**STUDENT MENTAL HEALTH DATA ANALYSIS**

## **Overview**

File Source: StudentDepressionAnalysisDataset.csv

**Analysis Finding**  
The Data Shows **moderate correlations** between financial stress and depression, academic pressure and suicidal thoughts, and a **weak negative correlation** between study satisfaction and CGPA.

We discovered there are **1,727 students** that are at risk. The accounted student are having **high academic pressure** with **low sleep duration, high financial stress** and experiencing **depression**

### **Dataset Dimension**

* Id
* Gender
* Age
* City
* Profession
* Academic Pressure
* CGPA
* Study Satisfaction
* Sleep Duration
* Dietary Habits
* Degree
* Have you ever had suicidal thoughts?
* Work Pressure
* Job Satisfaction
* Work/Study Hours
* Financial Stress
* Family History of Mental Illness
* Depression

### **Business Request for Data Analysis**

**Client/Stakeholder:** University Counseling and Student Services Department

**Business Context:** The University is concerned about the growing mental health challenges among students, which are believed to be impacting academic performance, well-being, and overall satisfaction. The Student Services Department wants to assess key factors such as academic pressure, financial stress, dietary habits, and sleep patterns, and understand their correlation with mental health conditions like depression and suicidal thoughts.

The University has collected survey data from students, including their academic experiences, health behaviors, and mental health history. The goal is to use data analysis to gain insights that can help in designing better mental health programs, improving academic outcomes, and providing more targeted support to at-risk students.

**Specific Deliverables:**

* An Overview Dashboard on Excel (For Non-Techy User)
* Visual Correlation Analysis Report
* Risk Group Identification
* Recommendations for Action

### **Business Requirements**

Academic Performance KPIs

* Average Academic Pressure
* Average CGPA
* Average Study Satisfaction

Mental Health KPIs

* Suicidal Thoughts Percentage
* Depression Rate
* Family History Rate

Health and Well-being KPIs

* Average Sleep Duration
* Dietary Habits Percentage
* Work/Study Hours vs. Sleep Duration Ratio

Stress and Financial KPIs

* Average Financial Stress
* Financial Stress and Depression Correlation

Demographic KPIs

* By Gender
  + Suicidal Thoughts Percentage
  + Depression Rate
* By Age Group
  + Suicidal Thoughts Percentage
  + Depression Rate
* By Profession
  + Suicidal Thoughts Percentage
  + Depression Rate
* By Degree
  + Suicidal Thoughts Percentage
  + Depression Rate

Behavioral KPIs

* Academic Pressure vs. Suicidal Thoughts Correlation
* Study Satisfaction vs. CGPA Correlation

Risk Assessment KPIs

* High-Risk Students Count

### **Process for Data Analysis**

* Exploratory Data Analysis using Power Query
  + Clean and Transform Dataset
  + Do Initial Analysis by Pivot Tables
* Creating Data Visualization
  + Overview Dashboard
  + Detailed View Report

## **The Data Analysis Process**

* Exploratory Data Analysis using Power Query
* Transorm the Dataset by the following:
  + Filter Profession = *Student* Only
  + Remove **City, Work Pressure,** and **Job Satisfaction** Column
  + Add columns for **Family History of Mental Illness**, and **Suicidal Thoughts** as *1s and 0s*
  + Add columns for **Depression** as *Yes*/*No* values
  + Filter out “*Others*” in **Sleep Duration, Dietary Habit**
  + Filter dataset to include only Ages less or equal to *37*
  + Created Age Group with Rangesize = 5
  + Filter “0” in Study Satisfaction, CGPA, Academic Pressure, and *null* in Financial Stress
  + Remove Duplicates

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

**Initial Analysis**

Check all KPIs

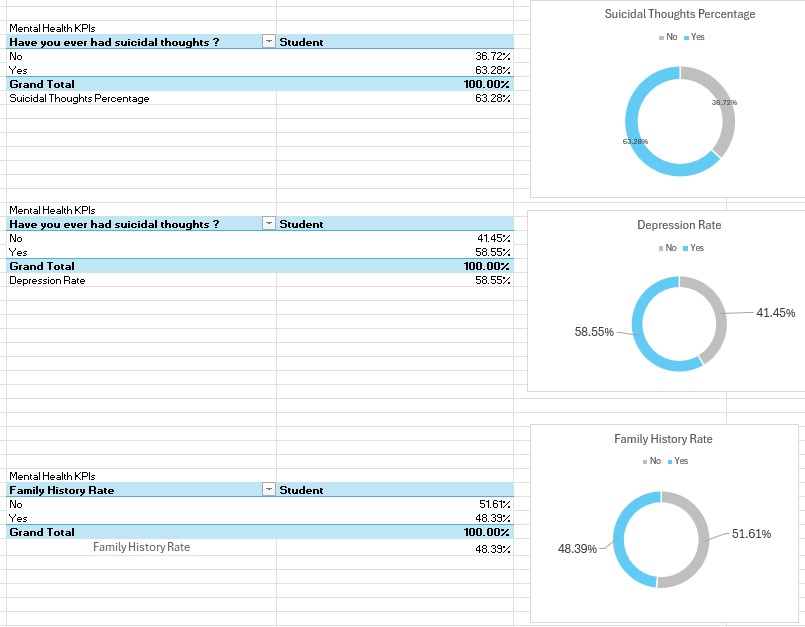
Academic Performance KPIs:

* Average Academic Pressure
* Average CGPA
* Average Study Satisfaction



Mental Health KPIs:

* Suicidal Thoughts Percentage
* Depression Rate
* Family History Rate



Health and Well-being KPIs:

* Sleep Duration Student Count
* Dietary Habits Student Count
* A blue and white rectangular object with black text

  Description automatically generatedWork/Study Hours vs. Sleep DurationA screenshot of a computer

  Description automatically generated

Stress and Financial KPIs:

* A screenshot of a computer

  Description automatically generatedAverage Financial Stress

Demographic KPIs

add slicer:

* Gender
* Degree
* Age Group

A screenshot of a computer

Description automatically generated

Correlation:

* Financial Stress and Depression Correlation
* Academic Pressure vs. Suicidal Thoughts Correlation
* Study Satisfaction vs. CGPA Correlation

Uses Spearman's Rank Correlation Formula

**Spearman’s Rank Correlation** is used to measure the strength and direction of the relationship between two variables with ordinal data. It doesn’t assume equal differences between ranks, making it suitable for non-normally distributed data. Ordinal data has a natural order, but the distance between categories isn't consistent or meaningful, like satisfaction ratings (1 = Low, 5 = High).

A black background with white text

Description automatically generated

**Formula Breakdown**

1. Rank the data:
   * Each data value in both variables is ranked (from 1 to n based on their size). If there are tied values (duplicates), they are assigned the average rank.
2. Calculate the rank differences di
   * For each pair of values, you calculate the difference between their ranks in the two variables, di=rank of variable 1−rank of variable 2
3. Square the rank differences:
   * For each pair, square the differences to eliminate negative values, i.e., di2
4. Sum of squared differences:
   * Sum all the squared differences, ∑ di2
5. Substitute in the formula:
   * Finally, substitute the sum of squared differences and the number of data points (n) into the formula. This provides the Spearman's rank correlation coefficient ρ

**Solve the Correlation:**

A screenshot of a computer

Description automatically generated

Risk Assessment KPIs

* At-Risk Students = *Students with****High Academic Pressure*** *+ Students with****High Financial Stress*** *+ Students with****Low Sleep Duration*** *+ Students with****Depression***

A graph with red lines

Description automatically generated with medium confidence

**Creating Visual Dashboard**

* Combine all Visuals and KPIs to create Visual Dashboard

A screenshot of a medical report

Description automatically generated