

Marketing Analytics Business Case

Harsh Wardhan Singh

Day-3

Sentiment Analysis

Sentiment analysis is a natural language processing (NLP) technique used to identify and quantify emotions in text. It helps businesses understand customer opinions, categorize feedback, and make data-driven decisions based on public sentiment and tone.



Steps Followed in Sentiment Analysis.

1. Imported required libraries(e.g., pandas, pyodbc, nltk, matplotlib)
2. Connected to SQL Server databaseUsed pyodbc to connect and fetch customer reviews from a SQL table.
3. Loaded data into a DataFrameUsed pandas.read_sql() to retrieve and explore the dataset.
4. Downloaded and initialized VADER sentiment analyzer
nltk.download('vader_lexicon')
and initialized SentimentIntensityAnalyzer().

```
import pandas as pd
from nltk.sentiment.vader import SentimentIntensityAnalyzer
import nltk

# Download VADER Lexicon
nltk.download('vader_lexicon')

# Initialize VADER
vad = SentimentIntensityAnalyzer()

# Define sentiment scoring function
def sentiment_score(review):
    sentiment = vad.polarity_scores(review)
    return sentiment['compound']

# Define sentiment categorization using rating and score
def categorize_sentiment(score, rating):
    if score > 0.05:
        if rating >= 4:
            return 'Positive'
        elif rating == 3:
            return 'Mixed Positive'
        else:
            return 'Mixed Negative'
    elif score < -0.05:
        if rating <= 2:
            return 'Negative'
        elif rating == 3:
            return 'Mixed Negative'
        else:
            return 'Mixed Positive'
    else:
        if rating >= 4:
            return 'Positive'
        elif rating <= 2:
            return 'Negative'
        else:
            return 'Neutral'
```

Steps Followed in Sentiment Analysis.

5. Calculated sentiment scoresApplied
sia.polarity_scores() to extract compound
sentiment scores for each review.

6. Classified sentiment categoriesUsed a
custom function combining sentiment score
and star rating to assign sentiment labels like
Positive, Negative, Mixed, etc.

7. Bucketed sentiment scoresGrouped
scores into ranges (e.g., 0.5 to 1.0) for
visualization and aggregation.

8. Exported final datasetSaved the enriched
DataFrame to a CSV for further use in
visualization tools like Power BI.

```
# Define sentiment bucket
def sentiment_bucket(score):
    if score >= 0.5:
        return '0.5 to 1.0'
    elif 0.0 <= score < 0.5:
        return '0.0 to 0.49'
    elif -0.5 <= score < 0.0:
        return '-0.49 to 0.0'
    else:
        return '-1.0 to -0.5'

# Apply functions to your DataFrame (assuming `df` already contains your data)
df['SentimentScore'] = df['ReviewText'].apply(sentiment_score)

df['SentimentCategory'] = df.apply(
    lambda row: categorize_sentiment(row['SentimentScore'], row['Rating']), axis=1)

df['SentimentBucket'] = df['SentimentScore'].apply(sentiment_bucket)

# Show results
print(df.head())

# Save to CSV
df.to_csv('fact_customer_reviews_with_sentiment.csv', index=False)

# Test example
print(vad.polarity_scores("I'm so happy"))
print(vad.polarity_scores("It is the worst product i have"))
```

```
[nltk_data] Downloading package vader_lexicon to
[nltk_data] C:\Users\Lenovo\AppData\Roaming\nltk_data...
[nltk_data] Package vader_lexicon is already up-to-date!
```

	ReviewID	CustomerID	ProductID	ReviewDate	Rating	\
0	1	77	18	2023-12-23	3	
1	2	80	19	2024-12-25	5	
2	3	50	13	2025-01-26	4	
3	4	78	15	2025-04-21	3	
4	5	64	2	2023-07-16	3	

	ReviewText	SentimentScore	SentimentCategory	\
0	Average experience, nothing special.	-0.3089	Mixed Negative	
1	The quality is top-notch.	0.0000	Positive	
2	Five stars for the quick delivery.	0.0000	Positive	
3	Good quality, but could be cheaper.	0.2382	Mixed Positive	

Result

```
1 use marketdb
2
3 SELECT * FROM customer_journey
4 select * from customer_reviews
5 SELECT * FROM customers
6 SELECT * FROM engagement_data
7 SELECT * FROM geography
8 SELECT * FROM products
```

	ReviewID	CustomerID	ProductID	ReviewDate	Rating	ReviewText
1	1	77	18	2023-12-23	3	Average experience, nothing special.
2	2	80	19	2024-12-25	5	The quality is top-notch.
3	3	50	13	2025-01-26	4	Five stars for the quick delivery.
4	4	78	15	2025-04-21	3	Good quality, but could be cheaper.
5	5	64	2	2023-07-16	3	Average experience, nothing special.
6	6	81	1	2025-12-21	4	Customer support was very helpful.
7	7	16	1	2024-01-29	3	Average experience, nothing special.
8	8	55	8	2024-08-15	5	The quality is top-notch.
9	9	3	13	2023-09-01	4	I love this product, will buy again!
10	10	78	6	2024-06-17	5	Excellent product, highly recommend!
11	11	54	9	2023-07-17	4	Great purchase, very satisfied.
12	12	77	2	2025-01-13	2	Product did not meet my expectations.
13	13	93	6	2024-08-20	5	Great purchase, very satisfied.
14	14	23	2	2025-06-16	4	Good quality, but could be cheaper.
15	15	29	11	2023-01-27	5	Great purchase, very satisfied.
16	16	95	17	2023-01-06	5	Excellent product, highly recommend!
17	17	88	4	2025-10-29	3	Not worth the money.
18	18	17	16	2024-11-23	4	Shipping was fast and the item was well-p...
19	19	34	13	2024-01-04	4	The quality is top-notch.
20	20	34	6	2023-04-29	1	I had a bad experience with this product.
21	21	22	15	2023-04-08	5	Amazing value for the price.
22	22	52	3	2025-10-30	4	The quality is top-notch.
23	23	69	18	2023-07-02	5	Five stars for the quick delivery.
24	24	100	4	2025-07-18	3	I love this product, will buy again!
25	25	55	20	2024-12-19	4	Shipping was fast and the item was well-p...

fact_customer_reviews_with_sentiment - Excel											shashwardhan1995@outlo
Formulas Data Review View Help Tell me what you want to do											
<div><div>11 A A</div><div><div>General</div><div>Conditional Formatting</div><div>Format as Table</div></div><div><div>Normal</div><div>Bad</div><div>Good</div><div>Neutral</div></div><div><div>Calculation</div><div>Check Cell</div><div>Explanatory...</div><div>Input</div></div><div><div>Insert</div><div>Delete</div><div>Format</div></div><div><div>AutoSum</div><div>Fill</div><div>Clear</div></div></div>											
ont Alignment Number Styles Cells											
	A	B	C	D	E	F	G	H	I	J	
1	ReviewID	Customer	ProductID	ReviewDate	Rating	ReviewText	SentimentScore	SentimentCategory	SentimentBucket		
2	1	77	18	12/23/2023	3	Average experience, nothing special.	-0.3089	Mixed Negative	-0.49 to 0.0		
3	2	80	19	12/25/2024	5	The quality is top-notch.	0	Positive	0.0 to 0.49		
4	3	50	13	1/26/2025	4	Five stars for the quick delivery.	0	Positive	0.0 to 0.49		
5	4	78	15	4/21/2025	3	Good quality, but could be cheaper.	0.2382	Mixed Positive	0.0 to 0.49		
6	5	64	2	7/16/2023	3	Average experience, nothing special.	-0.3089	Mixed Negative	-0.49 to 0.0		
7	6	81	1	12/21/2025	4	Customer support was very helpful.	0.6997	Positive	0.5 to 1.0		
8	7	16	1	1/29/2024	3	Average experience, nothing special.	-0.3089	Mixed Negative	-0.49 to 0.0		
9	8	55	8	8/15/2024	5	The quality is top-notch.	0	Positive	0.0 to 0.49		
10	9	3	13	9/1/2023	4	I love this product, will buy again!	0.6696	Positive	0.5 to 1.0		
11	10	78	6	6/17/2024	5	Excellent product, highly recommend!	0.7773	Positive	0.5 to 1.0		
12	11	54	9	7/17/2023	4	Great purchase, very satisfied.	0.8016	Positive	0.5 to 1.0		
13	12	77	2	1/13/2025	2	Product did not meet my expectations.	0	Negative	0.0 to 0.49		
14	13	93	6	8/20/2024	5	Great purchase, very satisfied.	0.8016	Positive	0.5 to 1.0		
15	14	23	2	6/16/2025	4	Good quality, but could be cheaper.	0.2382	Positive	0.0 to 0.49		
16	15	29	11	1/27/2023	5	Great purchase, very satisfied.	0.8016	Positive	0.5 to 1.0		
17	16	95	17	1/6/2023	5	Excellent product, highly recommend!	0.7773	Positive	0.5 to 1.0		
18	17	88	4	10/29/2025	3	Not worth the money.	-0.1695	Mixed Negative	-0.49 to 0.0		
19	18	17	16	11/23/2024	4	Shipping was fast and the item was well	0	Positive	0.0 to 0.49		
20	19	34	13	1/4/2024	4	The quality is top-notch.	0	Positive	0.0 to 0.49		
21	20	34	6	4/29/2023	1	I had a bad experience with this product	-0.5423	Negative	-1.0 to -0.5		
22	21	22	15	4/8/2023	5	Amazing value for the price.	0.7351	Positive	0.5 to 1.0		
23	22	52	3	10/30/2025	4	The quality is top-notch.	0	Positive	0.0 to 0.49		
24	23	69	18	7/2/2023	5	Five stars for the quick delivery.	0	Positive	0.0 to 0.49		
25	24	100	4	7/18/2025	3	I love this product, will buy again!	0.6696	Mixed Positive	0.5 to 1.0		
26	25	55	20	12/19/2024	4	Shipping was fast and the item was well	0	Positive	0.0 to 0.49		
27	26	76	11	8/21/2024	4	Five stars for the quick delivery.	0	Positive	0.0 to 0.49		
28	27	56	6	11/24/2025	4	The quality is top-notch.	0	Positive	0.0 to 0.49		
29	28	76	8	5/27/2025	5	Five stars for the quick delivery.	0	Positive	0.0 to 0.49		
30	29	38	13	5/31/2023	4	I love this product, will buy again!	0.6696	Positive	0.5 to 1.0		
31	30	84	6	3/18/2024	4	Five stars for the quick delivery.	0	Positive	0.0 to 0.49		
32	31	93	4	8/25/2025	3	Good quality, but could be cheaper.	0.2382	Mixed Positive	0.0 to 0.49		
33	32	91	1	7/15/2023	4	The quality is top-notch.	0	Positive	0.0 to 0.49		
34	33	30	2	2/21/2023	4	Shipping was fast and the item was well	0	Positive	0.0 to 0.49		
35	34	81	15	3/30/2024	5	Excellent product, highly recommend!	0.7773	Positive	0.5 to 1.0		

Thank You

- [Sentiment Analysis Excel](#)