```
# @title import required libraries
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

from textblob import TextBlob
import nltk
nltk.download('vader_lexicon')
from nltk.sentiment import SentimentIntensityAnalyzer
```

```
[nltk_data] Downloading package vader_lexicon to /root/nltk_data...
```

```
# @title Load the Dataset
from google.colab import files
uploaded = files.upload()

df = pd.read_csv("/content/student_feedback_dataset.csv")
df.head()
```

```
<IPython.core.display.HTML object>
```

```
Saving student_feedback_dataset.csv to student_feedback_dataset
(1).csv
```


```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 1223,\n  \"fields\":
[\n    {\n      \"column\": \"student_id\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 1223,\n
\"samples\": [\n          \"STU0662\",\n          \"STU0221\",\n
\"STU0156\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"department\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 7,\n        \"samples\":
[\n          \"CIVIL\",\n          \"BIO\",\n          \"MBA\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"subject_name\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 15,\n        \"samples\": [\n
\"Poetry\",\n          \"Marketing\",\n          \"Concrete Tech\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"feedback_text\",\n
\"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 1223,\n        \"samples\": [\n          \"The
generation send music natural in pull drive fine.\",\n
\"Perhaps model detail himself indicate idea they.\",\n
\"Field project seek however huge during.\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"preprocessed_text\",\n
\"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 1223,\n        \"samples\": [\n          \"the
```

```
generation send music natural in pull drive fine\",\n
\"perhaps model detail himself indicate idea they\",\n
\"field project seek however huge during\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"sentiment_label\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 3,\n        \"samples\": [\n
\"negative\",\n          \"positive\",\n          \"neutral\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"sarcasm_flag\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
0,\n        \"min\": 0,\n        \"max\": 1,\n
\"num_unique_values\": 2,\n        \"samples\": [\n          1,\n
0\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"emotion_tag\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 6,\n        \"samples\":
[\n          \"anger\",\n          \"sadness\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"feedback_type\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 3,\n        \"samples\": [\n
\"Infrastructure\",\n          \"Academic\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"subject_specific_context\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 3,\n        \"samples\": [\n
\"Project-Based\",\n          \"Lab-Intensive\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"response_required\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
0,\n        \"min\": 0,\n        \"max\": 1,\n
\"num_unique_values\": 2,\n        \"samples\": [\n          0,\n
1\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    }\n  ]\
n}","type":"dataframe","variable_name":"df"}
```


```
# @title Basic dataset info
print("Shape:", df.shape)
df.info()
df.describe(include='all')
```

```
Shape: (1223, 11)
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1223 entries, 0 to 1222
Data columns (total 11 columns):
 #   Column                   Non-Null Count  Dtype
---  ------                   --------------  -----
 0   student_id               1223 non-null   object
 1   department               1223 non-null   object
```

```
 2   subject_name               1223 non-null   object
 3   feedback_text              1223 non-null   object
 4   preprocessed_text          1223 non-null   object
 5   sentiment_label            1223 non-null   object
 6   sarcasm_flag               1223 non-null   int64
 7   emotion_tag                1223 non-null   object
 8   feedback_type              1223 non-null   object
 9   subject_specific_context   1223 non-null   object
 10  response_required          1223 non-null   int64
dtypes: int64(2), object(9)
memory usage: 105.2+ KB
```


```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 11,\n  \"fields\": [\n{\n      \"column\": \"student_id\",\n      \"properties\": {\n\"dtype\": \"category\",\n        \"num_unique_values\": 3,\n\"samples\": [\n          \"1223\",\n          \"STU1223\",\n\"1\"\n        ],\n        \"semantic_type\": \"\",\n\"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"department\",\n      \"properties\": {\n        \"dtype\": \"category\",\n        \"num_unique_values\": 4,\n        \"samples\": [\n          7,\n          \"187\",\n          \"1223\"\n        ],\n\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"subject_name\",\n\"properties\": {\n        \"dtype\": \"category\",\n\"num_unique_values\": 4,\n        \"samples\": [\n          15,\n\"101\",\n          \"1223\"\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n\"column\": \"feedback_text\",\n      \"properties\": {\n\"dtype\": \"category\",\n        \"num_unique_values\": 3,\n\"samples\": [\n          \"1223\",\n          \"Oil several consider two national you with receive surface.\",\n          \"1\"\n        ],\n        \"semantic_type\": \"\",\n\"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"preprocessed_text\",\n      \"properties\": {\n        \"dtype\": \"category\",\n        \"num_unique_values\": 3,\n        \"samples\": [\n          \"1223\",\n          \"oil several consider two national you with receive surface\",\n          \"1\"\n        ],\n\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"sentiment_label\",\n\"properties\": {\n        \"dtype\": \"category\",\n\"num_unique_values\": 4,\n        \"samples\": [\n          3,\n\"489\",\n          \"1223\"\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n\"column\": \"sarcasm_flag\",\n      \"properties\": {\n\"dtype\": \"number\",\n        \"std\": 432.32143439959196,\n\"min\": 0.0,\n        \"max\": 1223.0,\n\"num_unique_values\": 5,\n        \"samples\": [\n0.13409648405560098,\n          1.0,\n          0.34089534635190644\n],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n}\n    },\n    {\n      \"column\": \"emotion_tag\",\n
```

\"properties\": {\n   \"dtype\": \"category\",\n
\"num_unique_values\": 4,\n   \"samples\": [\n   6,\n
\"209\",\n   \"1223\"\n   ],\n   \"semantic_type\":
\"\",\n   \"description\": \"\"\n   }\n   },\n   {\n
\"column\": \"feedback_type\",\n   \"properties\": {\n
\"dtype\": \"category\",\n   \"num_unique_values\": 4,\n
\"samples\": [\n   3,\n   \"431\",\n   \"1223\"\n
],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n
}\n   },\n   {\n   \"column\": \"subject_specific_context\",\n
\"properties\": {\n   \"dtype\": \"category\",\n
\"num_unique_values\": 4,\n   \"samples\": [\n   3,\n
\"425\",\n   \"1223\"\n   ],\n   \"semantic_type\":
\"\",\n   \"description\": \"\"\n   }\n   },\n   {\n
\"column\": \"response_required\",\n   \"properties\": {\n
\"dtype\": \"number\",\n   \"std\": 432.1868872895946,\n
\"min\": 0.0,\n   \"max\": 1223.0,\n
\"num_unique_values\": 5,\n   \"samples\": [\n
0.669664758789861,\n   1.0,\n   0.47052619024308223\n
],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n
}\n   }\n]\n}","type":"dataframe"}

```
# @title Data cleaning
df = df.drop_duplicates()

df = df.dropna(how='all')

if 'feedback' in df.columns:
    df['feedback'] = df['feedback'].fillna("")
else:
    print("⚠ Column named 'feedback' not found. Check CSV column
names.")
```

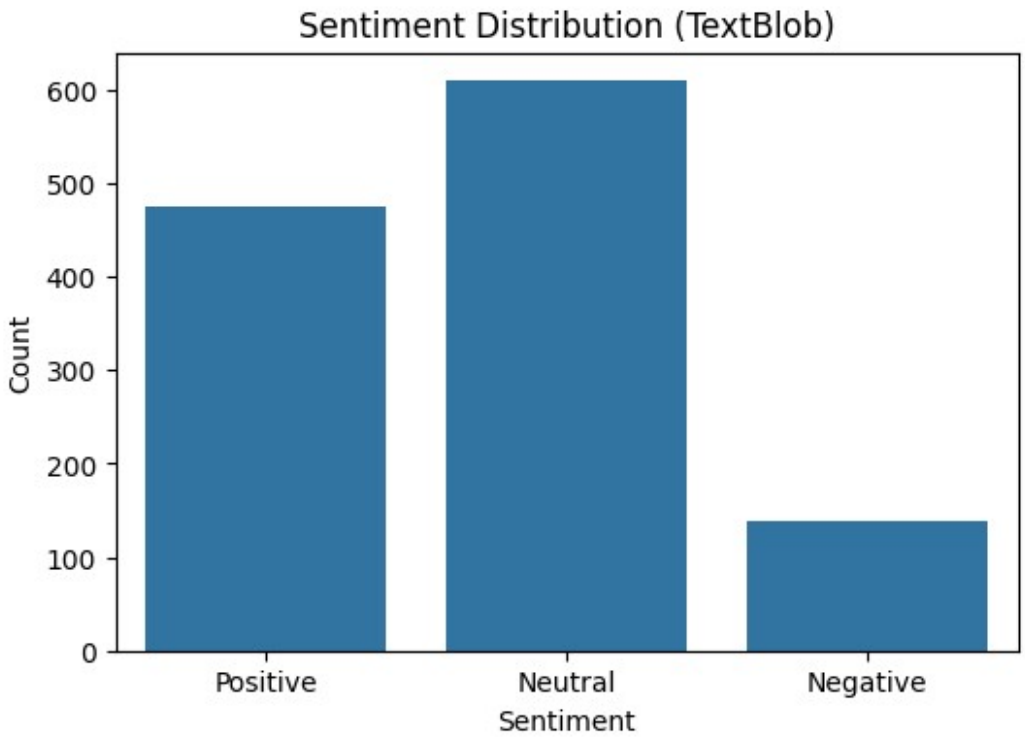
```
⚠ Column named 'feedback' not found. Check CSV column names.
```

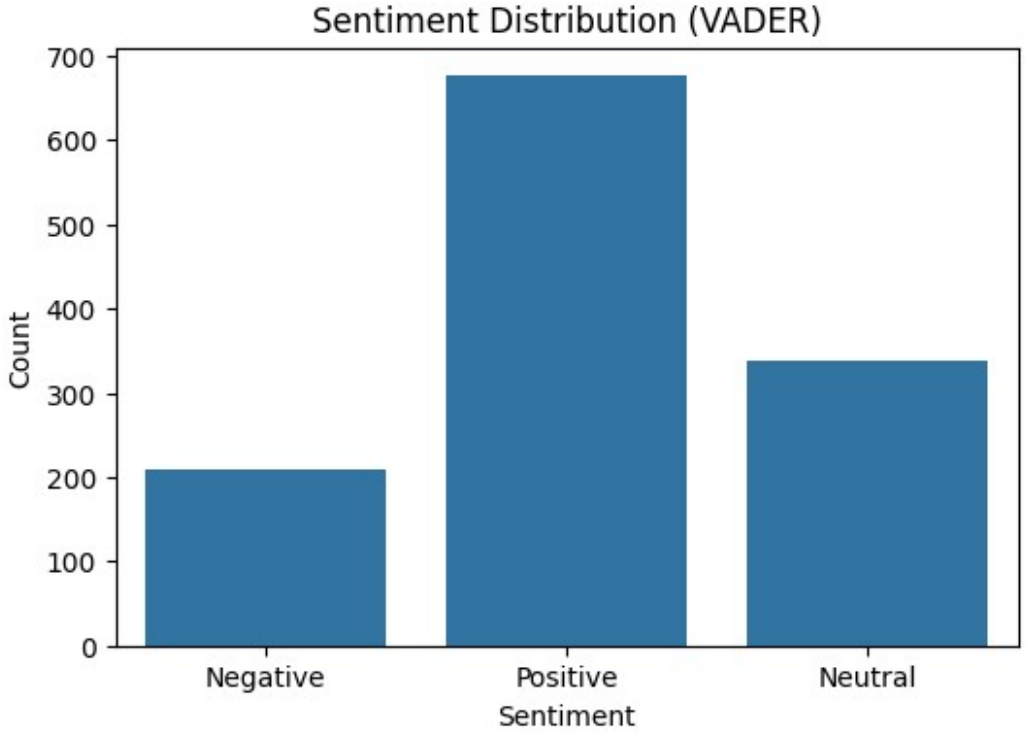
```
df.columns = df.columns.str.strip()
df = df.apply(lambda x: x.astype(str).str.strip() if x.dtype ==
"object" else x)

df.head()
```

{"summary":"{\n \"name\": \"df\",\n \"rows\": 1223,\n \"fields\":
[\n   {\n   \"column\": \"student_id\",\n   \"properties\": {\n
\"dtype\": \"string\",\n   \"num_unique_values\": 1223,\n
\"samples\": [\n   \"STU0662\",\n   \"STU0221\",\n
\"STU0156\"\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"department\",\n   \"properties\": {\n   \"dtype\":
\"category\",\n   \"num_unique_values\": 7,\n   \"samples\":
[\n   \"CIVIL\",\n   \"BIO\",\n   \"MBA\"\n
],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n

```
}\n    },\n    {\n      \"column\": \"subject_name\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 15,\n        \"samples\": [\n
\"Poetry\",\n          \"Marketing\",\n          \"Concrete Tech\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"feedback_text\",\n
\"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 1223,\n        \"samples\": [\n          \"The
generation send music natural in pull drive fine.\",\n
\"Perhaps model detail himself indicate idea they.\",\n
\"Field project seek however huge during.\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"preprocessed_text\",\n
\"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 1223,\n        \"samples\": [\n          \"the
generation send music natural in pull drive fine\",\n
\"perhaps model detail himself indicate idea they\",\n
\"field project seek however huge during\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"sentiment_label\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 3,\n        \"samples\": [\n
\"negative\",\n          \"positive\",\n          \"neutral\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"sarcasm_flag\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
0,\n        \"min\": 0,\n        \"max\": 1,\n
\"num_unique_values\": 2,\n        \"samples\": [\n          1,\n
0\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"emotion_tag\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 6,\n        \"samples\":
[\n          \"anger\",\n          \"sadness\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"feedback_type\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 3,\n        \"samples\": [\n
\"Infrastructure\",\n          \"Academic\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"subject_specific_context\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 3,\n        \"samples\": [\n
\"Project-Based\",\n          \"Lab-Intensive\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"response_required\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
0,\n        \"min\": 0,\n        \"max\": 1,\n
\"num_unique_values\": 2,\n        \"samples\": [\n          0,\n
1\n        ],\n        \"semantic_type\": \"\",\n
```

```
\"description\": \"\"\n      }\n    }\n  ]\
n}","type":"dataframe","variable_name":"df"}

# @title Sentiment Analysis(TextBlob)
def get_textblob_sentiment(text):
    return TextBlob(text).sentiment.polarity

df["textblob_score"] =
df["feedback_text"].apply(get_textblob_sentiment)

def sentiment_label(score):
    if score > 0.1:
        return "Positive"
    elif score < -0.1:
        return "Negative"
    else:
        return "Neutral"

df["textblob_sentiment"] = df["textblob_score"].apply(sentiment_label)

# @title VADER Sentiment Analysis
sia = SentimentIntensityAnalyzer()

def vader_label(text):
    score = sia.polarity_scores(text)["compound"]
    if score > 0.05:
        return "Positive"
    elif score < -0.05:
        return "Negative"
    else:
        return "Neutral"

df["vader_sentiment"] = df["feedback_text"].apply(vader_label)

# @title Visualization of sentiment Result
plt.figure(figsize=(6,4))
sns.countplot(x=df["textblob_sentiment"])
plt.title("Sentiment Distribution (TextBlob)")
plt.xlabel("Sentiment")
plt.ylabel("Count")
plt.show()

plt.figure(figsize=(6,4))
sns.countplot(x=df["vader_sentiment"])
plt.title("Sentiment Distribution (VADER)")
plt.xlabel("Sentiment")
plt.ylabel("Count")
plt.show()
```






```
# @title WordCloud
!pip install wordcloud
from wordcloud import WordCloud
```

```
all_text = " ".join(df["feedback_text"].tolist())

wordcloud = WordCloud(width=800, height=400,
background_color="white").generate(all_text)

plt.figure(figsize=(12,6))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis("off")
plt.title("Most Common Words in Feedback")
plt.show()
```

```
Requirement already satisfied: wordcloud in
/usr/local/lib/python3.12/dist-packages (1.9.4)
Requirement already satisfied: numpy>=1.6.1 in
/usr/local/lib/python3.12/dist-packages (from wordcloud) (2.0.2)
Requirement already satisfied: pillow in
/usr/local/lib/python3.12/dist-packages (from wordcloud) (11.3.0)
Requirement already satisfied: matplotlib in
/usr/local/lib/python3.12/dist-packages (from wordcloud) (3.10.0)
Requirement already satisfied: contourpy>=1.0.1 in
/usr/local/lib/python3.12/dist-packages (from matplotlib->wordcloud)
(1.3.3)
Requirement already satisfied: cycler>=0.10 in
/usr/local/lib/python3.12/dist-packages (from matplotlib->wordcloud)
(0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in
/usr/local/lib/python3.12/dist-packages (from matplotlib->wordcloud)
(4.61.0)
Requirement already satisfied: kiwisolver>=1.3.1 in
/usr/local/lib/python3.12/dist-packages (from matplotlib->wordcloud)
(1.4.9)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.12/dist-packages (from matplotlib->wordcloud)
(25.0)
Requirement already satisfied: pyparsing>=2.3.1 in
/usr/local/lib/python3.12/dist-packages (from matplotlib->wordcloud)
(3.2.5)
Requirement already satisfied: python-dateutil>=2.7 in
/usr/local/lib/python3.12/dist-packages (from matplotlib->wordcloud)
(2.9.0.post0)
Requirement already satisfied: six>=1.5 in
/usr/local/lib/python3.12/dist-packages (from python-dateutil>=2.7-
>matplotlib->wordcloud) (1.17.0)
```

Most Common Words in Feedback


```
# @title Key Insights
print("\n□ Key Insights:")
print("Most common TextBlob sentiment:",
df["textblob_sentiment"].value_counts().idxmax())
print("Most common VADER sentiment:",
df["vader_sentiment"].value_counts().idxmax())
```

```
□ Key Insights:
Most common TextBlob sentiment: Neutral
Most common VADER sentiment: Positive
```

```
# @title Save Output File
df.to_csv("/content/student_feedback_dataset.csv", index=False)

print("\n□ File saved as: event_feedback_with_sentiment.csv")
```

```
□ File saved as: event_feedback_with_sentiment.csv
```