

# Student Depression Analysis



# Introduction



The involved dataset is aimed at analyzing, understanding, and predicting depression levels among students.

It may include features such as demographic information (age, gender), academic performance (grades, attendance), lifestyle habits (sleep patterns, exercise, social activities), mental health history, and responses to standardized depression scales.

# Data cleaning & Preprocessing

- - Checked for missing values
- - Verified duplicates
- - Dataset is mostly clean itself



# Data Exploration



Dataset loaded using Pandas

Explored structure using  
head(), info(), describe()

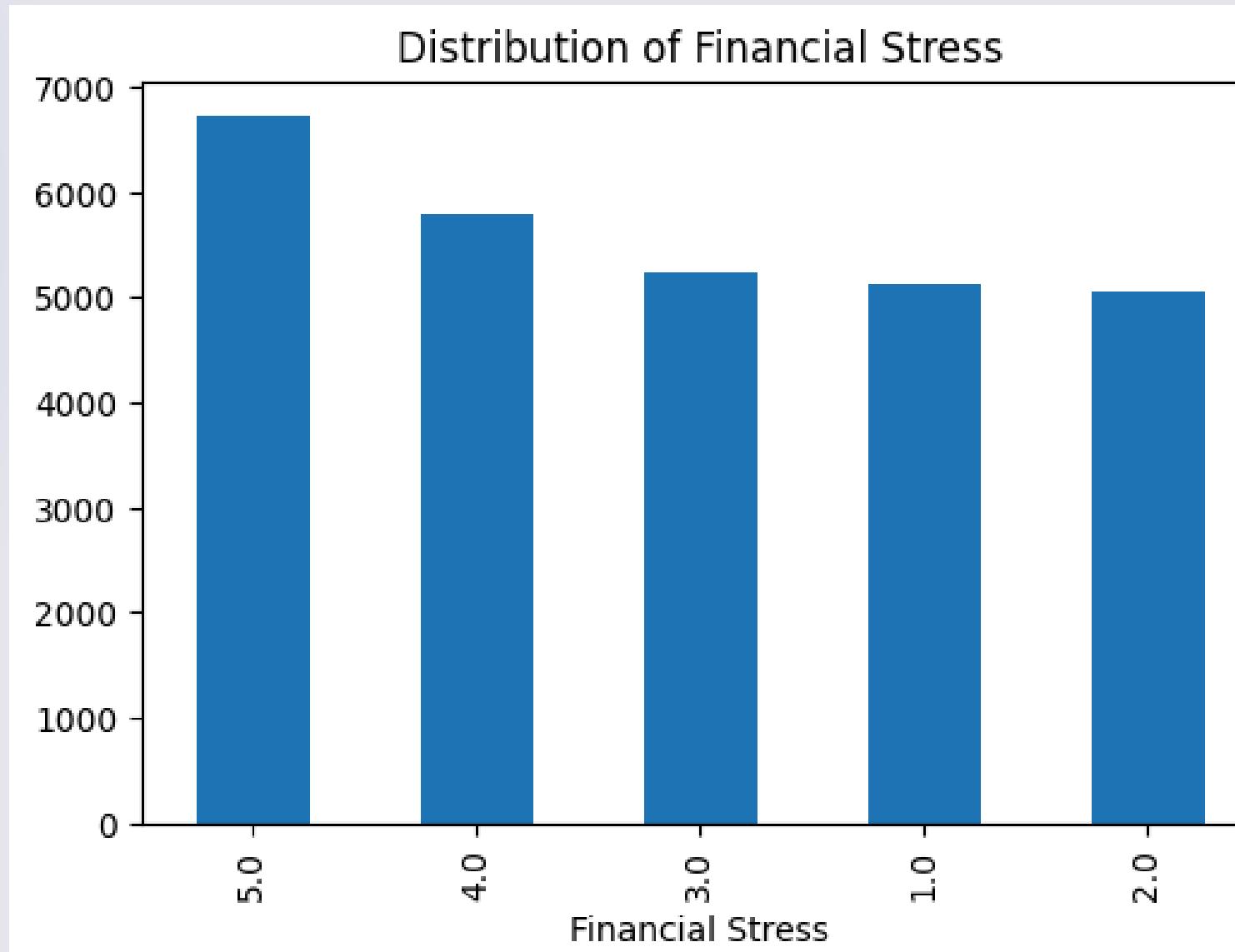
Checked feature distributions

# EDA on

- - Financial Stress distribution
- - Degree distribution
- - Depression prevalence among students



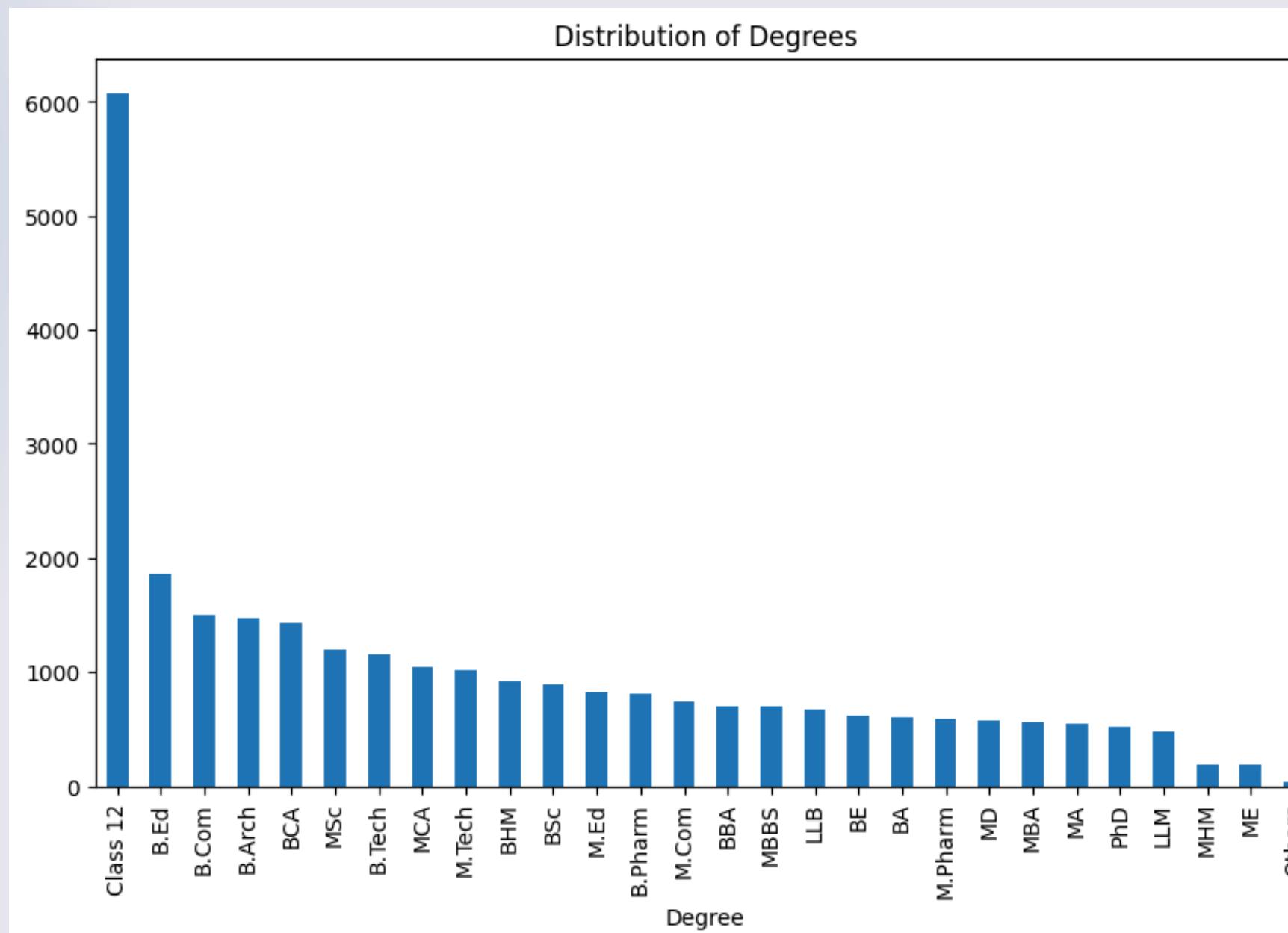
# Data Exploration



A higher proportion of students with financial stress may show symptoms of depression.



# Based on Degrees

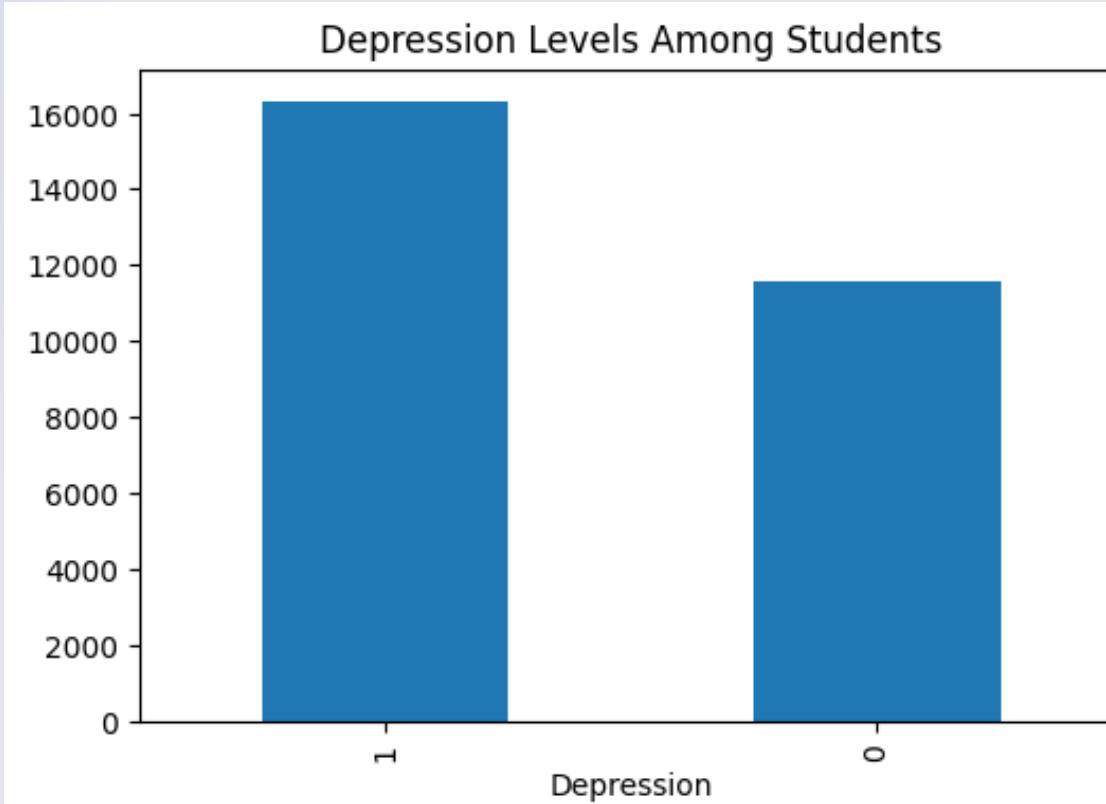


Students of class 12<sup>th</sup> show  
much more cases of depression  
than any other category

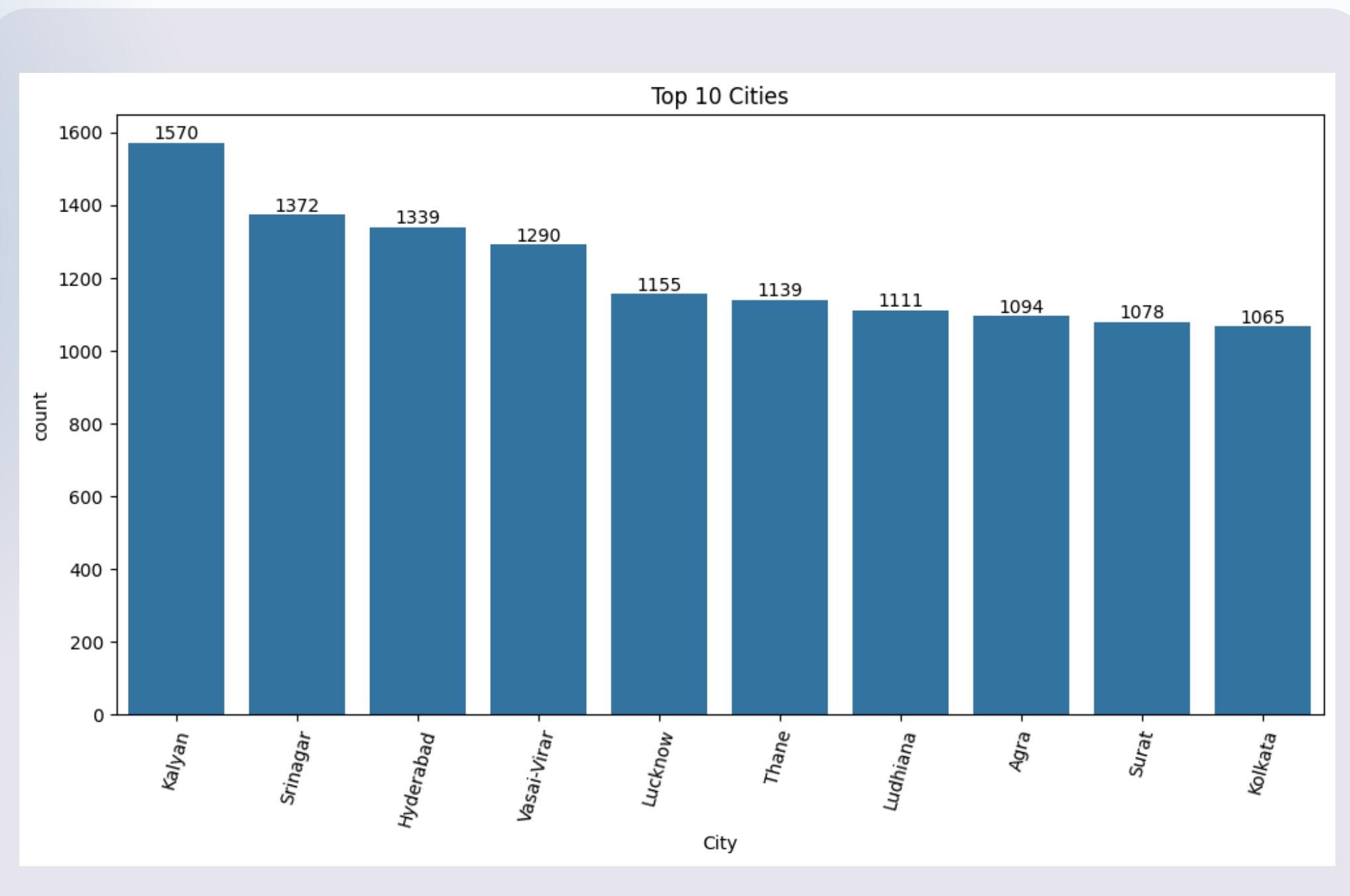
# Total



A total of 11563 students are found to have depression and is much greater in number compared to who don't



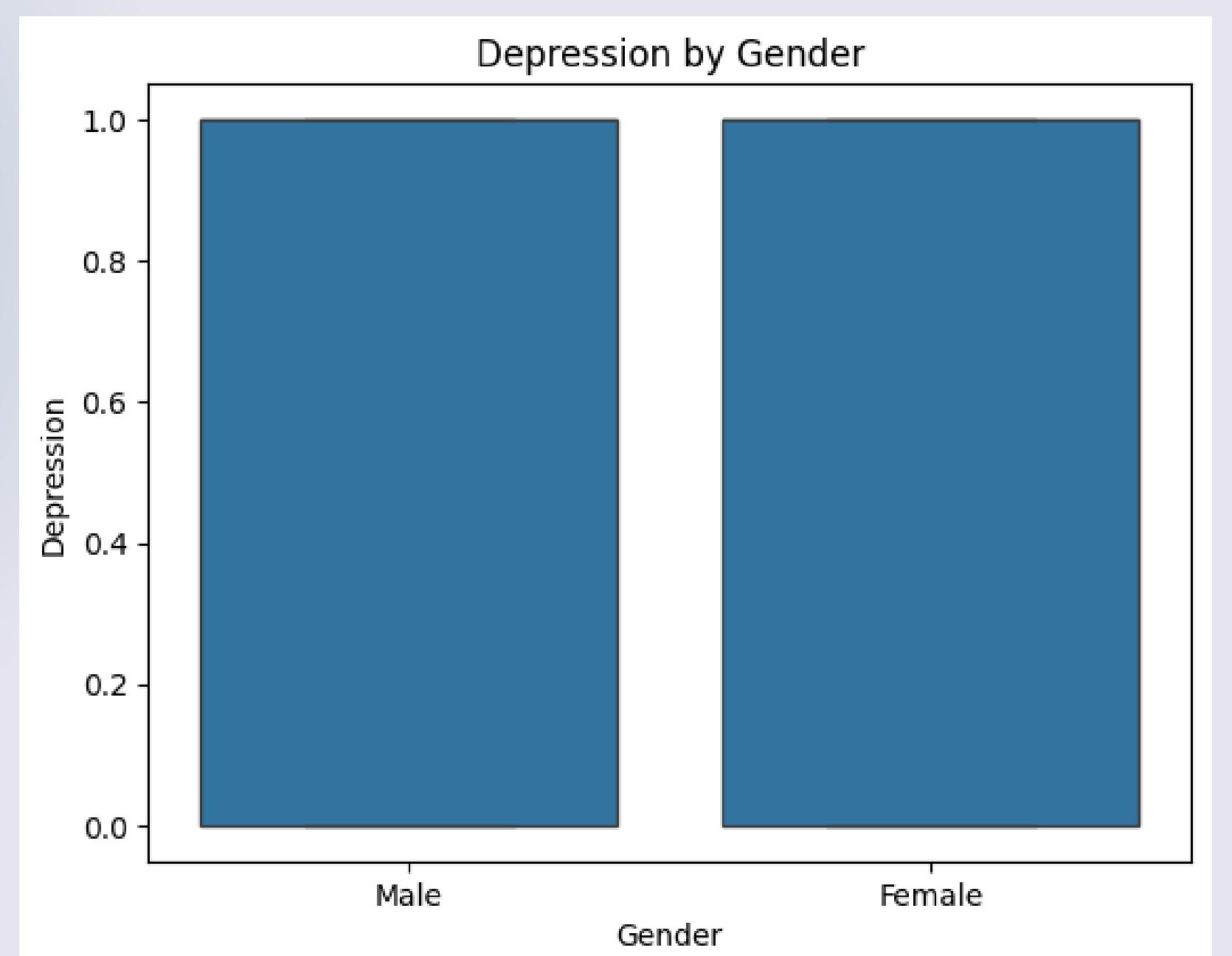
# Top 10 Cities



Kalyan appears to have highest no. of cases.



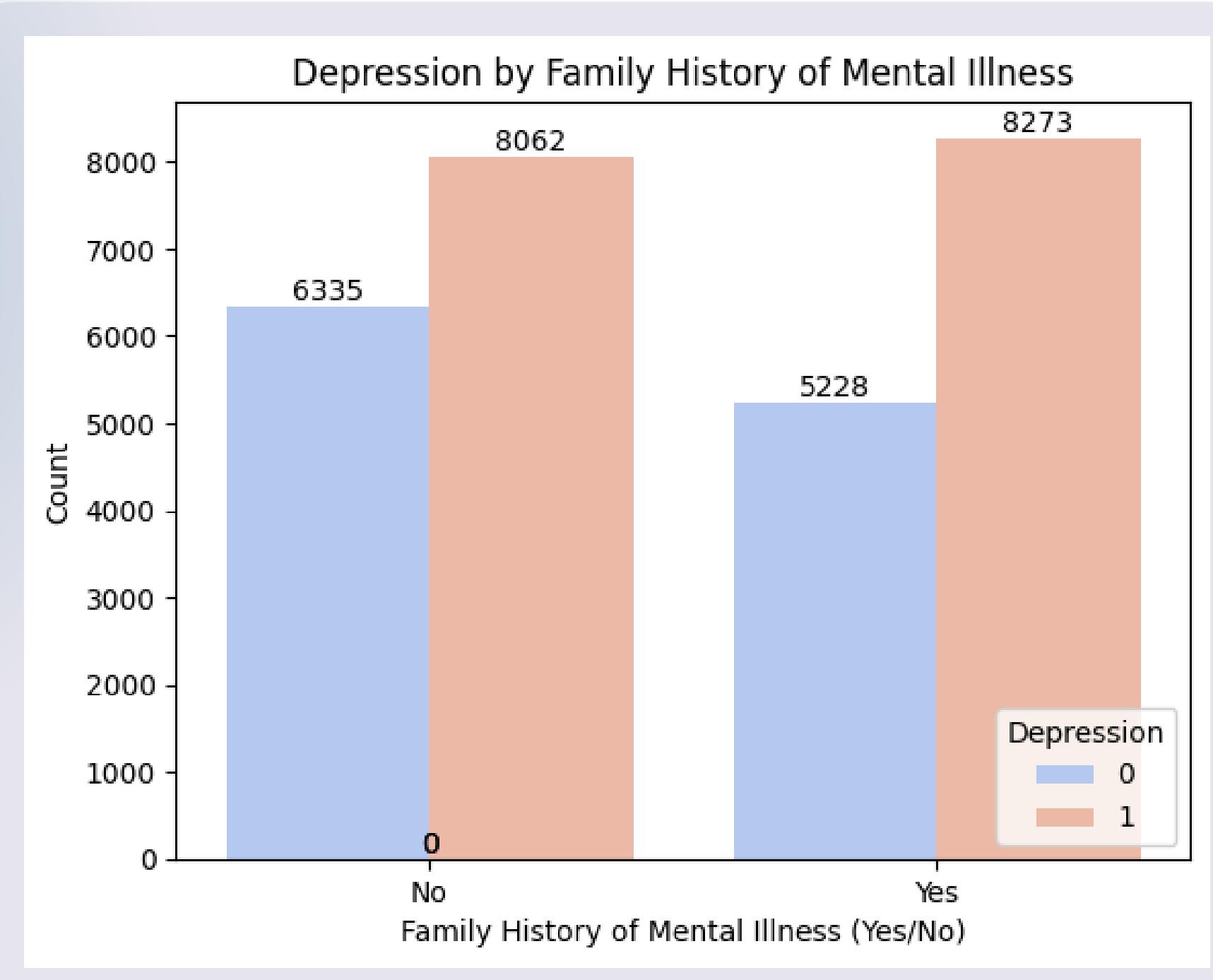
# Depression by Gender



Gender doesn't play an  
important role in this  
case



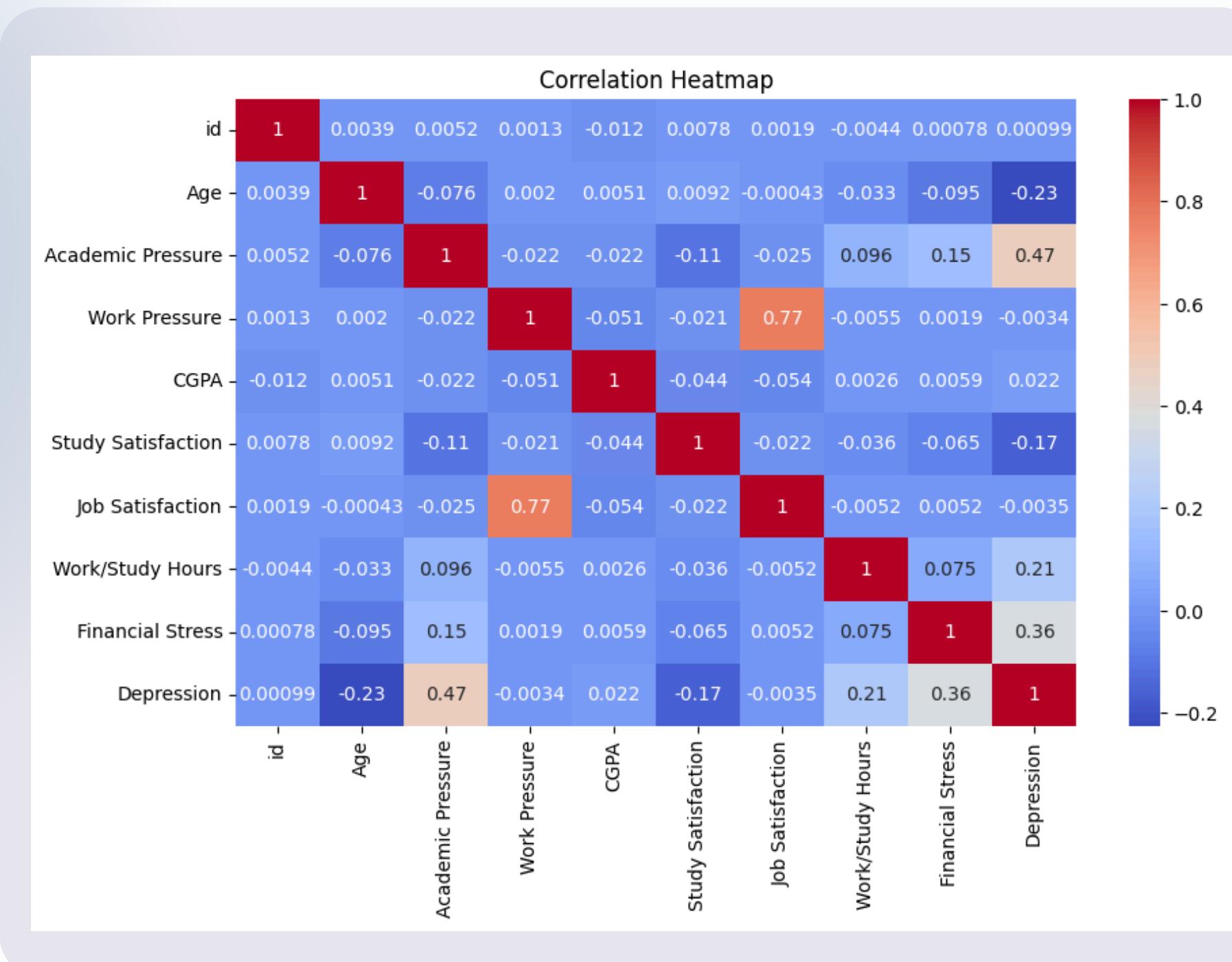
# Family History



Even Family History  
appears to be irrelevant.



# Correlation Heatmap



Features with higher correlation to Depression are potential predictors.

# Conclusion



- Depression among students is a significant concern, with over 11,500 cases identified in the dataset.
- Financial stress and academic pressure (especially in Class 12) are strong contributing factors.
- Geographical variation is visible, with Kalyan showing the highest reported cases.
- Gender and family history show little to no correlation, indicating that depression can affect all students regardless of background.
- Strongly correlated features identified in the analysis can be used as predictive markers to design early intervention systems.



# My Details



[GitHub](#)

[LinkedIn](#)

[param.wairagar@gmail.com](mailto:)