Data Science & Analytics - Internship

Intern's Name	Email ID
Sumaiya Tasnim	sumaiyaa.tasnim.18@gmail.com

Project Title: Student Event Feedback Analysis & Dashboard

Introduction:

Student feedback plays a crucial role in understanding how effectively campus events meet participants' expectations. Colleges conduct several events — from workshops and tech fests to cultural programs — and often collect feedback through Google Forms. However, these valuable insights remain underutilized.This project applies data science and sentiment analysis techniques to uncover meaningful patterns from student feedback. By analyzing both numeric ratings and textual comments, it aims to evaluate satisfaction levels, identify strengths and weaknesses, and suggest actionable improvements for future events.

Objectives:

The main objectives of the project are:

- To clean and preprocess raw feedback data collected through Google Forms.
- To analyze student ratings (1–5 or 1–10 scale) and compute overall satisfaction levels.
- To perform sentiment analysis on open-text feedback using VADER and TextBlob.
- To visualize satisfaction and sentiment trends using various charts and dashboards.
- To generate data-driven insights and recommendations that can improve event planning and delivery.

Assigned Dataset: *(Chosen)*

Raw Dataset	Cleaned Dataset Exported as
"student_feedback.csv"	"Student_Event_Feedback_Analysis.csv"

Tools Used:

Tool	Purpose
Jupyter Notebook (Python)	Code execution and data analysis
Python (pandas, numpy)	Data cleaning and manipulation
Matplotlib / Seaborn	Data visualization
TextBlob & VADER (NLP)	Sentiment analysis (textual feedback)
Power BI	Interactive dashboard creation
Excel	Data review and export validation

Expected Outcome:

- ✓ A clean and well-structured dataset ready for analysis.
- ✓ Identification of key satisfaction trends across different course aspects.
- ✓ Clear visualization of sentiment distribution and satisfaction levels.
- ✓ Insights into the most and least appreciated event elements.
- ✓ A Power BI dashboard summarizing student sentiment and satisfaction metrics.
- ✓ Actionable recommendations for improving student engagement and event quality.

My Github Link: https://github.com/sumaiya-tasnim-18

(All my projects are available on my Github Account)

Data Cleaning & Transformation Summary:

Task	Description / What Was Checked	Action / Correction	Result
Load Dataset	Imported student_feedback.csv into pandas	Verified successful load using df.head() and df.info()	Dataset loaded correctly with all columns present
Column Verification	Checked all column names and structure	Ensured 8 rating questions + feedback + sentiment columns exist	Columns correctly structured and renamed consistently
Data Type Check	Verified data types of each column	Converted all rating columns to int64, satisfaction to float64 where needed	All numeric and text columns have correct data types
Missing Values	Checked for nulls or blanks in all columns	No missing values found; dataset was complete	Dataset confirmed clean with 0 null entries
Duplicate Records	Checked for duplicate Student ID or entire row entries	No duplicates found; dataset unique per student	Ensured one feedback per student
Standardization of Ratings	Reviewed each rating column for consistency (scale 1–10)	Confirmed all ratings fall within valid range	Ratings standardized and valid
Derived Column – Overall_Satisfaction	Created average of 8 rating columns per student	Applied mean across aspect columns	Added column Overall_Satisfaction
Derived Column – Satisfaction_Level	Classified satisfaction as Low / Medium / High	Applied conditional bins based on average	Added column Satisfaction_Level
Text Feedback Cleaning	Checked for blank or invalid	Retained meaningful feedback; dropped empty strings	Clean Feedback_Comm

Task	Description / What Was Checked	Action / Correction	Result
	comments		ent column ready for sentiment analysis
Sentiment Analysis (VADER)	Applied VADER analyzer on feedback comments	Added Sentiment_Score and Sentiment_Label columns	Captured polarity-based sentiment for each comment
Sentiment Analysis (TextBlob)	Applied TextBlob for polarity & subjectivity	Added TB_Polarity, TB_Subjectivity, and TB_Sentiment_Label	Dual sentiment model comparison prepared
Export for Power BI	Prepared final structured dataset	Saved as Student_Event_Feedback_Analysis.csv	Clean, analysis-ready dataset exported successfully

Columns & Correct Data Type:

Column Name	Data Type	Column Name	Data Type	Column Name	Data Type
Student ID	int64	Well versed with the subject	int64	Explains concepts in an understandable way	int64
Use of presentations	int64	Degree of difficulty of assignments	int64	Solves doubts willingly	int64
Structuring of the course	int64	Provides support for students going above and beyond	int64	Course recommendation based on relevance	int64
Overall_Satisfaction	float64	Satisfaction_Level	object	Feedback_Comment	object
Sentiment_Score	float64	Sentiment_Label	object	TB_Polarity	float64
TB_Subjectivity	float64	TB_Sentiment_Label	object	—	—

Data & Sentiment Analysis Part

Rating Analysis:

Aspect	Average Rating	Insight
Well versed with the subject	7.50	Students rated this highest → teachers show strong subject mastery.
Explains concepts in an understandable way	6.08	Fairly positive, but some students struggle with clarity.

Aspect	Average Rating	Insight
Use of presentations	5.94	Mid-level satisfaction → visual aids are moderately effective.
Degree of difficulty of assignments	5.43	Lowest rating → assignments might feel too hard or unclear.
Solves doubts willingly	5.47	Indicates mixed experiences with doubt resolution.
Course recommendation based on relevance	5.60	Suggests the perceived relevance of courses could improve.

Correlation Insights:

Feedback Aspect	Correlation with Overall Satisfaction	Interpretation
Degree of difficulty of assignments	0.39	Strongest relationship — students' perception of assignment fairness has a big effect on satisfaction.
Solves doubts willingly	0.39	Equally important — accessible and supportive teachers drive satisfaction.
Course recommendation based on relevance	0.39	Courses that feel relevant strongly increase overall approval.
Provides support for students going above and beyond	0.38	Students value extra effort from instructors.
Structuring of the course	0.38	Well-organized courses link closely with satisfaction.
Explains concepts understandably	0.37	Clear teaching matters a lot, even if not the very top factor.
Well versed with the subject	0.20	Subject knowledge alone doesn't ensure satisfaction — delivery and support matter more.
Use of presentations	0.18	Visual aids help slightly but aren't a major driver.

Sentiment Analysis (Vader) Summary:

Sentiment	Approx. Count	Insight
Negative	350–400	A large portion of students expressed dissatisfaction in comments — aligns with "Low" or "Medium" ratings.
Neutral	300–350	Many comments are mixed or moderate — students see both positive and negative aspects.

Sentiment	Approx. Count	Insight
Positive	250–300	Fewer highly positive comments — only a minority of students strongly praised the course.

Sentiment Analysis (TextBlob) summary:

Column	Description
TB_Polarity	Sentiment score (-1 negative → +1 positive)
TB_Subjectivity	Measures how opinionated a comment is (0 = factual, 1 = highly subjective)
TB_Sentiment_Label	Categorized sentiment: Positive, Neutral, Negative

Note: For more detailed analysis, check the "sentiment_analysis_student_feedback.ipynb" file which in my github.

Data Dictionary for student_feedback.csv dataset after processing:

Column Name	Data Type	Description
Student ID	int64	Unique identifier for each student providing feedback
Well versed with the subject	int64	Rating (1–10) of how knowledgeable the instructor is about the subject
Explains concepts in an understandable way	int64	Rating (1–10) of clarity in teaching
Use of presentations	int64	Rating (1–10) of the effectiveness of visual aids or presentations
Degree of difficulty of assignments	int64	Rating (1–10) of how difficult students found assignments
Solves doubts willingly	int64	Rating (1–10) of instructor's responsiveness to questions
Structuring of the course	int64	Rating (1–10) of how well the course is organized
Provides support for students going above and beyond	int64	Rating (1–10) of additional support for motivated students
Course recommendation based on relevance	int64	Rating (1–10) of course applicability and relevance
Overall_Satisfaction	float64	Average rating across all 8 numeric feedback aspects per student
Satisfaction_Level	object	Categorized overall satisfaction (Low / Medium / High)
Feedback_Comment	object	Free-text comment provided by student

Column Name	Data Type	Description
Sentiment_Score	float64	VADER compound sentiment score (-1 to 1)
Sentiment_Label	object	VADER sentiment category: Positive / Neutral / Negative
TB_Polarity	float64	TextBlob polarity score (-1 to 1)
TB_Subjectivity	float64	TextBlob subjectivity score (0 to 1)
TB_Sentiment_Label	object	TextBlob sentiment category: Positive / Neutral / Negative

Insights & Recommendations:

Insight / Finding	Details	Recommendation
Top-Rated Aspects	Well versed with the subject, Explains concepts clearly, Use of presentations	Continue current teaching methods; maintain clarity and presentation use
Lowest-Rated Aspects	Degree of difficulty of assignments, Solves doubts willingly, Course recommendation based on relevance	Simplify assignments, provide more student support, clarify course relevance
Overall Satisfaction	Most students in Medium satisfaction, fewer High and some Low	Focus on improving course delivery and support to increase satisfaction
Sentiment Analysis Insights	Low satisfaction → Negative; Medium → Neutral; High → Positive; TextBlob may classify subtle negatives as Neutral	Pay attention to neutral comments for actionable improvements
Common Feedback Themes (Word Cloud)	Positive: good, clear, presentation; Negative: assignment, difficult, confusing, hard	Address recurring pain points: assignments and clarity of instructions
Correlation Insights	Positive correlation between ratings and sentiment; highest drivers: Well versed with subject & Explains clearly; assignment difficulty weak/negative correlation	Target low-correlation aspects for improvement (assignments, support)
Actionable Recommendations	- Assignments: simplify with examples - Student Support: spend more time solving doubts - Course Relevance: clarify outcomes - Instructional Clarity: maintain presentations - Follow-up Feedback: mid-course/post-assignment collection	Implement these steps to boost student satisfaction and learning outcomes

Student Event Feedback Analysis Dashboard Visualization Summary:

Visualization Title	Chart Type	DAX / Measure Used	Business Questions Addressed	Why It's Important / Insight
Ranking of Course Aspects by Average Rating	Funnel Chart	Average('Student_Event_Feedback_Analysis'[Well versed with the subject]) Average('Student_Event_Feedback_Analysis'[Explains concepts in an understandable way]) Average('Student_Event_Feedback_Analysis'[Use of presentations]) … (repeated for all 8 aspects)	Which course aspects are rated highest by students?	Identifies strengths and key drivers of overall satisfaction.
Sentiment vs. Satisfaction Matrix	Matrix	Count of Student ID per combination of Satisfaction_Level & Sentiment_Label: CALCULATE(COUNTROWS('Student_Event_Feedback_Analysis'), 'Student_Event_Feedback_Analysis'[Satisfaction_Level] = <Level>, 'Student_Event_Feedback_Analysis'[Sentiment_Label] = <Sentiment>)	How do textual sentiments (VADER) align with numeric satisfaction levels?	Quickly shows if students' ratings match their comments; identifies mismatches or neutral feedback in medium satisfaction.
Breakdown of Student Satisfaction Levels	Donut Chart	Low = CALCULATE(COUNTROWS('Student_Event_Feedback_Analysis'), 'Student_Event_Feedback_Analysis'[Satisfaction_Level] = "Low") Medium = CALCULATE(COUNTROWS('Student_Event_Feedback_Analysis'), 'Student_Event_Feedback_Analysis'[Satisfaction_Level] = "Medium") High = CALCULATE(COUNTROWS('Student_Event_Feedback_Analysis'), 'Student_Event_Feedback_Analysis'[Satisfaction_Level] = "High")	How many students fall into each satisfaction category?	Quick overview of overall satisfaction distribution.
Average Polarity Across Satisfaction	Waterfall Chart	Average TB_Polarity = AVERAGE('Student_Event_Feedback_Analysis'[TB_Polarity])	How does average polarity vary across Low /	Highlights sentiment differences among

Visualization Title	Chart Type	DAX / Measure Used	Business Questions Addressed	Why It's Important / Insight
Levels			Medium / High satisfaction?	satisfaction levels.
VADER vs TextBlob Sentiment Comparison	Stacked Bar Chart	Count Positive / Neutral / Negative for VADER and TB:CALCULATE(COUNTROWS('Student_Event_Feedback_Analysis'), 'Student_Event_Feedback_Analysis'[Sentiment_Label] = "Positive") (same for Neutral & Negative)CALCULATE(COUNTROWS('Student_Event_Feedback_Analysis'), 'Student_Event_Feedback_Analysis'[TB_Sentiment_Label] = "Positive")	Are VADER and TextBlob sentiment classifications consistent?	Ensures NLP models align and validates feedback analysis.
Cumulative Sentiment Across Feedbacks	Area Chart	Cumulative_Sentiment_Score = CALCULATE(SUM('Student_Event_Feedback_Analysis'[Sentiment_Score]), FILTER(ALL('Student_Event_Feedback_Analysis'), 'Student_Event_Feedback_Analysis'[Student ID] <= MAX('Student_Event_Feedback_Analysis'[Student ID])))	How does overall sentiment trend across feedbacks?	Shows sentiment progression and identifies dips or peaks over feedback sequence.

Student Event Feedback Analysis Dashboard KPI Summary:

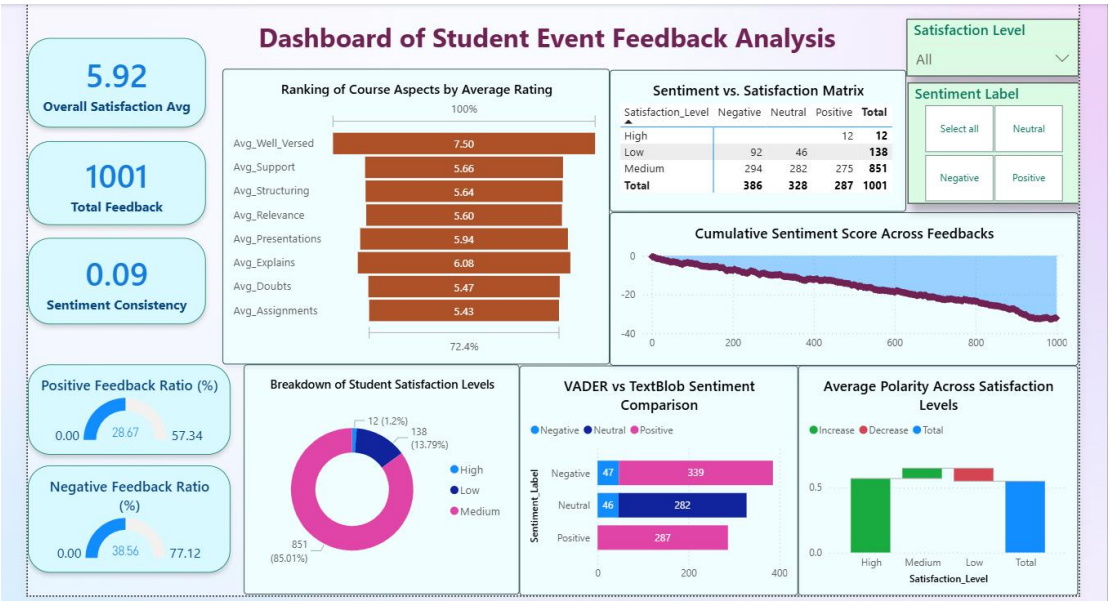
Column Name	DAX Formula	Business Question Addressed	Why It's Important / Insight
Overall Satisfaction Avg	Overall_Satisfaction_Avg = AVERAGE('Student_Event_Feedback_Analysis'[Overall_Satisfaction])	What is the average satisfaction across all students?	Provides a quick summary of overall course satisfaction.
Total Feedbacks	Total Feedbacks = COUNTROWS('Student_Event_Feedback_Analysis')	How many feedback responses were collected?	Measures engagement and dataset completeness.
Sentiment Consistency	Sentiment Consistency = ABS(AVERAGE('Student_Event_Feedback_Analysis'[Sentiment_Score]) -	How closely do VADER and TextBlob sentiments match?	Validates NLP analysis and ensures consistency of

Column Name	DAX Formula	Business Question Addressed	Why It's Important / Insight
	AVERAGE('Student_Event_Feedback_Analysis'[TB_Polarity]))		sentiment scoring.
Positive Feedback Ratio	Positive_Feedback_Ratio = DIVIDE(CALCULATE(COUNTROWS('Student_Event_Feedback_Analysis'), 'Student_Event_Feedback_Analysis'[Sentiment_Label] = "Positive"), COUNTROWS('Student_Event_Feedback_Analysis')) * 100	What proportion of feedback is positive?	Measures positivity of student feedback at a glance.
Negative Feedback Ratio	Negative_Feedback_Ratio = DIVIDE(CALCULATE(COUNTROWS('Student_Event_Feedback_Analysis'), 'Student_Event_Feedback_Analysis'[Sentiment_Label] = "Negative"), COUNTROWS('Student_Event_Feedback_Analysis')) * 100	What proportion of feedback is negative?	Quickly highlights areas of concern or dissatisfaction.

Student Event Feedback Analysis Dashboard Filters Summary:

Filter Name	Field Type	Purpose / What It Controls	Why It's Important / Insight
Satisfaction Level	Categorical (Text)	Allows users to filter and view data by satisfaction category — Low, Medium, or High	Helps understand how satisfaction impacts feedback ratings and sentiment trends, revealing which satisfaction groups need improvement.
Sentiment Label (VADER)	Categorical (Text)	Enables filtering feedback based on sentiment classification — Positive, Neutral, or Negative	Helps identify emotional tone of feedback and analyze the correlation between sentiment and overall satisfaction.

Final Interactive Dashboard of Student Event Feedback Analysis:



Conclusion:

The Student Event Feedback Analysis project successfully transformed raw student feedback into actionable insights. By combining quantitative ratings and qualitative comments, the analysis revealed key satisfaction drivers, pain points, and sentiment trends. Using both VADER and TextBlob sentiment models ensured reliable interpretation of feedback tone. The resulting Power BI dashboard provides a clear view of how students perceive course aspects, enabling data-driven improvements in academic experience and event quality.

Learning Outcomes:

- Understood how to clean and preprocess survey data effectively.
- Learned to calculate and interpret overall satisfaction scores.
- Applied TextBlob and VADER for text sentiment analysis.
- Visualized results using Power BI charts and KPIs.
- Compared sentiment results from multiple NLP models.
- Derived meaningful insights to support decision-making.
- Created a complete end-to-end data analytics workflow.

Skills Gained:

- ✓ Data Cleaning & Preprocessing
- ✓ Exploratory Data Analysis (EDA)
- ✓ Sentiment Analysis (VADER & TextBlob)
- ✓ Data Visualization in Power BI
- ✓ DAX Measure Creation
- ✓ Insight Generation & Interpretation
- ✓ Report Documentation & Presentation
- ✓ Analytical Thinking & Problem Solving