***Statistics Final Project Break Up (Preliminary)***

***1. Preliminary Work: Data Exploration and Descriptive Analysis***  🡪 Vero 2nd Dec Midnight

***a***. Explore Dataset:

- Understand the dataset structure, variables, and scope.

- Identify the timeline (2015-2022) and number of records (16,215) for Master's (12,091) and PhD (4,124) programs.

***b.*** Descriptive Analysis:

- Summarize key statistics (mean, variance) for applicant characteristics (nationality, age, gender, etc.).

- Utilize visualizations (histograms, box plots) to illustrate distribution patterns.

- Identify missing or inconsistent data for further handling.

***-*** Filter the Data points with respect to Fall Semester

***2. Preliminary Work: Data Cleaning and Variable Transformation*** -->Sharon 3rd Dec 2023

***a.*** Handle Missing Data:

- Implement imputation or removal strategies for missing values.

- Ensure data integrity through thorough data cleaning procedures.

***b.*** Variable Transformation:

- Convert categorical variables into dummy variables if necessary.

- Standardize numerical values for consistent analysis.

- Perform exploratory data analysis (EDA) to detect and address outliers.

***3. Preliminary Work: Correlational and Explanatory Analysis***   **🡪** Anik 4th December 2023

***a.*** Correlational Analysis:

- Explore relationships between variables, particularly age variance and degree type.

- Utilize visualizations (scatter plots, correlation matrices) to depict relationships.

***b***. Explanatory Analysis:

- Conduct initial exploratory regression analysis to identify factors influencing age variance.

- Evaluate the impact of degree type, previous course failure, and other variables on age variance.

***4. Preliminary Work: Statistical Tests and Initial Findings*** ->Babli 5th December 2023

***a.*** Statistical Tests:

- Apply Levene's test for homogeneity of variances to assess age variance differences between groups.

- Utilize chi-squared tests of independence to examine relationships, e.g., academic probation status and program application.

***b.*** Initial Findings:

- Summarize significant relationships or patterns observed.

- Provide recommendations for further analysis based on preliminary findings.

Version Control:  
RData and RMarkdown