**Healthcare Project –**

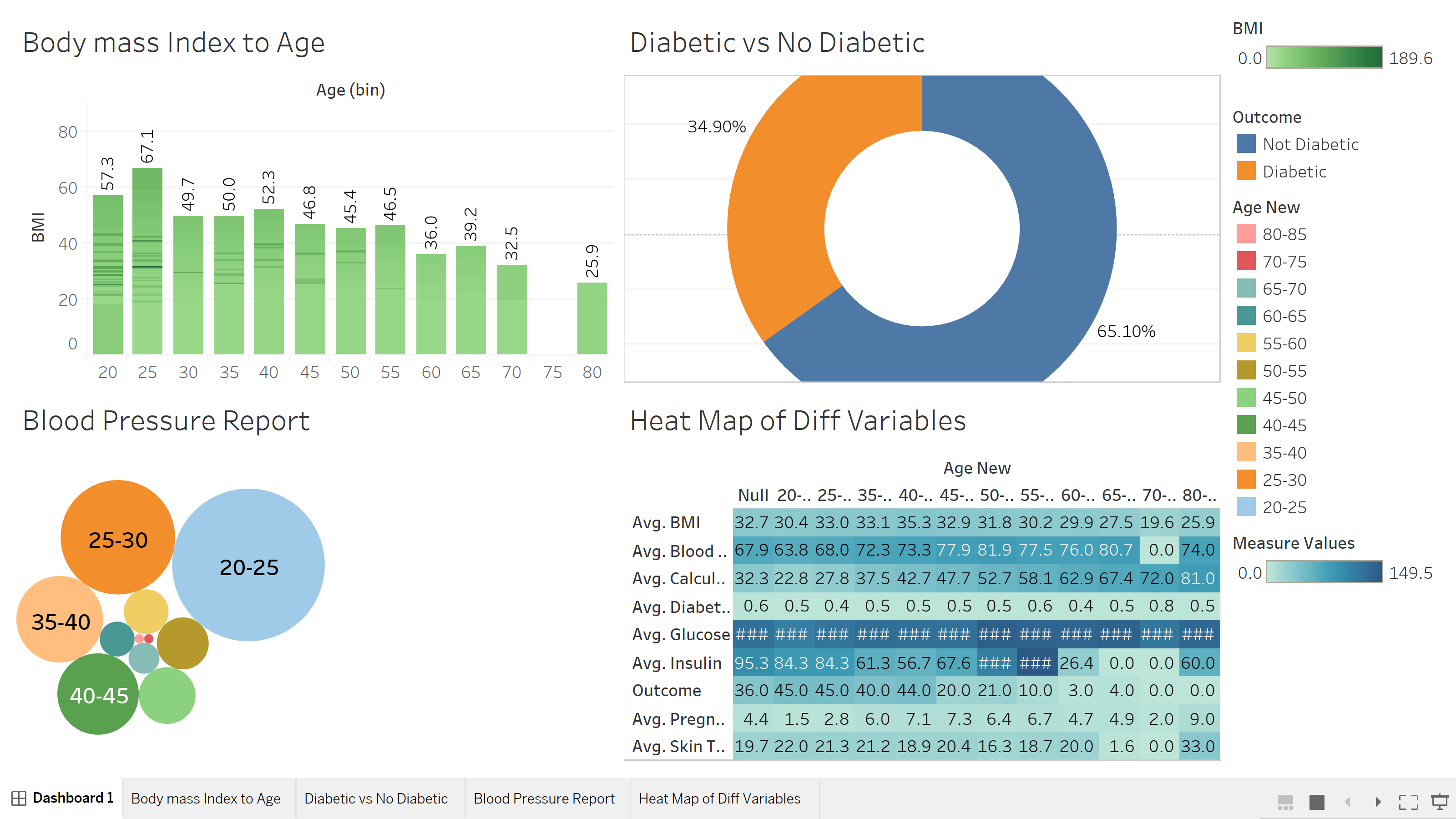
[CountMeIn (github.com)](https://github.com/CountMeIn)

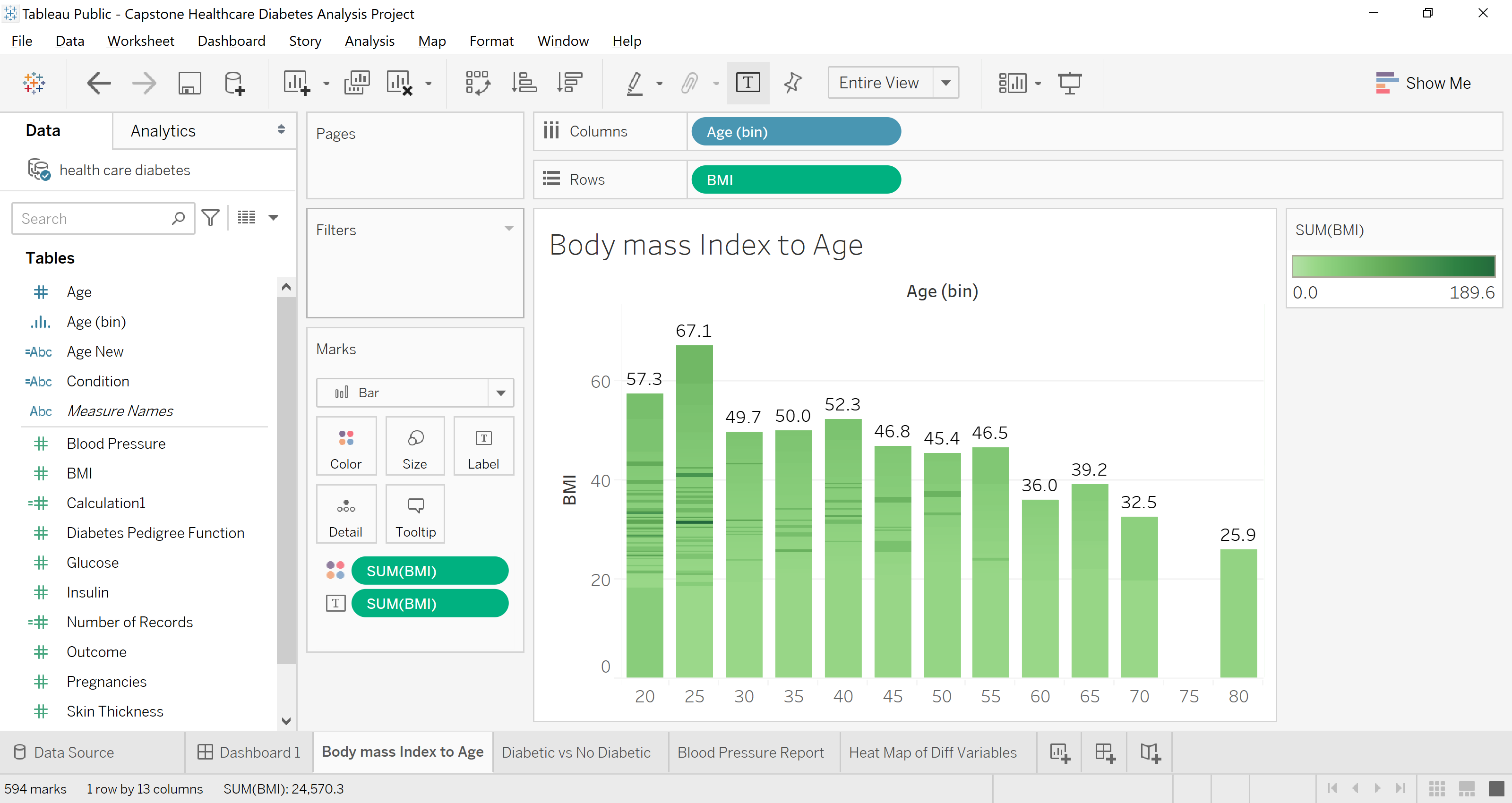
All activities have been completed and submitted as part of python code, results and code PDF and the Tableau outputs. I have attached all the outputs at the end of the file too.

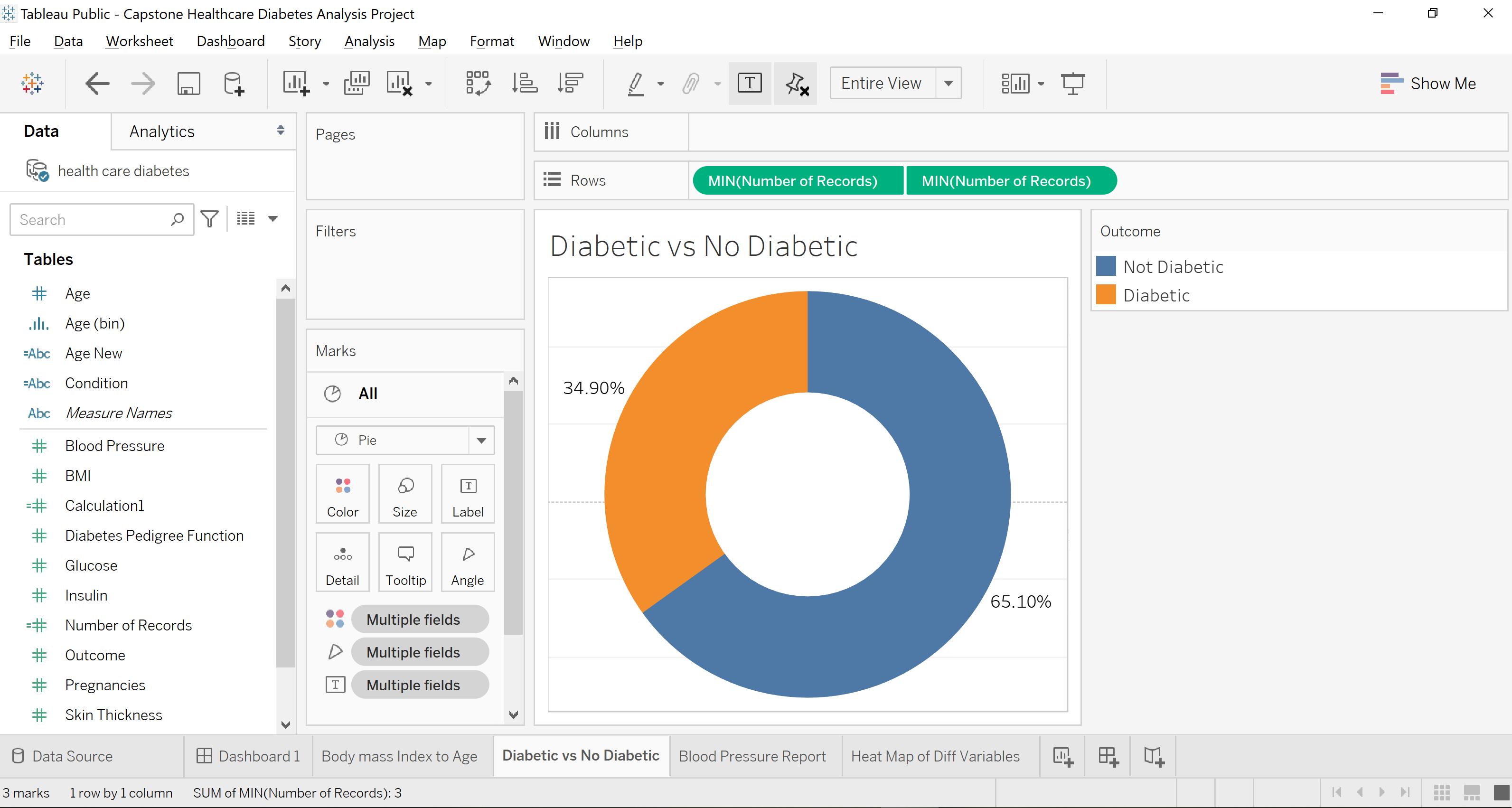
All tasks that have been completed are marked as **Done**

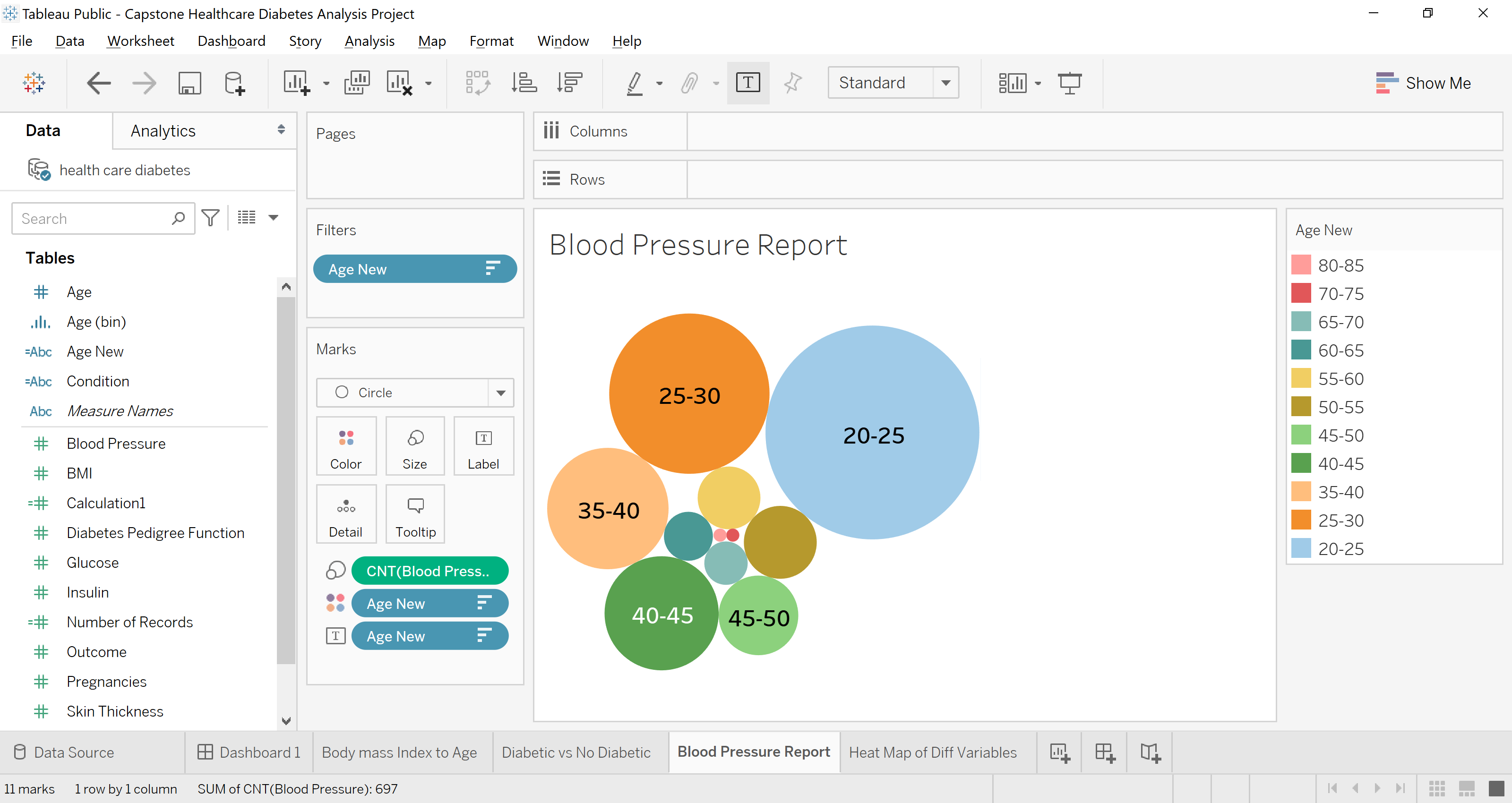
1. **Tableau Outputs**

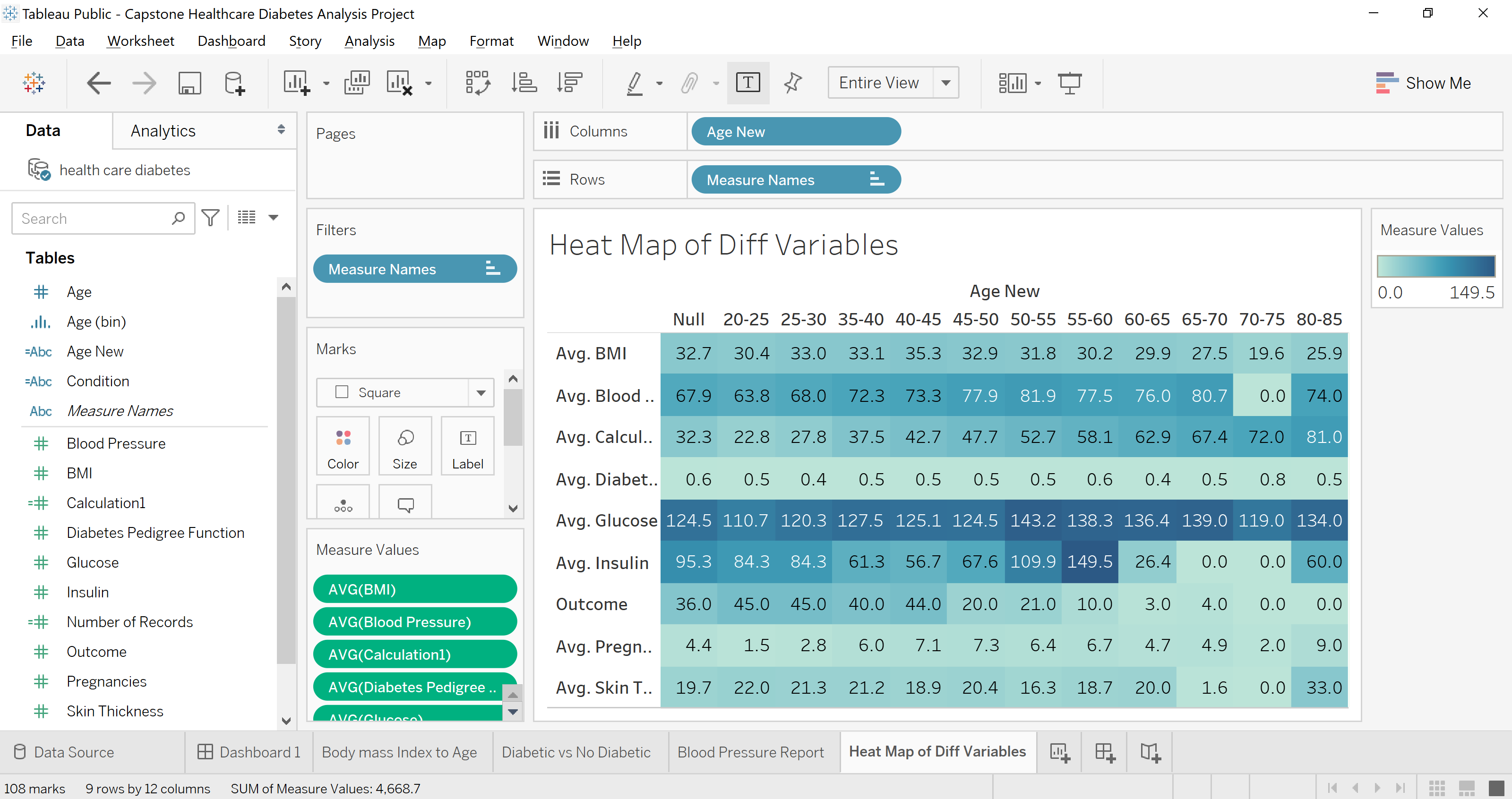
[Rajesh Nagarjunan - Profile | Tableau Public](https://public.tableau.com/profile/rajesh.nagarjunan#!/)



****

****

****

****

1. **Project Task and Completion Status**

### ****Project Task: Week 1****

**Data Exploration:**

1. Perform descriptive analysis. Understand the variables and their corresponding values. On the columns below, a value of zero does not make sense and thus indicates missing value: - **Done**

• Glucose

• BloodPressure

• SkinThickness

• Insulin

• BMI

2. Visually explore these variables using histograms. Treat the missing values accordingly. - **Done**

3. There are integer and float data type variables in this dataset. Create a count (frequency) plot describing the data types and the count of variables. - **Done**

### ****Project Task: Week 2****

**Data Exploration:**

1. Check the balance of the data by plotting the count of outcomes by their value. Describe your findings and plan future course of action. - **Done**

2. Create scatter charts between the pair of variables to understand the relationships. Describe your findings. - **Done**

3. Perform correlation analysis. Visually explore it using a heat map. - **Done**

### ****Project Task: Week 3****

**Data Modeling:**

1. Devise strategies for model building. It is important to decide the right validation framework. Express your thought process. - **Done**

2. Apply an appropriate classification algorithm to build a model. Compare various models with the results from KNN algorithm. - **Done**

**Project Task: Week 4**

**Data Modeling:**

1. Create a classification report by analyzing sensitivity, specificity, AUC (ROC curve), etc. Please be descriptive to explain what values of these parameter you have used. - **Done**

**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- **Done**

b. Scatter charts between relevant variables to analyze the relationships- **Done**

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

d. Heatmap of correlation analysis among the relevant variables- **Done**

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. - **Done**