Final Project

Intorduction to Data Science with Python

Luka Dumbadze, Nikoloz Kvinikadze, Ivane Urjumelashvili

Kutaisi International University

January 22, 2025

Contents

Project Goals

- **Objective:**
  - Understand the factors influencing depression among students.
  - Analyze relationships between lifestyle, academic performance, and mental health.
- **Key Deliverables:**
  - Clean and preprocess the dataset.
  - Conduct exploratory data analysis (EDA) with visualizations and statistical insights.

Dataset Overview

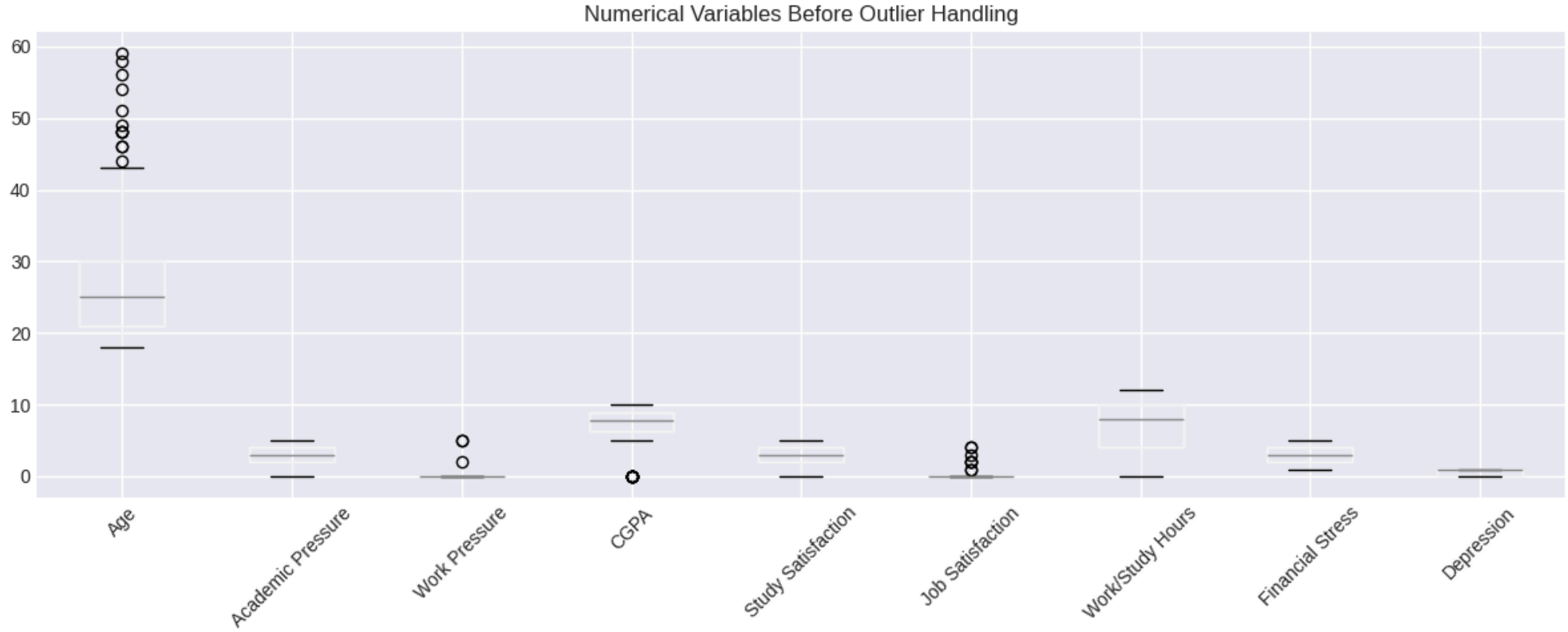
- **Dataset Description:**
  - Source: Provided dataset on student depression.
  - Rows: Number of students.
  - Columns: Features such as age, gender, CGPA, depression status, academic pressure, etc.
- **Target Variable:** Depression_Status (Binary: Yes/No)
- **Challenges:**
  - Missing values in some columns.
  - Presence of outliers in numerical data.

Data Preprocessings

- **Steps Taken:**
  - Handled missing values using median/mode imputation.
  - Identified and addressed outliers using the IQR method.
  - Encoded categorical variables (e.g., *Sleep Duration* mapped to numeric values).
- **Visuals:**
  - **before-and-after box plot** to show outlier handling.

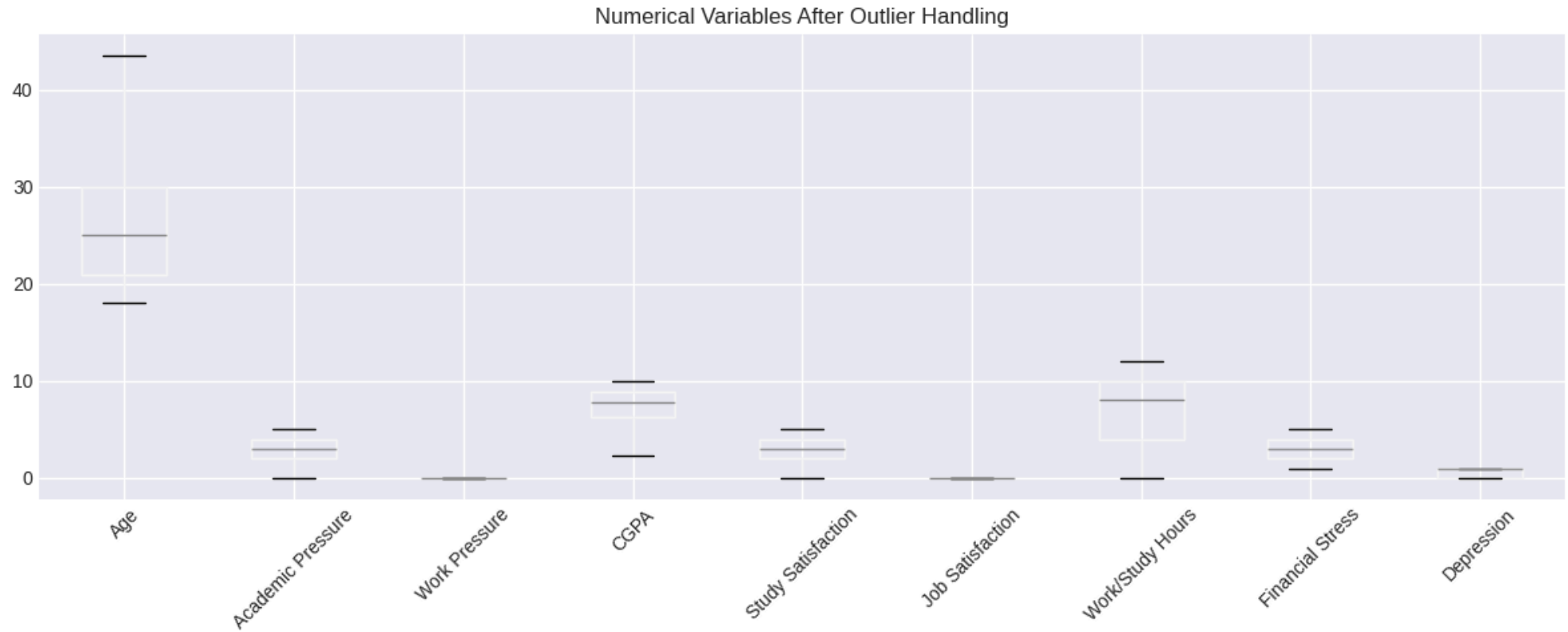
Data Preprocessings

before box plot



Data Preprocessings

before box plot



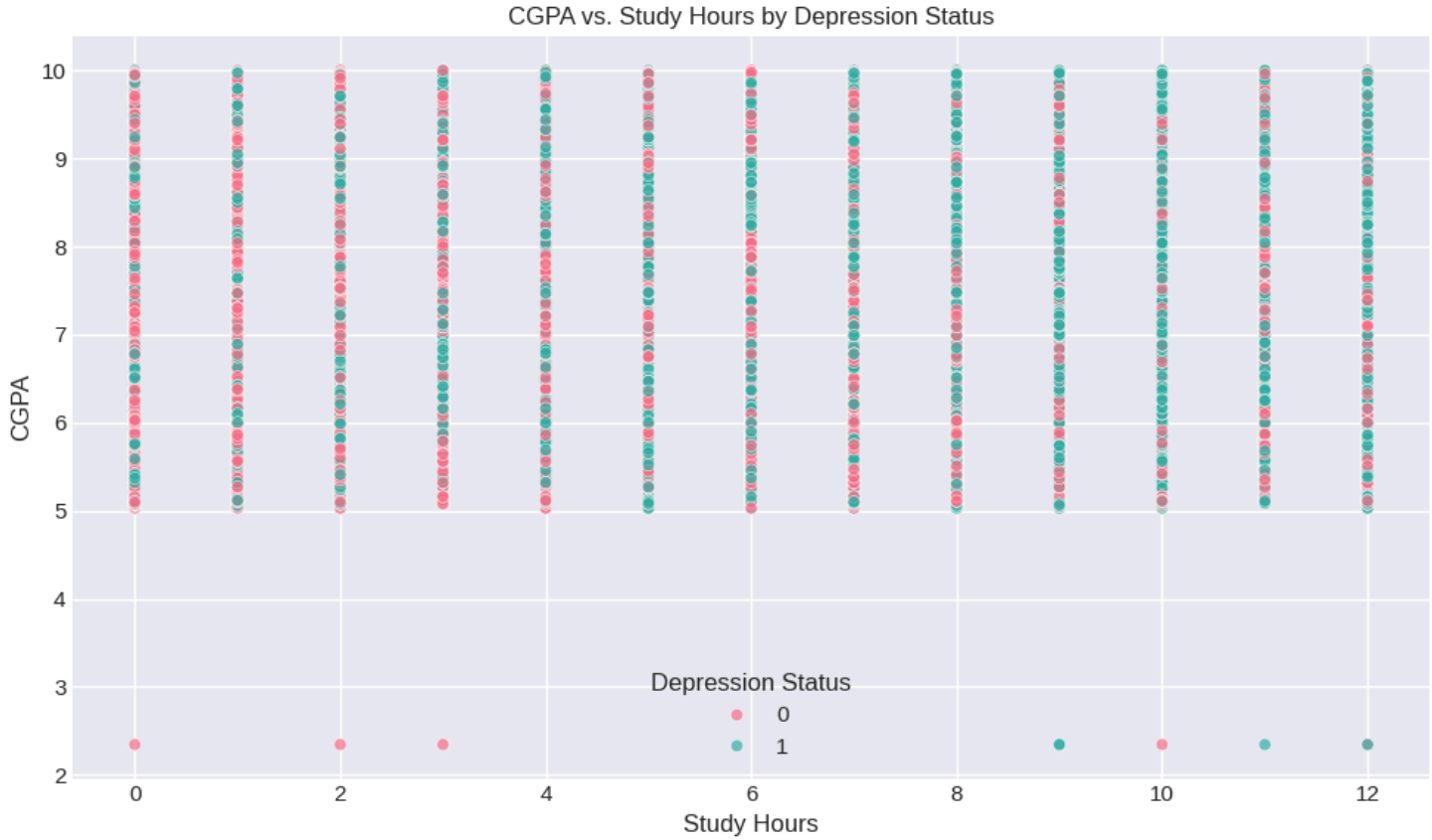
Exploratory Data Analysis (EDA)

- **Purpose:**
  - Discover patterns and trends in the data.
  - Visualize relationships between key features and depression status.
- **Tools Used:**

  Pandas, Seaborn, Matplotlib, and Scipy for statistical tests.

Key Findings from EDA

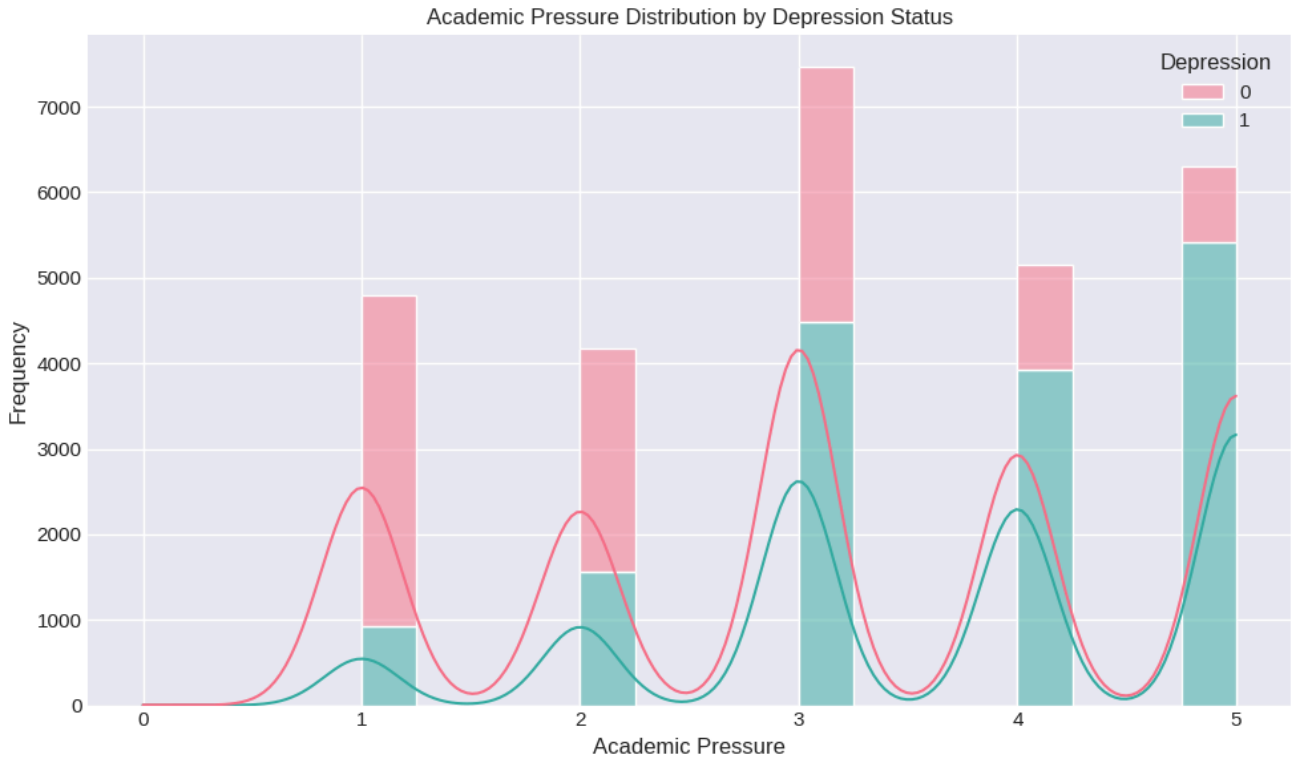
Students with lower CGPA tend to have higher depression rates.



Key Findings from EDA

Academic Pressure Distribution

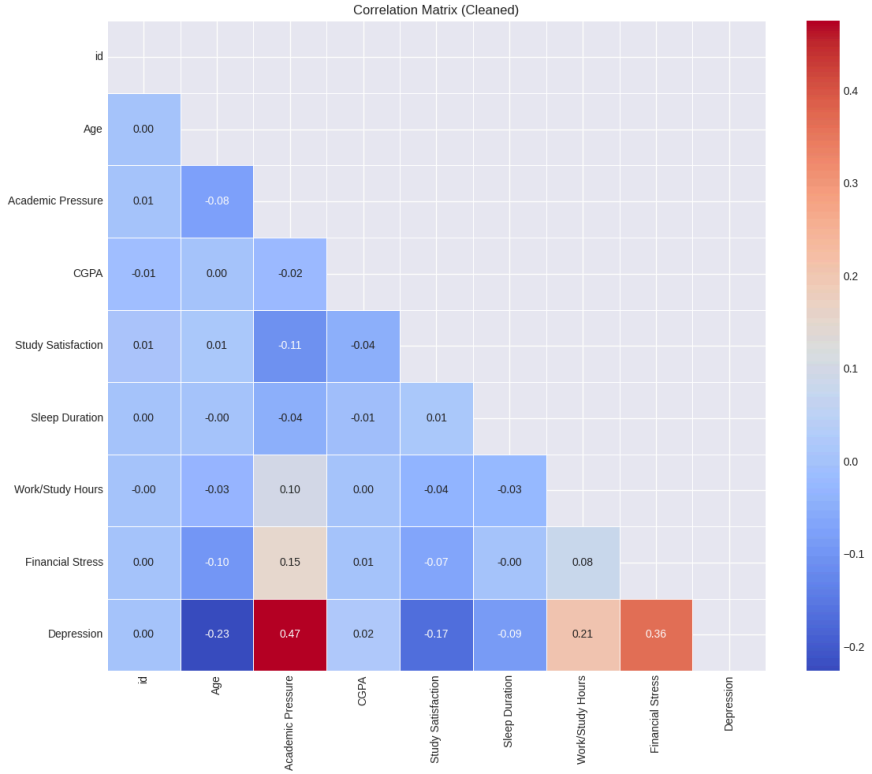
Students experiencing higher academic pressure are more likely to report depression.



Key Findings from EDA (continued)

Correlation Heatmap

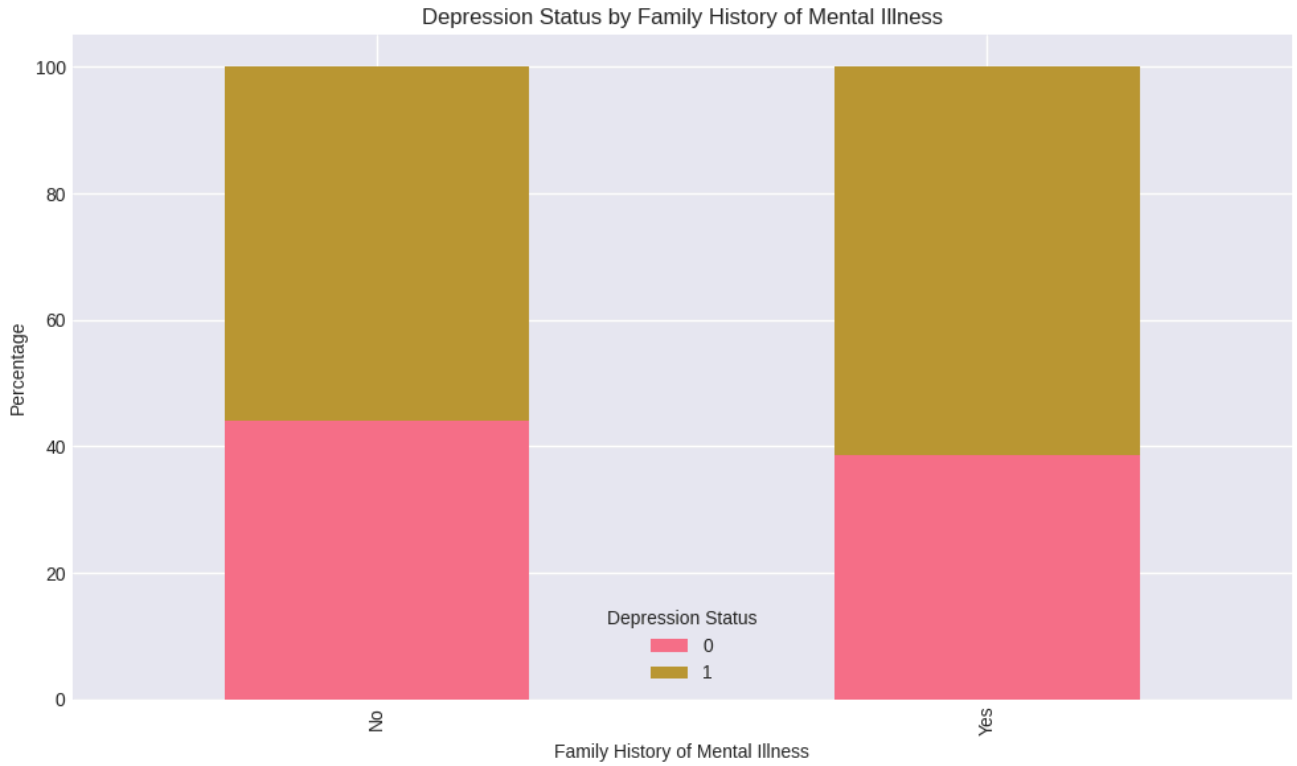
Strong correlations observed between depression and factors like academic pressure, work pressure, and financial stress.



Key Findings from EDA (continued)

Family History of Mental Illness

Students with a family history of mental illness have a higher likelihood of depression.



Statistical Insights

- **Risk Factors:**
  - **Mann-Whitney U Test:** Significant differences in academic pressure and CGPA between students with and without depression ($p < 0.05$).
  - **Chi-Square Test:** Significant association between depression and family history of mental illness.
- **Effect Sizes:**
  - Academic pressure shows a medium effect size, indicating a notable impact.

Conclusion

- **Summary of Findings:**
  - Academic and work pressure, along with financial stress, are key contributors to depression.
  - Family history of mental illness is a significant categorical factor.
- **Effect Sizes:**
  - Academic pressure shows a medium effect size, indicating a notable impact.

Encourage the audience to ask questions about the EDA process or findings.