

Vanessa Munoz

CIS 4150: Foundations of Business Intelligence

October 24th, 2021

Heart Failure Predictions

## **Introduction:**

In this project, I will conduct an analysis of heart failure predictions on many different people struggling with heart conditions. The overall main goal is not to only gain more knowledge about heart condition predictions but to also understand the highest and lowest risk factors on different people who have heart conditions. Every year, many healthcare organizations have been collecting data but many healthcare workers are unsure of how to utilize the information they have gathered. At the end of this project, I'll be showing a variety of different trends and visualizations by using 2 different datasets on heart failure predictions. The two datasets that I will be using will consist of different heart predictions from the month of November 2020 through the current month and year. Data analysis is an incredibly important aspect of business in many healthcare organizations. Being able to process and analyze data about the public can help healthcare workers make informed decisions that can create an incredible competitive advantage in the work field. In this case, having knowledge of the prediction and risk factors of heart failure can positively impact many patients that are struggling with heart conditions. As cardiovascular diseases rapidly increase every day, many people are unaware that the number 1 leading cause of death in America is heart failure. According to this article by Renee Twombly, "In 2002, 696,947 people in the United States died of heart disease, compared with 557,271 deaths from cancer. This is because heart disease is particularly lethal to those aged 85 and older (R. Thombly, 2005). With many healthcare industries gathering a tremendous amount of data, many healthcare workers have not yet embedded data visualizations within their analytical teams. Data analytics can be used by many healthcare organizations to discover viewer trends and to find out different information that'll be helpful to them in the near future. As technology advances, there are many different application tools that analysts can use

in creating dashboards and visualization charts. By using Tableau Prep Builder for data cleaning and analysis, many healthcare organizations can create many visualized data that'll show the predicted rate of patients with heart conditions.

**Objective:**

In this project, I will be designing different visualizations with Tableau Chart in predicting the survival rate for patients with heart failure. The two datasets will be obtained from the website called Kaggle. Then the datasets will be uploaded into Tableau where I will be able to map the data. And I will create different visualizations through a pie chart, bar chart, and line chart. The datasets will have over 300 rows of data, and it has 10-13 different columns. The datasets will also be both csv format files. As patients discover their risk factors for cardiovascular disease at an early age, many hospitals can gather and keep track of all patients' progress and outcomes. The process of having cardiovascular diseases usually evolves over decades, but it usually starts as early as our childhood. Throughout time people usually change their eating habits and also develop unhealthy behaviors at a young age, many people start to develop serious health conditions and diseases during their life. According to this article by Castelli M.D., "An increase in total to high-density lipoprotein cholesterol ratio, hypertension, cigarette smoking, excess weight, elevated blood sugar levels, lack of exercise, stress, electrocardiographic abnormalities, and other factors are associated with the development of these diseases" (W.P. Castelli, 1984). With many people struggling with cardiovascular diseases earlier in life, they are unaware of the risk factors for cardiovascular disease. By being aware of the different risk factors, many people can live a healthy lifestyle. It is important to improve the rate of heart disease at an early age because it'll be a tremendous impact on their life when they're older.

**Motivation:**

Heart disease can be predicted based on many symptoms such as age, gender, and blood type. As hospitals use data analysis, many healthcare facilities will be able to predict future diseases and improve diagnosis to reduce the death rate in heart patients. The purpose of this project is to introduce Tableau (visualization software) and provide impactful examples of Tableau. By creating visualizations in using Tableau Prep Builder, I will be able to find trends regarding the highest and lowest risk factors in heart failure prediction. Also, all visualizations will be available for students, healthcare workers, and also academic users. Hospitals would be able to give lectures and notes on many different heart failure predictions by utilizing their dataset. Hospitals would be able to identify patients' heart failures more quickly by keeping better track of their patient's health predictions. The primary purpose of the dataset implementation is to improve the patient's health predictions within all hospitals. The data visualizations in Tableau will have a positive impact on many healthcare organizations. And many patients will improve their recovery through heart diseases when they realize the highest number of risk factors in heart disease.

**Analytical Questions:**

Is there any relationship between gender and chest pain types for the patients who have heart disease?

Which gender is more likely to have heart disease?

Which work type is more likely to have a stroke?

Is there a relationship between gender, residence type, and smoking status?

**References:**

Twombly, R. (2005). Cancer surpasses heart disease as leading cause of death for all but the very elderly. *Journal of the National Cancer Institute*, 97(5), 330-331.

Castelli, W. P. (1984). Epidemiology of coronary heart disease: the Framingham study. *The American journal of medicine*, 76(2), 4-12.

**Dataset URL:**

<https://www.kaggle.com/fedesoriano/heart-failure-prediction>

<https://www.kaggle.com/andrewmvd/heart-failure-clinical-data>

## Datasets in Tableau:

| heart 12 fields   Filter Values...   |      |                |                     |         |               |
|--|------|----------------|---------------------|---------|---------------|
| Clear the check box to remove fields. You can also filter your data or change data types. <a href="#">Add a clean step</a> to view and clean data. |      |                |                     |         |               |
| Fields selected: 12 of 12  |      |                |                     |         |               |
| <input checked="" type="checkbox"/>  | Type | Field Name     | Original Field Name | Changes | Preview       |
| <input checked="" type="checkbox"/>  | #    | Age            | Age                 |         | 40, 49, 37    |
| <input checked="" type="checkbox"/>  | Abc  | Sex            | Sex                 |         | M, F          |
| <input checked="" type="checkbox"/>  | Abc  | ChestPainType  | ChestPainType       |         | ATA, NAP      |
| <input checked="" type="checkbox"/>  | #    | RestingBP      | RestingBP           |         | 140, 160, 130 |
| <input checked="" type="checkbox"/>  | #    | Cholesterol    | Cholesterol         |         | 289, 180, 283 |
| <input checked="" type="checkbox"/>  | #    | FastingBS      | FastingBS           |         | 0             |
| <input checked="" type="checkbox"/>  | Abc  | RestingECG     | RestingECG          |         | Normal, ST    |
| <input checked="" type="checkbox"/>  | #    | MaxHR          | MaxHR               |         | 172, 156, 98  |
| <input checked="" type="checkbox"/>  | Abc  | ExerciseAngina | ExerciseAngina      |         | N             |
| <input checked="" type="checkbox"/>  | #    | Oldpeak        | Oldpeak             |         | 0, 1          |
| <input checked="" type="checkbox"/>  | Abc  | ST_Slope       | ST_Slope            |         | Up, Flat      |
| <input checked="" type="checkbox"/>  | #    | HeartDisease   | HeartDisease        |         | 0, 1          |

| heart_failure_clinical_records_dataset 2 13 fields   Filter Values...  |      |                    |                          |         |                              |
|--|------|--------------------|--------------------------|---------|------------------------------|
| Clear the check box to remove fields. You can also filter your data or change data types. <a href="#">Add a clean step</a> to view and clean data. |      |                    |                          |         |                              |
| Fields selected: 13 of 13  |      |                    |                          |         |                              |
| <input checked="" type="checkbox"/>  | Type | Field Name         | Original Field Name      | Changes | Preview                      |
| <input checked="" type="checkbox"/>  | #    | age                | age                      |         | 75, 55, 65                   |
| <input checked="" type="checkbox"/>  | #    | anaemia            | anaemia                  |         | 0                            |
| <input checked="" type="checkbox"/>  | #    | creatinine_phos... | creatinine_phosphokinase |         | 582, 7,861, 146              |
| <input checked="" type="checkbox"/>  | #    | diabetes           | diabetes                 |         | 0                            |
| <input checked="" type="checkbox"/>  | #    | ejection_fraction  | ejection_fraction        |         | 20, 38                       |
| <input checked="" type="checkbox"/>  | #    | high_blood_pre...  | high_blood_pressure      |         | 1, 0                         |
| <input checked="" type="checkbox"/>  | #    | platelets          | platelets                |         | 265,000, 263,358.03, 162,000 |
| <input checked="" type="checkbox"/>  | #    | serum_creatini...  | serum_creatinine         |         | 1.9, 1.1, 1.3                |
| <input checked="" type="checkbox"/>  | #    | serum_sodium       | serum_sodium             |         | 130, 136, 129                |
| <input checked="" type="checkbox"/>  | #    | sex                | sex                      |         | 1                            |
| <input checked="" type="checkbox"/>  | #    | smoking            | smoking                  |         | 0, 1                         |
| <input checked="" type="checkbox"/>  | #    | time               | time                     |         | 4, 6, 7                      |
| <input checked="" type="checkbox"/>  | #    | DEATH_EVENT        | DEATH_EVENT              |         | 1                            |

**Data Table Descriptions:**

| <b>Field Names:</b> | <b>Names:</b>       | <b>Description:</b>  |
|---------------------|---------------------|--|
| Age                 | Ages                | The ages of patients with heart disease.   |
| Sex                 | Gender              | The gender of patients with heart disease.   |
| Diabetes            | Diabetes            | Patients with diabetes, and they're struggling with heart disease.   |
| High_Blood_Pressure | High Blood Pressure | Patients with high blood pressure, and they're struggling with heart disease.  |
| Exercise_Angina     | Exercise Angina     | Exercise-induced angina  |
| Year                | Years               | Years of patients who struggle with heart disease.   |
| Anaemia             | Anemia              | Patients who have produced fewer red blood cells than normal and they're struggling with heart disease   |
| Ejection_fraction   | Ejection Fraction   | Percentage of blood leaving the heart at each contraction  |
| MaxHR               | Maximum Heart Race  | Maximum heart rate achieved  |
| Cholesterol         | Cholesterol         | When cholesterol is high in your blood, it builds up in the walls of your arteries, causing a process called atherosclerosis which is a form of heart disease. |