Title of the Project:
"Heart Health Analytics: Predicting Clinical Outcomes in Heart Failure Patients"

Executive Summary:

This project aims to analyze the heart failure clinical dataset from Kaggle to identify key factors contributing to patient mortality. Using RStudio, we'll apply statistical analysis and machine learning techniques to develop predictive models. Our goal is to uncover insights that can help improve patient care and inform clinical decision-making.

Project Deliverables:

- Comprehensive analytical report including statistical analysis and model findings.
- A set of predictive models for patient outcomes.
- Code repository with documented R scripts.
- Final presentation outlining key insights and recommendations.

Project Acceptance Criteria:
Completion of all analyses by the specified deadline.
Clear, comprehensive documentation of all code and methodologies used.

Project Exclusions:
Analysis of treatment efficacy or drug interactions.
Long-term patient follow-up data post-hospital discharge.

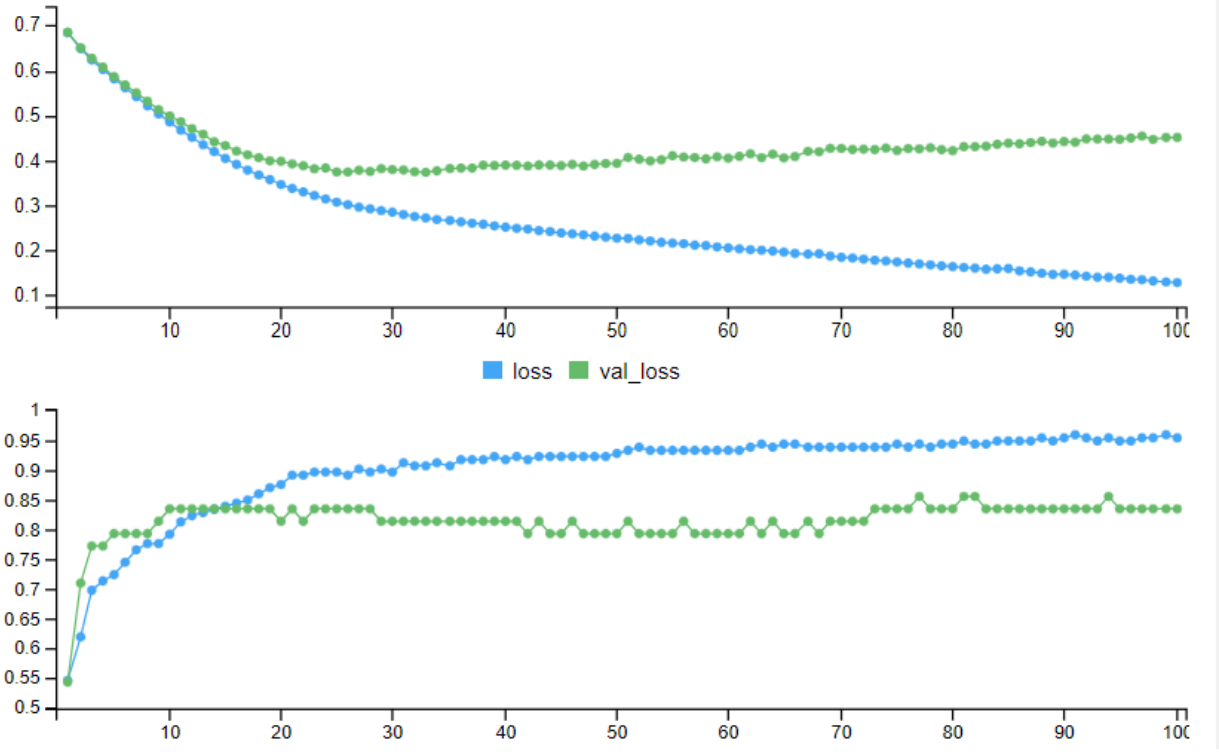
Project Constraints:
Data: Limited to the variables available in the Kaggle dataset.
Software: Analysis constrained to tools available in RStudio.

Project Schedule:
Week 1-2: Data acquisition and initial exploration.
Week 3-4: Statistical analysis and model development.
Week 6: Writing the analytical report.

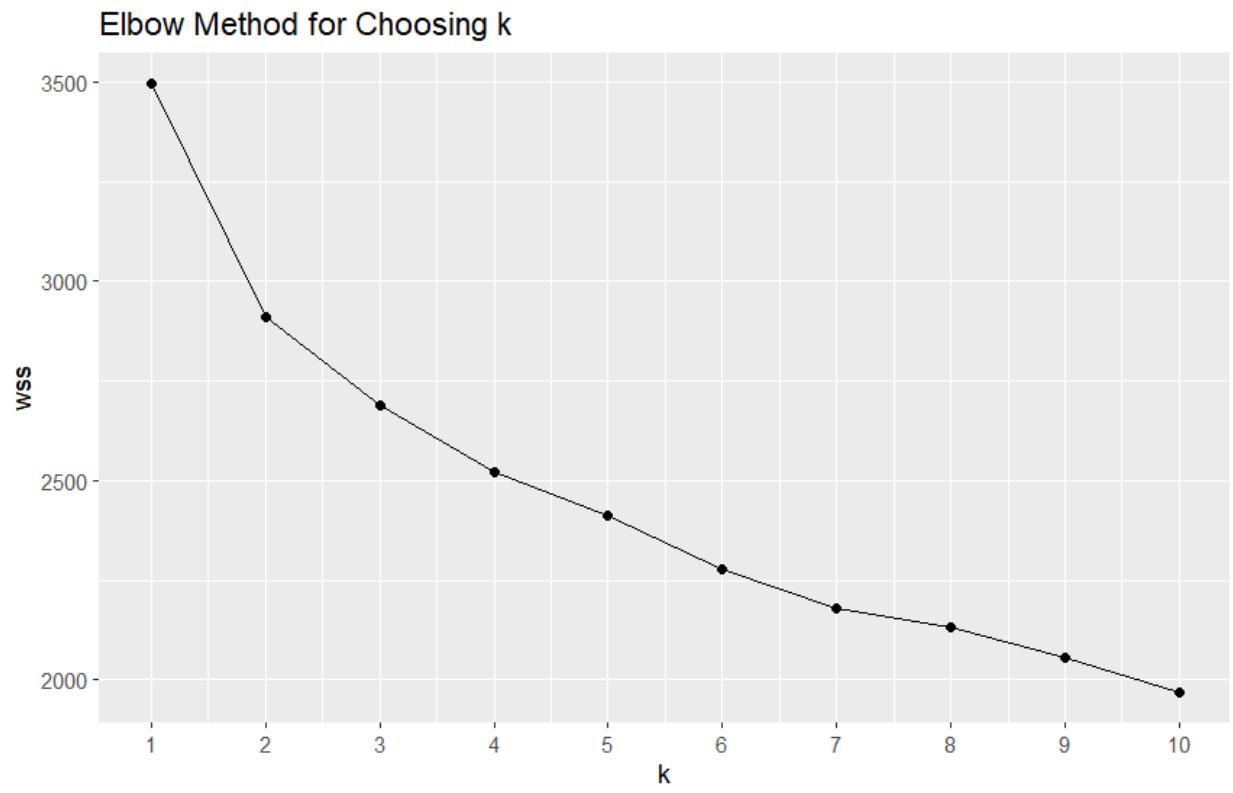
Project Resources:
Dataset: Kaggle Heart Failure Clinical Data.
Paper/Journals/Articles: Access to JSTOR and IEEE Xplore for relevant literature.

Risk and Issue Management Plan:

- Analisis SWOT
- 5W 1H
- 5 Why
- Others

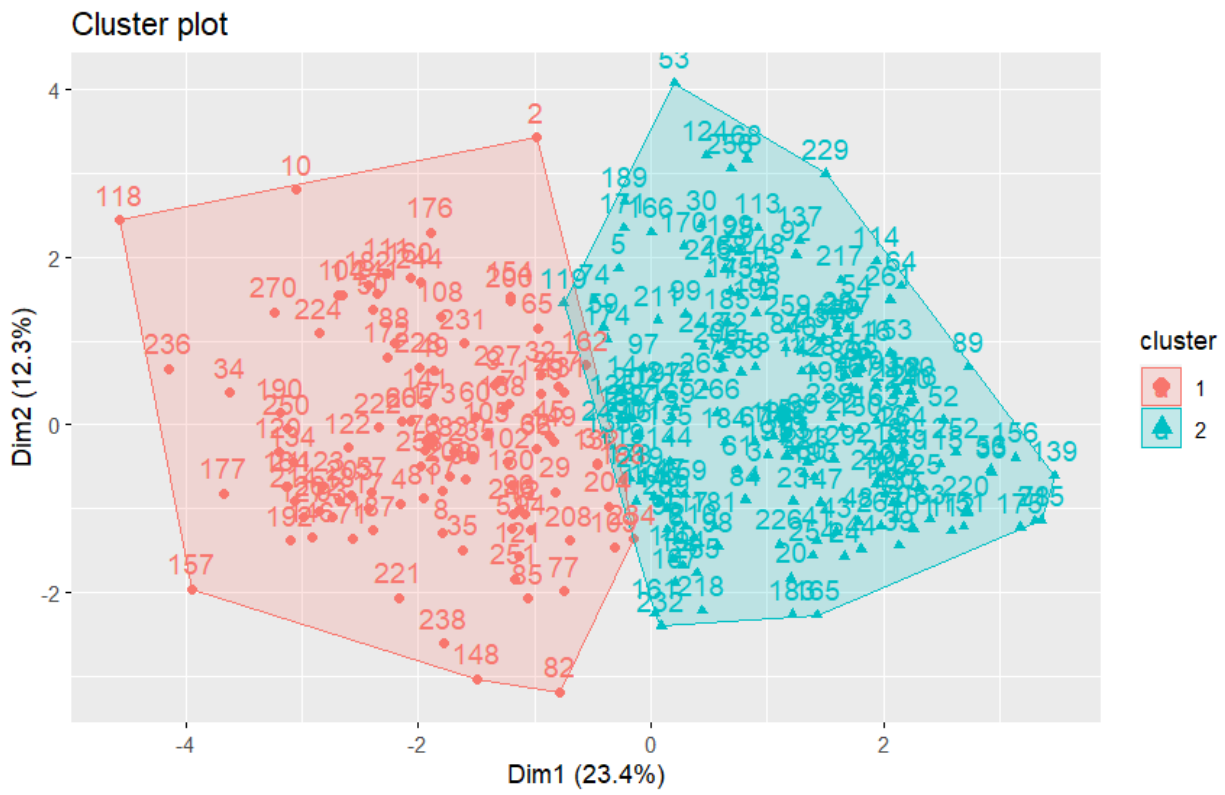



```
[1] "Test accuracy: 0.733333349227905"
```



Cluster Summary

```
 cluster      age       sex       cp trestbps     chol       fbs    restecg
thalach     exang
1       1 58.28713 0.8217822 2.752475 134.9703 258.8614 0.1782178 1.2376238
132.5644 0.6831683
2       2 52.13018 0.5917160 1.828402 129.1775 244.1598 0.1301775 0.8934911
159.9053 0.1183432
    oldpeak     slope        ca     thal    target
1 1.8554455 0.9603960 1.2574257 2.504950 0.8712871
2 0.5686391 0.3609467 0.3195266 1.414201 0.1893491
```



Cluster 1 Analysis

- Age: The average age in this cluster is approximately 58.3 years, indicating a predominance of older individuals.
- Sex: With a mean value of 0.82 (assuming 1 represents male and 0 female), this cluster is predominantly male.
- Chest Pain Type (cp): A higher average value (2.75), suggesting more severe types of chest pain are common in this group.
- Resting Blood Pressure (trestbps): The mean resting blood pressure is around 135 mmHg, which is on the higher side.
- Cholesterol (chol): The average cholesterol level is about 259 mg/dl, indicating elevated levels.
- Fasting Blood Sugar (fbs): The mean value is 0.18, suggesting that higher fasting blood sugar levels might be more common in this group.
- Resting ECG (restecg): The average value is around 1.24, which might indicate some abnormalities in resting electrocardiographic results.
- Maximum Heart Rate (thalach): The average maximum heart rate is approximately 132 bpm, which is relatively lower compared to cluster 2.
- Exercise Induced Angina (exang): A higher mean value of 0.68 indicates a greater prevalence of exercise-induced chest pain.

- ST Depression (oldpeak): The average value is around 1.86, indicating notable ST depression during exercise.
- Slope of ST Segment (slope): The mean value is close to 1, which might suggest a flat or unsloping ST segment.
- Number of Major Vessels (ca): With an average of 1.26, this suggests a higher prevalence of visible major vessels in angiographic tests.
- Thalassemia (thal): The mean value of 2.50 could indicate a higher prevalence of abnormal thalassemia results.
- Target: The mean value of 0.87 suggests that a high proportion of individuals in this cluster are likely to have heart disease.

Cluster 2 Analysis

- Age: The average age is about 52.1 years, indicating a relatively younger group.
- Sex: With a mean value of 0.59, this cluster has a more balanced distribution of males and females, slightly leaning towards male.
- Chest Pain Type (cp): The average value is 1.83, indicating less severe chest pain types compared to cluster 1.
- Resting Blood Pressure (trestbps): The mean value is around 129 mmHg, which is lower than cluster 1.
- Cholesterol (chol): The average cholesterol level is about 244 mg/dl, which is lower than cluster 1.
- Fasting Blood Sugar (fbs): A mean value of 0.13 suggests lower fasting blood sugar levels compared to cluster 1.
- Resting ECG (restecg): The mean value is 0.89, indicating fewer ECG abnormalities.
- Maximum Heart Rate (thalach): A higher average of approximately 160 bpm suggests better exercise tolerance.
- Exercise Induced Angina (exang): A lower mean value of 0.12 indicates less frequent exercise-induced chest pain.
- ST Depression (oldpeak): The average value is around 0.57, indicating less ST depression.
- Slope of ST Segment (slope): The mean value is around 0.36, suggesting more upsloping ST segments.
- Number of Major Vessels (ca): An average of 0.32 indicates fewer visible vessels in angiographic tests.
- Thalassemia (thal): The average value of 1.41 suggests fewer thalassemia abnormalities.
- Target: The mean value of 0.19 indicates that a lower proportion of individuals in this cluster are likely to have heart disease.

Overall Analysis

- Cluster 1 seems to represent older individuals with more severe symptoms and risk factors for heart disease. This group shows higher values in variables traditionally associated with higher heart disease risk.
- Cluster 2 consists of relatively younger individuals with fewer symptoms and risk factors, suggesting a lower risk of heart disease.