Part 1 - Basic Understanding

• Question: What information is contained in the dataset? List the variables and describe what each one represents.

Variable Name		Description
Student_ID	-	Unique ID of each student
Name	-	Full name of the student
Gender	-	Gender of the student (Male/Female)
State	-	Home state of the student
Major	-	Student's field of study
GPA	-	Grade Point Average (Academic performance indicator)
Credits_Completed	-	Total number of academic credits completed
Scholarship	-	Whether the student has a scholarship (Yes/No)
Part_Time_Job	-	Whether the student has a part-time job (Yes/No)

Explanation: The dataset provides personal and academic details of students, including their gender, state, academic performance (GPA), field of study, and financial/work background. These variables help in analyzing student trends, academic progress, and demographic distribution. Whether the student has a part-time job (Yes/No) |

Descriptive Statistics:

- **Question: Calculate the average GPA and the average number of credits completed. What do these averages tell us about the student population?**

  ### Answer:

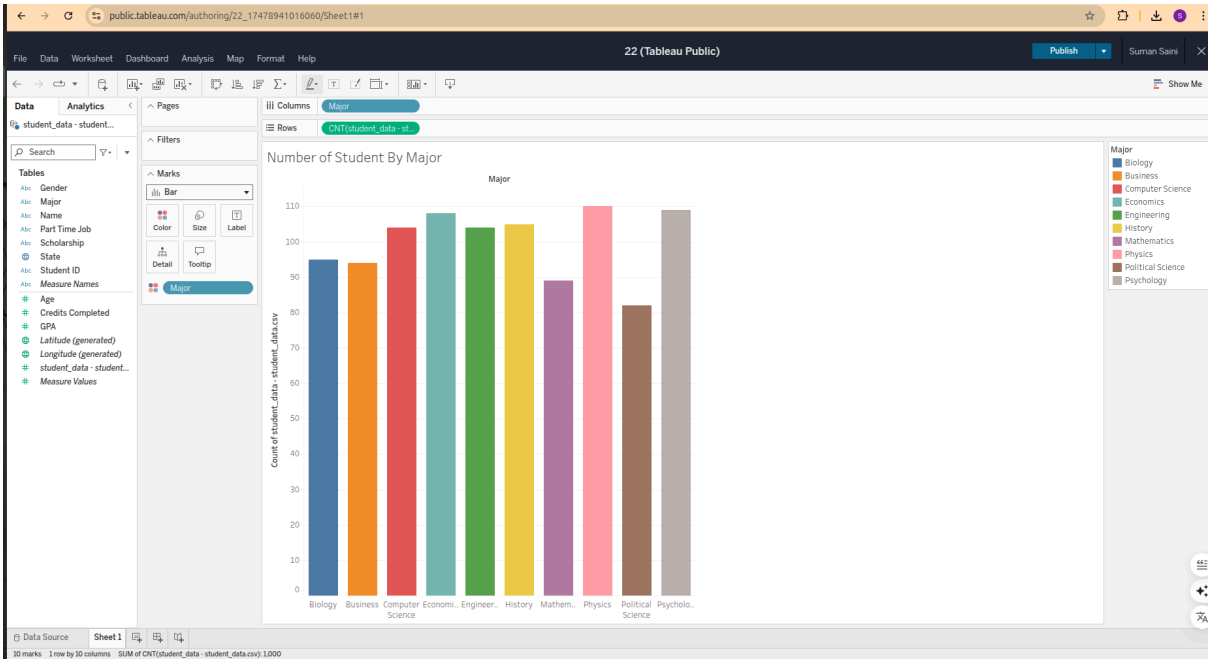
- **Average GPA**: 3.01
- **Average Credits Completed**: 26.6

These averages suggest that most students are in the early stages of their studies and are maintaining a B-grade average.

Data Distribution:

● Question: Create a bar chart showing the number of students in each major. What can you infer from the chart?

Answer:

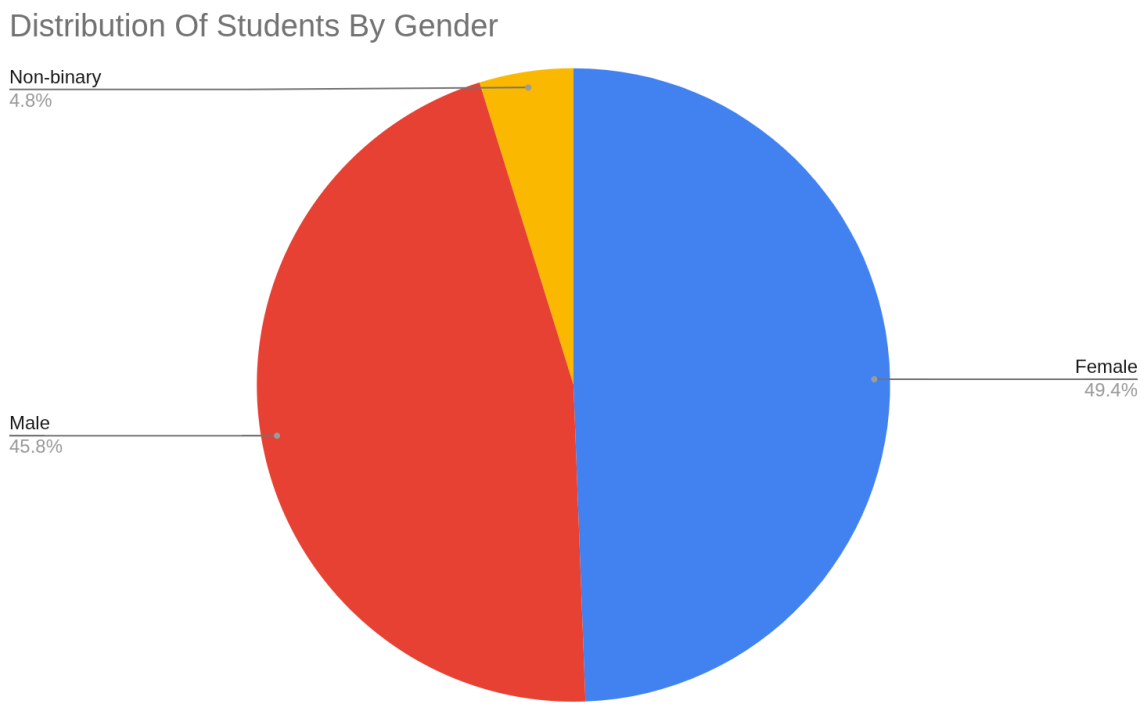


Observation: This bar chart helps visualize how many students are enrolled in each major. The taller the bar, the more students are pursuing that field. According to the chart, Mathematics has the highest number of students (110), followed by Psychology (109) and Economics (108). The least number of students are in Political Science (82). This shows that some majors are more popular or have more enrollment than others.

Gender Analysis:

- **Question: Create a pie chart showing the distribution of male and female students. What does this tell you about the gender balance in the dataset?**

Answer:

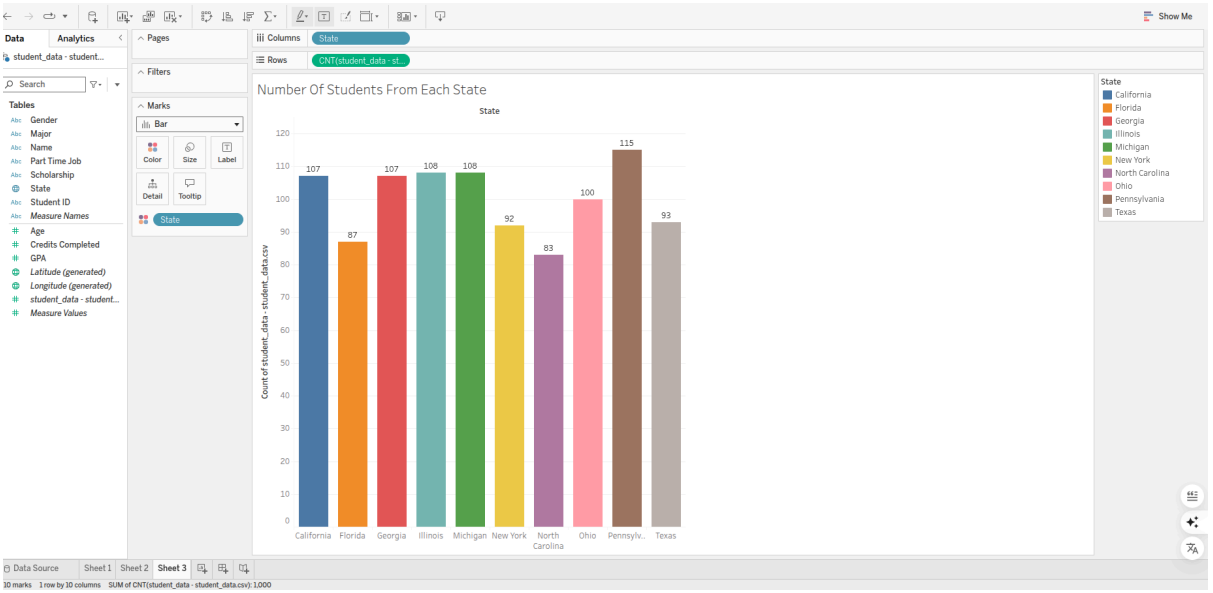


Observation: This pie chart shows the gender distribution of students using the count of Student ID. The slices represent the number of Female, Male, and Non-binary students. Female students are slightly more in number (49.4%), followed by Male (45.8%), and Non-binary (4.8%). This chart gives a clear visual representation of gender diversity among students.

State Distribution:

- **Question: Create a bar chart showing the number of students from each state. How does the state distribution look?**

Answer:



Observation: This bar chart displays the number of students from each state. The states with the highest number of students are Pennsylvania (115), Illinois (108), and Michigan (108), indicating a strong student representation from these regions. California and Georgia also have 107 students each, showing similar enrollment levels.

On the other hand, North Carolina has the lowest number of students (83), followed by Florida (87). These states have a relatively smaller representation in the dataset.

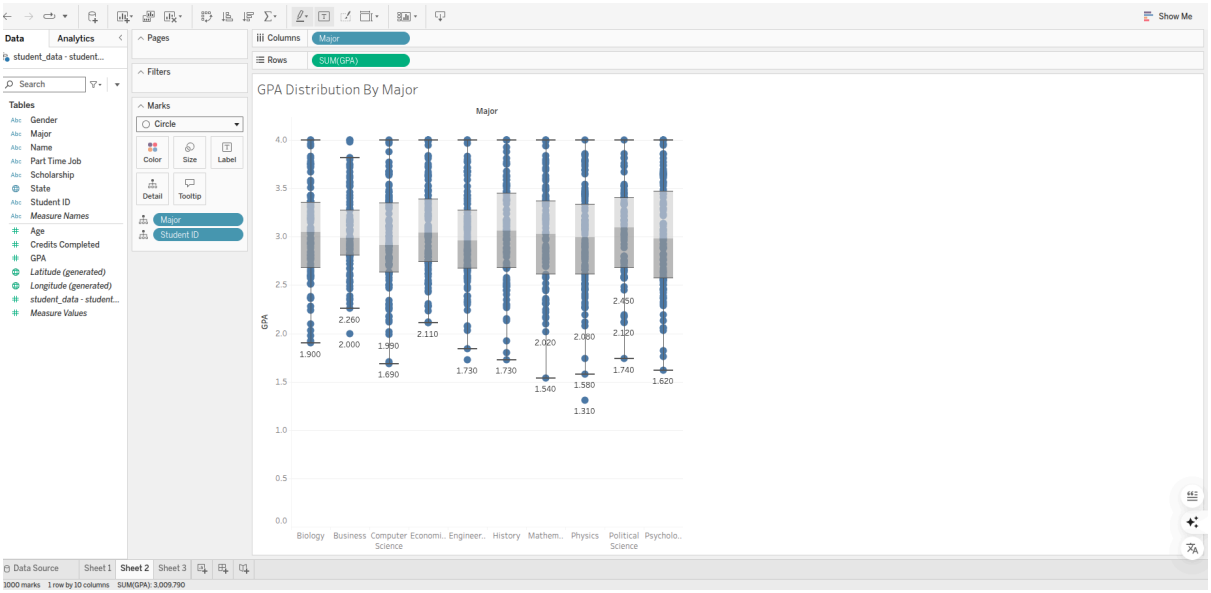
This chart helps visualize geographic diversity and can assist in identifying which states contribute most to the student population.

Part 2 - More Complex Visualizations

GPA and Major Analysis:

- **Question: Create a box plot to show the distribution of GPA for each major. What insights can you gain about the GPA distribution across different majors?**

Answer:



Observation: **T**his box plot displays the GPA distribution across different majors, highlighting how academic performance varies by field of study. Each box represents the middle 50% of GPA values (interquartile range), with the line inside showing the median GPA.

Key insights:

- Biology, Business, Computer Science, and Engineering have relatively higher median GPAs, indicating stronger academic performance in these fields.
- Political Science and Physics show wider boxes and more scattered points, suggesting greater variation in student performance and potential outliers.
- Some majors like History and Mathematics have tightly packed boxes, indicating more consistent performance among students.

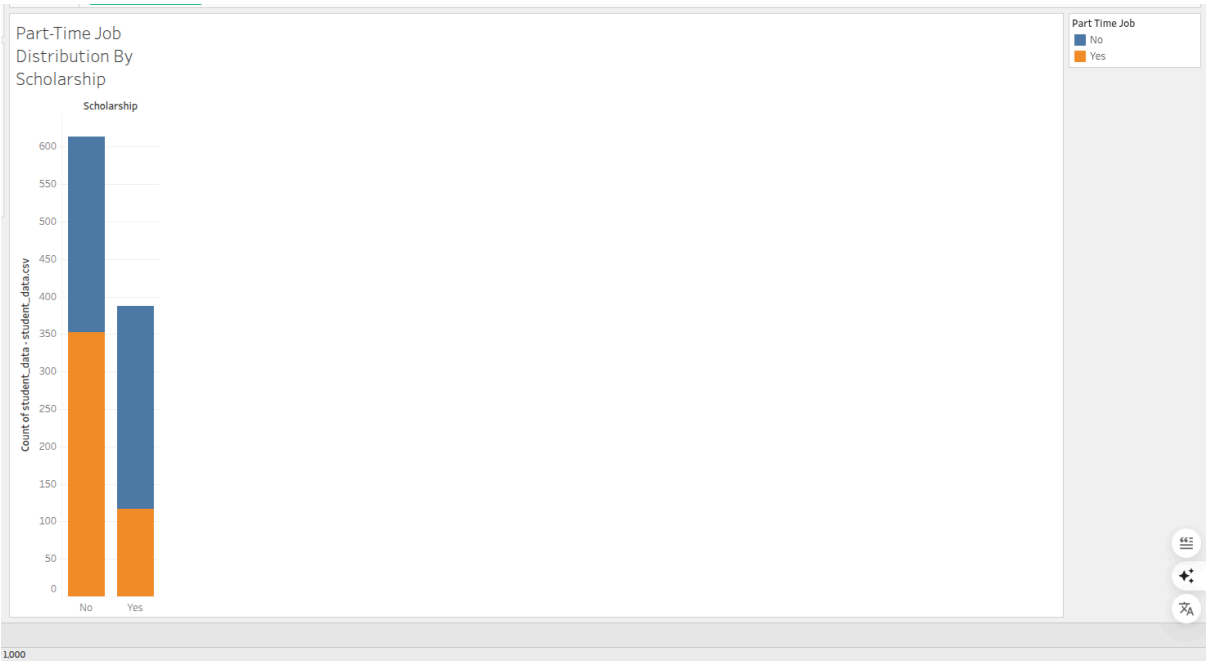
● The presence of outlier dots outside the whiskers shows individual students performing significantly above or below the typical range for their major.

This visualization is useful for comparing academic consistency and identifying majors where students might need more academic support.

Scholarship and Part-Time Job Analysis:

● **Question: Create a stacked bar chart showing the relationship between having a scholarship and having a part-time job. What patterns do you observe?**

Answer:



Observation: This stacked bar chart illustrates the relationship between scholarship status and part-time job engagement among students.

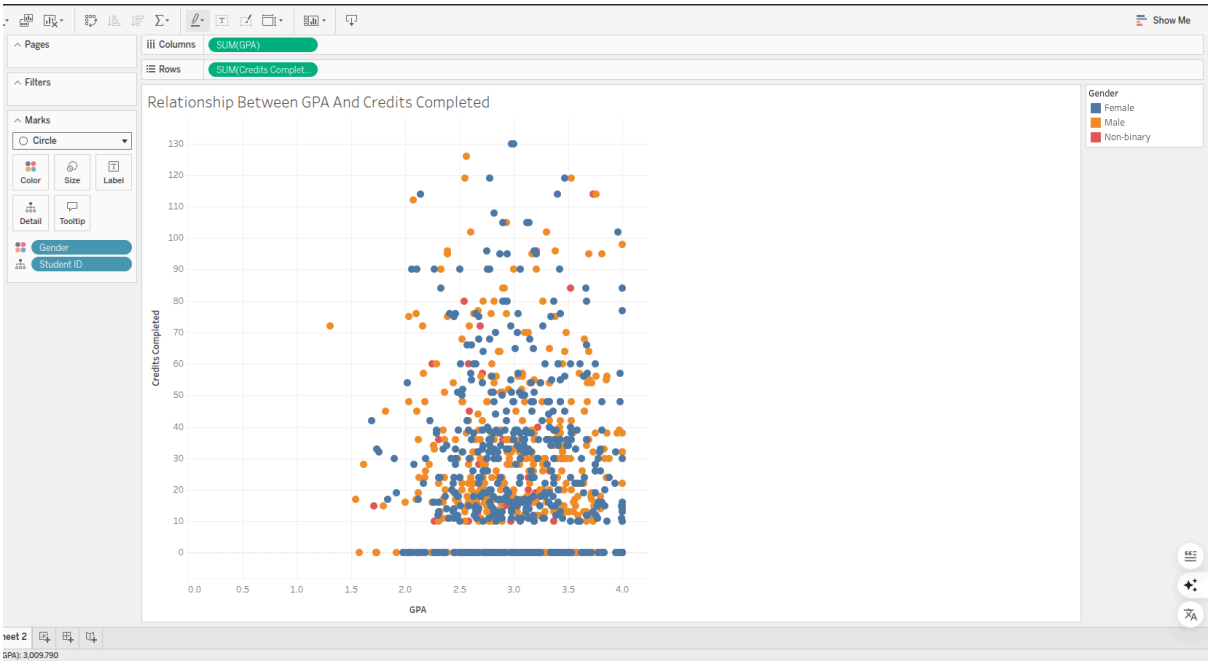
● Among students without scholarships:
  ○ A large number of students have part-time jobs, likely due to financial need.
  ○ The “Yes” (part-time job) segment is taller here.
● Among students with scholarships:
  ○ A smaller proportion have part-time jobs.

○ Many students do not work part-time, possibly because financial aid reduces the need.

GPA and Credits relationship:

● Question: Create a scatter plot to examine the relationship between GPA and credits completed. What does the plot tell you about the relationship between the two variables?

Answer:



Observation: This scatter plot examines how students' academic performance (GPA) relates to the number of academic credits they have completed. Each dot represents a student, color-coded by gender.

● Spread of Data: The points are widely spread across both axes, indicating no strong linear relationship between GPA and credits completed.
● Consistent GPA Across Credits: Students with both high and low GPAs are present across various credit levels — whether they've completed few credits or many.

- Cluster Around Mid-GPA: A large cluster of students appears around a GPA of 2.5 to 3.5, suggesting that most students fall within an average performance range.
- Zero Credit Students: There are a number of students who have a GPA but show zero credits completed, which could suggest recent enrollees, transfer credit issues, or data anomalies.
- Gender Insight: No noticeable trend differences are visible between Male, Female, and Non-binary students in terms of GPA vs credits