Dataset link-

https://www.kaggle.com/datasets/riyasinghrathore/iit-students-depression-data

# Demographic Analysis

1. What is the gender distribution of the students?

Answer: This query will show the count of students grouped by gender.

Query:

```
SELECT Gender, COUNT(*) AS Count
FROM dataset
GROUP BY Gender;
```

Description: This query groups the students by gender and counts the number of students in each group.

2. What is the average age of the students?

Answer: This query will calculate the average age of the students in the dataset.

Query:

```
SELECT AVG(Age) AS Average_Age
FROM dataset;
```

Description: This query calculates the average age of all students in the dataset.

3. How does the age distribution vary by gender?

Answer: This query provides statistical details (min, max, mean, and standard deviation) of age grouped by gender.

Query:

```
SELECT Gender, MIN(Age) AS Min_Age, MAX(Age) AS Max_Age, AVG(Age) AS Average_Age,
STDDEV(Age) AS Age_StdDev
FROM dataset
GROUP BY Gender;
```

Description: This query calculates the minimum, maximum, average, and standard deviation of age for each gender group.

# Mental Health Analysis

4.What percentage of students experience frequent thoughts or ideas of guilt?

Answer: This query shows the percentage of students experiencing different levels of guilt.

Query:

```
SELECT guilt_feelings, (COUNT(*) * 100.0 / (SELECT COUNT(*) FROM dataset)) AS Percentage
FROM dataset
GROUP BY guilt_feelings;
```

Description: This query calculates the percentage of students in each category of guilt feelings.

5.How do students rate their mood and outlook on life?

Answer: This query provides the percentage distribution of students' mood and outlook ratings.

Query:

```
SELECT mood_outlook, (COUNT(*) * 100.0 / (SELECT COUNT(*) FROM dataset)) AS
Percentage
FROM dataset
GROUP BY mood_outlook;
```

Description: This query calculates the percentage of students in each mood and outlook category.

6.How many students feel they have a clear understanding of their current situation and circumstances?

Answer: This query shows the percentage of students who feel they have a clear understanding of their current situation.

Query:

```
SELECT understanding_situation, (COUNT(*) * 100.0 / (SELECT COUNT(*) FROM dataset)) AS
Percentage
FROM dataset
GROUP BY understanding_situation;
```

Description: This query calculates the percentage of students in each category of understanding their situation.

7.What is the overall sense of self-worth and confidence among students?

Answer: This query provides the percentage distribution of students' self-worth and confidence ratings.

Query:

```
SELECT self_worth_confidence, (COUNT(*) * 100.0 / (SELECT COUNT(*) FROM dataset)) AS
Percentage
```

```
FROM dataset
GROUP BY self_worth_confidence;
```

Description: This query calculates the percentage of students in each self-worth and confidence category.

8.How many students have experienced difficulties with sleep?

Answer: This query shows the percentage of students who have experienced various levels of sleep difficulties.

Query:

```
SELECT sleep_difficulties, (COUNT(*) * 100.0 / (SELECT COUNT(*) FROM dataset)) AS Percentage
FROM dataset
GROUP BY sleep_difficulties;
```

Description: This query calculates the percentage of students in each category of sleep difficulties.

# Academic Performance and Activities

9.What is the distribution of CGPA among students?

Answer: This query shows the count of students in each CGPA range.

Query:

```
SELECT cgpa, COUNT(*) AS Count
FROM dataset
GROUP BY cgpa
ORDER BY cgpa;
```

Description: This query groups students by their CGPA and counts the number of students in each CGPA range.

10.What is the placement/internship status of the students?

Answer: This query provides the percentage distribution of students' placement/internship status.

Query:

```
SELECT placement_status, (COUNT(*) * 100.0 / (SELECT COUNT(*) FROM dataset)) AS Percentage
FROM dataset
GROUP BY placement_status;
```

Description: This query calculates the percentage of students in each placement/internship status category.

11.How does daily screen time affect students' mental health?

Answer: This query shows the percentage distribution of students' daily screen time.

Query:

```
SELECT screen_time, (COUNT(*) * 100.0 / (SELECT COUNT(*) FROM dataset)) AS Percentage
FROM dataset
GROUP BY screen_time;
```

Description: This query calculates the percentage of students in each screen time category.

12.Is there a correlation between CGPA and depression index?

Answer: This query calculates the average depression index for each CGPA range.

Query:

```
SELECT cgpa, AVG(depression_index) AS Average_Depression_Index
FROM dataset
GROUP BY cgpa;
```

Description: This query calculates the average depression index for students grouped by their CGPA.

13.How do students perceive their current financial situation?

Answer: This query shows the percentage distribution of students' financial situation perceptions.

Query:

```
SELECT financial_situation, (COUNT(*) * 100.0 / (SELECT COUNT(*) FROM dataset)) AS
Percentage

FROM dataset

GROUP BY financial_situation;
```

Description: This query calculates the percentage of students in each financial situation category.

# Social Comparison and Personal Perception

14.How many students feel their friends have better internship/placement prospects?

Answer: This query shows the percentage of students who feel their friends have better prospects.

Query:

```
SELECT better_prospects, (COUNT(*) * 100.0 / (SELECT COUNT(*) FROM dataset)) AS Percentage

FROM dataset

GROUP BY better_prospects;
```

Description: This query calculates the percentage of students who feel their friends have better prospects.

15.How do students describe their level of creativity and intelligence compared to others?

Answer: This query provides the percentage distribution of students' self-reported creativity and intelligence levels.

Query:

```
SELECT creativity_intelligence, (COUNT(*) * 100.0 / (SELECT COUNT(*) FROM dataset)) AS Percentage

FROM dataset

GROUP BY creativity_intelligence;
```

Description: This query calculates the percentage of students in each creativity and intelligence category.

# Detailed Cross-Analysis

16.Is there a gender difference in the experience of guilt feelings?

Answer: This query shows the count of students experiencing different levels of guilt, grouped by gender.

Query:

```
SELECT Gender, guilt_feelings, COUNT(*) AS Count
FROM dataset
GROUP BY Gender, guilt_feelings;
```

Description: This query groups students by gender and guilt feelings and counts the number of students in each group.

17.Is there a correlation between age and mood/outlook on life?

Answer: This query calculates the average mood/outlook rating for each age group.

Query:

```
SELECT Age, AVG(mood_outlook) AS Average_Mood_Outlook
FROM dataset
GROUP BY Age;
```

Description: This query groups students by age and calculates the average mood/outlook rating for each group.

18.How does the level of social connection affect the sense of self-worth and confidence?

Answer: This query calculates the average self-worth and confidence rating for each level of social connection.

Query:

```
SELECT social_connection, AVG(self_worth_confidence) AS Average_Self_Worth_Confidence
FROM dataset
GROUP BY social_connection;
```

Description: This query groups students by their level of social connection and calculates the average self-worth and confidence rating for each group.

19.Is there a relationship between difficulties with sleep and the level of anxiety?

Answer: This query calculates the average anxiety level for each category of sleep difficulties.

Query:

```
SELECT sleep_difficulties, AVG(anxiety_level) AS Average_Anxiety_Level
FROM dataset
```

```
GROUP BY sleep_difficulties;
```

Description: This query groups students by their level of sleep difficulties and calculates the average anxiety level for each group.

20.How do changes in physical movements correlate with feelings of being overwhelmed?

Answer: This query calculates the average feeling of being overwhelmed for each category of physical movements.

Query:

```
SELECT physical_movements, AVG(feeling_overwhelmed) AS
Average_Feeling_Overwhelmed

FROM dataset

GROUP BY physical_movements;
```

Description: This query groups students by their level of physical movements and calculates the average feeling of being overwhelmed for each group.

21.What is the relationship between financial situation and depression index?

Answer: This query calculates the average depression index for each financial situation category.

Query:

```
SELECT financial_situation, AVG(depression_index) AS Average_Depression_Index

FROM dataset

GROUP BY financial_situation;
```

Description: This query groups students by their financial situation and calculates the average depression index for each group.

# Lifestyle and Hobbies

22.What are the most common hobbies among students?

Answer: This query shows the count of students for each hobby.

Query:

```
SELECT hobbies, COUNT(*) AS Count
FROM dataset
GROUP BY hobbies
ORDER BY Count DESC;
```

Description: This query groups students by their hobbies and counts the number of students in each hobby category, ordering the results by count in descending order.

23.How do different hobbies impact students' mental health?

Answer: This query calculates the average depression index for each hobby.

Query:

```
SELECT hobbies, AVG(depression_index) AS Average_Depression_Index
FROM dataset
GROUP BY hobbies;
```

Description: This query groups students by their hobbies and calculates the average depression index for each hobby category.

# Advanced Analysis

24.Perform a cluster analysis to identify different student mental health profiles (requires advanced SQL, often done in combination with machine learning tools)

Answer: Assuming we have created clusters using an external tool and saved the cluster labels.

Query:

```
SELECT cluster_label, AVG(depression_index) AS Average_Depression_Index,
AVG(anxiety_level) AS Average_Anxiety_Level

FROM dataset

GROUP BY cluster_label;
```

Description: This query groups students by their cluster label and calculates the average depression index and anxiety level for each cluster.

25.Identify the most significant predictors of depression index using regression analysis (usually done in statistical software)

Answer: This query prepares data for regression analysis.

Query:

```
SELECT guilt_feelings, mood_outlook, self_worth_confidence, sleep_difficulties,
depression_index

FROM dataset;
```

Description: This query selects relevant columns to be used in a regression analysis to identify predictors of depression index.

# Additional Queries

26.What is the average screen time by gender?

Answer: This query calculates the average screen time for each gender.

Query:

```
SELECT Gender, AVG(screen_time) AS Average_Screen_Time
FROM dataset
GROUP BY Gender;
```

Description: This query groups students by gender and calculates the average screen time for each gender.

27.How many students report high confidence levels?

Answer: This query counts the number of students with high confidence levels.

Query:

```
SELECT COUNT(*) AS High_Confidence_Count
FROM dataset
WHERE self_worth_confidence = 'High';
```

Description: This query counts the number of students who have reported high confidence levels.

28.What is the average age of students experiencing severe sleep difficulties?

Answer: This query calculates the average age of students with severe sleep difficulties.

Query:

```
SELECT AVG(Age) AS Average_Age
FROM dataset
WHERE sleep_difficulties = 'Severe';
```

Description: This query calculates the average age of students who have reported severe sleep difficulties.

29.What is the percentage of students with different placement statuses?

Answer: This query shows the percentage of students for each placement status.

Query:

```
SELECT placement_status, (COUNT(*) * 100.0 / (SELECT COUNT(*) FROM dataset)) AS
Percentage
FROM dataset
GROUP BY placement_status;
```

Description: This query calculates the percentage of students in each placement status category.

30.Is there a correlation between screen time and sleep difficulties?

Answer: This query calculates the average sleep difficulties for each screen time category.

Query:

```
SELECT screen_time, AVG(sleep_difficulties) AS Average_Sleep_Difficulties
FROM dataset
GROUP BY screen_time;
```

Description: This query groups students by their screen time and calculates the average sleep difficulties for each screen time category.

31.What is the gender distribution among different CGPA ranges?

Answer: This query shows the count of students for each CGPA range, grouped by gender.

Query:

```
SELECT Gender, cgpa, COUNT(*) AS Count
FROM dataset
GROUP BY Gender, cgpa;
```

Description: This query groups students by their gender and CGPA range, and counts the number of students in each group.

32.What is the impact of financial situation on anxiety levels?

Answer: This query calculates the average anxiety level for each financial situation category.

Query:

```
SELECT financial_situation, AVG(anxiety_level) AS Average_Anxiety_Level
FROM dataset
GROUP BY financial_situation;
```

Description: This query groups students by their financial situation and calculates the average anxiety level for each group.

33.Is there a relationship between age and depression index?

Answer: This query calculates the average depression index for each age group.

Query:

```
SELECT Age, AVG(depression_index) AS Average_Depression_Index
FROM dataset
GROUP BY Age;
```

Description: This query groups students by their age and calculates the average depression index for each age group.

34.How many students report frequent feelings of guilt by age group?

Answer: This query shows the count of students with frequent feelings of guilt, grouped by age.

Query:

```
SELECT Age, COUNT(*) AS Count
FROM dataset
WHERE guilt_feelings = 'Frequent'
GROUP BY Age;
```

Description: This query groups students by their age and counts the number of students who report frequent feelings of guilt in each age group.

35.What is the distribution of anxiety levels by gender?

Answer: This query shows the count of students for each anxiety level, grouped by gender.

Query:

```
SELECT Gender, anxiety_level, COUNT(*) AS Count
FROM dataset
GROUP BY Gender, anxiety_level;
```

Description: This query groups students by their gender and anxiety level, and counts the number of students in each group.

36.What is the average CGPA by gender?

Answer: This query calculates the average CGPA for each gender.

Query:

```
SELECT Gender, AVG(cgpa) AS Average_CGPA
FROM dataset
GROUP BY Gender;
```

Description: This query groups students by their gender and calculates the average CGPA for each gender.

37.What is the impact of social connection on depression index?

Answer: This query calculates the average depression index for each level of social connection.

Query:

```
SELECT social_connection, AVG(depression_index) AS Average_Depression_Index
FROM dataset
GROUP BY social_connection;
```

Description: This query groups students by their level of social connection and calculates the average depression index for each group.

38.What is the relationship between financial situation and mood/outlook on life?

Answer: This query calculates the average mood/outlook rating for each financial situation category.

Query:

```
SELECT financial_situation, AVG(mood_outlook) AS Average_Mood_Outlook
FROM dataset
GROUP BY financial_situation;
```

Description: This query groups students by their financial situation and calculates the average mood/outlook rating for each group.

39.How many students report low self-worth and confidence levels by age?

Answer: This query shows the count of students with low self-worth and confidence levels, grouped by age.

Query:

```
SELECT Age, COUNT(*) AS Count
FROM dataset
WHERE self_worth_confidence = 'Low'
GROUP BY Age;
```

Description: This query groups students by their age and counts the number of students who report low self-worth and confidence levels in each age group.

40.What is the impact of daily screen time on anxiety levels?

Answer: This query calculates the average anxiety level for each screen time category.

Query:

```
SELECT screen_time, AVG(anxiety_level) AS Average_Anxiety_Level
FROM dataset
GROUP BY screen_time;
```

Description: This query groups students by their screen time and calculates the average anxiety level for each screen time category.