Data Visualization in Cardiology

Vijayalakshmi S
Division of Mathematics
School of Advanced Sciences
Vellore Institute of Technology
Chennai, India
vijayalakshmi.s2021@vitstudent.ac.in

Venkatalakshmi S
Division of Mathematics
School of Advanced Sciences
Vellore Institute of Technology
Chennai, India
venkatalakshmi.2021@vitstudent.ac.in

Dr. Kalyani Desikan
Professor
Division of Mathematics
School of Advanced Sciences
Vellore Institute of Technology
Chennai, India
kalyanidesikan@vit.ac.in

***Abstract* — With the widespread use of the internet and global communication systems in today's digital world, information is rapidly growing. Healthcare systems are also undergoing digital transformations due to increased use of healthcare information systems, electronic medical records, wearable, smart, and handheld devices, among other things. These digital transformations generate a large amount of data. Big data plays a vital role in healthcare, especially in the field of cardiology. The problem with such data is that it takes a lot of time to analyze and identify the disease accurately. Visualizations help in understanding and analyzing the data better.**

In this paper we apply different data visualization techniques in the field of cardiology, with a specific focus on heart failure and valve replacement, to get a good understanding of the data and predict the disease. The benefits of the study are to make it easier for predicting the survival of the patients. Data set with features such as age, gender, blood pressures, sugar level, type of heart disease is used in our study. R programming is used to visualize the data and predict heart disease.

Keywords— Healthcare, Cardiology Big data, Data Visualization Techniques

I. INTRODUCTION

Exploratory Data Analysis (EDA) is a pre-processing procedure that allows you to better understand your data. Although there are several ways and stages for completing EDA, the most of them are narrowly focused, focusing on either visualization or dispersion, and are thus incomplete. Univariate data analyzes only one data column at a time, whereas multivariate data examines more than two variables. Extract averages, mean, minimum, and maximum values, among other things, to gain a better understanding of variables. Identify data mistakes, outliers, and missing values. Visualize data in graphs such as box plots, scatter plots, histograms, and so on to find trends.

II. HEART DISEASE

Visualizing cardiac disease is the most difficult task in medicine, and it is based on things such as the patient's physical examination, symptoms, and signs [1-3]. The smoking habit, family history of diseases, blood pressure, and working environment are all factors that influence heart disease. It is widely regarded as the most lethal disease in human history. In particular, in this type of disease, the heart is unable to deliver the necessary amount of blood to the remaining organs of the human body in order for them to perform their normal functions [4]. Physical bodily weakness, inappropriate breathing, swollen feet, and other signs of heart disease are only a few. The approaches are necessary for detecting difficult heart disorders that carry a significant risk of impacting human life [5].

A range of signs, such as age, gender, pulse rate, and so on, can be used to predict heart disease. Disease prediction, improved diagnosis, symptom analysis, delivering appropriate medications, enhancing quality of care, lowering expenses, extending life duration, and lowering the mortality rate of cardiac patients are all aided by data analysis in healthcare.

Heart failure frequently develops after the heart has been damaged or weakened by other conditions. Heart failure, on the other hand, can occur if the heart becomes too stiff.

Heart failure and aging - It is usually affected by the oldest. It has no effect on age.

III. DATASET

Cardiovascular diseases (CVDs) are the leading cause of death worldwide, killing an estimated 17.9 million people each year, accounting for 31% of all deaths. CVDs are a common cause of heart failure, and this dataset contains 12 variables that can be used to visualize heart failure mortality. Those suffering heart disease or those who are at high heart disease risk (due to the presence from one or many risk factors such as high, diabetes, hyperlipidemia, or pre-existing illness) require early identification and care. The research is based on data for heart failure that is publicly available. Age, sex, high blood pressure, diabetes, anaemia, creatinine phosphokinase, ejection fraction, platelets, serum creatinine, serum sodium,

and smoking are among the 300 entries in the dataset. We can find the relationships using this dataset while retaining the age and sex as the primary attributes. For instance, we can visualize the chance of survival of an heart failure person with anaemia

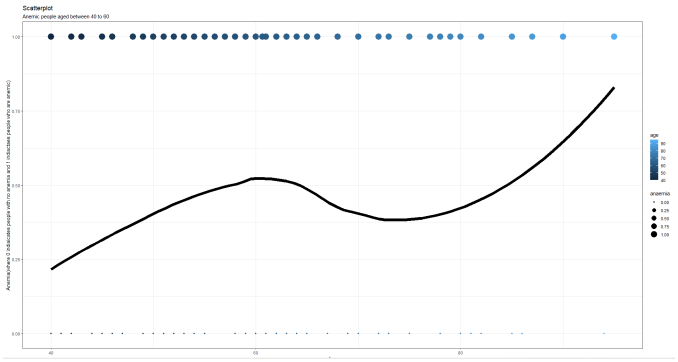

Fig 1.1

As shown in fig. 1.1, the amount of anemia rises with age. People between the ages of 50 and 60 have higher levels of anemia in their bodies.

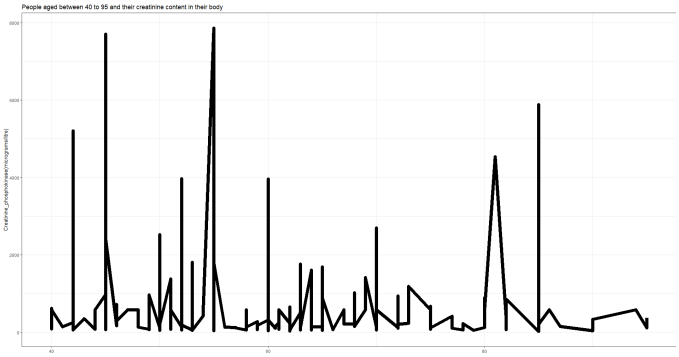

Fig 1.2

Adults between the ages of 40 and 50 had a larger creatinine content in their bodies than people between the ages of 50 and 60, according to fig 1.2 from the dataset we have.

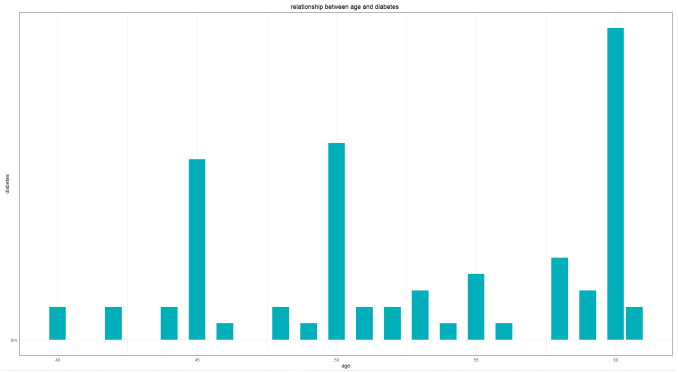

Fig 1.3

Between the ages of 40 and 60, those with diabetes and non-diabetes have the same ratio, as shown in fig 1.3.

age	sex	high_bloo	DEATH_EV
49	0	1	0
50	0	1	1
50	0	1	1
53	0	1	1
50	0	1	1
55	0	1	0
59	0	1	1
58	0	1	0
60	0	1	0
60	0	1	0
60	0	1	0
60	0	1	0
46	0	1	1
46	0	1	0
40	0	1	0
45	0	1	0
50	0	1	0
49	0	1	0
48	0	1	1
54	0	1	1
55	0	1	0
50	0	1	0
45	0	1	0

table 1.4

According to table. 1.4, those between the ages of 40 and 60 have higher blood pressure than those above the age of 60.

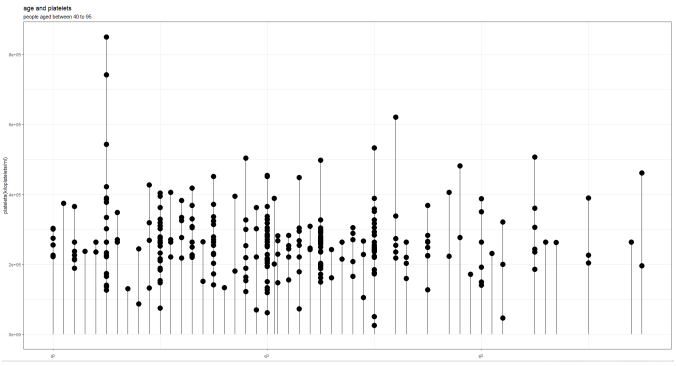

Fig 1.5

According to fig 1.5, the platelet count is higher between the ages of 40 and 60 than between the ages of 60 and 95.

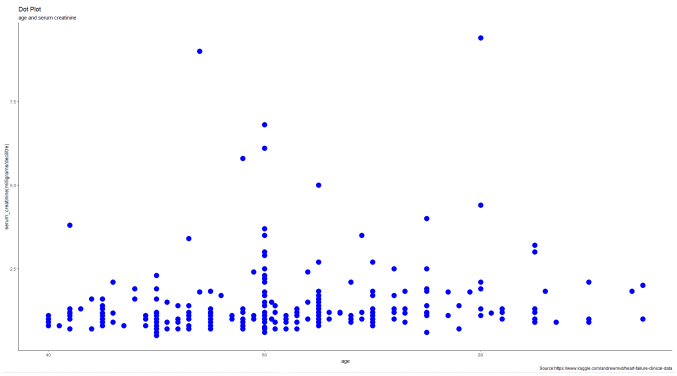

Fig 1.6

According to fig. 1.6, the serum creatinine level in the body declines with age.

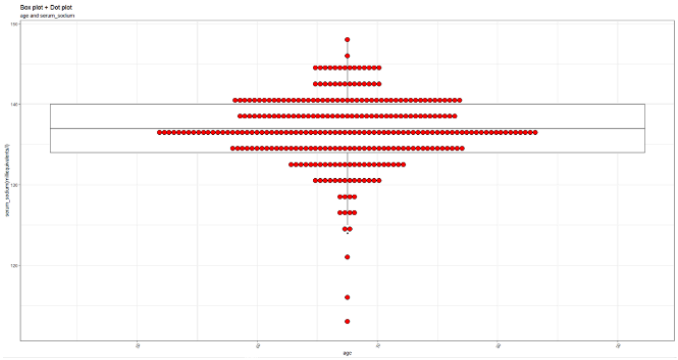

Fig 1.7

According to fig. 1.7, the greatest blood sodium levels are found between the ages of 50 and 80, with levels ranging between 130 and 140.

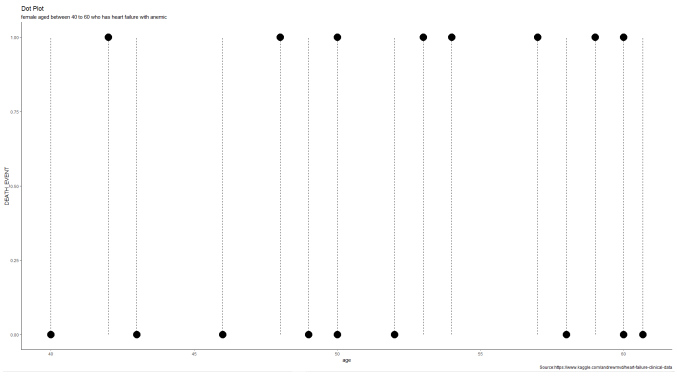

Fig 1.8

According to fig. 1.8, females between the ages of 40 and 60 who have heart failure and anaemic have roughly similar chances of survival.

age	anaemia	sex	DEATH_EV
49	1	0	0
48	1	0	1
50	1	0	1
50	1	0	0
60	1	0	1
57	1	0	1
53	1	0	1
60	1	0	1
42	1	0	1
60	1	0	1
59	1	0	1
60	1	0	0
58	1	0	0
60	1	0	0
60	1	0	0
43	1	0	0
58	1	0	0
46	1	0	0
52	1	0	0
50	1	0	1
49	1	0	0
60.667	1	0	0
50	1	0	0
50	1	0	0
40	1	0	0
60	1	0	0
48	1	0	1
54	1	0	1
50	1	0	0

Tab. 1.9

According to tab. 1.8, females between the ages of 40 and 60 who have heart failure but are not anaemic have the highest probability of success.

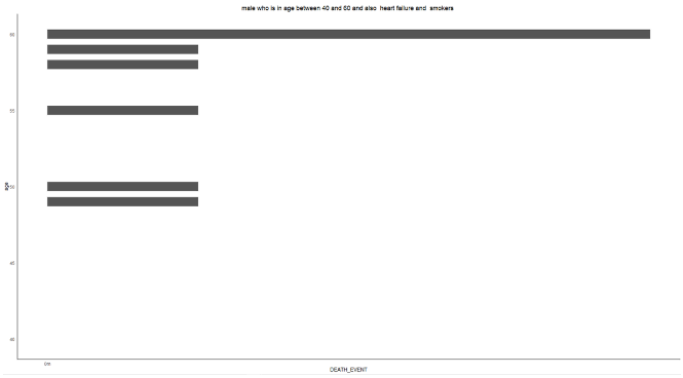

Fig 1.10

Figure 1.10 displays males who smoke and have heart failure. It has been shown that people between the ages of 45 and 60 have a 65% probability of surviving.

Fig 1.11

Figure 1.11 displays men aged 40 to 60 who have heart failure but are not anaemic. According to the graph above, those between the ages of 55 and 60 have a lower chance of surviving.

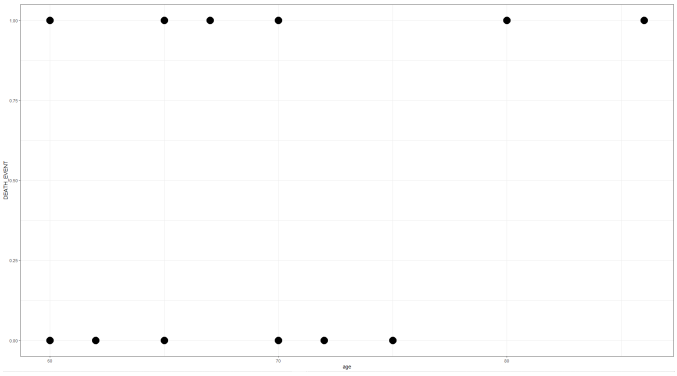

Fig 1.12

Figure 1.12 depicts anaemic ladies who also have heart failure. When compared to those above the age of 75, those between the ages of 60 and 75 have the maximum chance of survival.

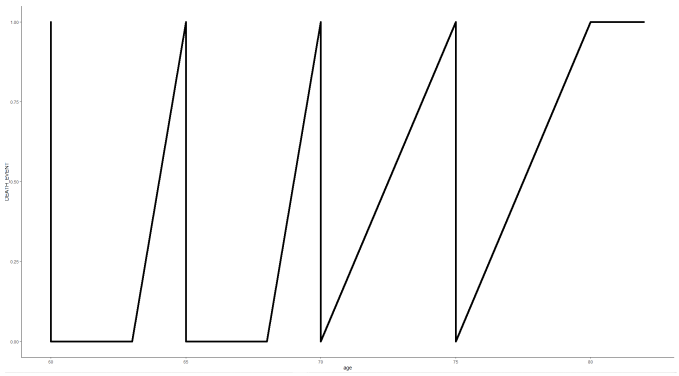

Fig 1.13

Fig 1.13 depicts females aged 60 to 95 who have diabetes and heart failure. Those above the age of 70 have the lowest chance of survival when compared to those under the age of 70.

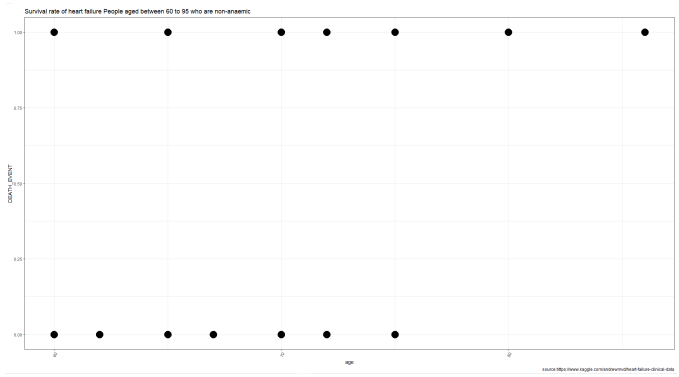

Fig 1.14

Figure 1.14 displays females aged 60 to 95 who have heart failure but are not anaemic. Those between the ages of 60 and 75 have the best chance of survival when compared to those above the age of 75.

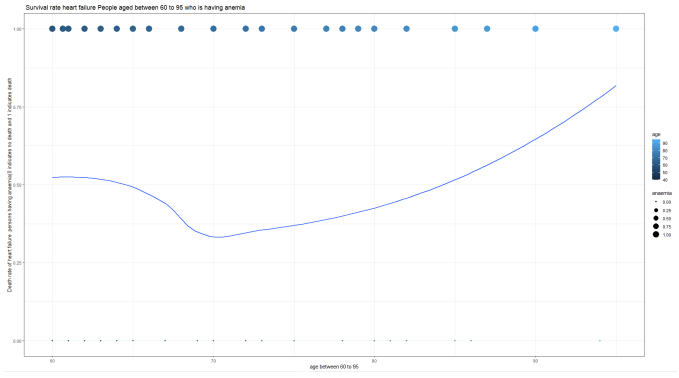

Fig 1.15

Females between the ages of 60 and 95 with heart failure and anaemia have roughly equal survival rates, as shown in fig 1.15.

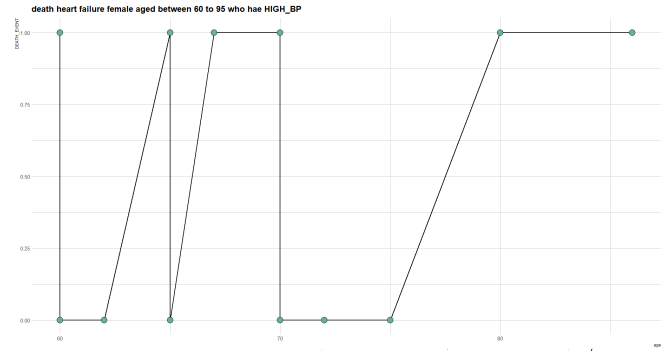

Fig 1.16

Figure 1.16 depicts heart failure females aged 60 to 95 with high blood pressure.Those between the ages of 60 and 70 have an equal chance of survival, but those between the ages of 70 and 80 have a higher survival rate, and those over 80 have a steadily increasing death rate..

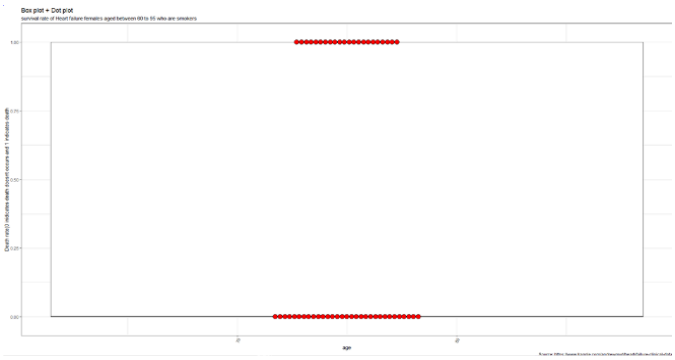

Fig 1.17

Female smokers with heart failure between the ages of 60 and 90 are depicted in Figure 1.17. Those aged 70 to 80 have a better probability of surviving.

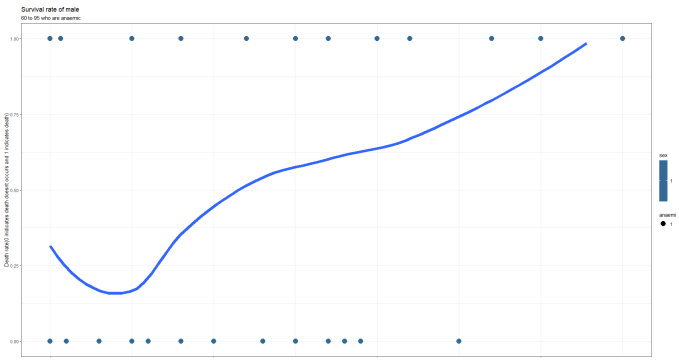

Fig 1.18

Men aged 60 to 90 who are anaemic and have heart failure are depicted in Figure 1.18. When compared to those between the ages of 60 and 80, those over the age of 80 have a lower survival chances.

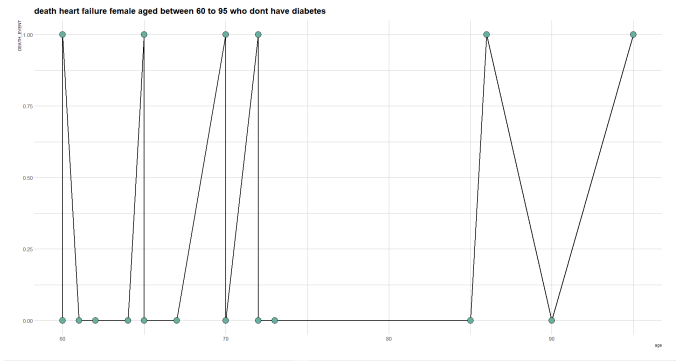

Fig 1.19

Females between the ages of 60 and 90 who are non-diabetic and have heart failure are depicted in Figure 1.19. Those above the age of 80 have an equal chance of survival, whereas those between the ages of 60 and 80 have a better chance.

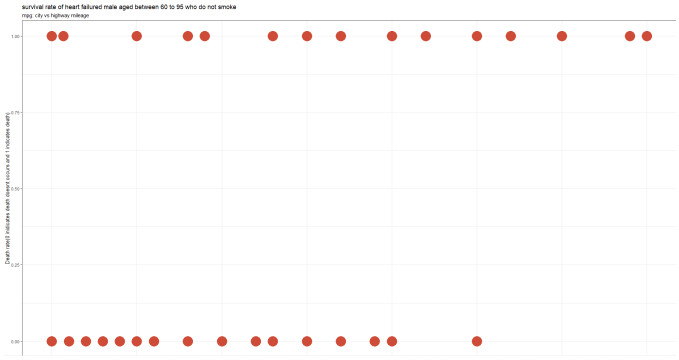

Fig 1.20

Figure 1.20 depicts male smokers aged 60 to 90 who have heart failure. Those aged 60 to 75 have a high chance of survival; those over 75 have a gradually decreasing chance of survival.

CONCLUSION

Based on the foregoing observations, we may conclude that by using visualisation, we can make it easier for doctors to detect cardiac ailments at an early stage. This research uses visualisation techniques to depict the chances of survival of patients with heart failure, as well as other factors such as anaemia, diabetes, high blood pressure, and so on. This type of exploratory data analysis may be used to make future predictions using a large number of attributes.

ACKNOWLEDGEMENTS

We'd like to take this occasion to extend our heartfelt gratitude to our project leader, Dr. Kalyani Desikan, Professor, School of Advanced Sciences, for her unwavering support and helpful advice during the project.

We are grateful to Dr. Mahalakshmi S, Professor and Dean of the School of Advanced Sciences at VIT Chennai, for offering the school's facilities and unwavering support to our research.

We would like to convey our gratitude to Dr. Muthunagai, our Department Head, for her assistance during this project.

We'd want to take this occasion to thank the entire faculty of the institution for their help and advice.

We are grateful to our parents, family, and friends for their unwavering support throughout our project and for allowing us to pursue this programme at such a prominent university.

REFERENCES

[1] V. Manikantan & S.Latha,"Predicting the Analysis of Heart Disease Symptoms Using Medicinal Data Mining Methods", International Journal on Advanced Computer Theory and Engineering, Volume-2, Issue-2, pp.5-10, 2013.

[2] Dr.A.V.Senthil Kumar, "Heart Disease Prediction Using Data Mining preprocessing and Hierarchical Clustering", International Journal of Advanced Trends in Computer Science and Engineering, Volume-4, No.6, pp.07-18, 2015.

[3] Uma.K, M.Hanumanthappa, "Heart Disease Prediction Using Classification Techniques with Feature Selection Method", Adarsh Journal of Information Technology, Volume-5, Issue-2, pp.22-29, 2016

[4] A. L. Bui, T. B. Horwich, and G. C. Fonarow, "Epidemiology and risk profile of heart failure," Nature Reviews Cardiology, vol. 8, no. 1, pp. 30–41, 2011.

[5]J.Mourão-Miranda,A.L.W.Bokde,C.Born,H.Hampel,and M. Stetter, "Classifying brain states and determining the discriminating activation patterns:supportvectormachineon functionalMRIdata,"NeuroImage,vol.28,no.4,pp.980–995, 2005.