Below are detailed **SQL question ideas**, categorized for your **portfolio project**, each aimed to showcase a specific **data analysis concept** (aggregations, joins, conditions, window functions, etc.). You can implement these with SQLite, PostgreSQL, MySQL, or any engine of your choice and document the queries and findings on your GitHub.

1. Descriptive Statistics

Q1. What is the average daily social media usage by academic level?

Q2. What is the average mental health score by gender?

Q3. Count how many students use each social media platform most frequently.

Q4. What's the average addiction score for students who report that social media affects their academic performance?

2. Behavioral Insights

Q5. What is the correlation between average daily usage and mental health score? (Use Python visualization for this, SQL for data prep)

Q6. What is the average sleep hours for students addicted (score > 7) to social media vs those not addicted?

Q7. How does relationship status influence conflicts over social media?

3. Demographic Breakdown

Q8. Which country has the highest average addiction score?

Q9. Show the average mental health score by country and gender.

Q10. Count of students from each academic level by country.

4. Impact Analysis

Q11. Compare average academic-affecting users vs non-affecting in terms of:

- Sleep hours
- Usage hours
- Addiction score

Q12. For each academic level, find the percentage of students who said social media affects their academics.

5. Advanced SQL (Window Functions / Ranking)

Q13. Rank students within their academic level by addiction score.

Q14. For each country, find the student with the highest daily usage.

Q15. Use a window function to calculate the average addiction score within each academic level.

6. Custom Segmentation

Q16. Categorize students into "Low", "Moderate", and "High" social media users based on Avg_Daily_Usage_Hours:

- Low: < 3 hrs
- Moderate: 3–6 hrs
- High: > 6 hrs

  Then show how these categories correlate with mental health scores.

7. Hypothetical Dashboards

You can simulate or actually build a dashboard in **Looker Studio, Tableau, or Power BI** using SQL as backend.

Metrics to visualize:

- Usage hours by platform
- Sleep vs addiction score scatter plot
- Platform usage by gender
- Academic performance vs addiction levels

Next Steps for GitHub Documentation

1. Create a `README.md` with:

   a. Project title

   b. Overview and dataset description

   c. Tools used

   d. Objective

2. Create folders for:

   a. `/sql_queries` — with `.sql` files or `.md` documentation

   b. `/visualizations` — graphs or dashboards

   c. `/insights_summary` — your interpretations and key findings

3. Optional:

   a. Include a Jupyter Notebook for data exploration with SQL + pandas.

   b. Build a `.csv` cleaner and uploader with Python.

Would you like me to generate the actual SQL code for any of these questions or help you set up the GitHub documentation structure?