# Predictions in Hospital Mortality

BAIS Baddies Inc.

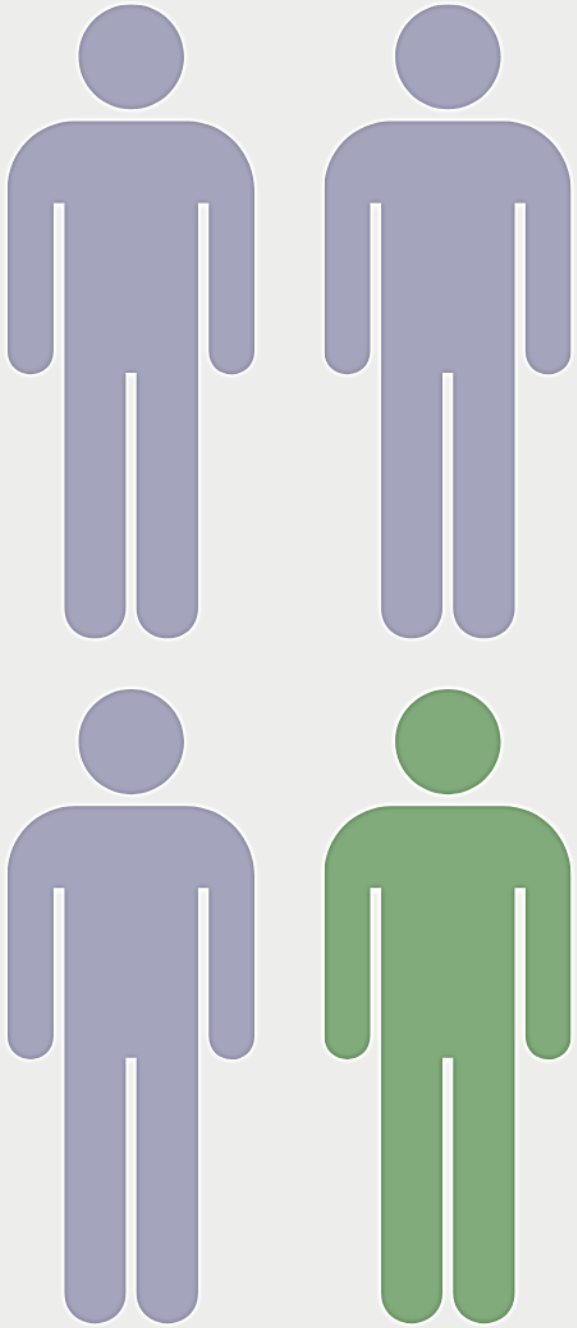
Lauren Jamison | Brittany Fleury | Natalie McCormick | Sophie Selk

# OUTLINE

| | |
|---|---|
| **Background** | The impact of cardiac arrests in America and risk-adjusted mortality positive effects on paient outcome |
| **Business Goal** | Actionable goal for Hospitals to increase survival rates |
| **Data Overview** | How to predict mortality, a description of hospital dataset, and data preparation process |
| **Exploratory Analysis** | Practical charts, correlations , and discoveries in dataset |
| **Methods and Evaluation** | Synopsis of models used for analysis and evaluations metric |
| **Conclusion and Future Outlook** | Summary of dataset analysis and recommendations for hospitals |

# Background

One person dies every **19 Seconds** in the United States from heart attacks



659,000 people in the United States die of heart attacks

OR

**1 in every 4 deaths**

# The implementation of Risk-Adjusted Mortality decreases the overall mortality rates in Hospitals

## Risk-Adjusted Mortality

Ability to prioritize patients based on characteristics that predict risk of death

1.3%

Implement of risk-adjusted mortality caused absolute reduction in mortality

# Business Goal

**Predict the mortality rate of patients who enter the ICU (Intensive Care Units) due to heart failure to provide doctors with information about what patients they need to prioritize in order to reduce patient-in-house mortality rates**

## Potential Clients:

- Doctor and Hospital Owners
- Identify features that take priority of those suffering from heart failure
- Apply Risk-Adjusted practice to Hospitals to improve mortality rates

# Data Mining Goal

To determine what features are most important in determining a patient's chance of in-house mortality

## Supervised Learning

Training data includes both the input and result

## Classification

Positive or Negative Outcome of Mortality

## Target Variable

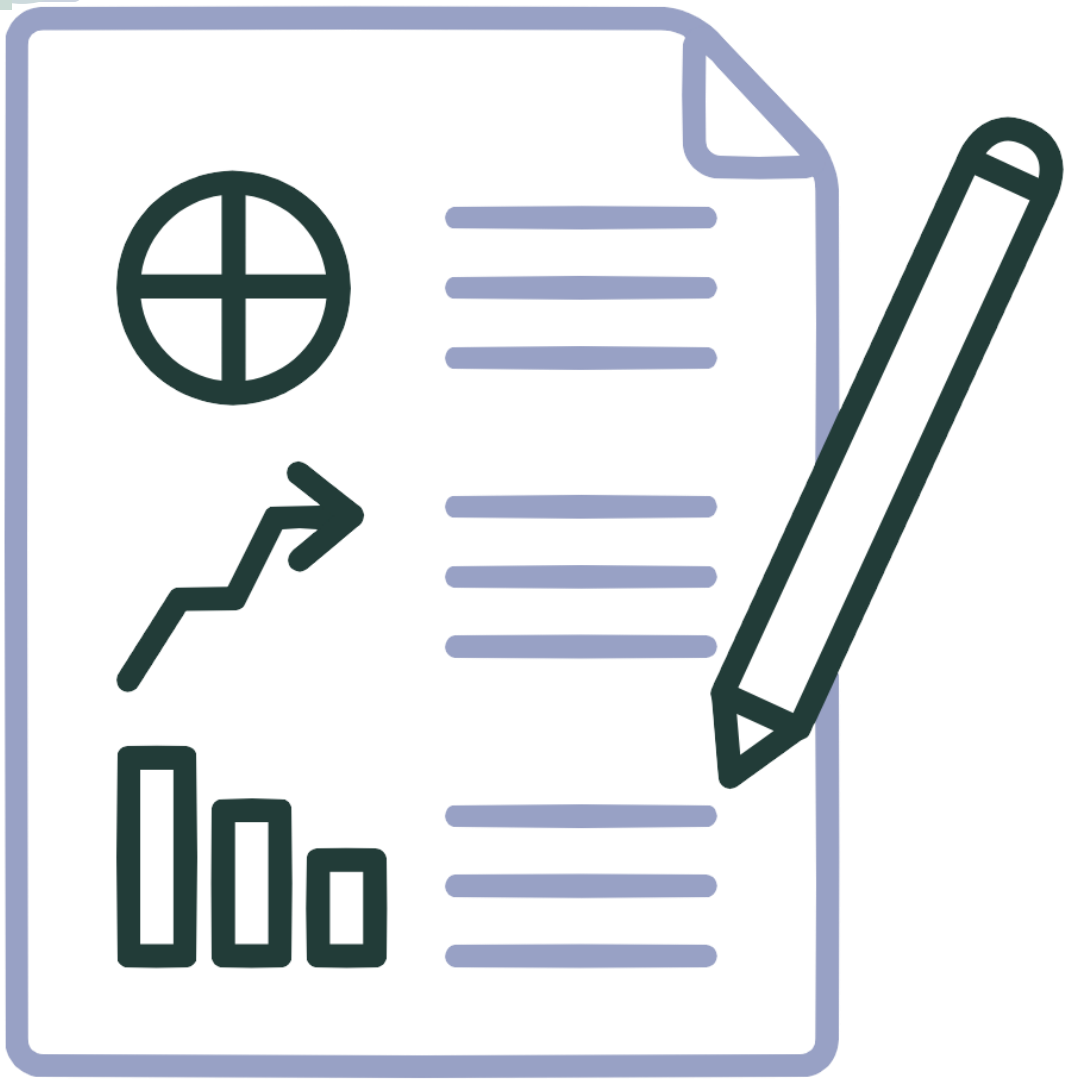
Outcome:
Mortality after cardiac arrest in ICU

0 = Alive
1 = Death

# Dataset Overview

| | |
|---|---|
| Instances | 1,176 |
| Categorical Features | 7 |
| Numeric Features | 10 |
| Target Variable % | 159/1176 = 13.52% |

# Data Preparation



## Important Features

- Random forest of all 50 features to select 15 most important other than ID and Target Variable

## Missing variables

- Diagnosed and assessed missing variables to create fuller dataset
- Deleting all observations with missing

## Transformed Features

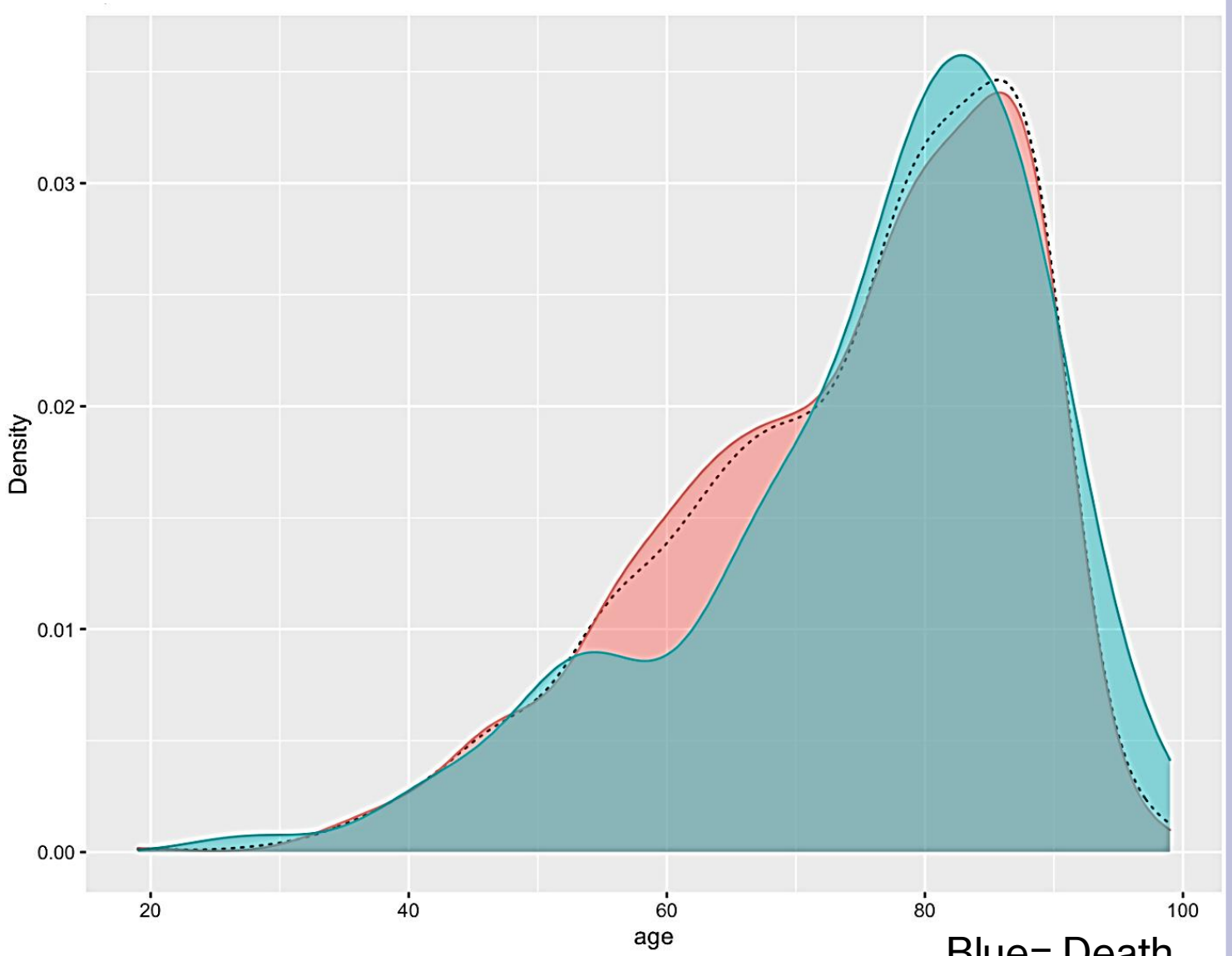
- Transformed Categorical features as indicator variables

## Recoded Features

- Numerical features recoded to mean
- Categorical features into mode


# Exploratory Analysis

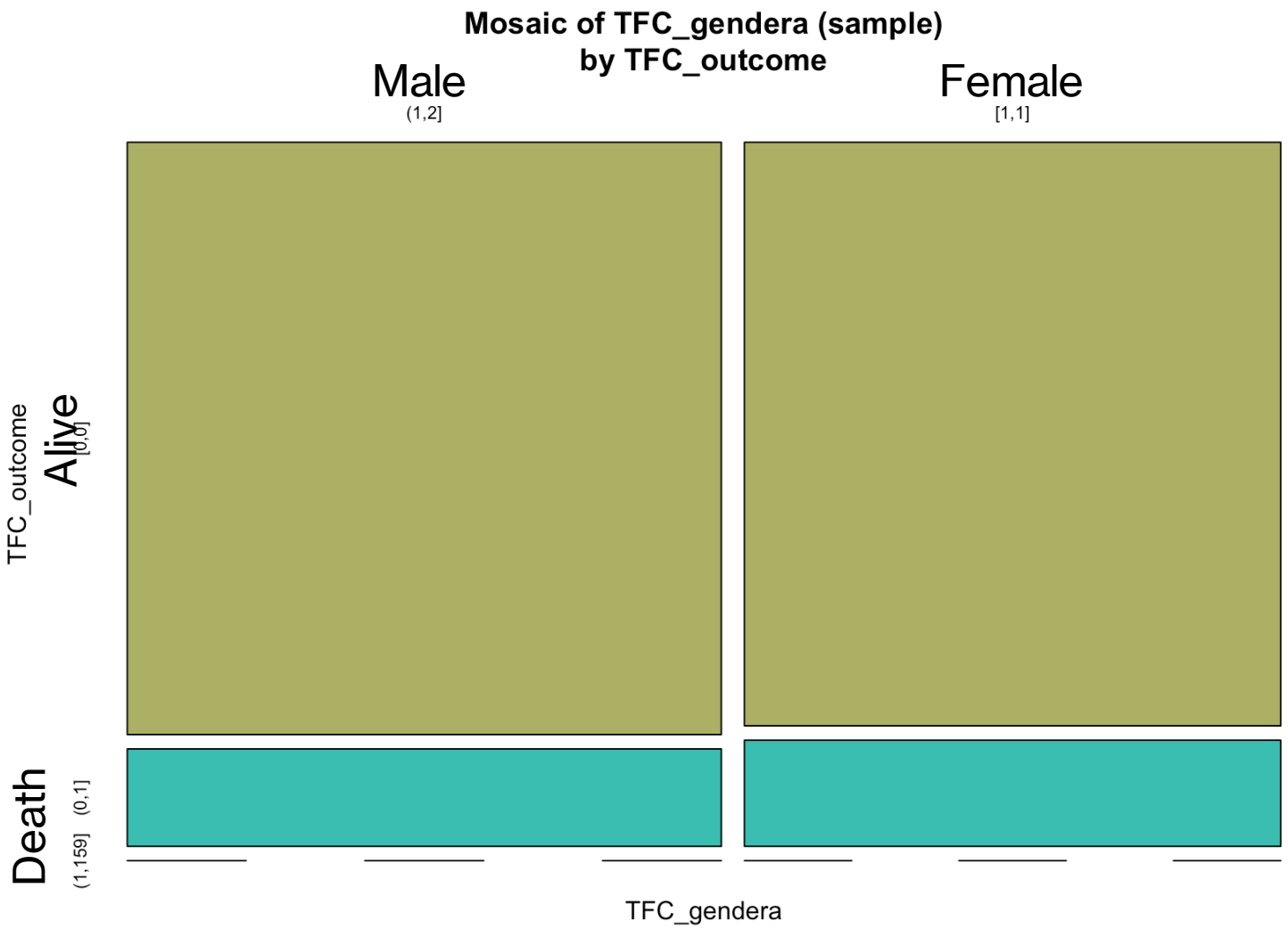
# Age and Mortality Histogram




Blue= Death
Red= Survived

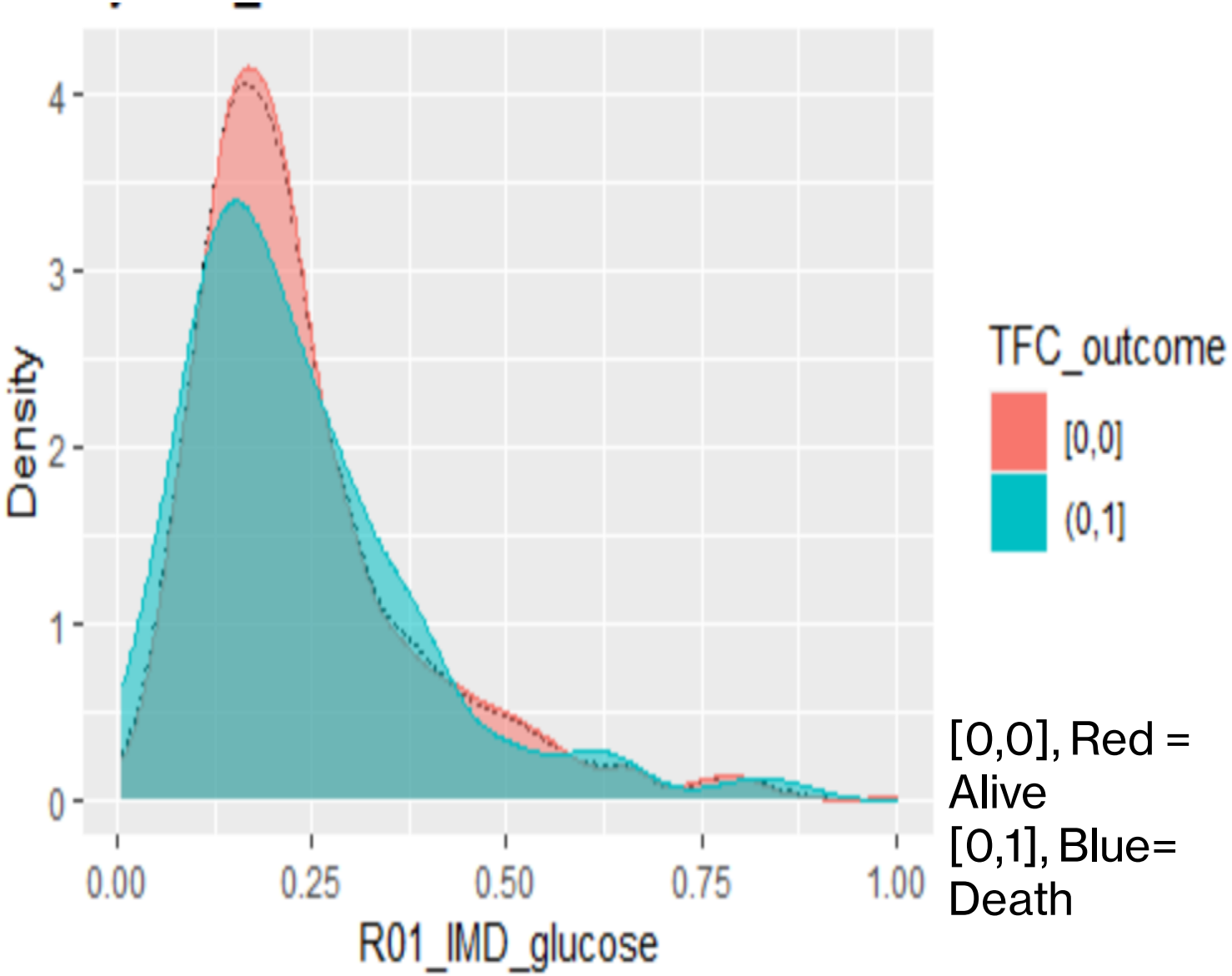
**Ages from 70- to 90-year-olds would be more likely to be in the ICU for heart failure and have a higher risk of death**

# Gender and Mortality Outcome




**Females have a slightly higher risk in morality than males when they are hospitalized for heart failure**

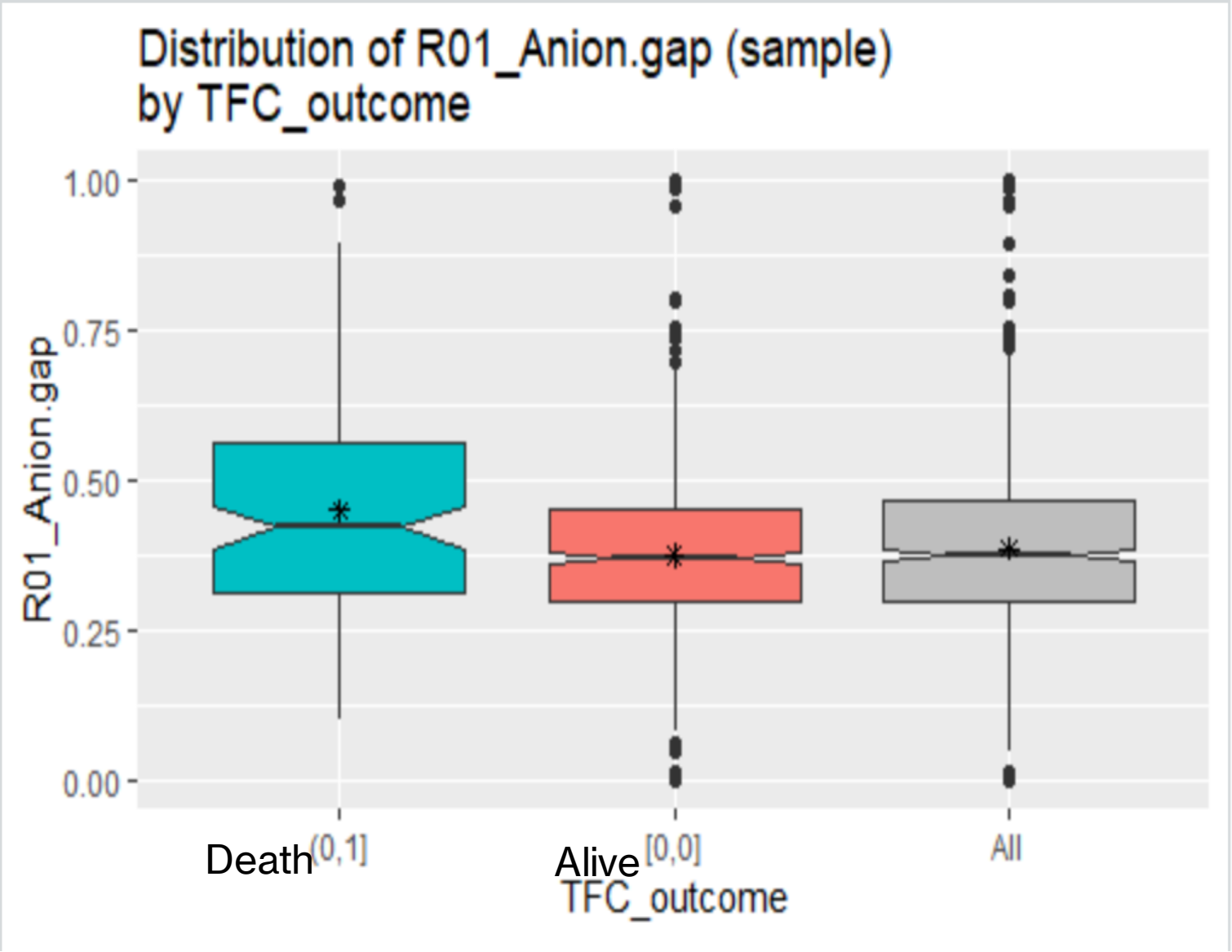
# Glucose Levels and Mortality




[0,0], Red = Alive
[0,1], Blue= Death

**Majority of patients with lower glucose levels have a decreased risk of death**

# Anion Gap and Mortality



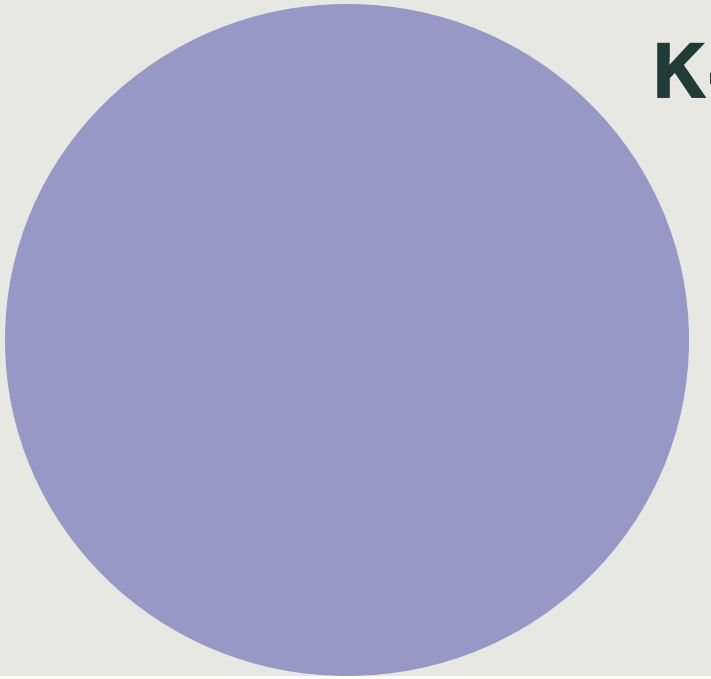

If the patient had a higher anion gap, there was a higher chance of death.

Normal levels for anion gap are 3 to 10 mEq/L.

Anion gap= if too high or too low, it may be a sign of a disorder in your lungs, kidneys, or other organ systems.

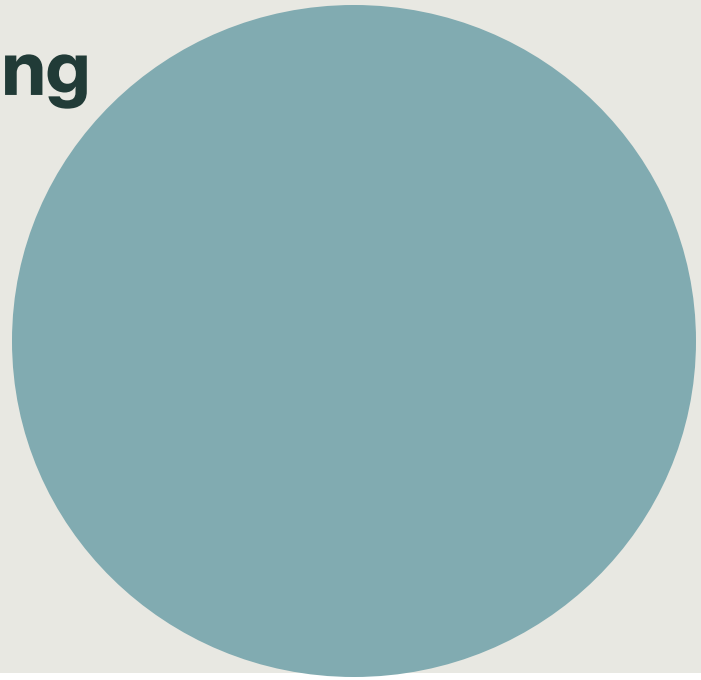
# K-Means Clustering

**2 Segments**



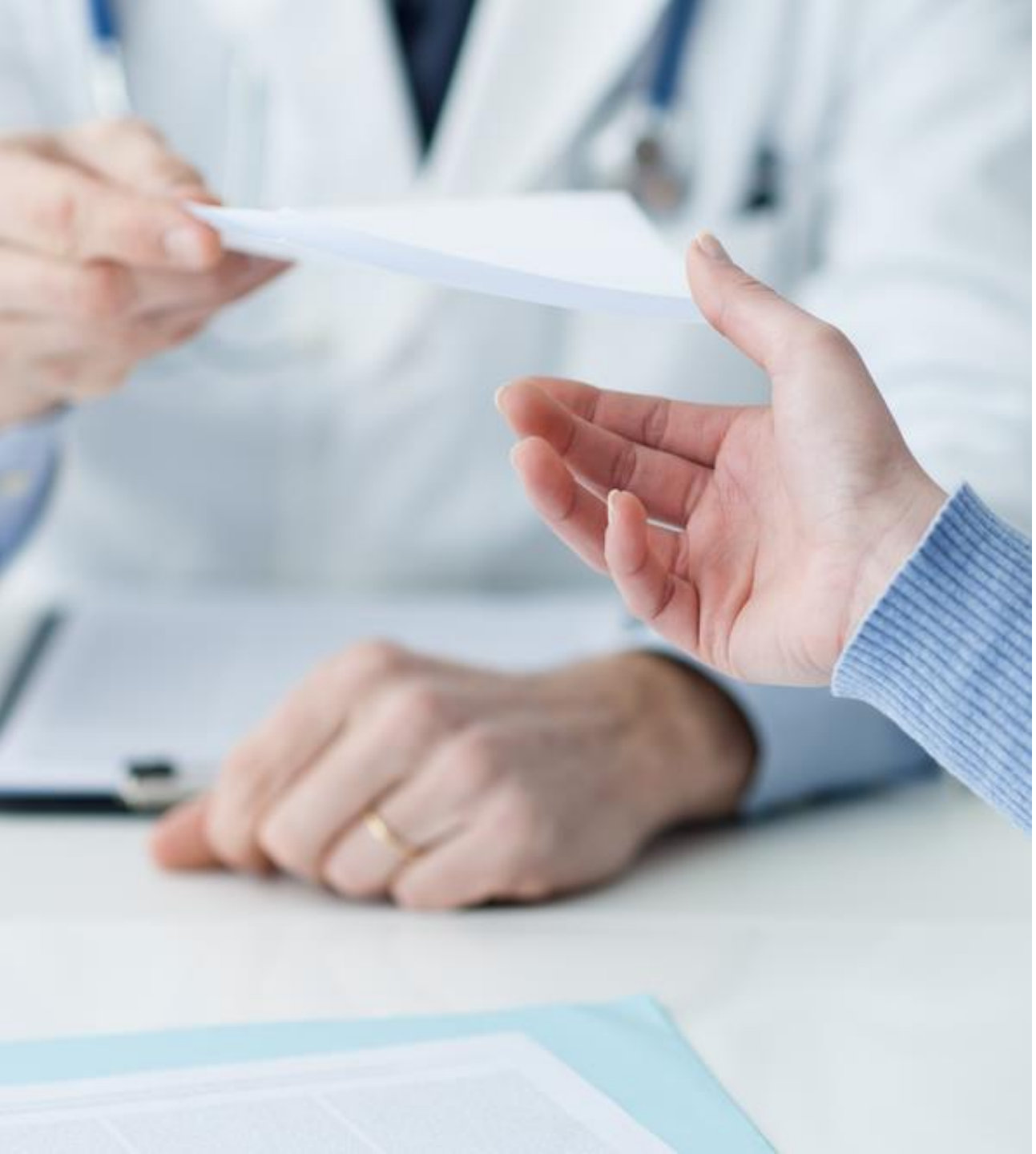
## Cluster 1

- Size: 497
- Male
- Average Age: 67
- More Likely to be Hypertensive
- Not Diabetic
- Not Likely to have Hyperlipemia or have renal failure



## Cluster 2

- Size: 472
- Female
- Average Age: 70
- Less likely to be Hypertensive
- Diabetic
- More likely to have Hyperlipemia and have renal failure



# Method and Evaluation

## All Classification Models

Logistic Regression, Decision Tree, SVM, Random Forest, AdaBoost, Gradient Boost and ANN.

## Most Robust Evaluation AUC

Robustness of AUC and Ease of Interpretation

## Cross-Validation

All models with random seeds 42

| Models | Cross-Validated AUC |
|---|---|
| Gradient | 0.9280 |
| Ada | 0.9264 |
| **Logistic Regression** | **0.9315** |
| SVM: Radial | 0.9056 |
| SVM: Polynomial | 0.9202 |
| Random Forest | 0.9352 |
| ANN | 0.7779 |
| Decision Tree | 0.8710 |

## Recommendations

Increasing attention to patients with the following factors upon admittance to the ICU will lead to a decrease of in-hospital mortality rates:

| | | | |
|---|---|---|---|
| Patient has Atrial Fibrillation | Increased Anion Gap | High Heart Rate | Increased Respiratory Rate |
| High BMI | Ages 70-90 | Hypertensive (High Blood Pressure) | Patient is Diabetic |

# Next Steps

## Data Gathering

- Use long term data (only was collected for 24 hours)
- Gather data from more hospitals

## Feature Selection

- Select different features
  - Only used 15 out of 50 features

## Increase Scope

- Gather information on other leading causes of death

# Questions

# Appendix

# Logistic Regression Model

```
Coefficients:
                             Estimate Std. Error
(Intercept)                -23.61499  803.33400
TFC_gender(1,2]             -0.32522    0.26674
TFC_hypertensive(0,1]        0.89988    0.29649
TFC_atrialfibrillation(0,1] 19.61208  803.33279
TFC_diabetes(0,1]            0.13893    0.27819
TFC_Hyperlipemia(0,1]       -0.04524    0.27025
TFC_Renal.failure(0,1]      -0.04445    0.27842
R01_age                      1.10779    0.87425
R01_Platelets               -3.52567    1.42079
R01_Anion.gap                3.07078    1.21316
R01_Bicarbonate             -1.95177    1.16022
R01_IMD_heart.rate           2.25119    0.91381
R01_IMD_Respiratory.rate     1.99165    1.01762
R01_IMD_Urine.output        -3.26022    1.19276
R01_IMD_glucose             -0.51234    0.90125
R01_IMD_TNM_BMI              1.77399    0.51467
                           z value Pr(>|z|)
(Intercept)                 -0.029 0.976549
TFC_gender(1,2]             -1.219 0.222752
TFC_hypertensive(0,1]        3.035 0.002404 **
TFC_atrialfibrillation(0,1]  0.024 0.980523
```