

# College Event Feedback Analysis

## Data Science & Analytics

Turning student feedback into actionable insights using ratings and sentiment analysis.



**Tech Stack:** Python · pandas · Matplotlib · Seaborn · TextBlob · NLP

**Platform:** Google Colab

**Created by:** Mayank Jindal



# Project Objective



College events such as workshops, seminars, cultural fests, and hackathons regularly collect student feedback. However, this feedback is often underutilised and limited to numerical ratings.

The objective of this project is to analyse student event feedback using data science techniques and Natural Language Processing (NLP) to:

**Understand overall student satisfaction**

**Identify patterns across event types and departments**

**Extract sentiment from textual feedback**

**Provide actionable recommendations for improving future events**

This project demonstrates how structured ratings and unstructured text can be combined to support data-driven decision-making in educational institutions.

# Dataset Description

A realistic, simulated dataset was created to mirror feedback collected via Google Forms after college events.

**1,000** **6**  
**Student responses** **Key columns**

Each row represents one student's feedback after attending an event

## Key Columns:

- Event Name
- Event Type (Workshop, Seminar, Cultural, Tech Fest)
- Department
- Rating (1–5 scale)
- Feedback Comment (text)
- Would Recommend (Yes / No)

## Sample Data

index	Event_Name	Event_Type	Department	Rating	Feedback_Comment	Would_Recommend
0	AI Workshop	Workshop	Electronics	1	Too long and unengaging	No
1	Entrepreneurship Seminar	Seminar	Civil	2	Poor organisation	No
2	AI Workshop	Workshop	Computer Science	4	Very informative and well organised	Yes
3	AI Workshop	Workshop	Civil	2	Needs better planning	No
4	Tech Symposium	Seminar	Management	4	Well structured and engaging	Yes

The dataset is designed to reflect real-world student feedback behavior, including positive, neutral, and negative responses.

# Tools & Technologies Used

This project was implemented using the following tools:



## Google Colab

Cloud-based Python environment



## Pandas & NumPy

Data cleaning and manipulation



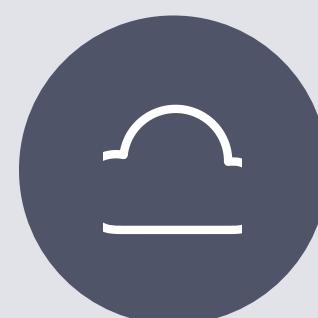
## Matplotlib & Seaborn

Data visualization



## TextBlob

Sentiment analysis using NLP



## WordCloud

Visual analysis of frequent complaint keywords

These tools enabled efficient handling of both numerical and textual data.

# Data Cleaning & Preparation

```
# Dataset info
df.info()

...
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 6 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Event_Name       1000 non-null    object  
 1   Event_Type       1000 non-null    object  
 2   Department       1000 non-null    object  
 3   Rating           1000 non-null    int64  
 4   Feedback_Comment 1000 non-null    object  
 5   Would_Recommend  1000 non-null    object  
dtypes: int64(1), object(5)
memory usage: 47.0+ KB
```

```
# Basic text cleaning
df['Feedback_Comment'] = df['Feedback_Comment'].str.lower()
df['Feedback_Comment'] = df['Feedback_Comment'].str.replace('[^a-zA-Z ]', '', regex=True)

df[['Feedback_Comment']].head()
```

	Feedback_Comment	grid icon
0	too long and unengaging	bar chart icon
1	poor organisation	bar chart icon
2	very informative and well organised	
3	needs better planning	
4	well structured and engaging	

Before analysis, the dataset was thoroughly inspected and prepared to ensure reliability.

## Key steps included:

01

02

03

Verifying dataset structure and dimensions

Checking for missing and duplicate values

Cleaning textual feedback

Cleaning textual feedback by:

- Converting text to lowercase
- Removing special characters and noise

Text cleaning is essential because sentiment analysis models are sensitive to inconsistencies and irrelevant characters.



# Exploratory Data Analysis

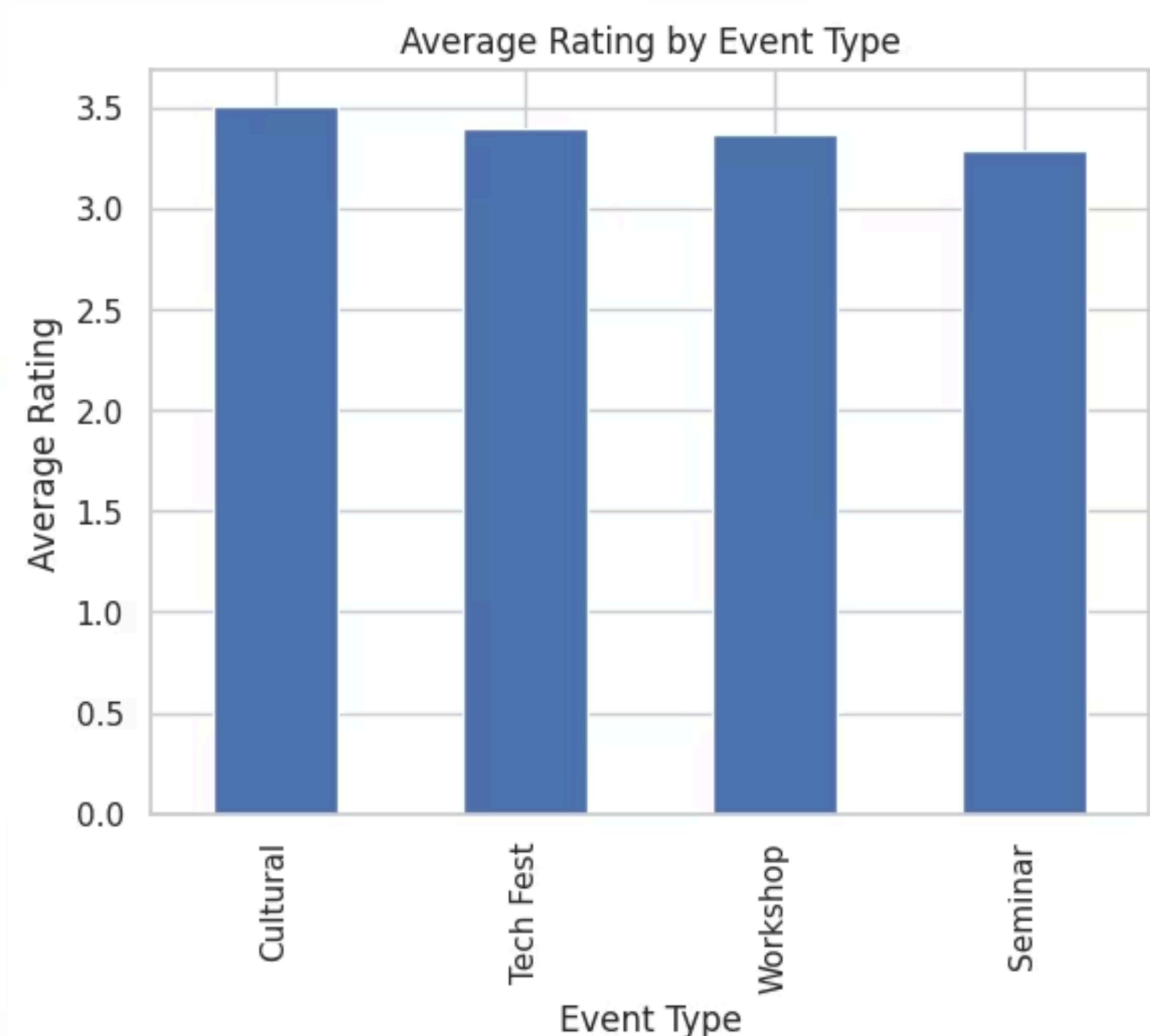
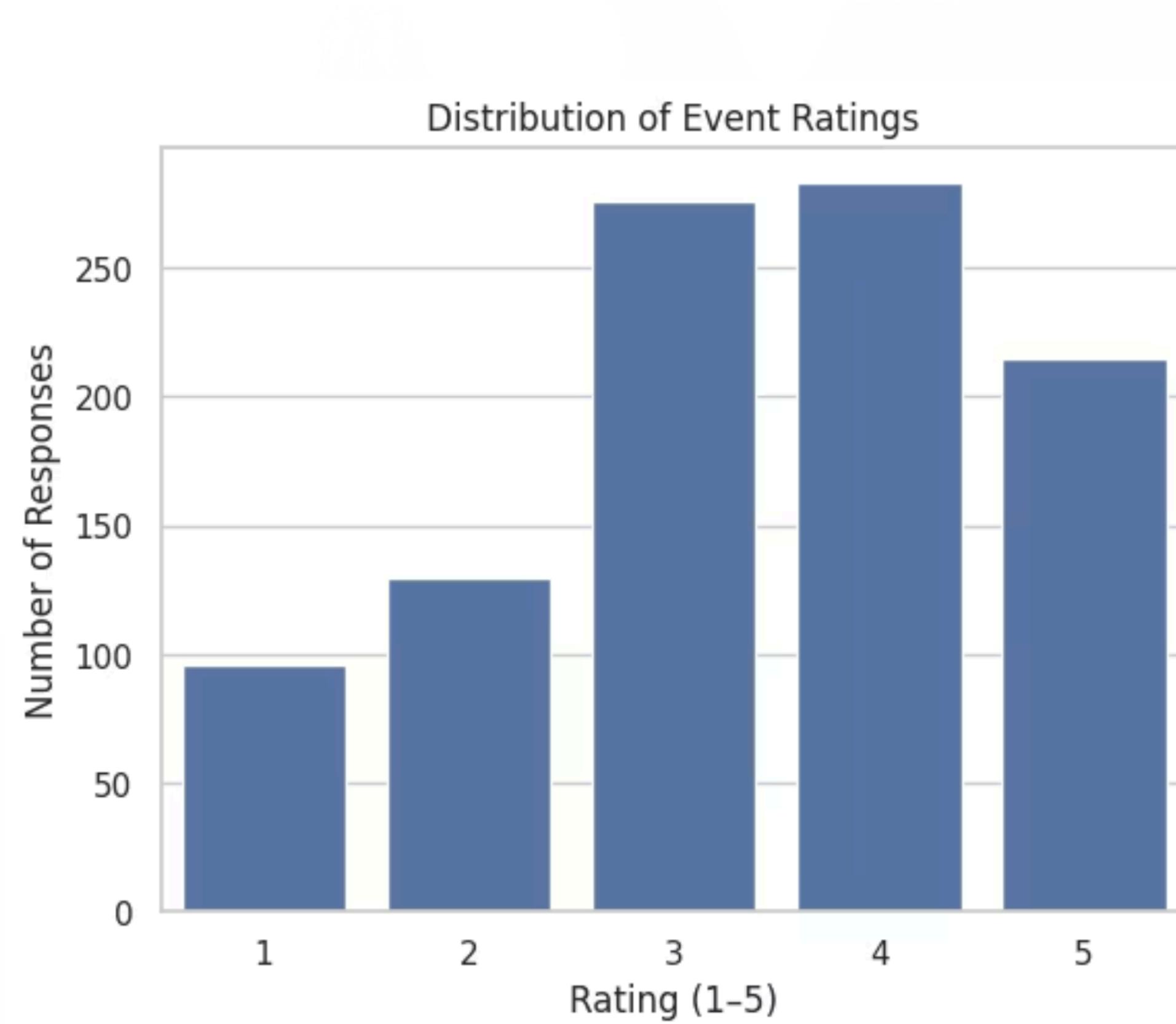
Exploratory Data Analysis was conducted to understand student satisfaction trends.

## Key findings:

Rating 4 is the most frequent score, indicating general satisfaction

Extreme negative ratings are relatively low

Average ratings vary across event types



When analysed by event type, **cultural events** consistently achieve higher average ratings, suggesting stronger emotional engagement compared to purely academic events.

# Sentiment Analysis of Feedback

To analyze textual feedback, sentiment polarity was calculated using **TextBlob**.

Each feedback comment was classified into:

1

**Positive**

expresses satisfaction or appreciation

2

**Neutral**

emotionally weak or generic comments

3

**Negative**

expresses dissatisfaction or complaints

```
def get_sentiment(text):
    return TextBlob(text).sentiment.polarity

df['Sentiment_Score'] = df['Feedback_Comment'].apply(get_sentiment)
df[['Feedback_Comment', 'Sentiment_Score']].head()

Feedback_Comment  Sentiment_Score
0      too long and unengaging     -0.125
1          poor organisation     -0.400
2  very informative and well organised      0.200
3           needs better planning      0.500
4      well structured and engaging      0.400

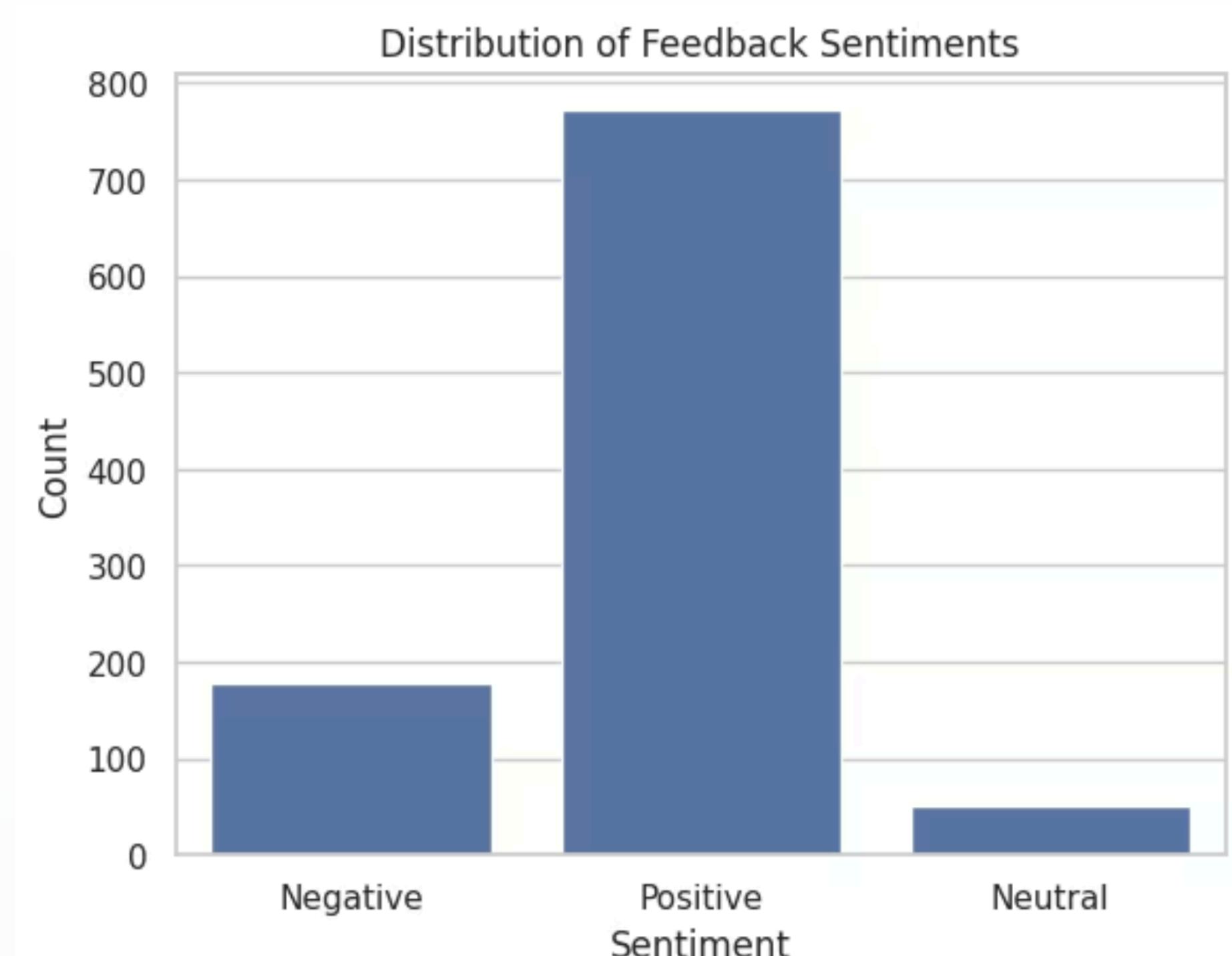
def sentiment_label(score):
    if score > 0:
        return "Positive"
    elif score < 0:
        return "Negative"
    else:
        return "Neutral"

df['Sentiment_Label'] = df['Sentiment_Score'].apply(sentiment_label)

df['Sentiment_Label'].value_counts()

count
Sentiment_Label
Positive      772
Negative      178
Neutral       50

dtype: int64
```



Sentiment analysis enables deeper insight into student opinions that numerical ratings alone cannot capture.

# Relationship Between Ratings and Sentiment

A comparison between ratings and sentiment scores was performed to validate student behaviour.



## Key observations:



Higher ratings generally align with positive sentiment

Some high ratings still show neutral sentiment due to brief or generic feedback

Sentiment provides context behind numerical satisfaction scores

This confirms that sentiment analysis is a necessary complement to rating-based evaluation.

# ⚠ Common Complaints in Negative Feedback

To identify recurring issues, a word cloud was generated from negative feedback comments.



Frequent complaint themes include:

1

Boring or unengaging content

2

Poor organisation

3

Long or tiring sessions

4

Overall average experience

These insights highlight that dissatisfaction is driven more by **content quality and engagement** than logistical factors.

# Conclusion & Recommendations

## Key Conclusions

- Overall student sentiment is positive
- Cultural events receive the highest satisfaction and sentiment scores
- Engagement quality is the primary driver of negative feedback

## Recommendations

**Introduce more interactive elements in workshops and seminars**

**Reduce session length and improve event structure**

**Use sentiment analysis regularly to monitor feedback trends**

## Final Note

This project demonstrates how data science and NLP can transform raw student feedback into meaningful insights that help institutions continuously improve campus events.