Final Project Contribution Grade

Each group is allocated five points per person (10 points total for 2-person groups and 15 points total for 3-person groups) that can be distributed among your group members to reflect individual contributions to the project. The max total points any one person can receive is 10 points. **See example allocations below for 2-person and 3-person groups:**

- Equal distribution:

Person	Points
1	5
2	5

Person	Points
1	5
2	5
3	5

- Unequal distribution (half point increments are fine):

Person	Points
1	7
2	3

Person	Points
1	8
2	3.5
3	3.5

Instructions:

1. Meet with your group members and allow each person to share their contributions to the final project.
2. **Respectfully** discuss and consider the contributions made by you and your group members.
3. **Every group member should mention their contribution in detailed manner.**
4. Decide on an appropriate point allocation for the contribution component of the project grade.
5. Once a **unanimous decision** has been reached, complete the form on the next page.
6. Have someone from the group turn in the signed paper to me in the final class.
7. If you cannot reach a unanimous decision, please e-mail me for further instructions.

Group #: 3

Full name: Aditya Baxi

Points awarded: 5

Signature: 

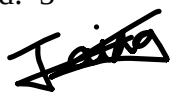
Full name: Ivan Francis

Points awarded: 5

Signature: 

Full name: Jainam Jagani

Points awarded: 5

Signature: 

Make sure that the total points allocated sum to 10 for 2-person groups and 15 for 3-person groups.

Each group member should fill out the following pages with their own contribution to the project in detail.

Full Name: Aditya Baxi

Contribution:

I made significant contributions to the project in various aspects. Firstly, I played a crucial role in finding and selecting the appropriate dataset for our project on Hepatitis C. Through thorough research and evaluation, I helped my group in finding a reliable dataset that provided the necessary variables and observations to conduct a comprehensive analysis. This dataset formed the backbone of our project and laid the foundation for our subsequent analyses.

Furthermore, I took the lead in fitting the K-nearest neighbors (KNN) model as part of our predictive modeling approach. I diligently preprocessed the data, including handling missing values, feature scaling, and data partitioning. I then implemented the KNN algorithm, The KNN model provided valuable insights into the patterns and relationships within the data, allowing us to make predictions and draw conclusions regarding Hepatitis C.

In addition to my contributions in finding the dataset and fitting the KNN model, I actively participated in preparing a substantial portion of the project report. I played a key role in analyzing and interpreting the results obtained from the KNN model, identifying important features, and discussing the implications of our findings. I ensured that the report was comprehensive, well-structured, and effectively communicated the key insights derived from our analysis.

Overall, my contributions in finding the dataset, fitting the KNN model, and actively participating in the project report have been instrumental in advancing our understanding of Hepatitis C. Through collaborative efforts, we have made significant progress in shedding light on the factors influencing the disease and providing valuable insights for further research and decision-making related to this specific dataset.

Full Name: Ivan Francis

Contribution:

I played a key role in performing basic summary statistics and generating visualizations to gain initial insights into the dataset. By analyzing descriptive statistics and creating visual representations such as histograms, scatter plots, and box plots, I helped in understanding the distribution and relationships of the variables, identifying any outliers or patterns, and providing a comprehensive overview of the data.

Furthermore, I actively participated in finding the appropriate dataset for the project. Through thorough research and evaluation, I sourced a high-quality dataset that contained relevant information and variables required for the analysis of Hepatitis C. This involved assessing the dataset's reliability, data quality, and compatibility with the project's objectives.

Moreover, I contributed significantly to the fitting of the random forest model. I played a crucial role in implementing the random forest algorithm, training the model, and evaluating its performance. Additionally, I generated the variable importance graph, which provided valuable insights into the significance of different predictors in predicting Hepatitis C. By identifying the most influential variables, we gained a deeper understanding of the key factors contributing to the occurrence of the disease.

Overall, my contributions encompassed conducting basic summary statistics, visualizations, identifying a suitable dataset, and playing a pivotal role in fitting the random forest model and interpreting its results. These contributions have been instrumental in advancing the project's objectives and providing valuable insights into the dataset and its relation to Hepatitis C while also understanding key statistical modelling techniques and how to tune them in order to fit our objective.

Full Name: Jainam Jagani

Contribution:

In this project, my contributions were diverse and encompassed various aspects of the research. Firstly, I played a crucial role in the initial phase by actively searching and finding the appropriate dataset for the analysis of Hepatitis C. This involved extensive research and evaluation of various sources to ensure the dataset's relevance, reliability, and comprehensiveness.

Furthermore, I took the lead in analyzing the dataset and addressing the first and second research questions. Specifically, I focused on fitting the Multinomial Logistic classification model to investigate the relationship between different variables and Hepatitis C. This involved carefully selecting the relevant predictors, handling missing data, and interpreting the model results to gain meaningful insights into the factors associated with the disease.

Moreover, I played a pivotal role in the comparative analysis of different models by developing the Area Under the Curve (AUC) curve. This enabled us to evaluate and compare the performance of the random forest, KNN, and Multinomial Logistic classification models in predicting Hepatitis C. The AUC curve provided a comprehensive visualization of the models' predictive capabilities, aiding in the selection of the most accurate and reliable model.

Overall, my contributions in finding the dataset, conducting parts of the report, and fitting the Multinomial Logistic classification model, as well as developing the AUC curve for model comparison, significantly contributed to the project's success. These efforts collectively enhanced our understanding of statistical models and techniques.