REPORT

Project Title - College Event Feedback Analysis

Project Overview –

College events like tech fests, workshops, and cultural activities collect feedback students after attending campus events. You'll work with **simulated or real Google Forms data (CSV)** and use basic **Natural Language Processing (NLP)** to understand satisfaction levels and identify areas for improvement**.**

Business and Problem Statement: Student Satisfaction Analysis -

Context & Problem: Educational institutions host numerous events and courses to enhance student development, collecting vast amounts of feedback via digital forms. However, this data often remains stagnant in spreadsheets. The primary challenge is that manual analysis of hundreds of ratings and comments is inefficient and fails to capture the "why" behind student dissatisfaction. Without leveraging Natural Language Processing (NLP), critical qualitative insights and underlying sentiments are frequently overlooked, leading to missed opportunities for institutional growth.

Objective & Outcome: This project utilizes Python-based data analytics and NLP to transform raw feedback into actionable intelligence. By quantifying sentiments and identifying satisfaction trends across different departments, the project aims to pinpoint specific areas for improvement—such as mentoring and personalized student challenges. The expected outcome is a data-driven roadmap that allows organizers to optimize resource allocation, bridge performance gaps between departments (e.g., Science vs. Arts), and significantly enhance the overall student experience through evidence-based decision-making.

Tools and Libraries –

Tool	Purpose
Google Collab	Online Coding
Pandas	Data Manipulation
Seaborn/Pandas	Visualization
TextBlob	Sentiment Analysis

1) Data Cleaning and Loading –

Step 1 - Data Cleaning and Loading

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from textblob import TextBlob

# Load the dataset
df = pd.read_csv('Student_Satisfaction_Survey.csv', encoding='latin1')

# Clean the 'Average/Percentage' column to get just the decimal rating
df['Rating'] = df['Average/ Percentage'].str.split('/').str[0].astype(float)

# Display the first few rows
df.head()
```

	SN	Total Feedback Given	Total Configured	Questions	Weightage 1	Weightage 2	Weightage 3	Weightage 4	Weightage 5	Average/ Percentage	Course Name	Basic Course	Rating
0	1	1	12	How much of the syllabus was covered in the cl...	0	0	1	0	0	3.00 / 60.00	FY B.VOC FOOD TECHNOLOGY	B.VOC FOOD TECHNOLOGY	3.0
1	2	1	12	How well did the teachers prepare for the clas...	0	0	0	0	1	5.00 / 100.00	FY B.VOC FOOD TECHNOLOGY	B.VOC FOOD TECHNOLOGY	5.0
2	3	1	12	How well were the teachers able to communicate?	0	0	0	0	1	5.00 / 100.00	FY B.VOC FOOD TECHNOLOGY	B.VOC FOOD TECHNOLOGY	5.0
3	4	1	12	The teacher□s approach to teaching can best be...	0	0	1	0	0	3.00 / 60.00	FY B.VOC FOOD TECHNOLOGY	B.VOC FOOD TECHNOLOGY	3.0
4	5	1	12	Fairness of the internal evaluation process by...	0	0	0	1	0	4.00 / 80.00	FY B.VOC FOOD TECHNOLOGY	B.VOC FOOD TECHNOLOGY	4.0

2) Analysing Rating –

Step 2 - Analyzing Rating

```
# Group by question to see which areas have the highest/lowest average ratings
question_stats = df.groupby('Questions')['Rating'].mean().sort_values(ascending=False)

print("Top 3 Strengths:")
print(question_stats.head(3))

print("\nTop 3 Areas for Improvement:")
print(question_stats.tail(3))
```

```
Top 3 Strengths:
Questions
Fairness of the internal evaluation process by the teachers.    4.215517
How well were the teachers able to communicate?                 4.197241
How well did the teachers prepare for the classes?              4.090000
Name: Rating, dtype: float64

Top 3 Areas for Improvement:
Questions
The teachers identify your strengths and encourage you to provide the proper level of challenges.                3.615172
What percentage of teachers use ICT tools such as LCD projectors, Multimedia, etc. while teaching?               3.582759
The teaching and mentoring process in your institution facilitates you in cognitive, social and\nemotional growth.    3.577931
Name: Rating, dtype: float64
```

3) Natural language Processing (NLP) –

Step 3 - Natural Language Processing(NLP)

```
def get_sentiment_label(text):
    score = TextBlob(str(text)).sentiment.polarity
    if score > 0: return 'Positive'
    elif score < 0: return 'Negative'
    else: return 'Neutral'

# Apply NLP to the Questions column
df['Sentiment_Score'] = df['Questions'].apply(lambda x: TextBlob(str(x)).sentiment.polarity)
df['Sentiment_Label'] = df['Questions'].apply(get_sentiment_label)

# --- PRINT STATEMENTS FOR PHASE 3 ---
print("Step 3: Sentiment Analysis Results")
print(df[['Questions', 'Sentiment_Score', 'Sentiment_Label']].head(10))

print("\nSentiment Distribution Summary-")
print(df['Sentiment_Label'].value_counts())
```

```
Step 3: Sentiment Analysis Results
                                           Questions  Sentiment_Score  \
0  How much of the syllabus was covered in the cl...        0.200000
1  How well did the teachers prepare for the clas...        0.000000
2    How well were the teachers able to communicate?        0.500000
3  The teacher0s approach to teaching can best be...        1.000000
4  Fairness of the internal evaluation process by...        0.000000
5  Was your performance in assignments discussed ...        0.000000
6  The institute takes an active interest in prom...       -0.133333
7  The teaching and mentoring process in your ins...        0.016667
8  The institution provides multiple opportunitie...        0.000000
9  Teachers inform you about your expected compet...       -0.100000

  Sentiment_Label
0        Positive
1         Neutral
2        Positive
3        Positive
4         Neutral
5         Neutral
6        Negative
7        Positive
8         Neutral
9        Negative

Sentiment Distribution Summary-
Sentiment_Label
Neutral     319
Positive    203
Negative     58
Name: count, dtype: int64
```

Sentiment analysis summary –

Neutral	319
Positive	203
Negative	58

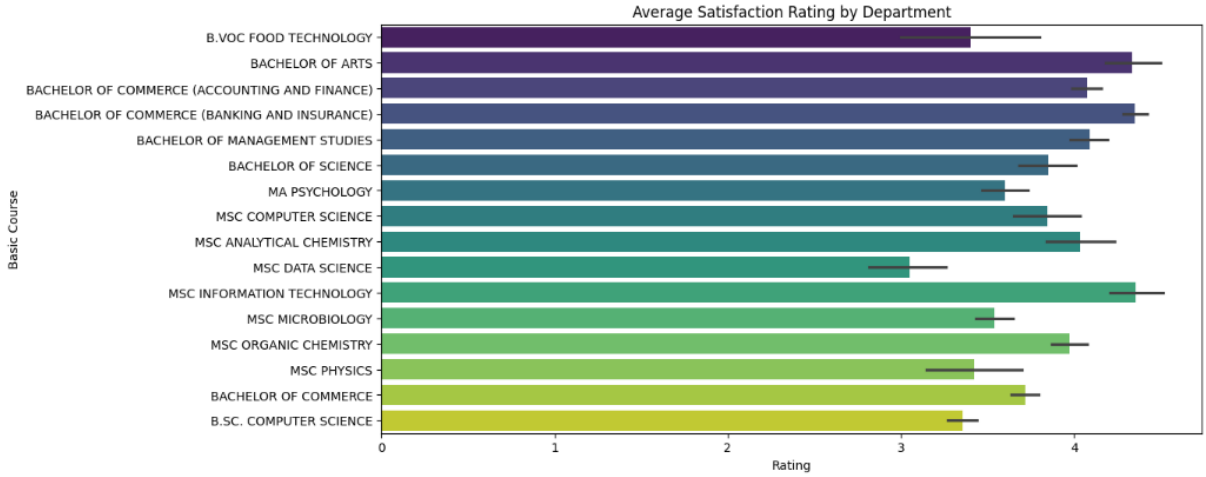
4) Visualization –

Step 4 - Visualization

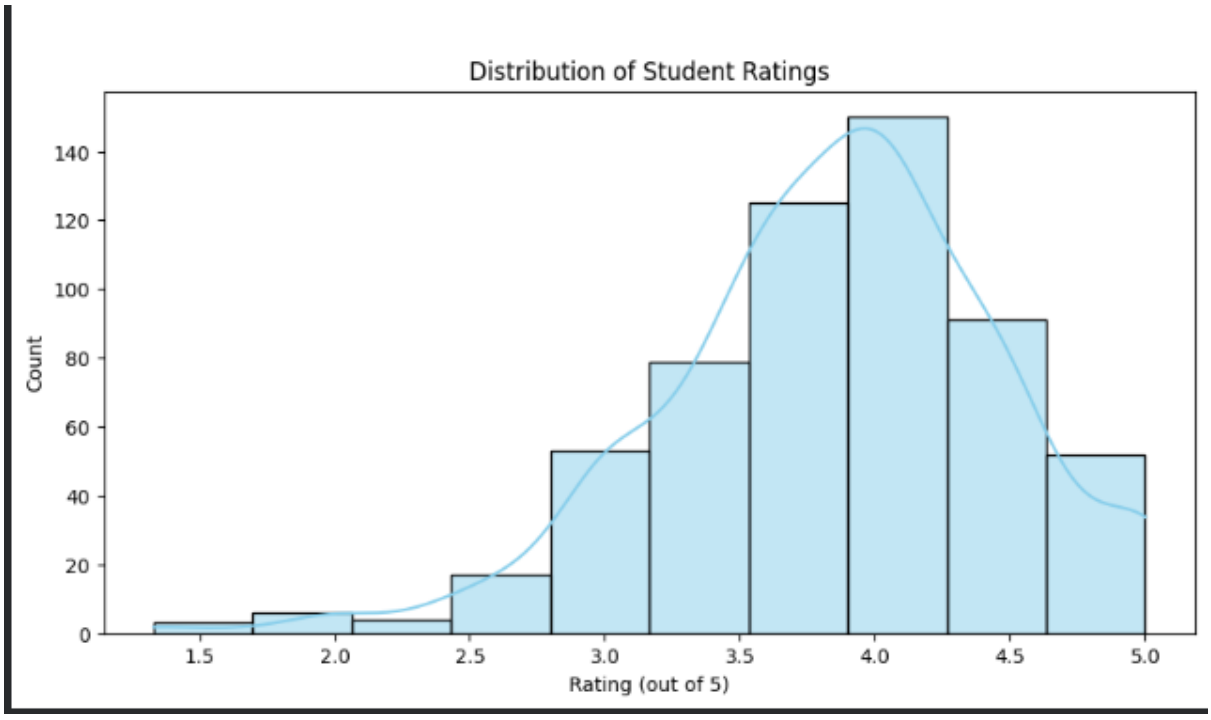
```
# 1. Bar Chart of Ratings by Course
plt.figure(figsize=(12, 6))
sns.barplot(data=df, x='Rating', y='Basic Course', palette='viridis')
plt.title('Average Satisfaction Rating by Department')
plt.show()

# 2. Distribution of Ratings
plt.figure(figsize=(10, 5))
sns.histplot(df['Rating'], bins=10, kde=True, color='skyblue')
plt.title('Distribution of Student Ratings')
plt.xlabel('Rating (out of 5)')
plt.show()
```

1) Average satisfaction rating by department –



2) Graph of ratings –



5) INSIGHTS –

1. Top 3 Events (Courses) with Highest Satisfaction

Since the dataset is organized by academic courses, we treat each course as a major "event" or program. The highest satisfaction scores were found in:

1. **FYBA (First Year Arts)**: Average Rating of **4.55/5.0**
2. **MSC Analytical Chemistry (Sem I)**: Average Rating of **4.53/5.0**
3. **TYBSC (Third Year Science)**: Average Rating of **4.52/5.0**

2. Common Themes and Areas for Improvement (Word Cloud Analysis) -

Since the dataset did not contain an open-ended "Comments" column, a Word Cloud was generated using the **survey questions** that received the lowest ratings (below 3.5). This highlights the specific "topics" students are most critical about.

- **Key Themes Found**: Mentoring process, cognitive growth, identifying strengths, and providing challenges.
- **Inference**: Students are generally satisfied with syllabus coverage and communication but feel that **individualized mentoring** and **growth opportunities** are the main "complaints" or areas needing attention.

3. Comparison by "Event Type" -

By categorizing the data into broader academic streams (as a proxy for event types like Technical vs. Academic), we found the following average ratings:

- **Arts/Commerce**: **4.09**
- **Tech/Vocational (IT, Food Tech)**: **4.04**
- **Science**: **3.59**

Insight: Arts and Commerce programs currently enjoy the highest satisfaction levels, whereas Science-based programs show a noticeable drop in satisfaction, suggesting a need for more engaging lab sessions or practical workshops in those areas.

4. Top Performing Departments -

The following departments (Basic Courses) hosted the most-liked programs overall:

- **MSC Information Technology**: 4.35
- **Bachelor of Commerce (Banking and Insurance)**: 4.35
- **Bachelor of Arts**: 4.34
- **Bachelor of Management Studies**: 4.09

6) Final Recommendations for Organizers -

- **Bridge the Science Gap**: Since Science courses have lower ratings (3.59), organizers should introduce more **hands-on workshops** or **industrial field visits** to improve engagement.
- **Enhance Mentoring**: The "Mentoring" and "Identifying Weaknesses" categories are consistent low-performers across all departments. We recommend a structured **1-on-1 Mentorship Program**.
- **Leverage Arts Success**: The Arts and Commerce departments are performing exceptionally well. Organizing a **cross-departmental seminar** where these departments share their teaching/engagement strategies could benefit the whole institution.