



# Exploring age, gender and health conditions as predictors of mortality in heart failure patients

# Meet Our Team



**Team  
Leader**

**Boikokobetso  
Ramashija**

**Project  
Manager**

**Kagiso Morele**

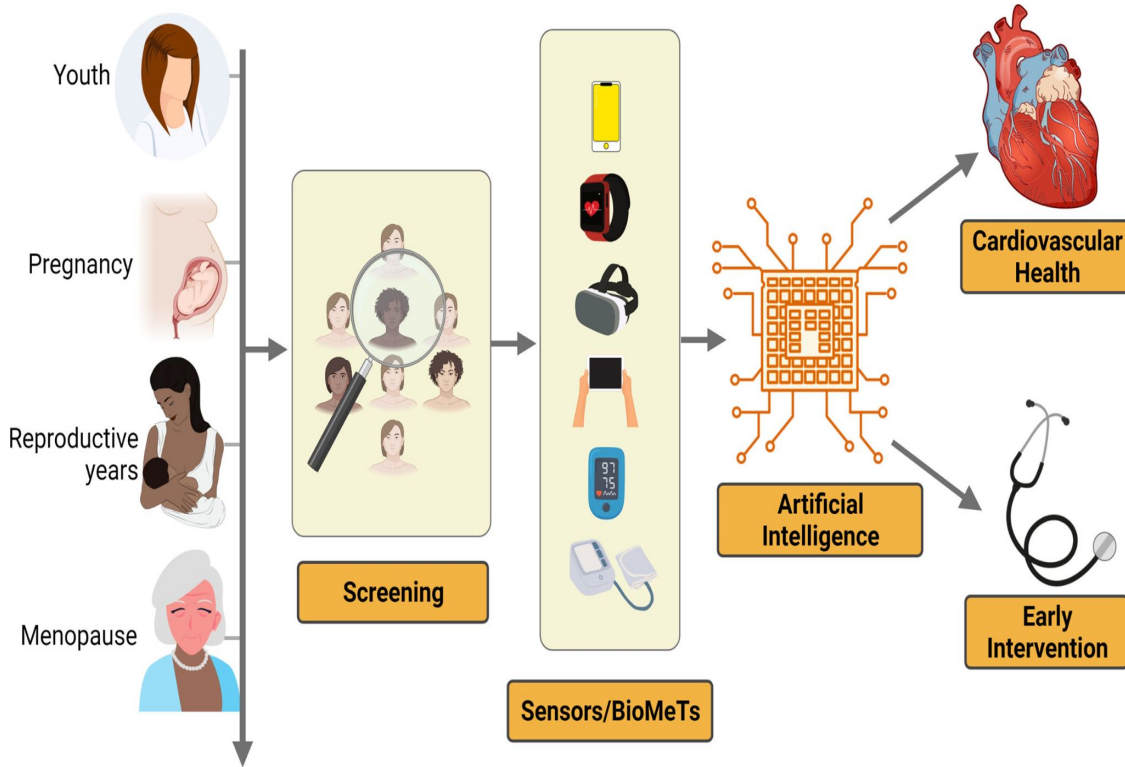
**Member**

**Tracy Nyoni**

**Member**

**Lebogang  
Swaratlhe**

# INTRODUCTION



**Cardiovascular Diseases (CVDs):** Leading cause of death worldwide. Heart Failure: Major contributor to CVD mortality.

**Challenge:** Identifying high-risk patients early on could save lives.

**Focus:** Investigating how factors like age, gender, and pre-existing conditions (hypertension, diabetes) affect mortality in heart failure patients.

**Future Goal:** To develop a machine learning model to help predict mortality risk, enabling healthcare providers to prioritize high-risk



### Risk Factors for HFpEF

Lack of exercise



Sleep apnea



Obesity



Hormonal  
differences



Smoke and/or  
drug use



Diabetes



Chronic  
kidney  
disease



High blood  
pressure  
and CAD



Age



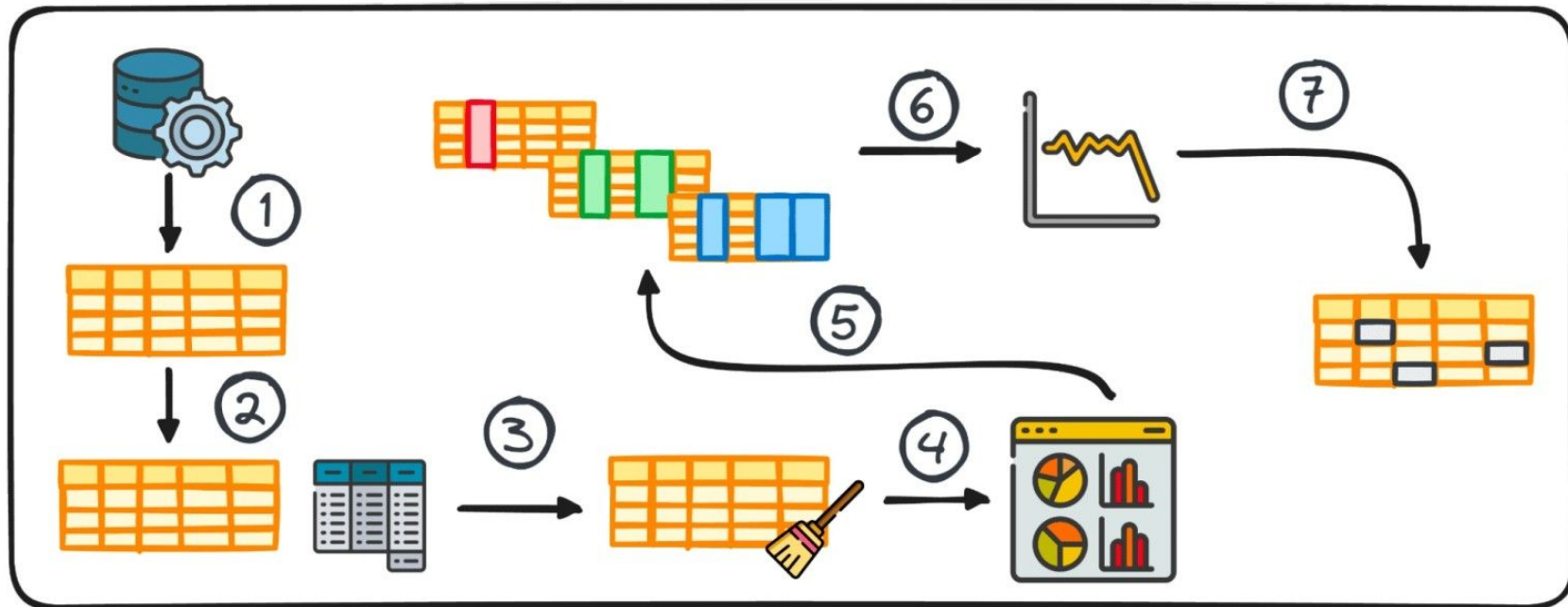
Heart failure, a serious result of heart disease, is a major cause of death worldwide. This project looks at how age, gender, and other health problems impact death rates in heart failure patients. The goal is to find useful insights to help doctors better care for these patients and prevent serious outcomes



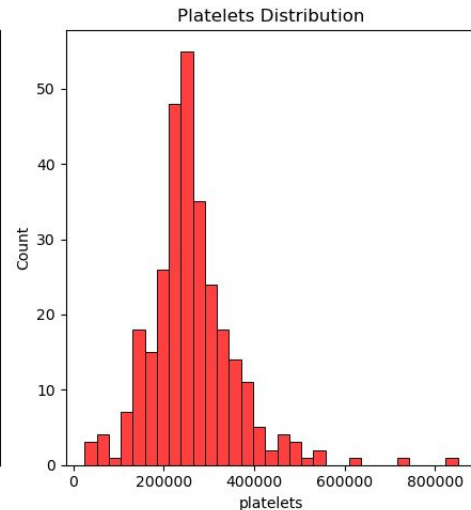
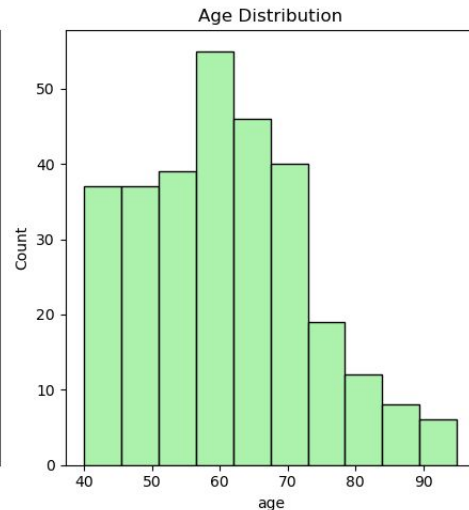
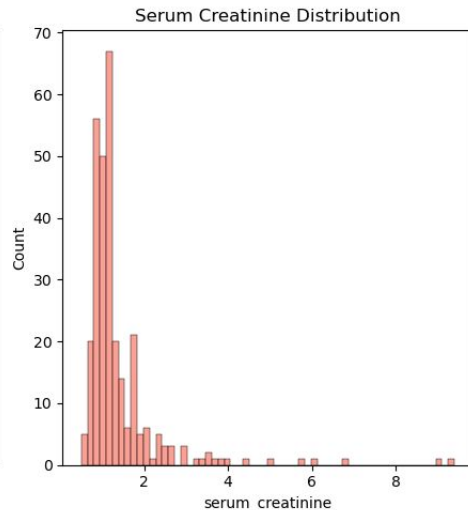
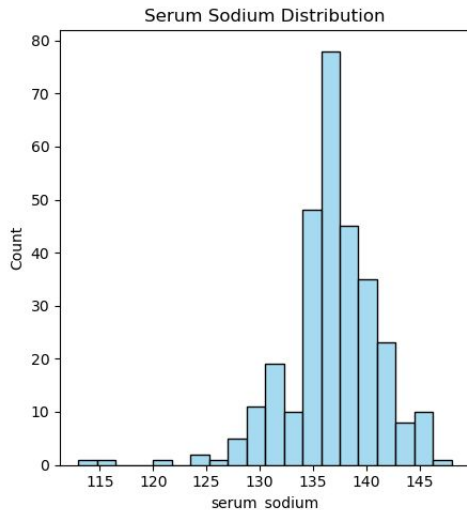
1. To analyze how age and gender impact the prevalence of health conditions (e.g., diabetes, hypertension, anemia) in heart failure patients.
2. To assess mortality risk across different demographic groups and health conditions.



# Data Processing



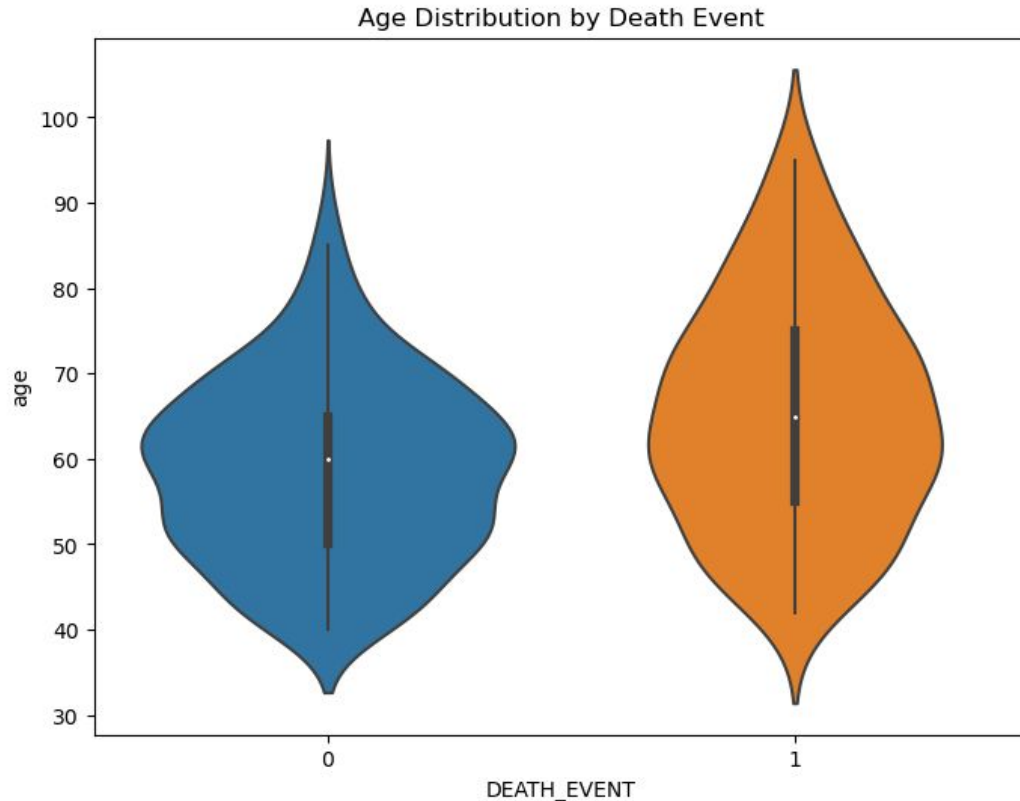
# Data Distribution Insights in Heart Failure Patients



Data analysis shows that serum sodium is left-skewed meaning that majority of the patients have higher serum values, generally high serum levels, factors like diet, lifestyle, or treatment adherence might be influencing these high levels, serum creatinine has outliers with a long tail, it is also right skewed, and from the age distribution we observe that it is slightly right skewed signaling that majority of our population are people younger than 70 years and above 40 years of age.



# Violin plot comparison of age by DEATH



- The data shows that 0 indicates survival and 1 indicates death, with the average age of death around 60-70, similar to survivors, but with higher values in the 75th percentile and some reaching 90-100, likely influenced by coexisting diseases.

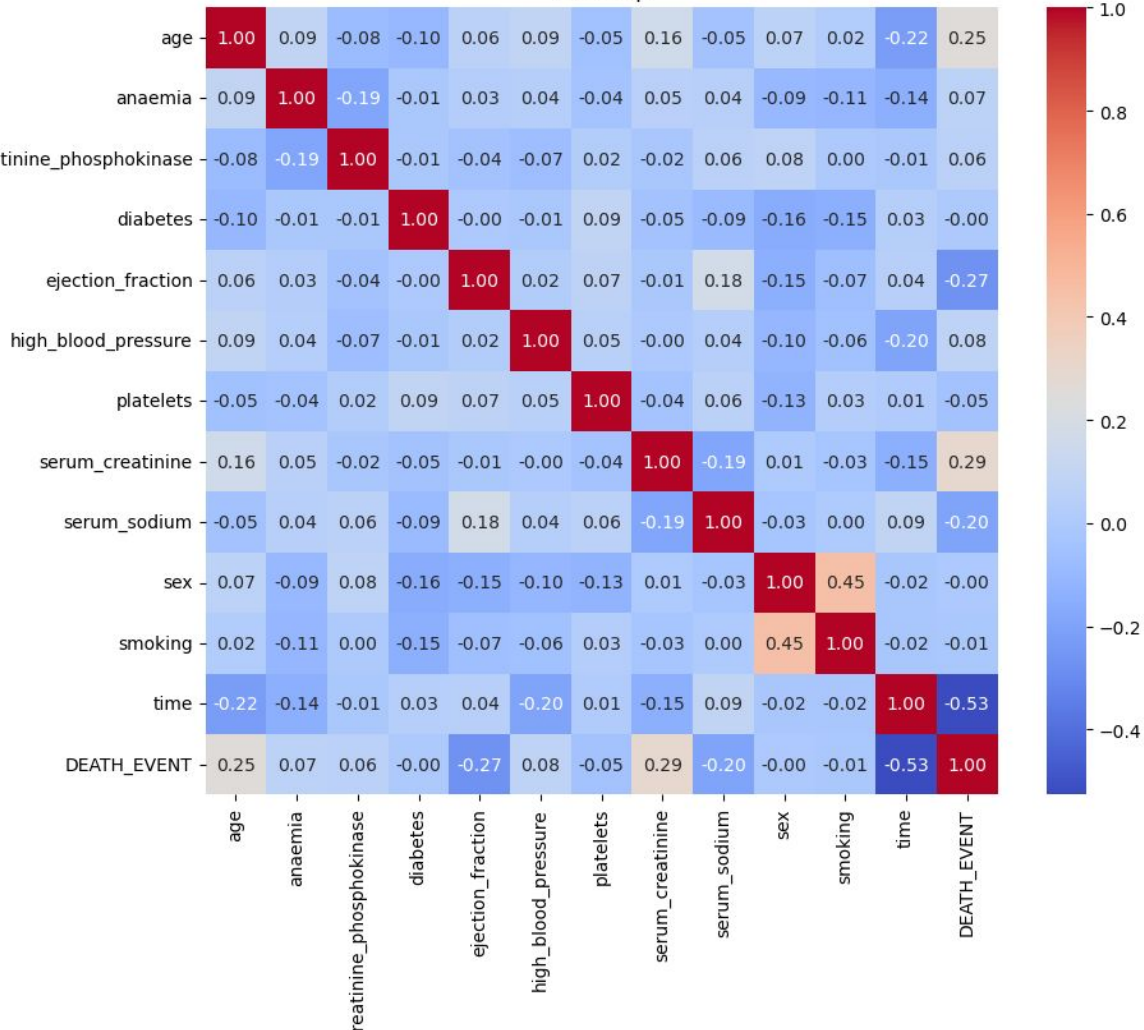




## Visualizing the correlation matrix

- The heatmap shows no correlation between DEATH\_EVENT and most health conditions, except a moderate negative correlation (-0.53) between time and DEATH\_EVENT, suggesting limited number of diseases have an impact on mortality

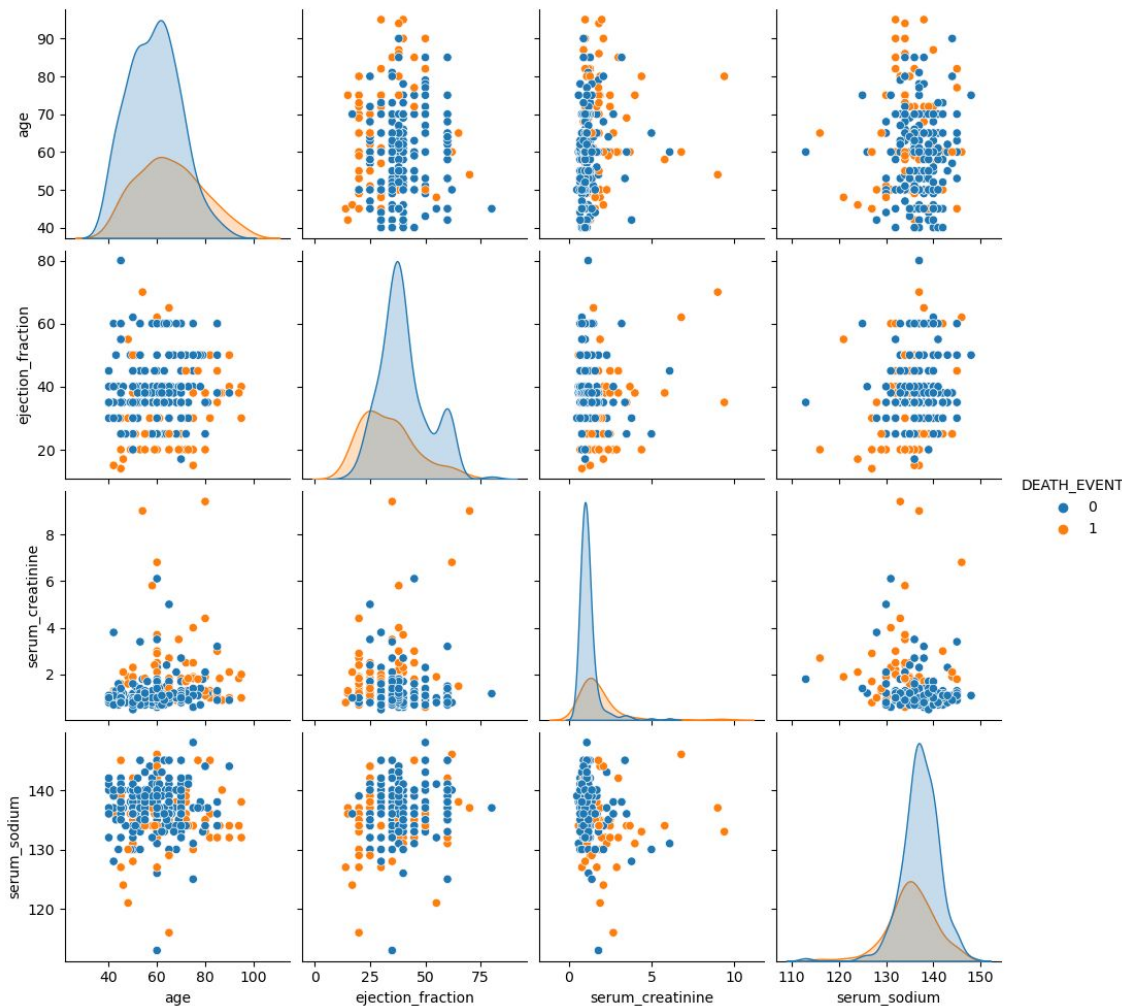
HeatMap



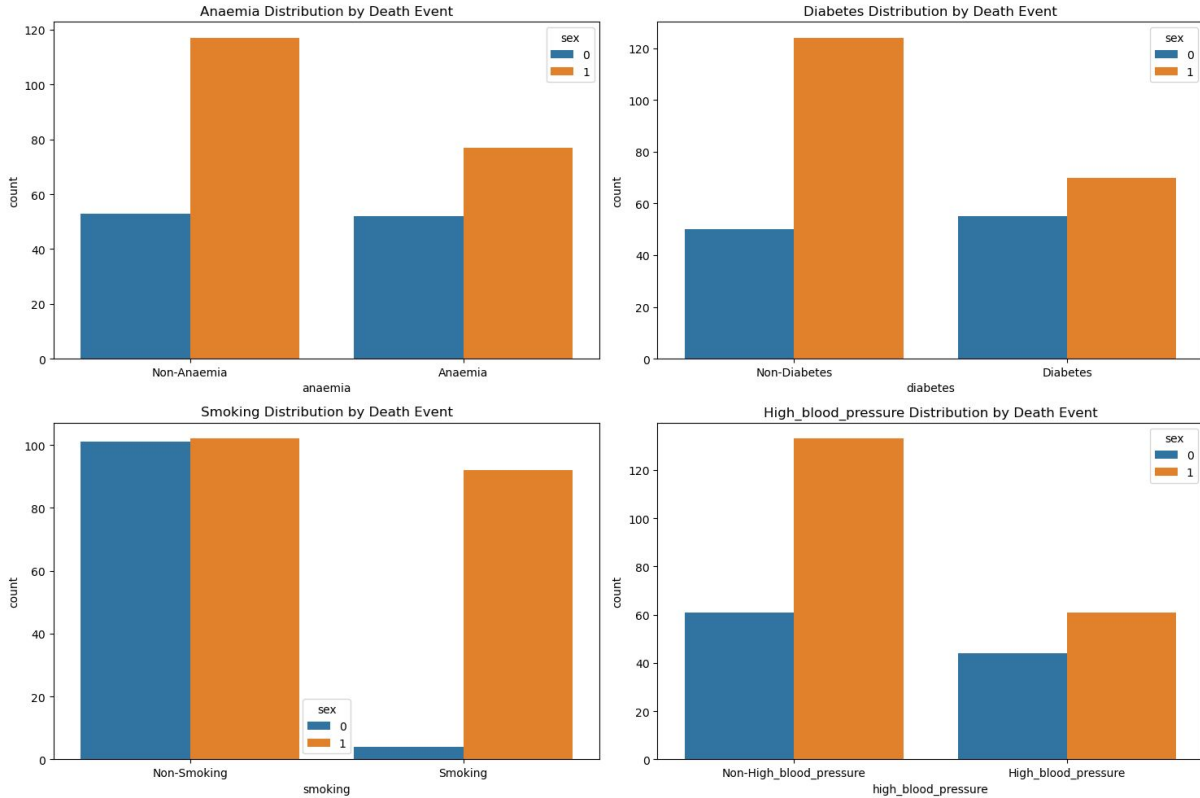


# Risk Factor Analysis with Multivariate Plot

In the pairplot, we visualise the correlation matrix. We can see from the pairplots, the scatterplots do not show a relation between variables, just as we saw in the heatmap, very low correlations. On the diagonal pairplots we see the distribution for each variable



# Checking the countplot of binary variables



- Bar graphs show that patients without anemia, diabetes, high blood pressure, or smoking have a higher likelihood of DEATH\_EVENT, and males have a higher incidence of heart failure-related death than females. From the smoking distribution by DEATH EVENT bar plots we can see more males smoke compared to females also.



# Conclusion and Recommendation

## Summary of Key Findings:

- When setting out on this project analysis, the initial thought of the team was that the data would support, what we thought to be general knowledge: That being sick with other illnesses and/or engaging in activities detrimental to your health, such as smoking would significantly impact chances of one dying heart failure. The story that this data tells quickly disproved that theory. In our research we found that:
- There is little to no correlation between being pre-diagnosed with illnesses such as, *Anaemia, Hypertension (High Blood Pressure), Diabetes and Hybernatriemia (High Sodium levels)* and experiencing the death event, caused by heart failure.
- There is however a relatively significant relationship between age and acquiring these illnesses as well as age the death event itself occurring.

## Implications of the Findings:

Looking at the results of this study, significant insights could be provided to various sectors in order to decrease the number of people dying from heart failure, as well as to provide better services to ill patients.

Below are two examples:

- Insurance policies companies could use the information/results provided to tailor-make and improve the insurance package models offered to their ill patients.
- Doctors and biomedical-engineers could divert their resources and attention to conducting studies on older patients with heart-related issues in order to come up with new medical technologies that assist in early intervention and personalized treatment strategies

**Reflection on the Data Source and Quality:**

- We were fortunate enough to find a high quality dataset that was from a fairly reputable data source. The data had a vast variety, that satisfied the purpose of the project.



# References

1. [https://www.kaggle.com/datasets/andrewmvd/heart-failure-clinical-data?select=heart\\_failure\\_clinical\\_records\\_dataset.csv](https://www.kaggle.com/datasets/andrewmvd/heart-failure-clinical-data?select=heart_failure_clinical_records_dataset.csv)
2. World Health Organization (WHO) 2021, *Cardiovascular diseases (CVDs)*, WHO, viewed [date you accessed],  
<https://www.who.int/health-topics/cardiovascular-diseases>.
3. Savarese, G. & Lund, L.H. 2017, 'Global public health burden of heart failure', *Cardiac Failure Review*, vol. 3, no. 1, pp. 7–11. doi:10.15420/cfr.2016:25:2.
4. Kannel, W.B. & McGee, D.L. 1979, 'Diabetes and cardiovascular risk factors: the Framingham study', *Circulation*, vol. 59, no. 1, pp. 8–13. doi:10.1161/01.cir.59.1.8.