnalysisHealthCareP...
Diabetic Pregnancies
2496 marks
2 rows by 5 columns
Anand Jha