Name: Aishwarya Wankhade
Data Analysis Domain Ip: 6202

Below are the steps followed to prepare the dataset for analysis. The responses to the questions are based on the provided dataset and the reference to the Questionnaire in Problem Statement.

1. The dataset was loaded into Power BI and transformed for analysis.
2. Values in the **'First Name'** and **'College Name'** columns were capitalized to maintain consistency.
3. Null and blank entries were removed from the **'College Name'** column.
4. Duplicate records were eliminated based on a combination of **'Email'** and **'College Name'**.
5. The column **'How did you come to know about this event?'** was split using the leftmost delimiter **'|'**.
6. The value **'Others'** in the **'How did you come to know about this event?'** column was replaced with **'College'**.
7. A new column, **'Family Income in LPA'**, was created based on the **'Family Income'** column, where:
8. '2' represents to '0–2 Lakhs'
9. '5' represents to '2–5 Lakhs'
10. '7' represents to '5–7 Lakhs'
11. '9' represents to '7 Lakhs +'
12. A new column, **'13Q CGPA Range'**, was created from the **'CGPA'** column with ranges **'6–8.5'** and **'8.6–10'**.
13. Created a new column '17Q CGPA Range' using the 'CGPA' column, with values '6–6.9', '7–7.9', '8–8.9', and '9–9.9'.
14. Leading and trailing white spaces were trimmed from the **'Leadership Skills'** column.

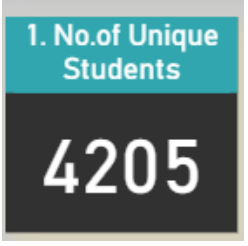
Basic Section

Successfully answered 8 of the 9 questions in accordance with the provided options

1. How many unique students are included in the dataset?

- The **'Email ID'** field was added to the Fields pane of a Card visualization and aggregated using the **Count** function. The **'Designation'** field was then applied as a visual-level filter, selecting categories such as Students, B.Sc IT, B.Tech, BE, BE IT, BE CS, MBA Student, M.Tech, MCA, PGDM Marketing Student, and Diploma in Electrical Engineering. This filtering and aggregation process identified **4,205 unique student records** in the dataset.

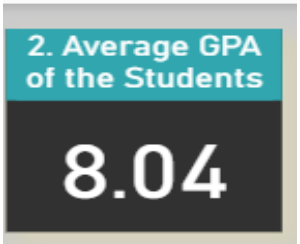
Screenshot of Visuals:



2. What is the average GPA of the students?

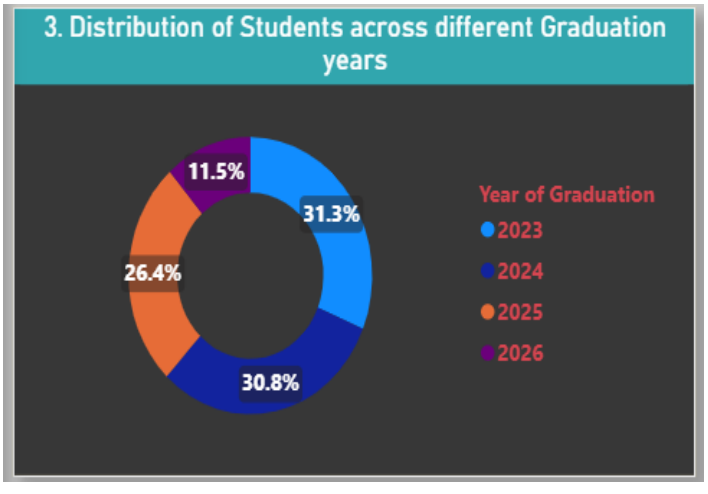
- The 'CGPA' field was added to the Fields pane of a Card visualization and aggregated using the Average function. The 'Designation' field was then applied as a visual-level filter, including categories such as Students, B.Sc IT, B.Tech, BE, BE IT, BE CS, MBA Student, M.Tech, MCA, PGDM Marketing Student, and Diploma in Electrical Engineering. This analysis resulted in an average CGPA of 8.04 for the selected student group**.**

Screenshot of Visuals:



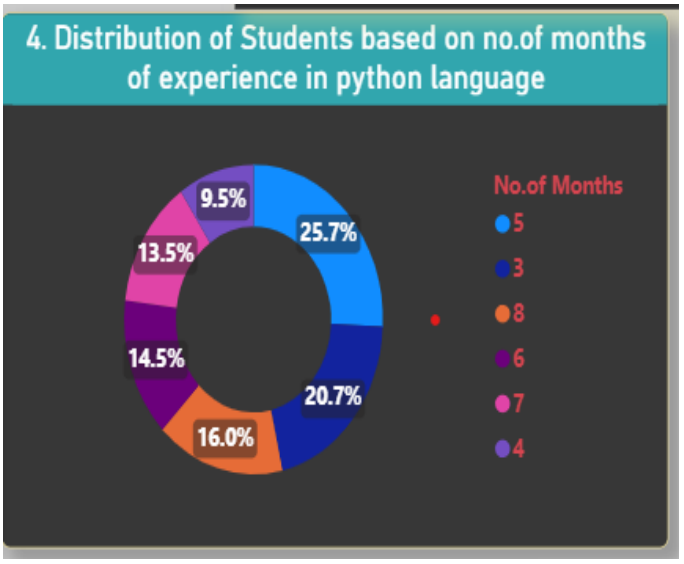
3. What is the distribution of students across different graduation years?

- The **'Year of Graduation'** field was added to the **Legend** and the **'Email ID'** field to the **Values** section of a Donut visualization, with the **Count** function applied to Email ID. The **'Designation'** field was then used as a visual-level filter, including categories such as Students, B.Sc IT, B.Tech, BE, BE IT, BE CS, MBA Student, M.Tech, MCA, PGDM Marketing Student, and Diploma in Electrical Engineering. The analysis showed that **31.3%** of students graduated in **2023**, **30.8%** in **2024**, **26.4%** in **2025**, and **11.5%** in **2026**.



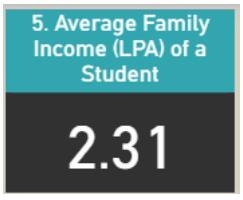
4. What is the distribution of student's experience with Python programming?

- The 'Experience with Python (Months)' field was added to the Legend and the 'Email ID' field to the Values section of a Donut chart, with the Count function applied to Email ID. The 'Designation' field was then used as a visual-level filter, including categories such as Students, B.Sc IT, B.Tech, BE, BE IT, BE CS, MBA Student, M.Tech, MCA, PGDM Marketing Student, and Diploma in Electrical Engineering. The analysis indicated that 20.7% of students had 3 months of Python experience, 9.5% had 4 months, 25.7% had 5 months, 14.5% had 6 months, 13.5% had 7 months, and 16% had 8 months of experience.



5. What is the Average Family Income of a Student?

- The **'Family Income in LPA'** field was added to the Fields pane of a Card visualization and aggregated using the **Average** function. The **'Designation'** field was then applied as a visual-level filter, including categories such as Students, B.Sc IT, B.Tech, BE, BE IT, BE CS, MBA Student, M.Tech, MCA, PGDM Marketing Student, and Diploma in Electrical Engineering. This analysis revealed that the **average family income of a student is 2.31 LPA**.



6.How does the GPA vary among different colleges? (Show top 5 results only)

- The **'College Name'** and **'CGPA'** fields were added to the **Columns** section of a Table visualization, with the **Average** function applied to CGPA. The table was then sorted in descending order based on the average CGPA. The results indicated minimal variation among the average GPAs of the top five colleges.

6. Top 5 Colleges by Average CGPA

College Name	Average of CGPA
Vidyalankar Institute Of Technology, Mumbai	8.13
Thakur Institute Of Management Studies, Career Development & Research - [Timscdr]	8.18
St Xavier'S College	8.18
B. K. Birla College Of Arts, Science & Commerce (Autonomous), Kalyan	8.15
Ap Shah Institute Of Technology	8.16

8. What is the average GPA for students from each city?

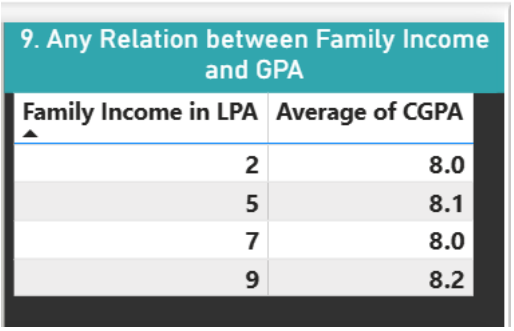
- The **'City'** and **'CGPA'** fields were added to the **Columns** section of a Table visualization, with the **Average** function applied to CGPA. The **'Designation'** field was then used as a visual-level filter, including categories such as Students, B.Sc IT, B.Tech, BE, BE IT, BE CS, MBA Student, M.Tech, MCA, PGDM Marketing Student, and Diploma in Electrical Engineering, and the results were sorted in descending order based on average CGPA. The analysis showed that the average GPAs of students across all cities ranged between **7.3 and 8.6**.

8. Average GPA of Students by each City

City	Average of CGPA
Kolhapur	8.60
Puri	8.55
Gurugram	8.53
Nizambad	8.49
Siwan	8.48
Sonipat	8.46
Haijipur	8.40
Srinagar	8.39
Hasan	8.38
Narwar	8.38
Wardha	8.37
Jammu	8.36
Una	8.34
Tirupati	8.34
Varanasi	8.34
Buldhana	8.33
Ahemdabad	8.33
Jaipur	8.32
Pune	8.32
Jhalwar	8.31
Aurangabad	8.29
Malda	8.29
Darbhanga	8.28
Sirsa	8.28
Gonda	8.27
Panji	8.27

9. Can we identify any relationship between Family Income and GPA?

- The 'Family Income in LPA' and 'CGPA' fields were added to the Columns section of a Table visualization, with the Average function applied to CGPA. The analysis indicated that the average GPA across all family income categories showed only minimal variation.



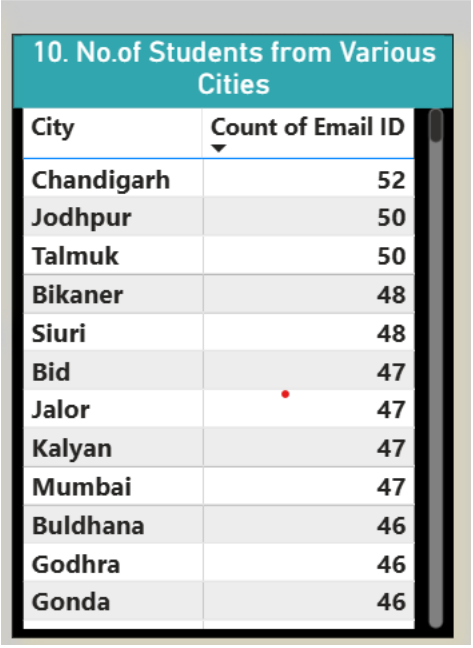
9. Any Relation between Family Income and GPA

Family Income in LPA	Average of CGPA
2	8.0
5	8.1
7	8.0
9	8.2

Moderate Section

10. How many students from various cities?

- The 'City' and 'Email ID' fields were added to the Columns section of a Table visualization, with the Count function applied to Email ID. The 'Designation' field was then used as a visual-level filter, including categories such as Students, B.Sc IT, B.Tech, BE, BE IT, BE CS, MBA Student, MTech, MCA, PGDM Marketing Student, and Diploma in Electrical Engineering. The analysis indicated that the number of students from each city ranged between 5 and 52.

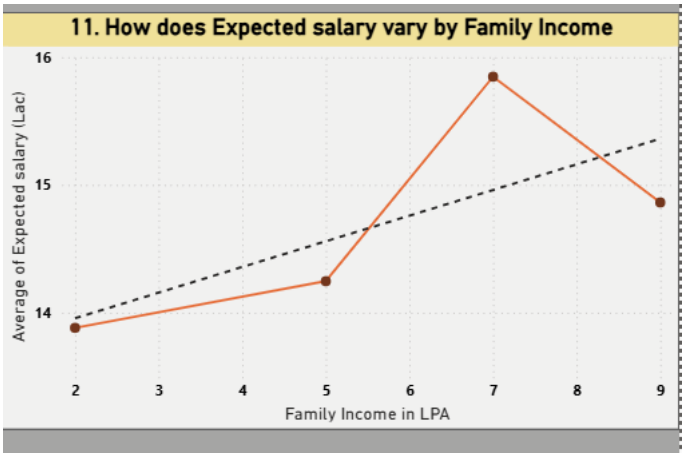


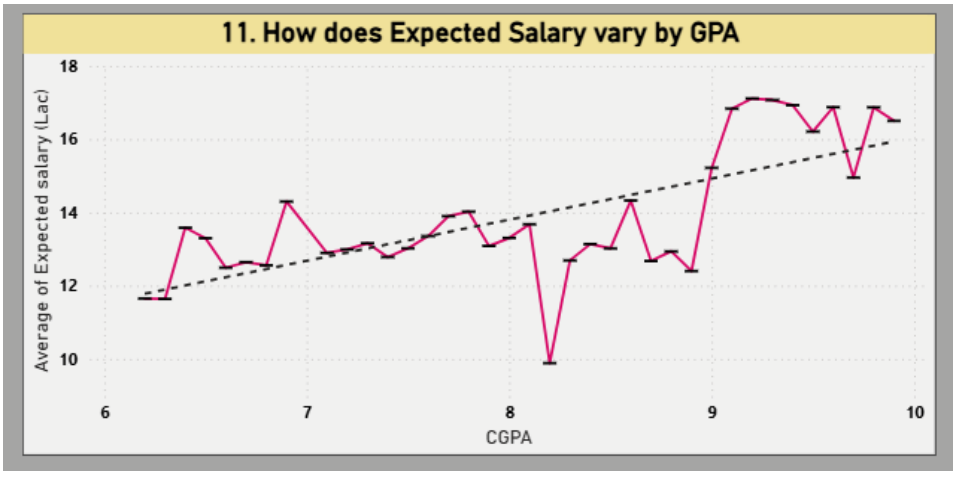
10. No.of Students from Various Cities

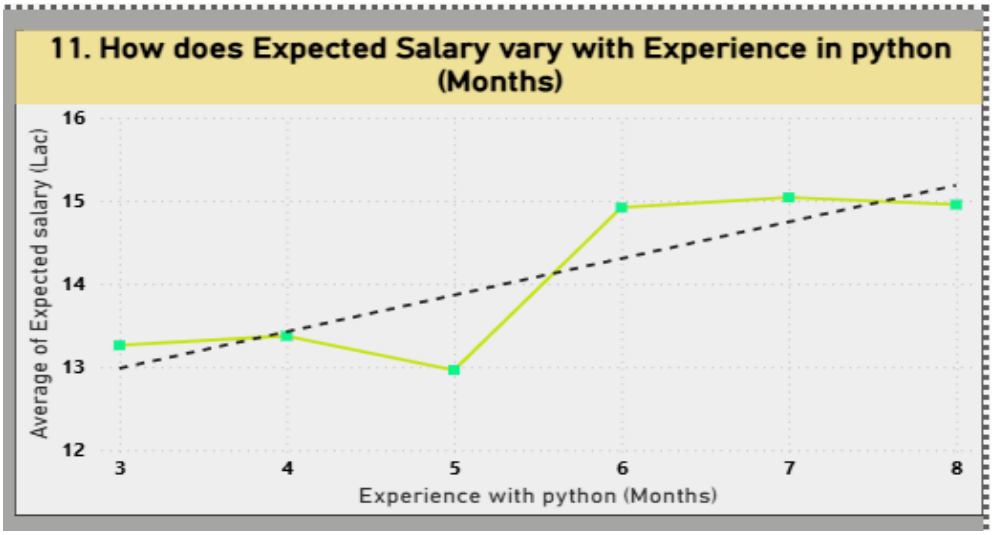
City	Count of Email ID
Chandigarh	52
Jodhpur	50
Talmuk	50
Bikaner	48
Siuri	48
Bid	47
Jalor	47
Kalyan	47
Mumbai	47
Buldhana	46
Godhra	46
Gonda	46

11. How does the expected salary vary based on factors like 'GPA', 'Family income', 'Experience with python (Months)'?

- All three visualizations corresponding to 11Q in Power BI display an upward trend in the trendlines. This indicates that, on average, expected salary increases with higher CGPA, family income, and experience with Python (months). The results suggest that stronger academic performance and greater technical experience are associated with higher salary expectations, while family income also shows an influence, peaking at mid-income levels before a slight decline.



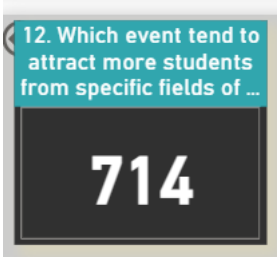




12. Which event tends to attract more students from specific fields of study?

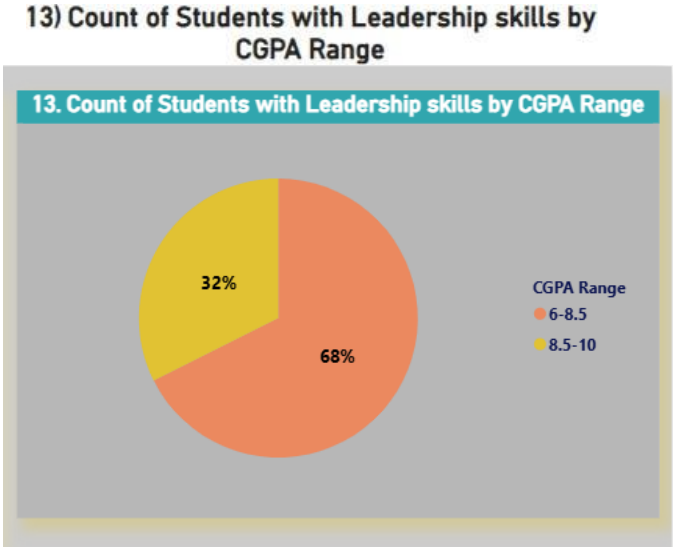
- The **'Events'** field was added to the **Small Multiples** section and the **'Email ID'** field to the **Data** section of a **Card (New)** visualization, with the **Count** function applied to Email ID. The **'Designation'** and **'Events'** fields were then applied as visual-level filters, including categories such as Students, B.Sc IT, B.Tech, BE, BE IT, BE CS, MBA Student, MTech, MCA, PGDM Marketing Student, and Diploma in Electrical Engineering. A **Top N** filter was applied on the **'Events'** field to highlight the most participated events.

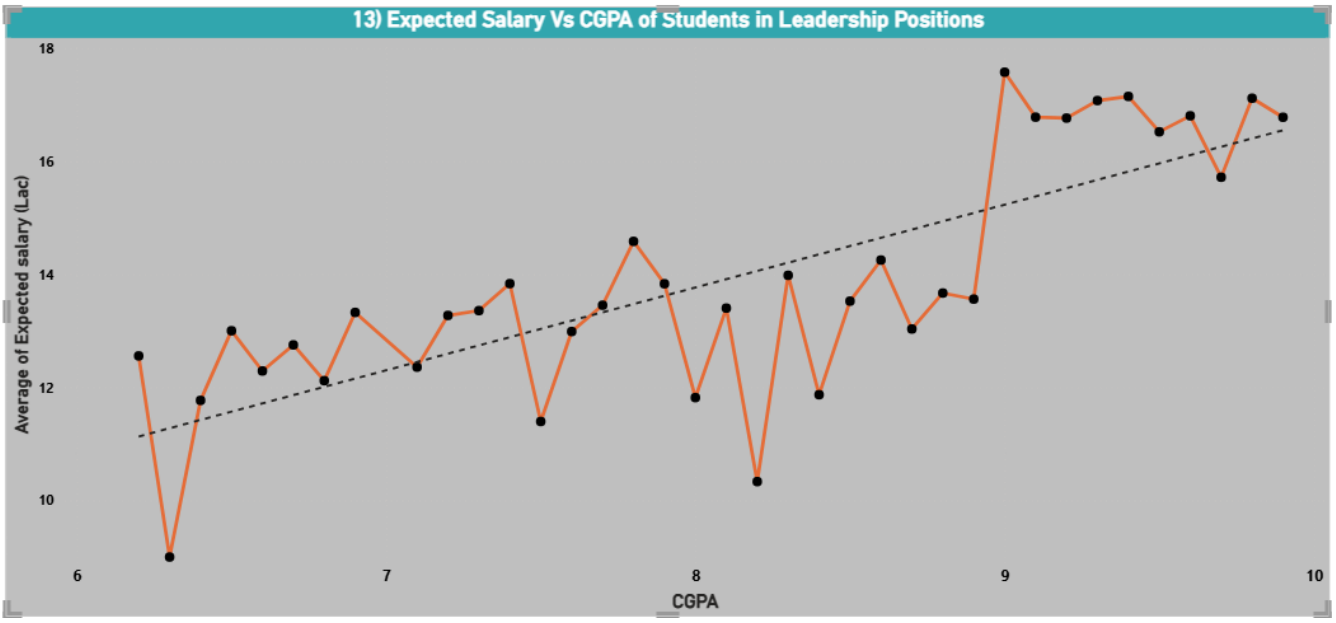
  Marketing Student', and 'Diploma in Electrical Engineering' and applied a Top N filter on the 'Events' field to display only the top 1 event based on the highest number of Email IDs. **This process revealed that the 'Product Design and Full Stack' event had the highest count of 714, indicating it attracted more students from specific fields of study.**



13. Do students in leadership positions during their college years tend to have higher GPAs or better expected salary?

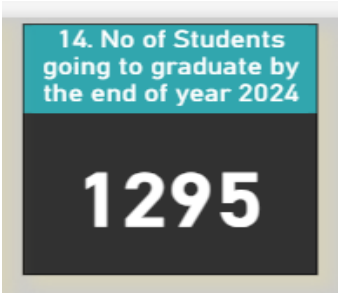
- Two visualizations were used to address 13Q, namely a pie chart and a scatter plot. The pie chart indicates that a larger proportion of students (68%) possess leadership skills with lower GPAs, compared to 32% of students who have leadership skills with higher GPAs. The scatter plot further shows that students with leadership skills and GPAs between 6 and 8.5 expect salaries ranging from 9 to 15 LPA, while those with leadership skills and GPAs between 8.5 and 10 expect salaries between 15 and 20 LPA. This suggests that students with leadership skills tend to expect higher salaries, even without maintaining the highest GPAs.





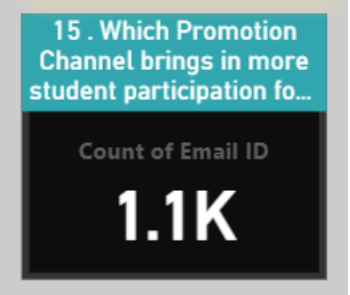
14. How many Students are graduating by the end of 2024?

- The **'Email ID'** field was added to the Fields pane of a Card visualization and aggregated using the **Count** function. The **'Designation'** and **'Year of Graduation'** fields were then applied as visual-level filters, including categories such as Students, B.Sc IT, B.Tech, BE, BE IT, BE CS, MBA Student, MTech, MCA, PGDM Marketing Student, and Diploma in Electrical Engineering, with the graduation year restricted to **2024**. This analysis revealed that **1,295 students** are graduating by the end of **2024**.



15. Which Promotion Channel brings in more student participation for the Event?

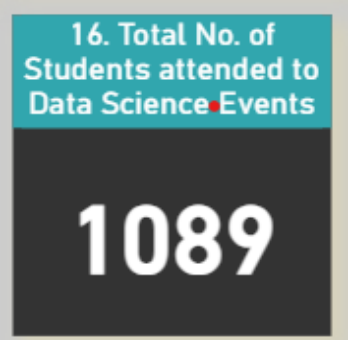
- The 'How did you come to know about this event?' field was added to the Small Multiples section and the 'Email ID' field to the Data section of a Card (New) visualization, with the Count function applied to Email ID. The 'Designation' and 'How did you come to know about this event?' fields were then applied as visual-level filters, including categories such as Students, B.Sc IT, B.Tech, BE, BE IT, BE CS, MBA Student, MTech, MCA, PGDM Marketing Student, and Diploma in Electrical Engineering, while excluding blank values. A Top N filter was applied to display only the top promotion channel based on the highest Email ID count. This analysis revealed that WhatsApp, with a count of 1,100 (1.1K), was the most effective channel for attracting student participation.



Whatsapp

16. Find the total number of Students who attended the events related to Data Science.

- The 'Email ID' field was added to the Fields pane of a Card visualization and aggregated using the Count function. The 'Designation', 'Attendee Status', and 'Events' fields were then applied as visual-level filters, including categories such as Students, B.Sc IT, B.Tech, BE, BE IT, BE CS, MBA Student, MTech, MCA, PGDM Marketing Student, and Diploma in Electrical Engineering in the Designation field, the 'Attending' category in the Attendee Status field, and only Data Science–related events in the Events field. This analysis revealed that 1,089 students attended only Data Science–related events.



17. Those who have high CGPA and more experience in language, those who have high expectations for salary? (Avg)

- The '17Q CGPA Range' and 'Experience with Python (Months)' fields were added to the X-axis of a scatter plot, while the 'Expected Salary (Lac)' field was placed on the Y-axis with the Average function applied. The visualization indicates that below a CGPA of 9, Python experience does not significantly increase salary expectations. However, students within the 9–10 CGPA range who have 6–8 months of Python experience tend to expect notably higher salaries.

Visual Screenshot:

17. Those who have high CGPA and more experience in language those who have high expectations for salary ? (Avg)