**DS Capstone Project 2- Healthcare Tableau Report**

**Problem Statement**

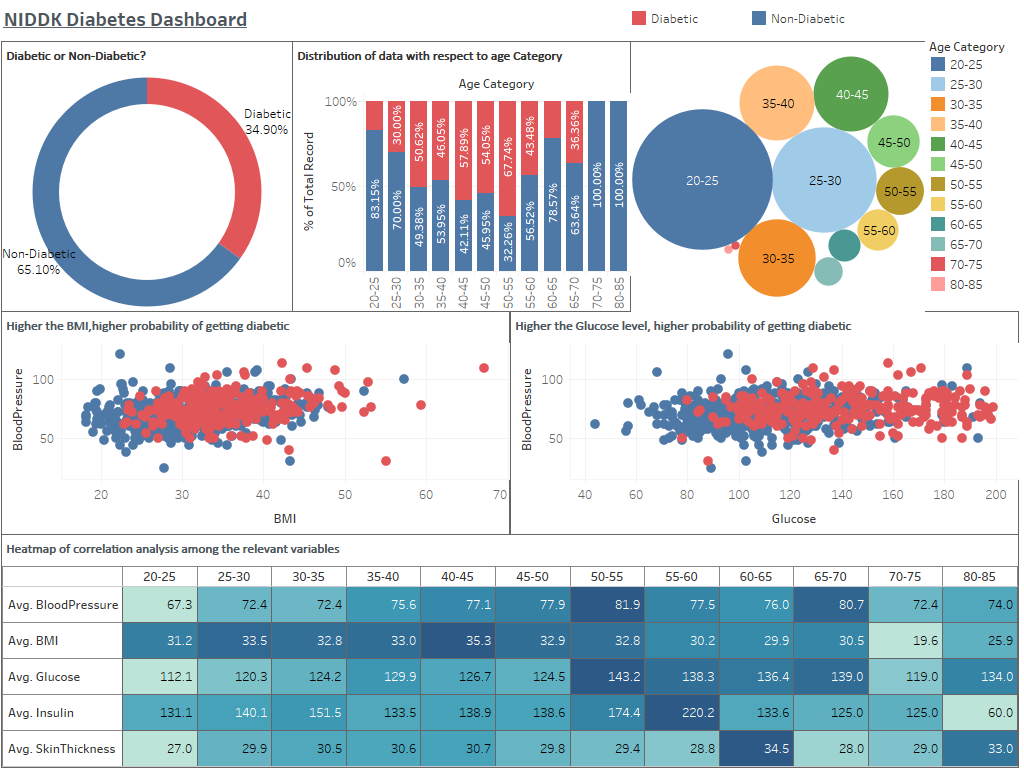
* NIDDK (National Institute of Diabetes and Digestive and Kidney Diseases) research creates knowledge about and treatments for the most chronic, costly, and consequential diseases.
* The dataset used in this project is originally from NIDDK. The objective is to predict whether or not a patient has diabetes, 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 consist 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.

**Project Task: Week 4**

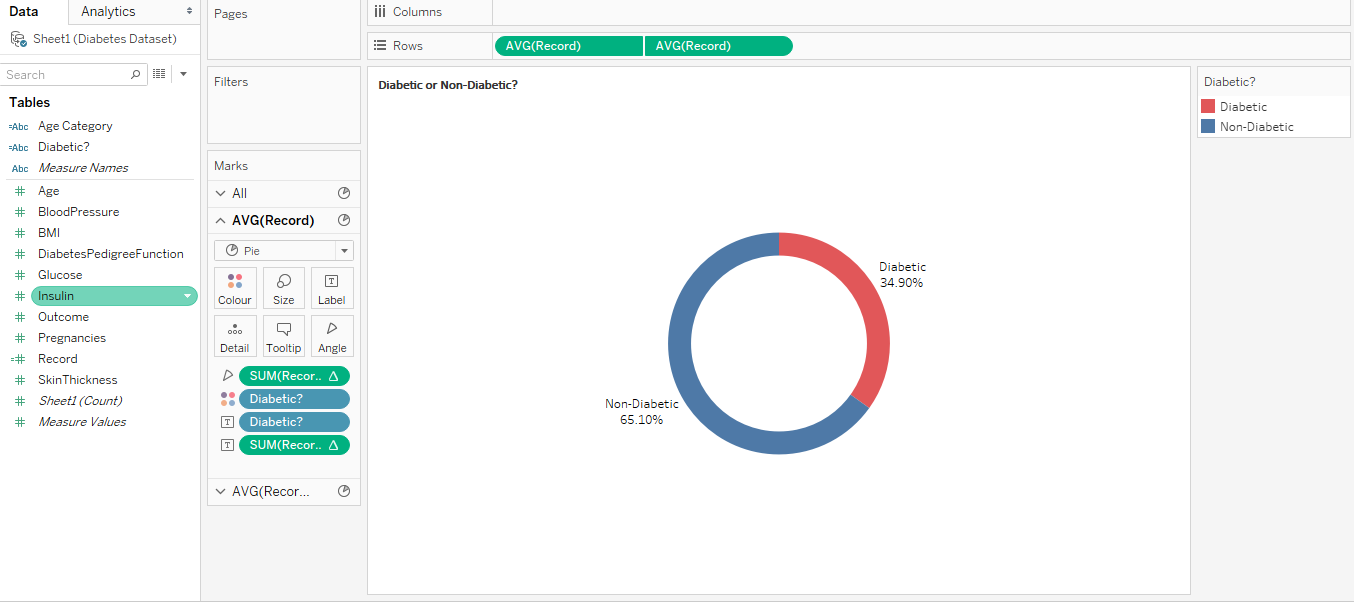
**Link :** [Project 2 - Diabetes - Pranay Dutta | Tableau Public](https://public.tableau.com/profile/pranay.dutta#!/vizhome/Project2-Diabetes/Dashboard?publish=yes)



**Data Reporting:**

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

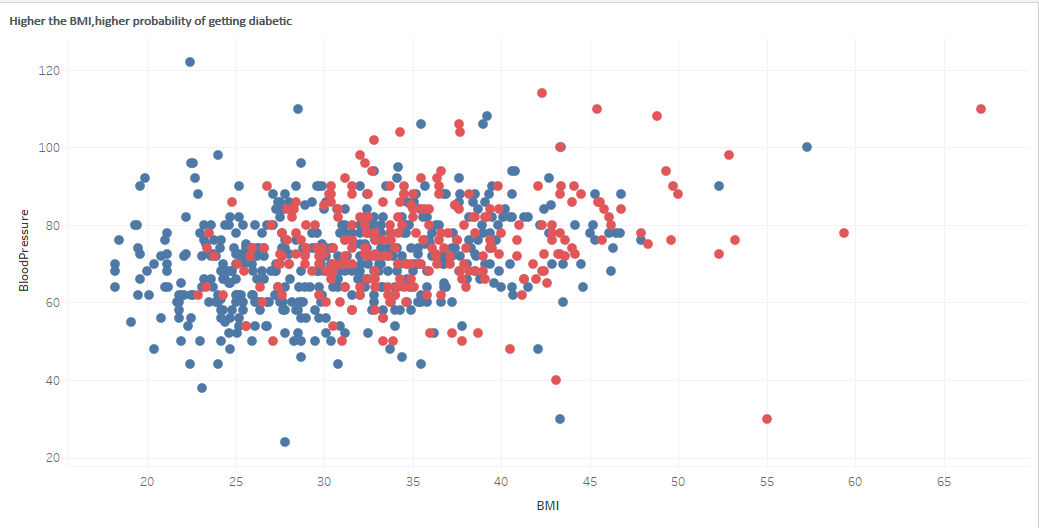
a. Pie chart to describe the diabetic or non-diabetic population



**Observations:** We have 34.9 % diabetic patient’s information.

b. Scatter charts between relevant variables to analyze the relationships

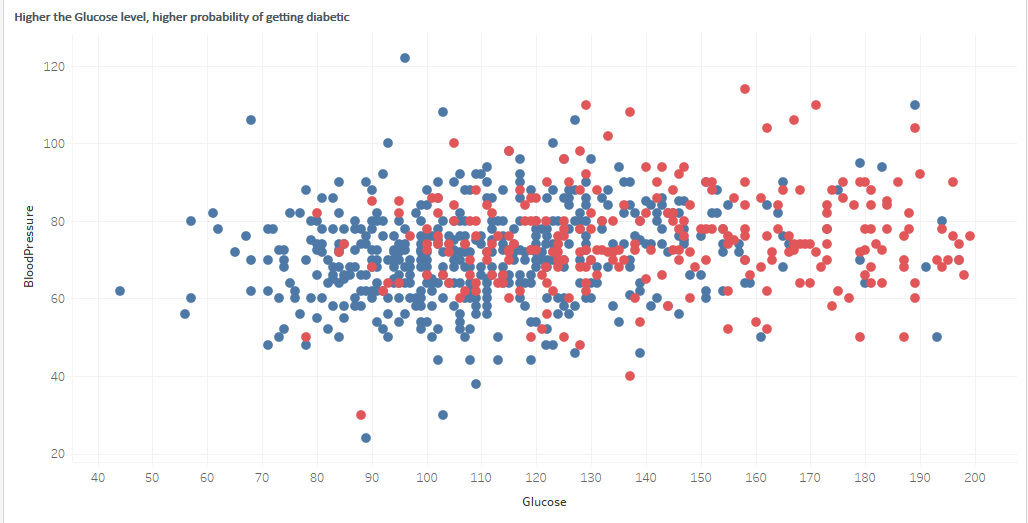
Scatter plot between BMI and Blood Pressure, Colour bifurcated by Outcome .:



**Observations:** From the above chart we can conclude that diabetic patients tend to have high BMI

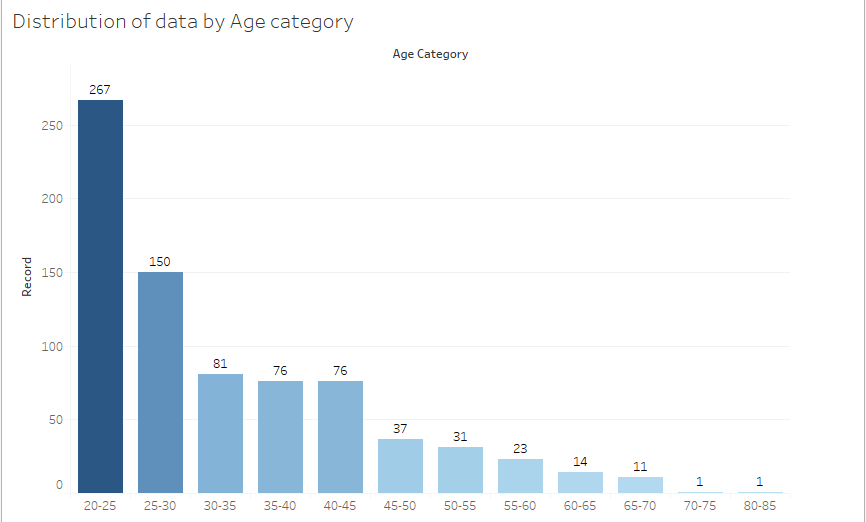
Scatter plot between Glucose and Blood Pressure, Colour bifurcated by Outcome .

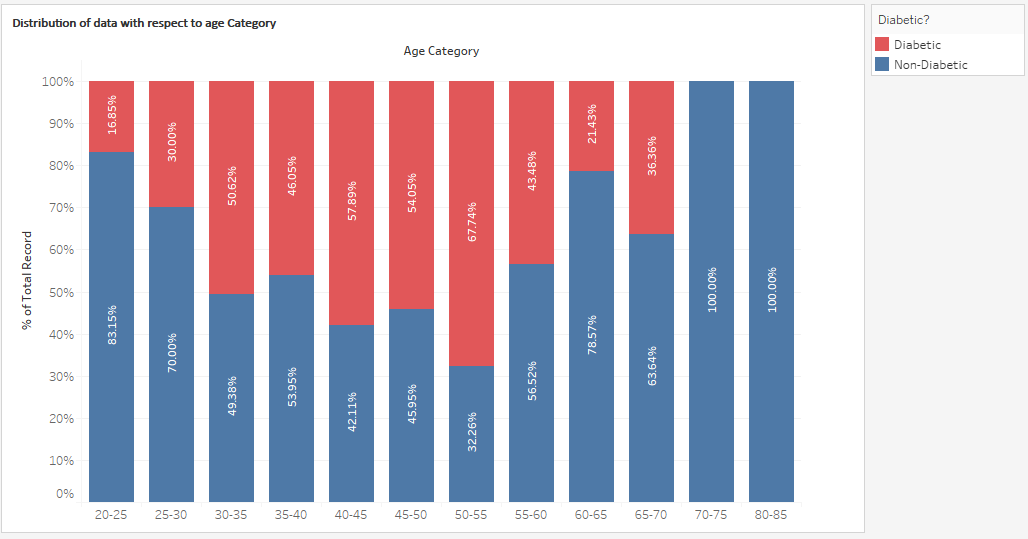




**Observations:** From the above chart we can conclude that diabetic patients tend to have high Glucose level

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

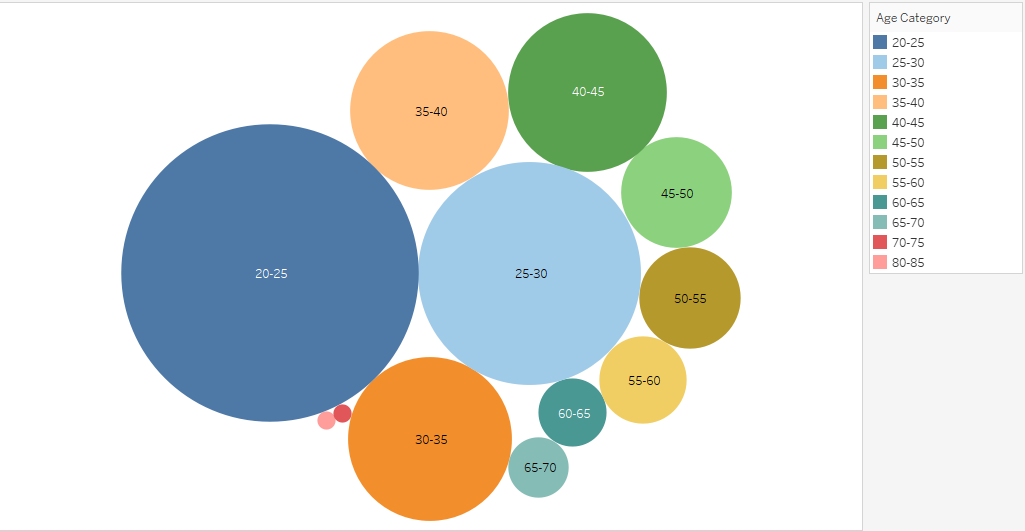
**Observations:** Out of 768 records we have 267 no of patients information within the age category 20-25

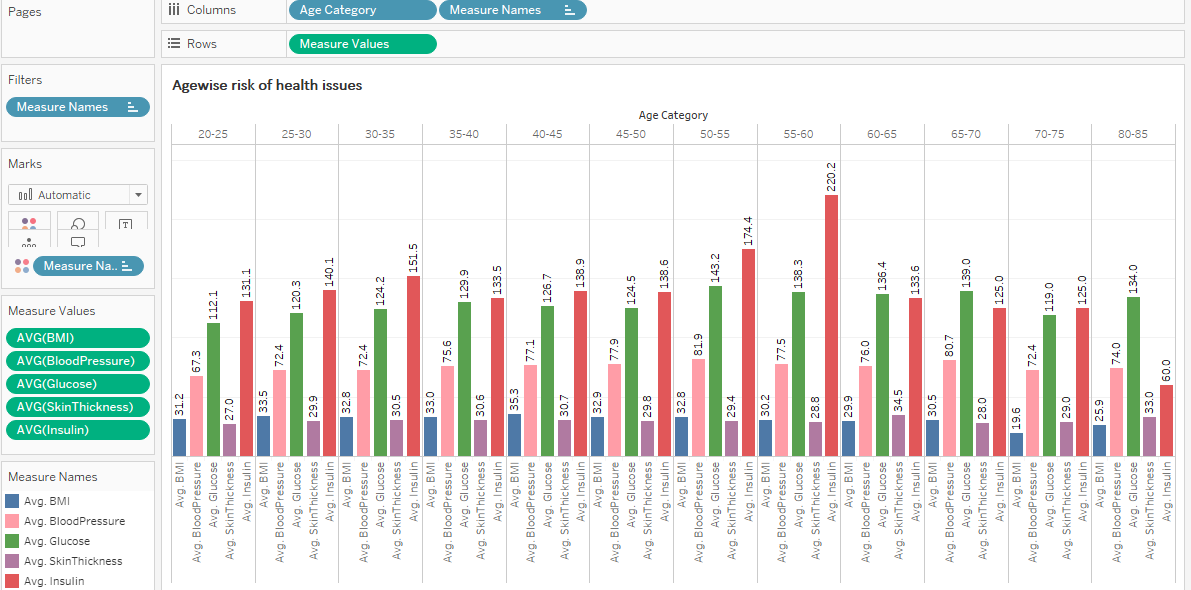


d. Heatmap of correlation analysis among the relevant variables



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.





**Observation :**

1. Glucose, blood pressure & insulin levels increase with age,
2. Skin thickness decreases with age and Skin thickness doesn't seem to be an indicator of any health issues