Appendix: Text Mining Analysis of *Na Minha Rua* App Reviews

1. Introduction

This appendix presents a comprehensive text mining analysis conducted on user reviews of the *Na Minha Rua* mobile application. The study aims to extract meaningful insights from user feedback collected from the Google Play Store and Apple App Store platforms. The objectives include understanding user experiences, identifying common issues, and assessing overall satisfaction with the application. The analysis also seeks to determine a "true rating" for the application by correlating sentiment derived from user comments with their provided star ratings. This was particularly important as discrepancies were observed, such as complaints being associated with high ratings.

2. Analysis Objectives

The text mining analysis was designed to achieve the following objectives:

1. Collection and aggregation of user reviews from multiple platforms.
2. Analysis of sentiment patterns in user feedback.
3. Identification of recurring themes and issues.
4. Evaluation of user satisfaction across different app versions.
5. Using machine learning models to derive a sentiment-based "true rating" of the application.

3. Technical Framework

3.1 Data Collection and Preparation

- Reviews were collected from the Google Play Store and Apple App Store using:
  - `google_play_scraper` : Extracts user feedback, ratings, and metadata from Android users.
  - `app_store_scraper` : Provides similar functionality for iOS user reviews.
- **Data preprocessing** steps included:
  - Cleaning and standardizing text by removing duplicates and ensuring consistency across platforms.
  - Translating non-English reviews to English using the `googletrans` library to facilitate uniform analysis.

3.2 Text Analysis Components

For text processing and sentiment analysis, the following tools were employed:

- `googletrans` : For translation of text into a common language.
- `wordcloud` : Visualized frequently occurring terms and themes in reviews.
- `transformers` : Utilized a RoBERTa-based sentiment analysis model for nuanced sentiment classification.

3.3 Visualization Components

- Tools such as **matplotlib**, **seaborn**, and **plotly** were employed to create:
  - Word clouds.
  - Sentiment trends over time.
  - Comparative visualizations of platform-based feedback.

4. Sentiment Analysis Results

4.1 TextBlob Sentiment Analysis

- Provided basic sentiment polarity scores as an initial benchmark.
- Limited in accuracy for Portuguese language content.

4.2 VADER Sentiment Analysis

- A more sophisticated approach adapted for social media-style content.
- Delivered compound scores and intensity analysis for multilingual data.

4.3 RoBERTa-Based Analysis

- The most sophisticated approach, leveraging the transformers architecture.
- Delivered the most accurate correlation between user sentiment and contextual nuances.

4.4 Comparative Rating Analysis

The analysis revealed significant differences between platforms and rating methodologies:

App Store vs Google Play Store Ratings:

- **Apple App Store**: 1.4 (True rating).
- **Google Play Store**: 2.5 (True rating).

Overall User Ratings:

- Combined user rating: 2.04.
  - Google Play Store users: 2.38.
  - Apple App Store users: 1.61.

RoBERTa Sentiment-Based Ratings:

- Overall sentiment-based rating: 1.93.

- Google Play Store: 2.10.
- Apple App Store: 1.71.

The RoBERTa-based analysis proved most effective in capturing sentiment nuances and aligning them with user-provided ratings while considering contextual factors in the reviews.

5. Key Findings

5.1 Platform Disparities

- **Apple users** were generally more critical, with ratings averaging between 1.4 and 1.71.
- **Google Play users** showed moderately higher satisfaction, with ratings averaging between 2.1 and 2.5.

5.2 Sentiment Correlation

- RoBERTa ratings closely aligned with user-provided scores.
- The model demonstrated superior handling of technical terminology and application-specific context.

5.3 Temporal Trends

- A decline in user satisfaction was observed over time.
- Negative sentiment spikes were often correlated with technical issues or poorly received app updates.

5.4 Key Problem Areas Identified

Note: To identify the key problem areas, we used a curated sample of the most negative and the most positive reviews derived from the dataset. These reviews were further analyzed using Claude, an advanced AI tool, to extract deeper insights into user feedback. This approach allowed for a more detailed examination of user sentiments, recurring issues, and actionable feedback.

1. **Login and Access Issues**:
   - Problems with credentials and recovery mechanisms prevented users from accessing the application.
2. **Photo Attachment Functionality**:
   - Users reported issues attaching images to complaints.
3. **GPS and Location Services**:
   - Dependence on GPS and inaccuracies in location tracking limited functionality.
4. **Issue Resolution Tracking**:
   - Cases were frequently marked as resolved without effective action.
5. **User Interface Concerns**:
   - The interface was criticized for being unintuitive and technically flawed.

6. **Customer Support Responsiveness**:
  - Users noted a lack of feedback and support for unresolved issues.

6. Conclusions and Recommendations

The analysis highlights significant areas for improvement in the *Na Minha Rua* application. By addressing technical issues, enhancing user interface design, and improving customer support, the application can potentially increase user satisfaction and functionality. The study also underscores the importance of sentiment-based metrics in evaluating user feedback, particularly for applications with significant public engagement.

Imports

In [7]:
```python
# Imports
!pip install google_play_scraper
!pip install app_store_scraper
from google_play_scraper import reviews, Sort
import time
import pandas as pd
import numpy as np
import json
!pip install googletrans==4.0.0-rc1
from googletrans import Translator
from textblob import TextBlob
!pip install vaderSentiment
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
!pip install transformers
from transformers import pipeline
```

```
Requirement already satisfied: google_play_scraper in c:\users\joaom\anaconda3\li
b\site-packages (1.2.7)
Collecting app_store_scraper
  Using cached app_store_scraper-0.3.5-py3-none-any.whl.metadata (5.1 kB)
Collecting requests==2.23.0 (from app_store_scraper)
  Using cached requests-2.23.0-py2.py3-none-any.whl.metadata (6.8 kB)
Requirement already satisfied: chardet<4,>=3.0.2 in c:\users\joaom\anaconda3\lib
\site-packages (from requests==2.23.0->app_store_scraper) (3.0.4)
Requirement already satisfied: idna<3,>=2.5 in c:\users\joaom\anaconda3\lib\site-
packages (from requests==2.23.0->app_store_scraper) (2.10)
Collecting urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 (from requests==2.23.0->app_st
ore_scraper)
  Using cached urllib3-1.25.11-py2.py3-none-any.whl.metadata (41 kB)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\joaom\anaconda3\lib
\site-packages (from requests==2.23.0->app_store_scraper) (2024.6.2)
Using cached app_store_scraper-0.3.5-py3-none-any.whl (8.3 kB)
Using cached requests-2.23.0-py2.py3-none-any.whl (58 kB)
Using cached urllib3-1.25.11-py2.py3-none-any.whl (127 kB)
Installing collected packages: urllib3, requests, app_store_scraper
  Attempting uninstall: urllib3
    Found existing installation: urllib3 2.2.2
    Uninstalling urllib3-2.2.2:
      Successfully uninstalled urllib3-2.2.2
  Attempting uninstall: requests
    Found existing installation: requests 2.32.3
    Uninstalling requests-2.32.3:
      Successfully uninstalled requests-2.32.3
Successfully installed app_store_scraper-0.3.5 requests-2.23.0 urllib3-1.25.11
```

```
ERROR: pip's dependency resolver does not currently take into account all the pac
kages that are installed. This behaviour is the source of the following dependenc
y conflicts.
anaconda-client 1.12.3 requires urllib3>=1.26.4, but you have urllib3 1.25.11 whi
ch is incompatible.
conda 24.5.0 requires requests<3,>=2.28.0, but you have requests 2.23.0 which is
incompatible.
conda-repo-cli 1.0.88 requires requests>=2.31.0, but you have requests 2.23.0 whi
ch is incompatible.
jupyterlab-server 2.25.1 requires requests>=2.31, but you have requests 2.23.0 wh
ich is incompatible.
sphinx 7.3.7 requires requests>=2.25.0, but you have requests 2.23.0 which is inc
ompatible.
```

Google Play Scraping

In [72]:
```python
# Define the app's unique identifier for Google Play
app_id_google_play = 'pt.cml.naminharualx'

# Initialize an empty list to store reviews with comments
all_reviews_with_comments_google_play = []

# Set initial values
continuation_token = None
sort_order = Sort.NEWEST  # You can use Sort.NEWEST or Sort.RATING

# Loop to fetch all reviews
while True:
    try:
        result, continuation_token = reviews(
```

```python
            app_id_google_play,
            lang='pt',   # Language of the reviews
            country='PT',   # Country of the reviews
            sort=sort_order,   # Sort order of reviews
            count=200,   # Number of reviews to fetch per request
            continuation_token=continuation_token
        )

        # Check if there are reviews fetched
        if not result:
            print("No more reviews to fetch.")
            break

        # Filter reviews to only include those with comments
        reviews_with_comments = [review for review in result if review.get('cont
        all_reviews_with_comments_google_play.extend(reviews_with_comments)

        # Print the number of reviews fetched so far
        print(f"Fetched {len(reviews_with_comments)} reviews with comments, Tota

        # To avoid hitting the server too quickly, add a small delay
        time.sleep(1)

    except Exception as e:
        print(f"Error occurred: {e}")
        break

# Print the total number of reviews fetched
print(f"Total reviews with comments fetched from Google Play Store: {len(all_rev

# Display a sample of the reviews with comments
for review in all_reviews_with_comments_google_play[:5]:  # Displaying first 5 r
    print(f"User: {review['userName']}")
    print(f"Rating: {review['score']}")
    print(f"Review: {review['content']}")
    print('-' * 40)
reviews_google  = pd.json_normalize(all_reviews_with_comments_google_play)
reviews_google.head()
```

```
Fetched 164 reviews with comments, Total: 164
No more reviews to fetch.
Total reviews with comments fetched from Google Play Store: 164
User: Daniel Calado
Rating: 5
Review: Muito prático e eficaz
----------------------------------------
User: Nuno Pereira
Rating: 1
Review: Não é possível submeter, e-mail inválido!
----------------------------------------
User: João Araujo
Rating: 2
Review: Impossível aceder APP...
----------------------------------------
User: Aze Teka
Rating: 1
Review: Aplicação apresenta o erro, "não consegue apresentar informação" não está
funcional.
----------------------------------------
User: Paulo Silva
Rating: 2
Review: Uma excelente ideia pessimamente implementada.
----------------------------------------
```

Out[72]:

	reviewId	userName	
0	afb2456d-e407-4c4c-af81-1b1461767b22	Daniel Calado	https://play-lh.googleusercontent.com/a-/A E63GdnDjqIhlcUW4GOsd4i2hi336iAvl_
1	2cb8ac76-23d2-4b3f-bd94-b4b715339b15	Nuno Pereira	https://play-lh.googleuserconte UjUkI8_exlaXHFy8iB4kYCnY8Ran8mOPBlz5SOC
2	59b240cb-19f3-4405-957e-af66c21d511c	João Araujo	https://play-lh.googleusercontent.com/a/ACg8ocISuQMPOc lzjFpMGjdDn7a0yuuEP-z
3	1f64ee5f-5683-4b96-bc98-26120b269252	Aze Teka	https://play-lh.googleuserconte UjWmbY2x5JGZCdoFhjHFls5WyjOe2Wj-GitU
4	2d3369eb-05d5-4a1b-9730-de32f76e0e5a	Paulo Silva	lh.googleusercontent.com/a/ACg8ocIYJgQVPNN8Z9_QrH_AFSDV2 jA5

Apple Store Scraping

In [73]:

```python
from app_store_scraper import AppStore

all_reviews_with_comments_aplle_store = AppStore(country='pt', app_name='naminha

all_reviews_with_comments_aplle_store.review(how_many=2000)
all_reviews_with_comments_aplle_store = pd.DataFrame(np.array(all_reviews_with_c
reviews_apple = all_reviews_with_comments_aplle_store.join(pd.DataFrame(all_revi
reviews_apple.head()
```

```
2024-07-03 14:55:37,463 [INFO] Base - Initialised: AppStore('pt', 'naminharua-l
x', 1186245468)
2024-07-03 14:55:37,463 [INFO] Base - Ready to fetch reviews from: https://apps.a
pple.com/pt/app/naminharua-lx/id1186245468
2024-07-03 14:55:40,426 [INFO] Base - [id:1186245468] Fetched 127 reviews (127 fe
tched in total)
```

Out[73]:

	date	review	rating	isEdited	title	userName	developerResponse
0	2024-05-18 10:12:14	A app encerra abruptamente quando se pretende reportar uma ocorrência mesmo na parte final, o que invalida o tempo dispêndido a preencher o campo da descrição.	1	False	A app não funciona - dispositivo móvel com iOS	Utilizador lx	NaN
1	2024-07-01 15:26:25	Estou com o mesmo problema já reportado por outro utilizador. A app apaga depois de finalizados os passos todos do report.	1	False	Bloqueio da app	Habitante da graça	NaN
2	2022-01-07 18:38:00	A ideia da App é genial e realmente funciona pois a resposta da CML até é bastante rápida após o pedido, mas a App é pouco intuitiva, bloqueia muito, tem imensas opções deidentificação do pedido que para o utilizador não é prático...e só serve para os trabalhadores da cml. Acontence imensas vezes bloquear no mapa e quando se volta para tras para desbloquear	2	False	Ideia genial mas bloqueia muito	alibabacas	NaN

	date	review	rating	isEdited	title	userName	developerResponse
		entra numa pagina inicial com a imagem do electrico com um texto em latim e já não se consegue avançar para a frente, tem que se esperar um dia e voltar a entrar!					
3	2024-06-27 09:49:16	Demora tanto tempo a preencher e no fim crasha e não permite submeter	1	False	Aplicação encerra e não permite submeter	RafaCT1	NaN
4	2024-06-26 22:08:53	A aplicação crasha quando se submete uma ocorrência. Não vale a pena instalar. Antes, quando funcionava, até era útil mas vem se degradando desde 2019.	1	False	Aplicação crasha Não vale a pena instalar	hajatento	NaN

Google Reviews Cleaning

```
In [75]:   # Create a new column 'app' to indicate the source of the reviews
           reviews_google['app'] = 'google'

           reviews_google = reviews_google.drop(columns=['reviewId', 'userImage', 'thumbsUp
           reviews_google = reviews_google.rename(columns={'content': 'review', 'score': 'r
```

Apple Reviews Cleaning

```
In [77]:   # Create a new column 'app' to indicate the source of the reviews
           reviews_apple['app'] = 'apple'

           reviews_apple = reviews_apple.drop(columns=['isEdited', 'developerResponse'])
           reviews_apple = reviews_apple.rename(columns={'review': 'review1'})
           reviews_apple["review"] = reviews_apple['title'].astype(str) +"-"+ reviews_apple
```

```
In [79]:   pd.set_option('display.max_colwidth', None)
           reviews_apple['review'].head()
```

```
Out[79]:  0
          A app não funciona - dispositivo móvel com iOS-A app encerra abruptamente quand
          o se pretende reportar uma ocorrência mesmo na parte final, o que invalida o te
          mpo dispêndido a preencher o campo da descrição.
          1
          Bloqueio da app-Estou com o mesmo problema já reportado por outro utilizador. A
          app apaga depois de finalizados os passos todos do report.
          2     Ideia genial mas bloqueia muito-A ideia da App é genial e realmente funcio
          na pois a resposta da CML até é bastante rápida após o pedido, mas a App é pouc
          o intuitiva, bloqueia muito, tem imensas opções deidentificação do pedido que p
          ara o utilizador não é prático…e só serve para os trabalhadores da cml. Aconten
          ce imensas vezes bloquear no mapa e quando se volta para tras para desbloquear
          entra numa pagina inicial com a imagem do electrico com um texto em latim e já
          não se consegue avançar para a frente, tem que se esperar um dia e voltar a ent
          rar!
          3
          Aplicação encerra e não permite submeter-Demora tanto tempo a preencher e no fi
          m crasha e não permite submeter
          4
          Aplicação crasha Não vale a pena instalar-A aplicação crasha quando se submete
          uma ocorrência. Não vale a pena instalar. Antes, quando funcionava, até era úti
          l mas vem se degradando desde 2019.
          Name: review, dtype: object
```

```python
In [80]:  reviews_apple = reviews_apple.drop(columns=['review1', 'title'])
```

Combining DataFrames

```python
In [81]:  # Combine all DataFrames into a single DataFrame
          reviews = pd.concat([reviews_google, reviews_apple], ignore_index=True)
          reviews.head(100)
```

Out[81]:

	userName	review	rating	date	app
0	Daniel Calado	Muito prático e eficaz	5	2024-07-01 19:42:33	google
1	Nuno Pereira	Não é possível submeter, e-mail inválido!	1	2024-06-24 17:00:57	google
2	João Araujo	Impossível aceder APP...	2	2024-06-15 15:29:01	google
3	Aze Teka	Aplicação apresenta o erro, "não consegue apresentar informação" não está funcional.	1	2024-06-13 14:06:26	google
4	Paulo Silva	Uma excelente ideia pessimamente implementada.	2	2024-06-10 09:01:38	google
...	...	...	...	...	...
95	Um usuário do Google	O maior problema é a aplicação nao dar para criar ocorrencias sem ter o GPS ligado o que a torna inutil talve seja este o objectivo da CML pois com o GPS ligado nem sempre a localização fica correcta e da para a sua correção. Fazer no desktop é mais difícil por causa da passagem das fotos do telemóvel uma opção de desactivar o GPS seria mais eficaz só hoje tenho 6 que pode ser que alguém as reporte ou que o sinal de trânsito caia em cima de um carro ou alguém e que a rotura de água aumente ainda mais incrível tal reapoata. Pois nem sempre o GPS indica o local exacto da ocorrência e mudando o local ela não aceita daí inútil mais uma vez	1	2020-01-03 18:01:39	google
96	Um usuário do Google	Uso bastante esta app. Infelizmente a CML dá como resolvidos problemas recorrentes, como é o caso dos contentores subterrâneos que se transformaram em lixeiras a céu aberto,; a falta de estacionamento exclusivo para os moradores no centro da centro da cidade cuja naior parte dos prédios não têm garagens (como por exemplo a zona 14); colocação indiscriminada de ciclovias e de passeios pedonais, retirando estacionamento, faixas de rodagem e lugares de estacionamento (ex: Av. Defensores de Chaves).	2	2019-12-28 16:29:04	google
97	Um usuário do Google	Muito complicado e lento	1	2019-10-12 15:12:01	google
98	Um usuário do Google	Não serve para nada. Diz que as credenciais são inválidas, não permite a recuperação e aparece erro do servidor	1	2019-10-06 05:35:35	google

	userName	review	rating	date	app
99	Um usuário do Google	Mto mau mesmo	1	2019-09-28 22:25:35	google

100 rows × 5 columns

Data Cleaning Steps

In [83]:
```python
# Concatenated DataFrame

# Step 1: Check for Missing Values
missing_values = reviews.isnull().sum()
print("Missing values in each column:\n", missing_values)

# Step 2: Verify Data Types and Convert if Necessary
print("\nData types before conversion:\n", reviews.dtypes)

# Convert 'date' column to datetime
reviews['date'] = pd.to_datetime(reviews['date'], errors='coerce')

print("\nData types after conversion:\n", reviews.dtypes)

# Step 3: Check for Duplicates
duplicates = reviews.duplicated().sum()
print("\nNumber of duplicate rows:", duplicates)

# Remove duplicates
reviews = reviews.drop_duplicates()

# Step 4: Standardize Column Names (if not already done)
reviews.columns = reviews.columns.str.lower()

print("\nColumn names after standardization:\n", reviews.columns)

# Display the first few rows of the cleaned DataFrame
print("\nSample of the cleaned DataFrame:\n", reviews.head())
```

```
Missing values in each column:
 userName    0
review       0
rating       0
date         0
app          0
dtype: int64

Data types before conversion:
 userName          object
review            object
rating             int64
date      datetime64[ns]
app               object
dtype: object

Data types after conversion:
 userName          object
review            object
rating             int64
date      datetime64[ns]
app               object
dtype: object

Number of duplicate rows: 0

Column names after standardization:
 Index(['username', 'review', 'rating', 'date', 'app'], dtype='object')

Sample of the cleaned DataFrame:
         username  \
0  Daniel Calado
1    Nuno Pereira
2     João Araujo
3       Aze Teka
4     Paulo Silva


review  \
0                                         Muito prático e
eficaz
1                          Não é possível submeter, e-mail inv
álido!
2                                         Impossível aceder
APP...
3  Aplicação apresenta o erro, "não consegue apresentar informação" não está func
ional.
4                          Uma excelente ideia pessimamente impleme
ntada.

   rating                date     app
0       5 2024-07-01 19:42:33  google
1       1 2024-06-24 17:00:57  google
2       2 2024-06-15 15:29:01  google
3       1 2024-06-13 14:06:26  google
4       2 2024-06-10 09:01:38  google
```

Translating Reviews

In [88]:
```python
# Initialize the translator
translator = Translator()

# Function to translate reviews
def translate_review(review):
    try:
        translated = translator.translate(review, src='pt', dest='en')
        return translated.text
    except Exception as e:
        return str(e)

# Apply the translation function to the 'review' column
reviews['review_english'] = reviews['review'].apply(translate_review)

# Display a sample of the DataFrame to check the translations
print(reviews[['review', 'review_english']].head())
```

Requirement already satisfied: hyperframe<6,>=5.2.0 in c:\users\joaom\anaconda3\l
ib\site-packages (from h2==3.*->httpcore==0.9.*->httpx==0.13.3->googletrans==4.0.
0-rc1) (5.2.0)
Requirement already satisfied: hpack<4,>=3.0 in c:\users\joaom\anaconda3\lib\site
-packages (from h2==3.*->httpcore==0.9.*->httpx==0.13.3->googletrans==4.0.0-rc1)
(3.0.0)

```
                                               review  \
0                                    Muito prático e
eficaz
1                          Não é possível submeter, e-mail inv
álido!
2                                      Impossível aceder
APP...
3  Aplicação apresenta o erro, "não consegue apresentar informação" não está func
ional.
4                          Uma excelente ideia pessimamente impleme
ntada.

                                               review_engli
sh
0                                    Very practical and effecti
ve
1                          It is not possible to submit, invalid emai
l!
2                                      Impossible to access app
...
3  Application presents the error, "can not present information" is not functiona
l.
4                          An excellent idea badly implemente
d.
```

Exploratory Data Analysis (EDA)

In [85]:
```python
import matplotlib.pyplot as plt
import seaborn as sns

# 1. Ratings Distribution
plt.figure(figsize=(10, 6))
sns.countplot(x='rating', data=reviews, palette='viridis', hue='rating', dodge=F
plt.title('Distribution of Ratings')
```

```python
plt.xlabel('Rating')
plt.ylabel('Count')
plt.legend([], [], frameon=False)
plt.show()
#plt.savefig('ratings_distribution.png')

# 2. Review Length Analysis
reviews['review_length'] = reviews['review'].apply(len)

plt.figure(figsize=(10, 6))
sns.histplot(reviews['review_length'], kde=True, color='blue')
plt.title('Distribution of Review Lengths')
plt.xlabel('Review Length (characters)')
plt.ylabel('Count')
plt.show()
#plt.savefig('review_length_distribution.png')

# Temporal Trends
plt.figure(figsize=(14, 7))
reviews.set_index('date').resample('D').size().plot(kind='line')
plt.title('Number of Reviews Over Time')
plt.xlabel('Date')
plt.ylabel('Number of Reviews')
plt.show()
#plt.savefig('reviews_over_time.png')
```

```
2024-07-03 14:55:40,563 [INFO] matplotlib.category - Using categorical units to p
lot a list of strings that are all parsable as floats or dates. If these strings
should be plotted as numbers, cast to the appropriate data type before plotting.
2024-07-03 14:55:40,567 [INFO] matplotlib.category - Using categorical units to p
lot a list of strings that are all parsable as floats or dates. If these strings
should be plotted as numbers, cast to the appropriate data type before plotting.
```


Distribution of Ratings

Distribution of Review Lengths

Number of Reviews Over Time
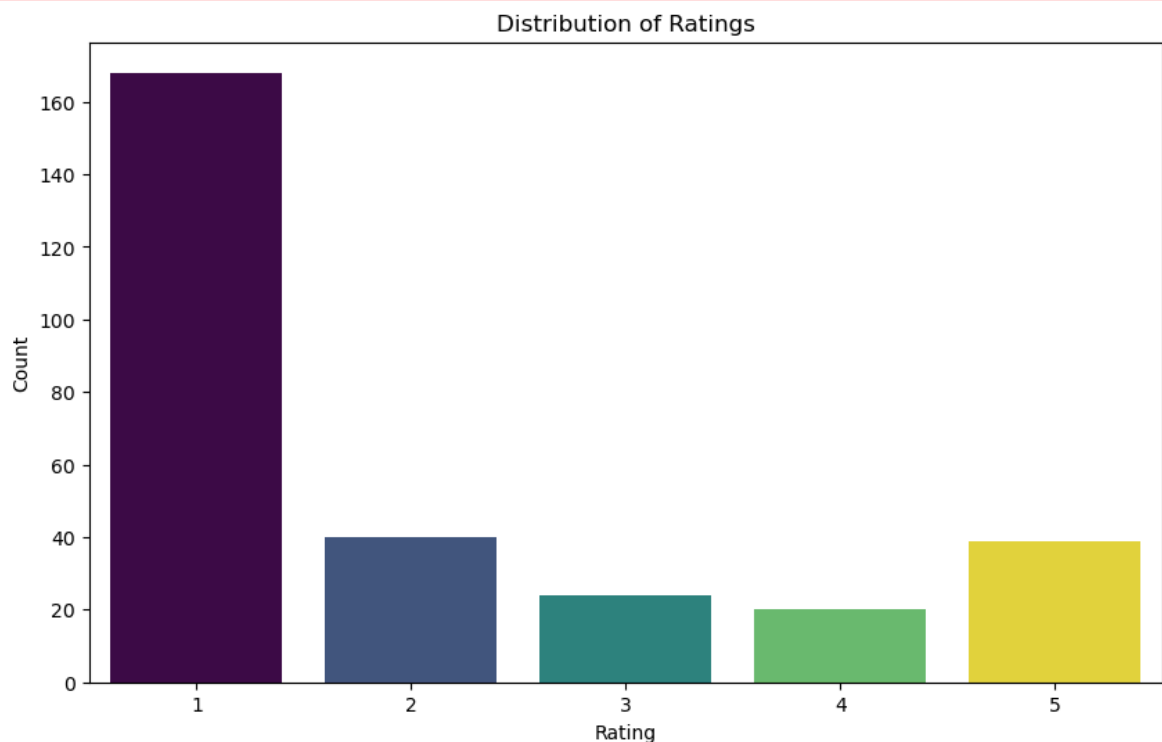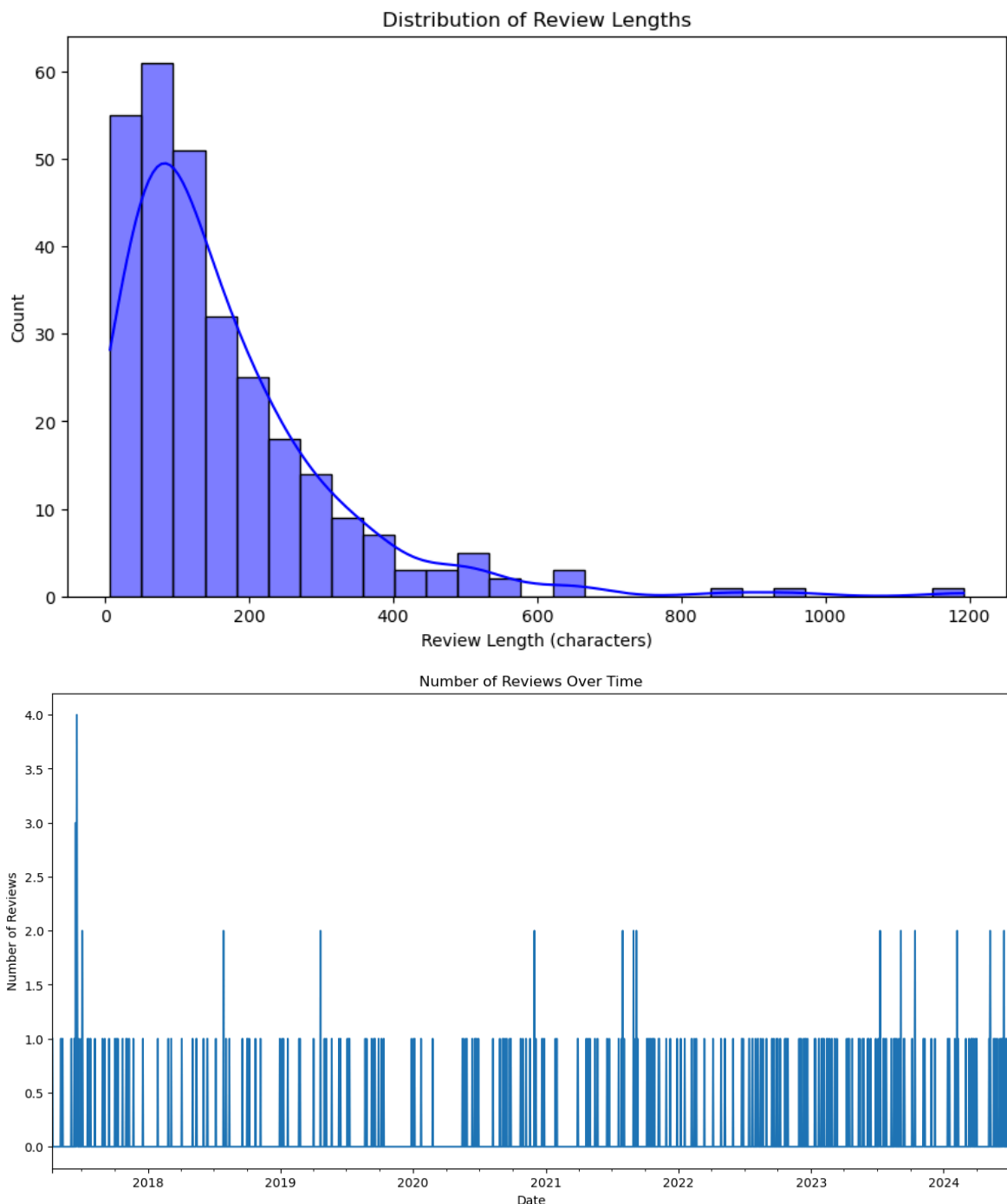
Text Blob Sentiment

Note: TextBlob provides a polarity score directly, but it does not provide separate negative, neutral, positive, or compound scores like VADER and Roberta.

```python
In [90]:  # Function to analyze sentiment
          def analyze_sentiment(review):
              analysis = TextBlob(review)
              # Get the polarity score
              polarity = analysis.sentiment.polarity
              # Classify the polarity
              if polarity > 0:
                  sentiment = 'Positive'
              elif polarity == 0:
                  sentiment = 'Neutral'
```

```
    else:
        sentiment = 'Negative'
    return sentiment, polarity

# Apply the sentiment analysis function to the translated reviews
reviews['textblob_sentiment'], reviews['textblob_polarity'] = zip(*reviews['revi

# Display a sample of the DataFrame to check the sentiment labels and polarity v
print(reviews[['review_english', 'textblob_sentiment', 'textblob_polarity']].hea
```

```
                                                         review_engli
sh  \
0                                      Very practical and effecti
ve
1                           It is not possible to submit, invalid emai
l!
2                                          Impossible to access app
...
3  Application presents the error, "can not present information" is not functiona
l.
4                                      An excellent idea badly implemente
d.

  textblob_sentiment  textblob_polarity
0           Positive           0.400000
1            Neutral           0.000000
2           Negative          -0.666667
3            Neutral           0.000000
4           Positive           0.150000
```

Plot Text blob Sentiment

In [92]:
```
# Display a sample of the DataFrame to check the sentiment labels and polarity v
print(reviews[['review', 'textblob_sentiment', 'rating']].head())
```

```
review  \
0                                              Muito prático e
eficaz
1                                 Não é possível submeter, e-mail inv
álido!
2                                            Impossível aceder
APP...
3  Aplicação apresenta o erro, "não consegue apresentar informação" não está func
ional.
4                                 Uma excelente ideia pessimamente impleme
ntada.

  textblob_sentiment  rating
0           Positive       5
1            Neutral       1
2           Negative       2
3            Neutral       1
4           Positive       2
```

Vader Sentiment Scoring - Valence Aware Dictionary and Sentiment Reasoner

In [95]:
```python
# Initialize the VADER sentiment analyzer
analyzer = SentimentIntensityAnalyzer()
```

In [96]:
```python
analyzer.polarity_scores('Very practical and effective')
```

Out[96]: `{'neg': 0.0, 'neu': 0.471, 'pos': 0.529, 'compound': 0.5209}`

In [97]:
```python
# Function to analyze sentiment using VADER
def vader_sentiment(review):
    vs = analyzer.polarity_scores(review)
    # Classify the compound score into categories
    if vs['compound'] >= 0.05:
        sentiment = 'Positive'
    elif vs['compound'] <= -0.05:
        sentiment = 'Negative'
    else:
        sentiment = 'Neutral'
    return sentiment, vs['neg'], vs['neu'], vs['pos'], vs['compound']

# Apply the VADER sentiment analysis function to the translated reviews
reviews['vader_sentiment'], reviews['vader_neg'], reviews['vader_neu'], reviews[

# Display a sample of the DataFrame to check the sentiment labels and polarity v
print(reviews[['review_english', 'vader_sentiment', 'vader_neg', 'vader_neu', 'v
```

```
                                             review_engli
sh  \
0                             Very practical and effecti
ve
1                      It is not possible to submit, invalid emai
l!
2                                   Impossible to access app
...
3  Application presents the error, "can not present information" is not functiona
l.
4                             An excellent idea badly implemente
d.

   vader_sentiment  vader_neg  vader_neu  vader_pos  vader_compound
0         Positive      0.000      0.471      0.529          0.5209
1          Neutral      0.000      1.000      0.000          0.0000
2          Neutral      0.000      1.000      0.000          0.0000
3         Negative      0.213      0.787      0.000         -0.4019
4         Positive      0.316      0.306      0.378          0.1531
```

Explanation:

- VADER Sentiment Analyzer: VADER is a lexicon and rule-based sentiment analysis tool specifically attuned to sentiments expressed in social media.

- Compound Score: The compound score is a normalized score ranging from -1 (most negative) to +1 (most positive). We use thresholds (e.g., >= 0.05 for positive, <=

-0.05 for negative) to classify sentiments.

In [100…
```python
# Visualisation to create customised sentiment rules
# Filter DataFrame for each sentiment category
positive_reviews = reviews[reviews['vader_sentiment'] == 'Positive']
negative_reviews = reviews[reviews['vader_sentiment'] == 'Negative']
neutral_reviews = reviews[reviews['vader_sentiment'] == 'Neutral']

# Save each category to a separate CSV file
positive_reviews.to_csv('positive_reviews.csv', index=False)
negative_reviews.to_csv('negative_reviews.csv', index=False)
neutral_reviews.to_csv('neutral_reviews.csv', index=False)

print("Separate CSV files have been saved for each sentiment category:")
print("- Positive reviews saved to 'positive_reviews.csv'")
print("- Negative reviews saved to 'negative_reviews.csv'")
print("- Neutral reviews saved to 'neutral_reviews.csv'")

# Ensure that review_english column contains strings
positive_reviews.loc[:, 'review_english'] = positive_reviews['review_english'].a
negative_reviews.loc[:, 'review_english'] = negative_reviews['review_english'].a
neutral_reviews.loc[:, 'review_english'] = neutral_reviews['review_english'].ast
```

```
Separate CSV files have been saved for each sentiment category:
- Positive reviews saved to 'positive_reviews.csv'
- Negative reviews saved to 'negative_reviews.csv'
- Neutral reviews saved to 'neutral_reviews.csv'
```

In [101…
```python
# Initialize the VADER sentiment analyzer
analyzer = SentimentIntensityAnalyzer()

# Function to analyze sentiment using VADER
def vader_sentiment(review):
    vs = analyzer.polarity_scores(review)
    # Classify the compound score into categories
    if vs['compound'] >= 0.05:
        sentiment = 'Positive'
    elif vs['compound'] <= -0.05:
        sentiment = 'Negative'
    else:
        sentiment = 'Neutral'
    return sentiment, vs['neg'], vs['neu'], vs['pos'], vs['compound']

# Apply the VADER sentiment analysis function to the reviews
reviews['vader_sentiment'], reviews['vader_neg'], reviews['vader_neu'], reviews[

# Function to map custom sentiment rules
def custom_sentiment(review, default_sentiment, neg, neu, pos, compound):
    # Define custom rules
    negative_phrases = [
        "not possible", "impossible", "error", "badly implemented",
        "no effective response", "crash", "doesn't work", "waste of time",
        "bugs", "poor", "mediocre"
    ]
    positive_phrases = [
        "excellent", "great", "love", "amazing", "fantastic",
        "very good", "well done", "works perfectly"
    ]
    neutral_phrases = [
        "okay", "average", "not bad", "decent"
```

```python
    ]

    # Check for negative phrases
    for phrase in negative_phrases:
        if phrase in review.lower():
            return 'Negative', neg, neu, pos, compound

    # Check for positive phrases
    for phrase in positive_phrases:
        if phrase in review.lower():
            return 'Positive', neg, neu, pos, compound

    # Check for neutral phrases
    for phrase in neutral_phrases:
        if phrase in review.lower():
            return 'Neutral', neg, neu, pos, compound

    # Return default sentiment if no custom rules apply
    return default_sentiment, neg, neu, pos, compound

# Apply custom rules to the VADER sentiment results
reviews[['vader_sentiment_updated', 'vader_neg_updated', 'vader_neu_updated', 'v
    lambda row: custom_sentiment(row['review_english'], row['vader_sentiment'],
)

# Display a sample of the DataFrame to check the final sentiment labels and comp
print(reviews[['review_english', 'vader_sentiment_updated', 'vader_neg_updated',
```

```
                                                        review_engli
sh  \
0                                    Very practical and effecti
ve
1                        It is not possible to submit, invalid emai
l!
2                                        Impossible to access app
...
3  Application presents the error, "can not present information" is not functiona
l.
4                                    An excellent idea badly implemente
d.

   vader_sentiment_updated  vader_neg_updated  vader_neu_updated  \
0                 Positive              0.000              0.471
1                 Negative              0.000              1.000
2                 Negative              0.000              1.000
3                 Negative              0.213              0.787
4                 Negative              0.316              0.306

   vader_pos_updated  vader_compound_updated
0              0.529                  0.5209
1              0.000                  0.0000
2              0.000                  0.0000
3              0.000                 -0.4019
4              0.378                  0.1531
```

```python
In [102…  # Display a sample of the DataFrame to check and compare the final sentiment lab
          print(reviews[['review_english', 'vader_sentiment', 'vader_sentiment_updated', '
```

```
                                                          review_engli
sh  \
0                                         Very practical and effecti
ve
1                               It is not possible to submit, invalid emai
l!
2                                         Impossible to access app
...
3  Application presents the error, "can not present information" is not functiona
l.
4                                         An excellent idea badly implemente
d.

   vader_sentiment vader_sentiment_updated  vader_compound  \
0         Positive                Positive          0.5209
1          Neutral                Negative          0.0000
2          Neutral                Negative          0.0000
3         Negative                Negative         -0.4019
4         Positive                Negative          0.1531

   vader_compound_updated
0                  0.5209
1                  0.0000
2                  0.0000
3                 -0.4019
4                  0.1531
```

In [104…

```python
# Create a bar plot with rating on the x-axis and vader_compound_updated on the
plt.figure(figsize=(10, 6))
ax = sns.barplot(data=reviews, x='rating', y='vader_compound_updated', palette='
ax.set_title('Compound Score by Apps Star Review')
ax.set_xlabel('Rating')
ax.set_ylabel('VADER Compound Score (Updated)')
plt.show()
```

```
2024-07-03 14:56:37,243 [INFO] matplotlib.category - Using categorical units to p
lot a list of strings that are all parsable as floats or dates. If these strings
should be plotted as numbers, cast to the appropriate data type before plotting.
C:\Users\joaom\AppData\Local\Temp\ipykernel_8608\1149685057.py:3: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v
0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effe
ct.

  ax = sns.barplot(data=reviews, x='rating', y='vader_compound_updated', palette
='viridis')
2024-07-03 14:56:37,250 [INFO] matplotlib.category - Using categorical units to p
lot a list of strings that are all parsable as floats or dates. If these strings
should be plotted as numbers, cast to the appropriate data type before plotting.
```

Compound Score by Apps Star Review
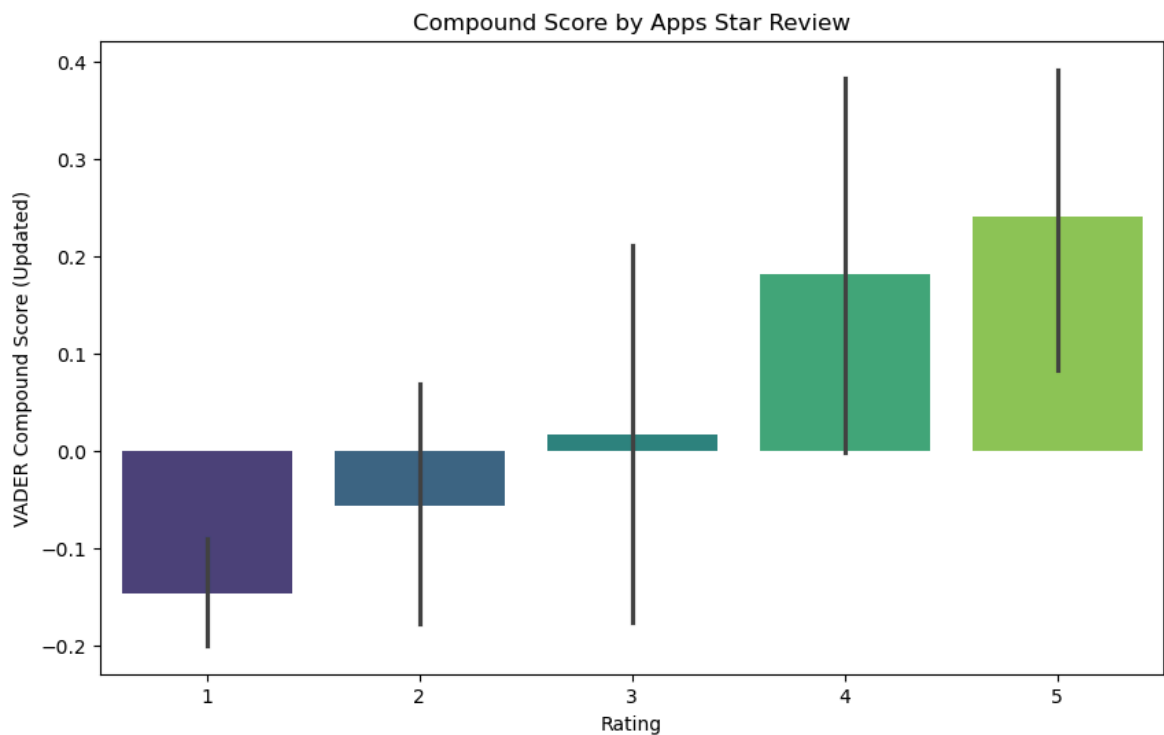
Plot Vader Results

```
In [105…  import matplotlib.pyplot as plt
          import seaborn as sns

          # Create a figure and three subplots
          fig, axs = plt.subplots(1, 3, figsize=(18, 6))

          # Plot the negative sentiment scores
          sns.barplot(data=reviews, x='rating', y='vader_neg_updated', palette='viridis',
          axs[0].set_title('Negative Sentiment Score by Rating')
          axs[0].set_xlabel('Rating')
          axs[0].set_ylabel('VADER Negative Score (Updated)')

          # Plot the neutral sentiment scores
          sns.barplot(data=reviews, x='rating', y='vader_neu_updated', palette='viridis',
          axs[1].set_title('Neutral Sentiment Score by Rating')
          axs[1].set_xlabel('Rating')
          axs[1].set_ylabel('VADER Neutral Score (Updated)')

          # Plot the positive sentiment scores
          sns.barplot(data=reviews, x='rating', y='vader_pos_updated', palette='viridis',
          axs[2].set_title('Positive Sentiment Score by Rating')
          axs[2].set_xlabel('Rating')
          axs[2].set_ylabel('VADER Positive Score (Updated)')

          # Adjust layout
          plt.tight_layout()

          # Show the plots
          plt.show()
```

```
2024-07-03 14:56:37,440 [INFO] matplotlib.category - Using categorical units to p
lot a list of strings that are all parsable as floats or dates. If these strings
should be plotted as numbers, cast to the appropriate data type before plotting.
C:\Users\joaom\AppData\Local\Temp\ipykernel_8608\3901910763.py:8: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v
0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effe
ct.

  sns.barplot(data=reviews, x='rating', y='vader_neg_updated', palette='viridis',
ax=axs[0])
2024-07-03 14:56:37,445 [INFO] matplotlib.category - Using categorical units to p
lot a list of strings that are all parsable as floats or dates. If these strings
should be plotted as numbers, cast to the appropriate data type before plotting.
2024-07-03 14:56:37,536 [INFO] matplotlib.category - Using categorical units to p
lot a list of strings that are all parsable as floats or dates. If these strings
should be plotted as numbers, cast to the appropriate data type before plotting.
C:\Users\joaom\AppData\Local\Temp\ipykernel_8608\3901910763.py:14: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v
0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effe
ct.

  sns.barplot(data=reviews, x='rating', y='vader_neu_updated', palette='viridis',
ax=axs[1])
2024-07-03 14:56:37,540 [INFO] matplotlib.category - Using categorical units to p
lot a list of strings that are all parsable as floats or dates. If these strings
should be plotted as numbers, cast to the appropriate data type before plotting.
2024-07-03 14:56:37,631 [INFO] matplotlib.category - Using categorical units to p
lot a list of strings that are all parsable as floats or dates. If these strings
should be plotted as numbers, cast to the appropriate data type before plotting.
C:\Users\joaom\AppData\Local\Temp\ipykernel_8608\3901910763.py:20: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v
0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effe
ct.

  sns.barplot(data=reviews, x='rating', y='vader_pos_updated', palette='viridis',
ax=axs[2])
2024-07-03 14:56:37,635 [INFO] matplotlib.category - Using categorical units to p
lot a list of strings that are all parsable as floats or dates. If these strings
should be plotted as numbers, cast to the appropriate data type before plotting.
```
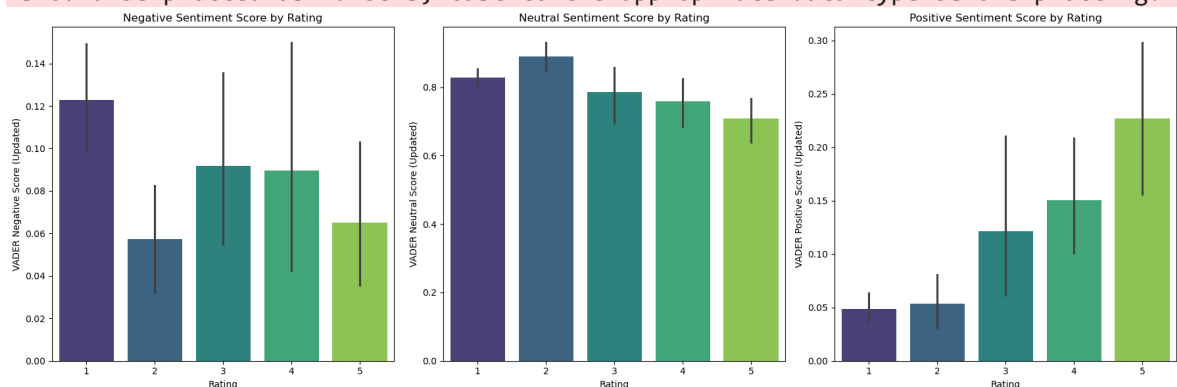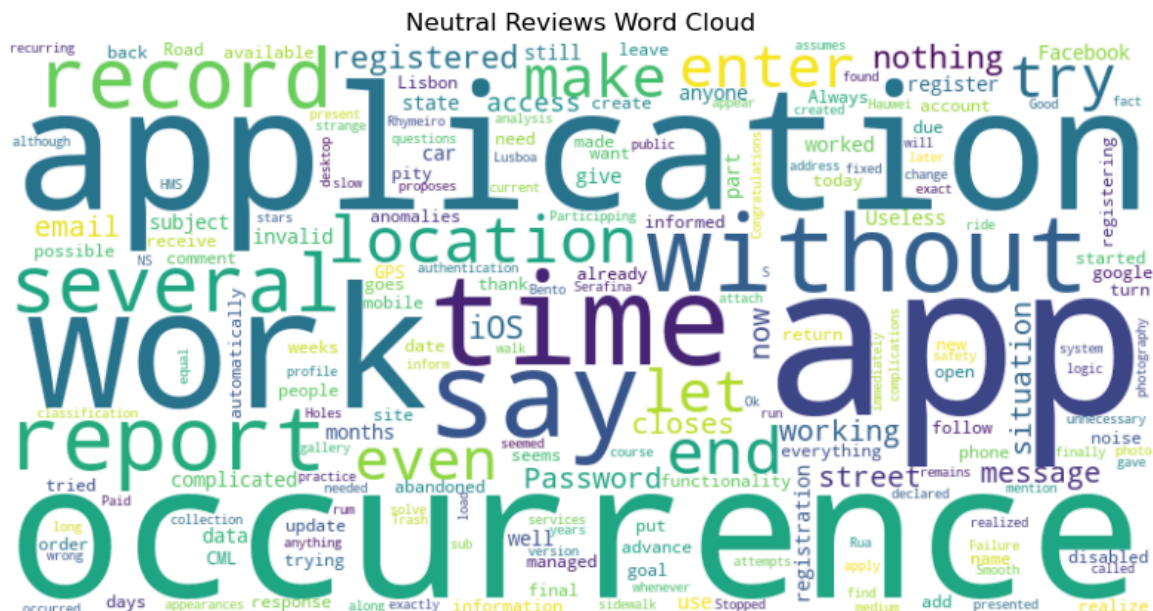


```python
from wordcloud import WordCloud

# Function to generate and display a word cloud
def plot_wordcloud(text, title):
    wordcloud = WordCloud(width=800, height=400, background_color='white').gener
```

```python
    plt.figure(figsize=(10, 6))
    plt.imshow(wordcloud, interpolation='bilinear')
    plt.title(title)
    plt.axis('off')
    plt.show()

# Generate word clouds for each sentiment
positive_reviews = ' '.join(reviews[reviews['vader_sentiment_updated'] == 'Posit
negative_reviews = ' '.join(reviews[reviews['vader_sentiment_updated'] == 'Negat
neutral_reviews = ' '.join(reviews[reviews['vader_sentiment_updated'] == 'Neutra

plot_wordcloud(positive_reviews, 'Positive Reviews Word Cloud')
plot_wordcloud(negative_reviews, 'Negative Reviews Word Cloud')
plot_wordcloud(neutral_reviews, 'Neutral Reviews Word Cloud')
```


Positive Reviews Word Cloud

Negative Reviews Word Cloud

Neutral Reviews Word Cloud

RoBERTa for Sentiment Analysis - Robustly optimized Bidirectional Encoder Representation from Transformers approach

Note:

- Roberta is a model trained of a large corpus of data.
- Roberta accounts for the words but also the context related to other words.
- In the transformers library by Hugging Face, the pipeline function for sentiment analysis automatically handles tokenization, so we do not need to tokenize the text manually before using it. The pipeline function takes care of all necessary preprocessing steps, including tokenization.

```python
In [109...  # Loading the sentiment analysis pipeline with a pre-trained RoBERTa model
           sentiment_analysis = pipeline("sentiment-analysis", model="cardiffnlp/twitter-ro
```

```
Requirement already satisfied: chardet<4,>=3.0.2 in c:\users\joaom\anaconda3\lib
\site-packages (from requests->transformers) (3.0.4)
Requirement already satisfied: idna<3,>=2.5 in c:\users\joaom\anaconda3\lib\site-
packages (from requests->transformers) (2.10)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in c:\user
s\joaom\anaconda3\lib\site-packages (from requests->transformers) (1.25.11)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\joaom\anaconda3\lib
\site-packages (from requests->transformers) (2024.6.2)
```

```python
In [110...  # Loading the sentiment analysis pipeline with a pretrained RoBERTa model
           sentiment_analysis = pipeline("sentiment-analysis", model="cardiffnlp/twitter-ro

           # Loading the CSV files
           positive_reviews = pd.read_csv('positive_reviews.csv')
           negative_reviews = pd.read_csv('negative_reviews.csv')
           neutral_reviews = pd.read_csv('neutral_reviews.csv')
```

```python
# Extract reviews for testing (first few lines)
positive_sample_reviews = positive_reviews['review_english'].head().tolist()
negative_sample_reviews = negative_reviews['review_english'].head().tolist()
neutral_sample_reviews = neutral_reviews['review_english'].head().tolist()

# Function to analyze sentiment and print results
def analyze_and_print_reviews(sample_reviews, category):
    results = sentiment_analysis(sample_reviews)
    print(f"Sentiment analysis for {category} reviews:")
    for review, result in zip(sample_reviews, results):
        print(f"Review: {review}\nSentiment: {result['label']}, Score: {result['

# Analyze and print sentiments
analyze_and_print_reviews(positive_sample_reviews, 'Positive')
analyze_and_print_reviews(negative_sample_reviews, 'Negative')
analyze_and_print_reviews(neutral_sample_reviews, 'Neutral')
```

Sentiment analysis for Positive reviews:
Review: Very practical and effective
Sentiment: LABEL_2, Score: 0.8686820268630981

Review: An excellent idea badly implemented.
Sentiment: LABEL_2, Score: 0.644403874874115

Review: When consulting an occurrence (for example, through the map), one of the
fields that should be visible is the date of creation of the occurrence.Unless th
e author includes the date in the description of the occurrence, we are unaware w
hen it was registered.
Sentiment: LABEL_1, Score: 0.7786567211151123

Review: I can't make the complaint I want- on the phone says it's unavailable
?????????
Sentiment: LABEL_0, Score: 0.944402277469635

Review: Occurrences of "public and noisy security" are not visible to other user
s, unlike other occurrences!
Sentiment: LABEL_0, Score: 0.4777948558330536

Sentiment analysis for Negative reviews:
Review: Application presents the error, "can not present information" is not func
tional.
Sentiment: LABEL_0, Score: 0.9097289443016052

Review: I got no effective response beyond "we will analyze your complaint."
Sentiment: LABEL_0, Score: 0.6881999373435974

Review: Close the cases without solving them.It is pure beam and waste of time.
Sentiment: LABEL_0, Score: 0.8699219822883606

Review: I can't participate occurrences ... does not accept photos
Sentiment: LABEL_0, Score: 0.9084978103637695

Review: The main failure is not in the app, is in the response of CML services to
occurrences
Sentiment: LABEL_0, Score: 0.8963856101036072

Sentiment analysis for Neutral reviews:
Review: It is not possible to submit, invalid email!
Sentiment: LABEL_0, Score: 0.9323604106903076

Review: Impossible to access app ...
Sentiment: LABEL_0, Score: 0.6867159008979797

Review: It has not been possible to access the occurred
Sentiment: LABEL_0, Score: 0.7232295274734497

Review: Holes in the without sidewalk ride that make people walk on Rua S.Bento,
along with NS 127/129 in Lisbon
Sentiment: LABEL_1, Score: 0.6460368633270264

Review: I can't change the address.App back back whenever I try to do it.
Sentiment: LABEL_1, Score: 0.6481189727783203

- LABEL_0 corresponds to Negative
- LABEL_1 corresponds to Neutral

- LABEL_2 corresponds to Positive

In [112...

```python
# Initialize the RoBERTa sentiment analysis pipeline
sentiment_analysis = pipeline("sentiment-analysis", model="cardiffnlp/twitter-ro

# Function to map RoBERTa labels to human-readable sentiments and extract scores
def map_label_to_sentiment(result):
    label_map = {
        "LABEL_0": "Negative",  # LABEL_0 corresponds to Negative
        "LABEL_1": "Neutral",   # LABEL_1 corresponds to Neutral
        "LABEL_2": "Positive"   # LABEL_2 corresponds to Positive
    }
    sentiment = label_map[result[0]['label']]
    score = result[0]['score']

    # Create separate scores for each sentiment based on the predicted sentiment
    if sentiment == 'Negative':
        neg_score = score
        neu_score = (1 - score) / 2
        pos_score = (1 - score) / 2
    elif sentiment == 'Neutral':
        neg_score = (1 - score) / 2
        neu_score = score
        pos_score = (1 - score) / 2
    else:
        neg_score = (1 - score) / 2
        neu_score = (1 - score) / 2
        pos_score = score

    return sentiment, neg_score, neu_score, pos_score

# Function to get sentiment using RoBERTa and map to human-readable labels and s
def get_roberta_sentiment(review):
    result = sentiment_analysis(review)
    return map_label_to_sentiment(result)

# Apply the RoBERTa sentiment analysis function to the reviews
reviews['roberta_sentiment'], reviews['roberta_neg'], reviews['roberta_neu'], re

# Display a sample of the DataFrame to check the sentiment labels and scores
print(reviews[['review_english', 'roberta_sentiment', 'roberta_neg', 'roberta_ne
```

```
                                                            review_engli
sh  \
0                                         Very practical and effecti
ve
1                                It is not possible to submit, invalid emai
l!
2                                           Impossible to access app
...
3  Application presents the error, "can not present information" is not functiona
l.
4                                          An excellent idea badly implemente
d.

   roberta_sentiment  roberta_neg  roberta_neu  roberta_pos
0           Positive     0.065659     0.065659     0.868682
1           Negative     0.932360     0.033820     0.033820
2           Negative     0.686716     0.156642     0.156642
3           Negative     0.909729     0.045136     0.045136
4           Positive     0.177798     0.177798     0.644404
```

Comparison and Conclusions

In [143...

```python
# Platform-wise Sentiment Distribution
plt.figure(figsize=(12, 6))
sns.countplot(x='vader_sentiment_updated', hue='app', data=reviews, palette='vir
plt.title('Sentiment Distribution by Platform')
plt.xlabel('Sentiment')
plt.ylabel('Count')
plt.legend(title='Platform')
plt.show()
```
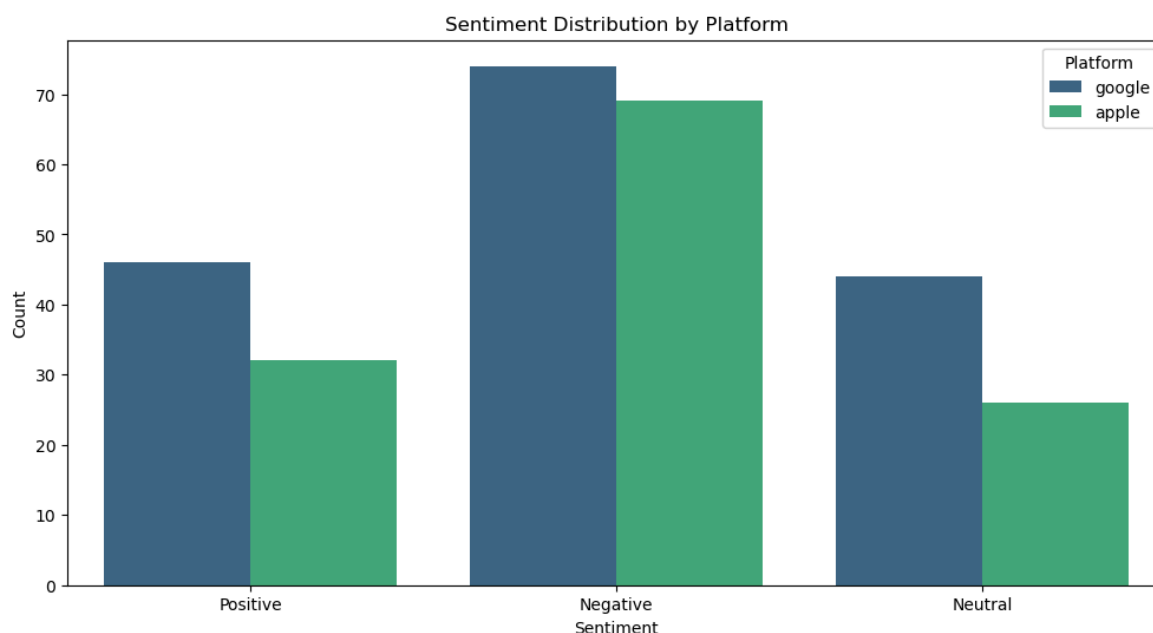

In [114...

```python
import matplotlib.pyplot as plt
import seaborn as sns

# 1. Visualize Sentiment Distribution
plt.figure(figsize=(10, 6))
sns.countplot(x='vader_sentiment_updated', data=reviews, palette='viridis')
plt.title('Distribution of Final Sentiments')
```

```python
plt.xlabel('Sentiment')
plt.ylabel('Count')
plt.show()

# 2. Sentiment Trends Over Time
plt.figure(figsize=(14, 7))
reviews.set_index('date').resample('M').vader_sentiment_updated.value_counts().u
plt.title('Sentiment Trends Over Time')
plt.xlabel('Date')
plt.ylabel('Count of Sentiments')
plt.show()
```

C:\Users\joaom\AppData\Local\Temp\ipykernel_8608\1299744140.py:6: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v
0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effe
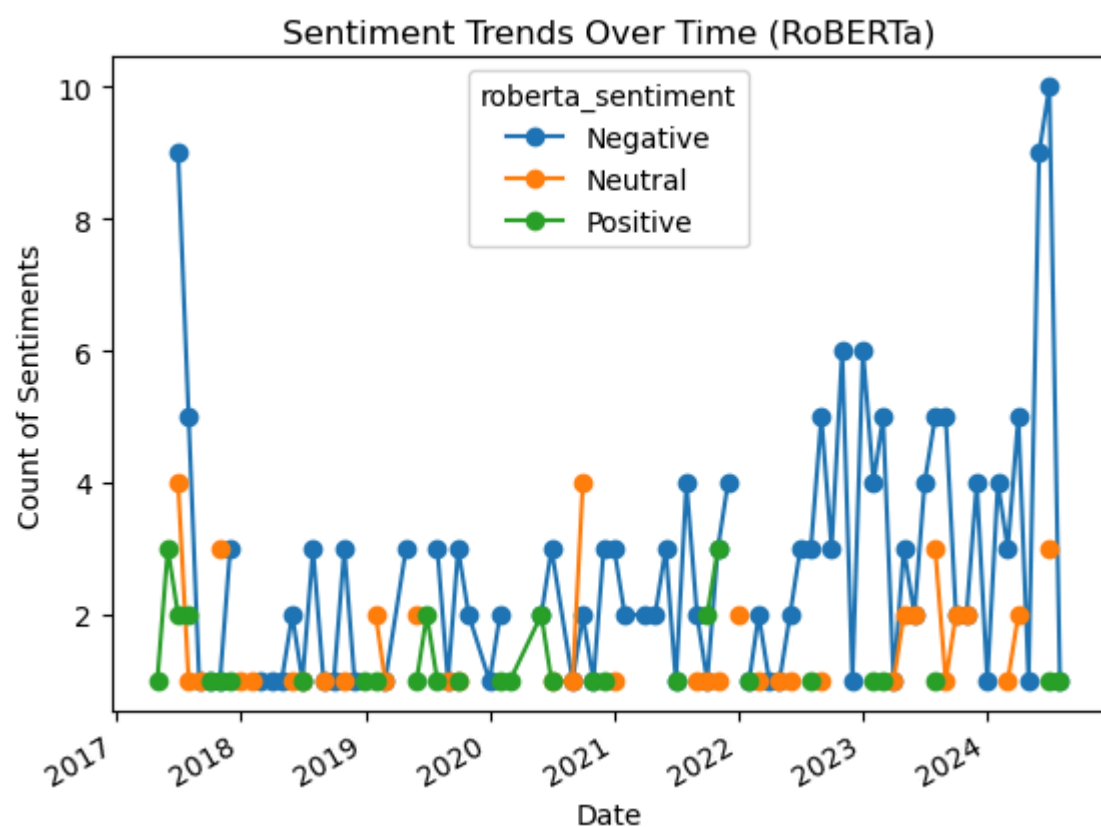ct.

  sns.countplot(x='vader_sentiment_updated', data=reviews, palette='viridis')

Distribution of Final Sentiments
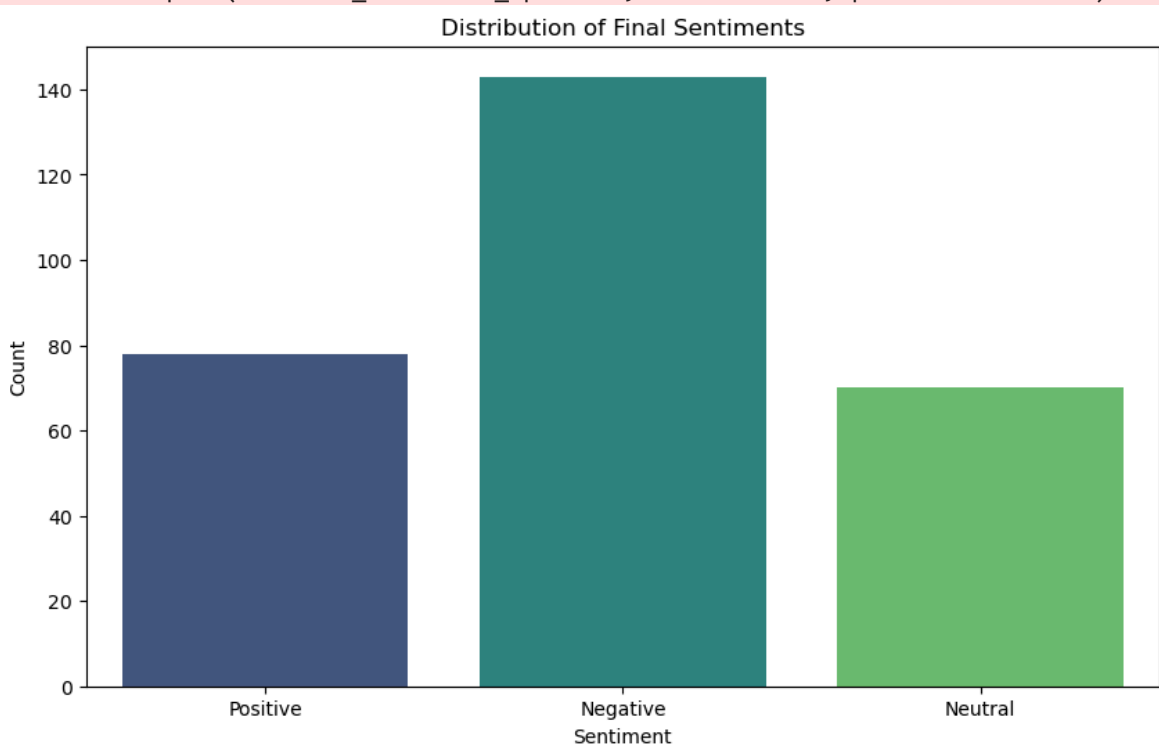
C:\Users\joaom\AppData\Local\Temp\ipykernel_8608\1299744140.py:14: FutureWarning:
'M' is deprecated and will be removed in a future version, please use 'ME' instea
d.
  reviews.set_index('date').resample('M').vader_sentiment_updated.value_counts().
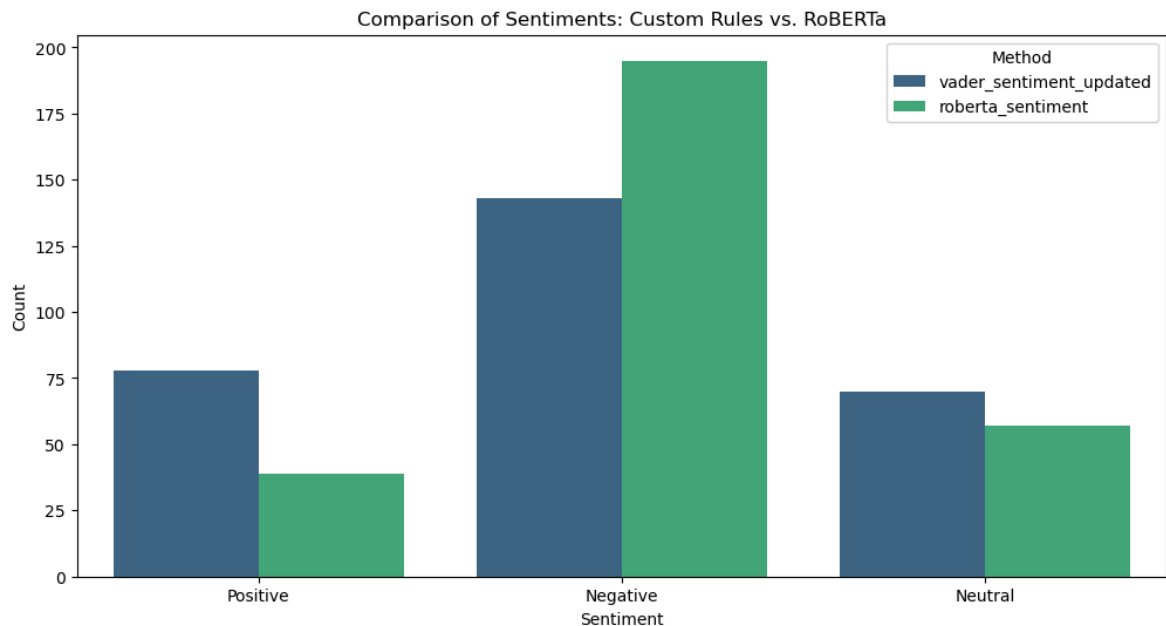unstack().plot(kind='line', marker='o')
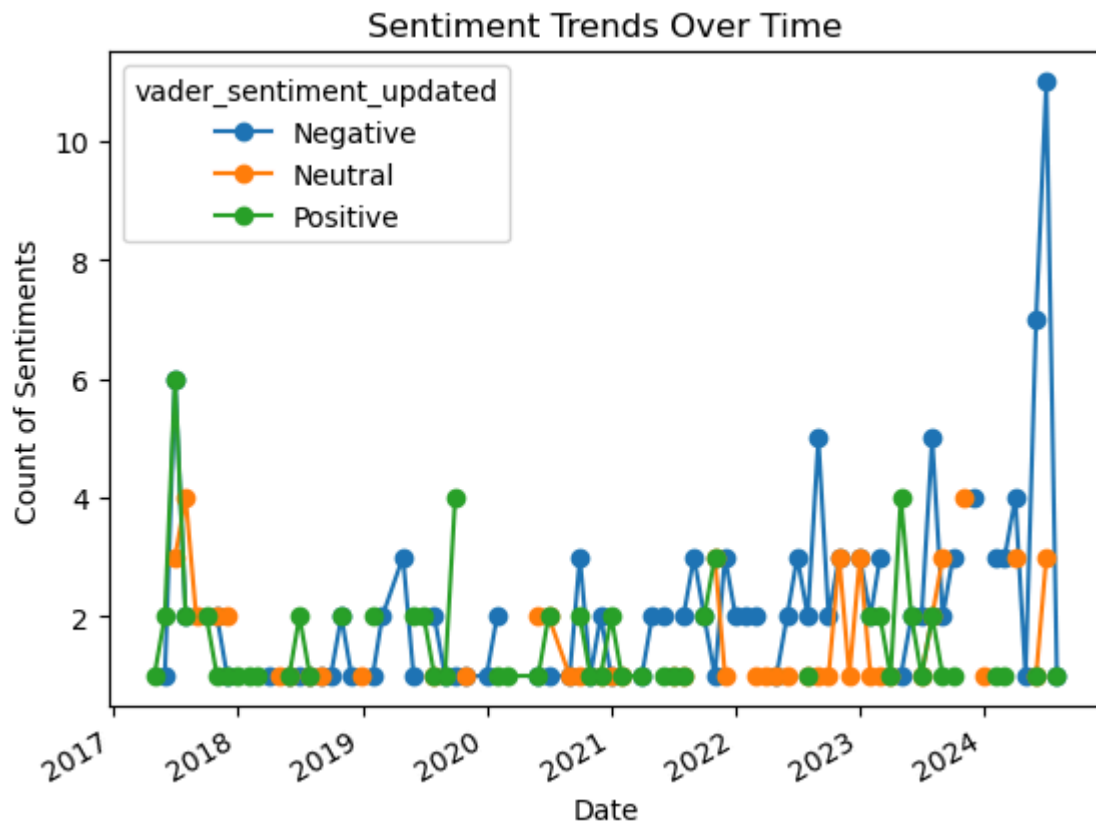
```
<Figure size 1400x700 with 0 Axes>
```

Sentiment Trends Over Time


```python
import matplotlib.pyplot as plt
import seaborn as sns

# 1. Visualize Sentiment Distribution
plt.figure(figsize=(10, 6))
sns.countplot(x='roberta_sentiment', data=reviews, palette='viridis')
plt.title('Distribution of RoBERTa-Derived Sentiments')
plt.xlabel('Sentiment')
plt.ylabel('Count')
plt.show()

# 2. Sentiment Trends Over Time
plt.figure(figsize=(14, 7))
reviews.set_index('date').resample('M').roberta_sentiment.value_counts().unstack
plt.title('Sentiment Trends Over Time (RoBERTa)')
plt.xlabel('Date')
plt.ylabel('Count of Sentiments')
plt.show()
```

```
C:\Users\joaom\AppData\Local\Temp\ipykernel_8608\1906245679.py:6: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v
0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effe
ct.

  sns.countplot(x='roberta_sentiment', data=reviews, palette='viridis')
```

Distribution of RoBERTa-Derived Sentiments
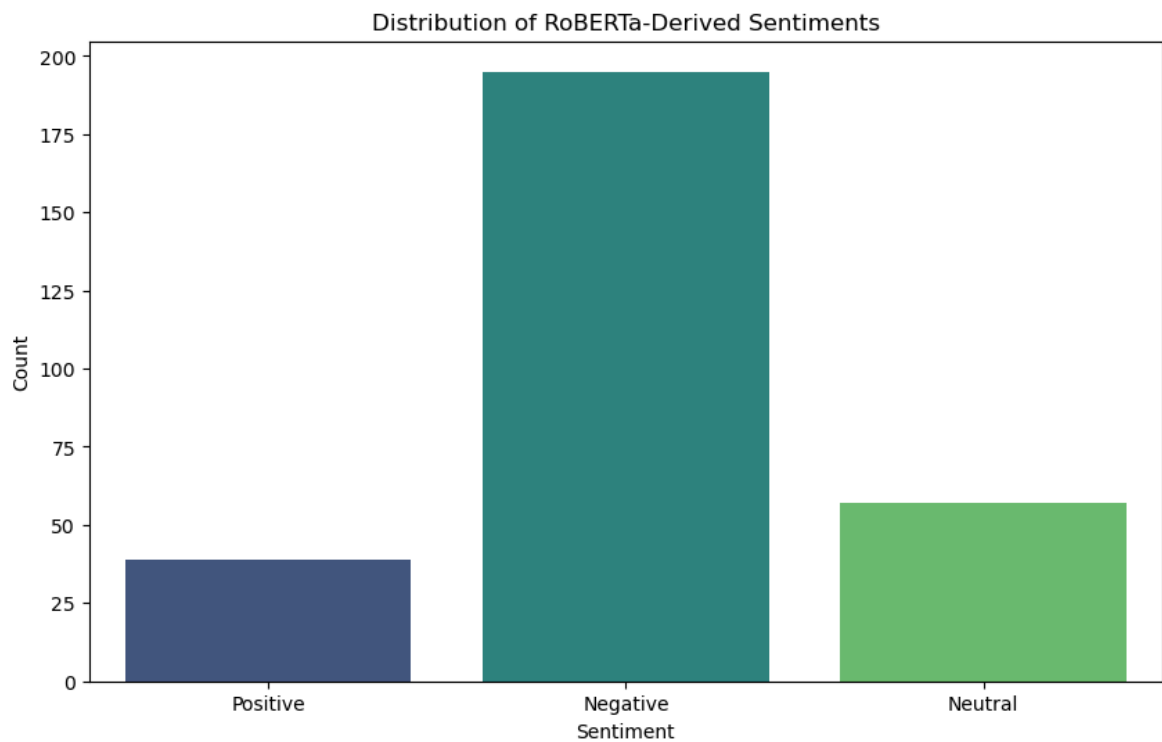

```
C:\Users\joaom\AppData\Local\Temp\ipykernel_8608\1906245679.py:14: FutureWarning:
'M' is deprecated and will be removed in a future version, please use 'ME' instea
d.
  reviews.set_index('date').resample('M').roberta_sentiment.value_counts().unstac
k().plot(kind='line', marker='o')
<Figure size 1400x700 with 0 Axes>
```

Sentiment Trends Over Time (RoBERTa)


```python
# Compare RoBERTa Sentiment with Custom Rules
comparison = reviews[['vader_sentiment_updated', 'roberta_sentiment']].copy()
comparison_melted = comparison.melt(var_name='method', value_name='sentiment')

plt.figure(figsize=(12, 6))
```

```python
sns.countplot(x='sentiment', hue='method', data=comparison_melted, palette='viri
plt.title('Comparison of Sentiments: Custom Rules vs. RoBERTa')
plt.xlabel('Sentiment')
plt.ylabel('Count')
plt.legend(title='Method')
plt.show()
```



```python
# Compare RoBERTa Sentiment with Custom Rules
comparison = reviews[['vader_sentiment_updated', 'roberta_sentiment']].copy()
comparison_melted = comparison.melt(var_name='method', value_name='sentiment')

plt.figure(figsize=(12, 6))
sns.countplot(x='sentiment', hue='method', data=comparison_melted, palette='viri
plt.title('Comparison of Sentiments: Custom Rules vs. RoBERTa')
plt.xlabel('Sentiment')
plt.ylabel('Count')
plt.legend(title='Method')
plt.show()
```



```python
# Compare RoBERTa and VADER Sentiments
comparison = reviews[['vader_sentiment_updated', 'roberta_sentiment']].copy()
```

```python
comparison_melted = comparison.melt(var_name='method', value_name='sentiment')

plt.figure(figsize=(12, 6))
sns.countplot(x='sentiment', hue='method', data=comparison_melted, palette='viri
plt.title('Comparison of Sentiments: RoBERTa vs. VADER')
plt.xlabel('Sentiment')
plt.ylabel('Count')
plt.legend(title='Method')
plt.show()
```
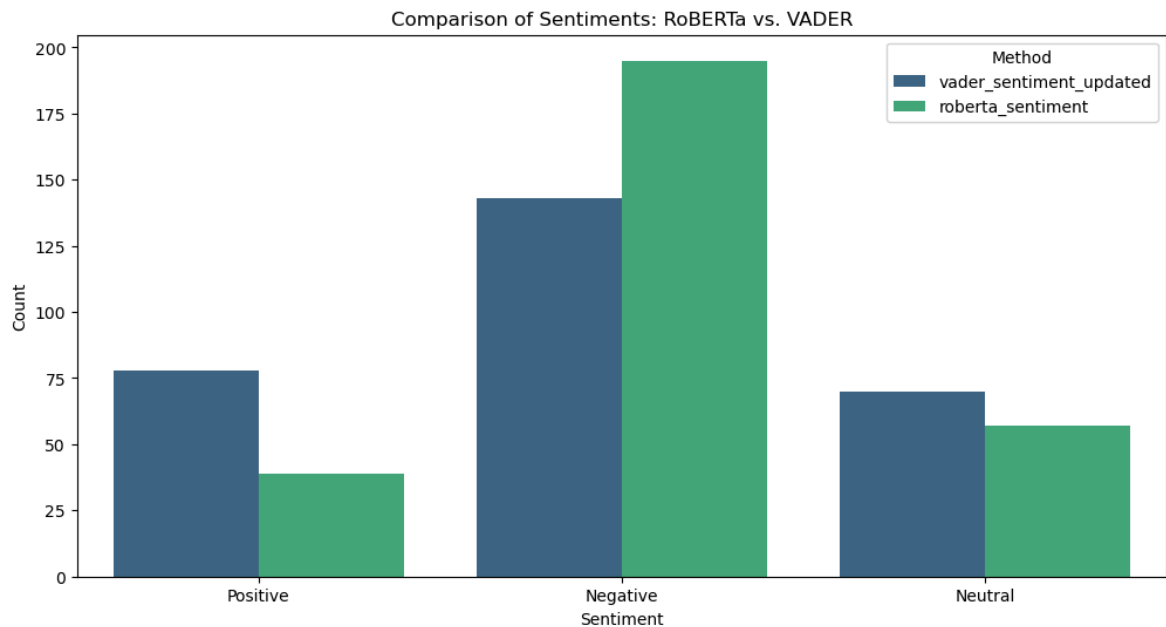


```python
import seaborn as sns
import matplotlib.pyplot as plt

# Select the relevant columns for comparison and include the rating column
pairplot_data = reviews[['vader_neg_updated', 'vader_neu_updated', 'vader_pos_up

# Create the pairplot with rating as the hue
sns.pairplot(pairplot_data, hue='rating', palette='viridis')
plt.suptitle('Pairplot Comparison of VADER and RoBERTa Sentiment Scores by Ratin
plt.show()
```

Pairplot Comparison of VADER and RoBERTa Sentiment Scores by Rating
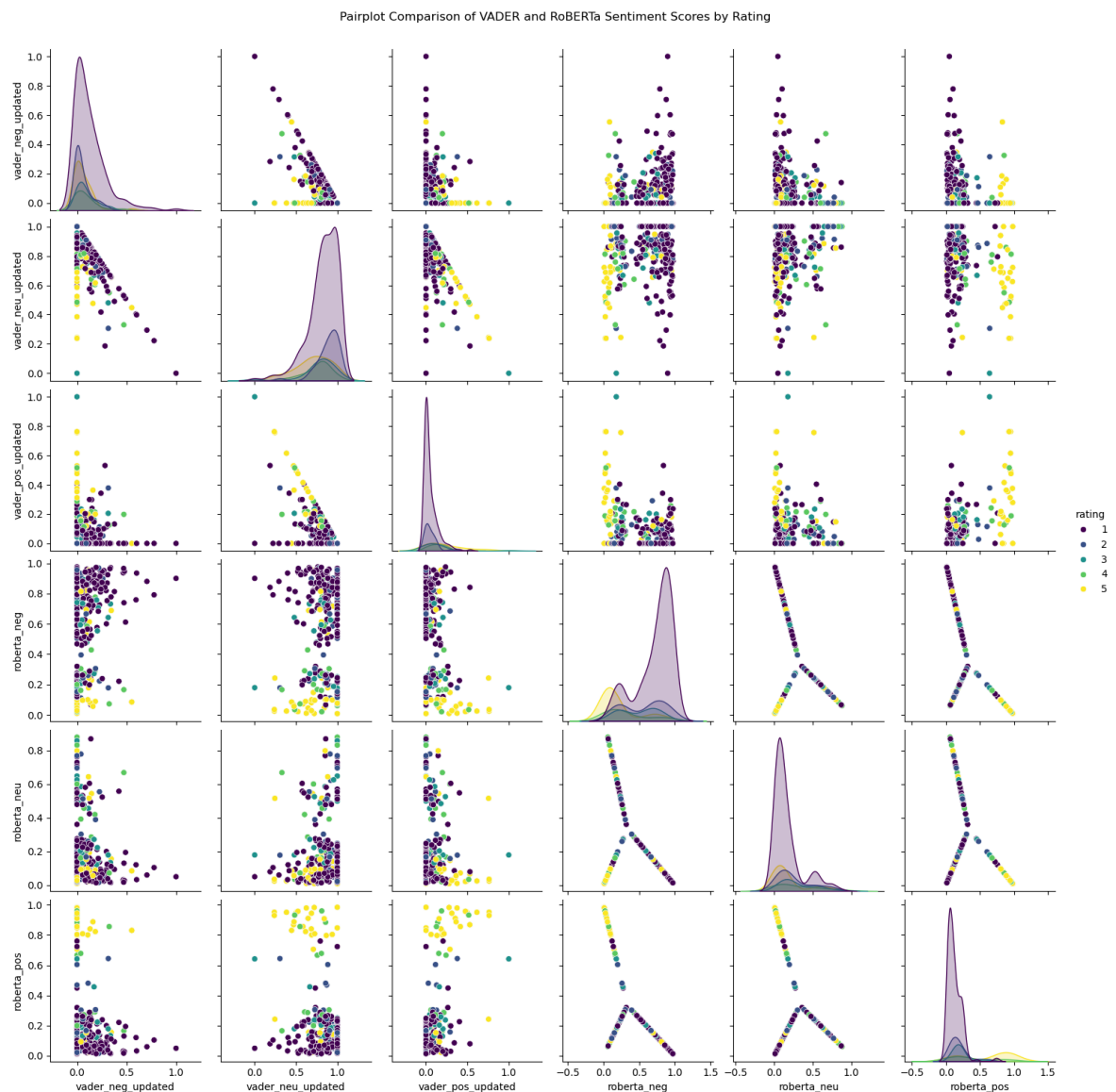
Review Examples:

Lets look at some examples where the model scoring and review score are different the most (1 star to 5 star reviews)

In [127...

```python
# Query for reviews with a rating of 5 and sort by RoBERTa negative score in des
top_discrepancy_reviews = reviews.query('rating == 5').sort_values('roberta_neg'

# Display the result
for index, row in top_discrepancy_reviews.iterrows():
    print(f"Original Review: {row['review']}")
    print(f"Translated Review: {row['review_english']}")
    print(f"RoBERTa Negative Score: {row['roberta_neg']}\n")
```

Original Review: Na minha rua anda um pit bul enorme e atira se as pessoas e aos animais e o dono ameaça as pessoas Boa da para resolver porque o animal ataca já liguei para a PSP e nada fazem até que algo grave aconteca
Translated Review: In my street a huge pit bul is and shoots if people and animals and the owner threatens the good people to solve because the animal attacks I already called PSP and do nothing until something serious happens
RoBERTa Negative Score: 0.9545252919197083

Original Review: 1) Na Praça José Queróz estão montados entradas de passageiros em dezenas de carreiras várias da carris outras muitas que vão para o exterior da cidade. tanto uma da carris como exteriores têm aqui o seu terminal. São centenas de pessoas que no passeio Este da Praça, junto das paragens passam. Porque é que neste local não é instalado serviços sanitários, convenientes se há espaço? Até onde existe ainda uma casa fechada, com ainda publicidade que foi de venda de jornais e revistas, não é alargada e podia ser aí. Quiosque São todos os dias dezenas de pessoas (até os motoristas das carreiras com terminal aqui) a urinar e defecar, junto das paredes dos prédios. É vergonhoso para quem aqui reside, (centenas de famílias) e turismo em geral (muito má imagem) na entrada duma das Rua principais do Parque das Nações, (R. João Pinto Ribeiro) 2) As limpezas diárias do local são raras, existe erva alta de aspeto vergonhoso, para quem passa por falta de higiene. e os caixotes do lixo são do antigos com tantos utilizadores, também se torna uma imagem muito desagradável. E porque se há espaço não são feitos os lixos subterrâneos. Cumprimentos um residente no 115, 4ºC Fernando Beato Marques
Translated Review: 1) José Queroz Square are set up passenger entries in dozens of careers several from Carris many others that go outside the city.Both from Carris and Outdoors have their terminal here.There are hundreds of people who on the ride this from the square, next to the stops.Why is not in this place sanitary services, convenient if there is space?As far as there is still a closed house, with still advertising that was the sale of newspapers and magazines, is not extended and could be there.Kiosk are dozens of people every day (even career drivers with terminal here) to urinate and defecate, next to the walls of the buildings.It is shameful for those who live here, (hundreds of families) and general tourism (very bad image) at the entrance of the main street of Parque das Nações, (R. João Pinto Ribeiro) 2) The daily cleaning of the place are rare, there isHigh herb of shameful appearance, for those who undergo lack of hygiene.And the rubbish crates are from the old ones with so many users, it also becomes a very unpleasant image.And because if there is space, underground waste is not made.Greetings a resident in 115, 4ºC Fernando Beato Marques
RoBERTa Negative Score: 0.8218367099761963

Original Review: Vandalizar propriedade-O bloco esquerda estava a pintar na entrada principal do instituto superior técnico o seu cartaz nos únicos muros limpos na entrada do Ist (frente para o ist lado direito).
Acho inadmissível que um partido político possa promover assim as suas ideologias estragando propriedade pública . Deviam ser obrigados a pintar a fachada !
(Tenho fotos que provam está ação)
Translated Review: Vandalizing Property The left block was painting at the main entrance of the Instituto Superior Technical, his poster on the only clean walls at the entrance of the IST (front to the right side).
I find it inadmissible for a political party to promote its ideologies spoiling public property.They should be required to paint the facade!
(I have photos that prove this action)
RoBERTa Negative Score: 0.8142930865287781

Original Review: Sou administrador do condominio da travessa da conceição da glória ,7 que tem cerca de 90 apartamentos e desde á 2 dias não é feita a recolha do lixo na travessa. Será que estão em greve ? ou então o que se passa ? por chamada telefónica para os serviços de recolha de lixo da cml ninguém atende
Translated Review: I am an administrator of the condominium of Travessa da Conceição da Gloria, 7 which has about 90 apartments and since 2 days is not made the g

arbage collection on the cross.Are they on strike?Or what's going on?by telephone call for CML's garbage collection services no one answers
RoBERTa Negative Score: 0.7171677947044373

Original Review: Funcional e útil no entanto sem seguimento practice-De início a app estava inoperacional, actualmente e após contacto pela equipa de programação da app (surpreendentemente agradável e atenta ás minhas questões) conseguiram ap rimorar a App até resolverem as questões por mim apresentadas. Infelizmente a nív el prático não funciona , as ocorrências quando são encaminhavas para as juntas d e freguesia, morrem lá ..... perpetua-se o problema. Tornando a APP um mero instr umento eleitoral... o que é pena pois até tinha potencial.
Translated Review: Functional and Useful However without follow-up on the start-u p the app was inoperative, currently and after contact with the app programming t eam (surprisingly pleasant and aware of my questions) managed to improve the app until the questions presented by me were resolved.Unfortunately at the practical level it does not work, the occurrences when they are sent to the parish council s, die there ..... the problem is perpetuated.Making the app a mere electoral ins trument ... which is pity because it even had potential.
RoBERTa Negative Score: 0.6943075656890869

Original Review: Não obtive resposta eficaz além de "iremos analisar a sua queixa ".
Translated Review: I got no effective response beyond "we will analyze your compl aint."
RoBERTa Negative Score: 0.6881999373435974

Original Review: Muito difícil no registo: >>Email já registado!!! >>Palavra Pass e incorrecta.... >>etc... Parece que a Ideia é desincentivar o Munícipe a partici par as ocorrências na via Pública, que é de todos e uma questão de Cidadania. Por que não um Email específico, como era no passado? Só a complicar??? Obrigado.
Translated Review: Very difficult in the record: >> Email already registered !!!> > Word Incorrects .... >> etc ... It seems that the idea is to discourage the cit izen to participate in the public road, which is everyone and a matter of citizen ship.Why not a specific email, as it was in the past?Just complicating ???Thanks.
RoBERTa Negative Score: 0.6138259172439575

Original Review: Desde há algumas semanas que não é possível nem na app nem no si te participar ocorrências. Antes funcionava muito bem mas desde um pouco antes da s eleições que deixou de funcionar. Update 28 de Novembro 2021: O problema só aco ntece quando se faz o login com as credenciais do facebook. Se entrarmos no site ou app com email e password já funciona.
Translated Review: Since a few weeks ago it is not possible in the app or on the site to participate in occurrences.It before it worked very well but since just b efore the elections it stopped working.UPDATE November 28, 2021: The problem only happens when you log in to Facebook credentials.If we enter the website or app wi th email and password already works.
RoBERTa Negative Score: 0.5339022278785706

Original Review: Prático eficaz.
Translated Review: Effective practical.
RoBERTa Negative Score: 0.2423071265220642

Original Review: Oh até parece que não há aplicações de location spoofer, em vez de estarem aí a reclamar com isso. Aposto que se fosse o jogo do Pokémon tinham s e logo lembrado disso mas pronto... Finalmente que a câmara de Lisboa descobre os smartphones, já não era sem tempo
Translated Review: Oh it seems that there are no location spoofer applications, i nstead of being there to complain about it.I bet if it was the Pokémon game they had soon remembered this but ready ... Finally, the Lisbon Chamber discovers the smartphones, it was no longer timeless

RoBERTa Negative Score: 0.24164289236068726

In [128...

```python
# Query for reviews with a rating of 5 and sort by RoBERTa positive score in des
top_discrepancy_reviews = reviews.query('rating == 5').sort_values('roberta_pos'

# Display the result
for index, row in top_discrepancy_reviews.iterrows():
    print(f"Original Review: {row['review']}")
    print(f"Translated Review: {row['review_english']}")
    print(f"RoBERTa Positive Score: {row['roberta_pos']}\n")
```

Original Review: Este Serviço da CM Lisboa, tem sido extraordinário é muito eficiente. Todas as ocorrências por mim endereçadas, são sempre executadas atempadamente. Parabéns a toda a equipa que faz deste Serviço "excelente". Teresa Sousa
Translated Review: This service of CM Lisbon has been extraordinary is very efficient.All occurrences addressed by me are always performed in time.Congratulations to the whole team that makes this service "excellent".Teresa Sousa
RoBERTa Positive Score: 0.9843660593032837

Original Review: Esta app está agora muito mais fácil de utilizar. Acessível e intuitiva. Parabéns. Esperemos que os serviços que os serviços disponibilizados possam passar a ser tão eficientes quanto esta APP.
Translated Review: This app is now much easier to use.Accessible and intuitive.Congratulations.We hope that the services that the services available can become as efficient as this app.
RoBERTa Positive Score: 0.9830625057220459

Original Review: Fantástica app. boa resposta às ocorrências colocadas. Parabéns.
Translated Review: Fantastic App.good response to the occurrences placed.Congratulations.
RoBERTa Positive Score: 0.9796355962753296

Original Review: Bastante útil, esperemos que os serviços da CML correspondam
Translated Review: Very useful, we hope that CML services will correspond
RoBERTa Positive Score: 0.9578760266304016

Original Review: Muito bom
Translated Review: Very good
RoBERTa Positive Score: 0.9518744945526123

Original Review: Muito positiva.O que reportei foi rápidamente atendido.
Translated Review: Very positive. I reported quickly.
RoBERTa Positive Score: 0.9516088366508484

Original Review: Muito boa....
Translated Review: Very good ....
RoBERTa Positive Score: 0.95113205909729

Original Review: Ferramenta super útil para os lisboetas. Parabéns-Simples e muito útil para podermos informar a câmara sobre as ocorrências com que nos deparamos na rua.
Translated Review: Super useful tool for Lisbon.Congratulations-Simple and very useful so that we can inform the House about the occurrences with which we come across the street.
RoBERTa Positive Score: 0.9449506998062134

Original Review: Muito útil.
Translated Review: Very useful.
RoBERTa Positive Score: 0.9302197098731995

Original Review: VALE A PENA-App super útil , fácil de navegar.. de todas as ocorrências q reportei 90% já foram resolvidas, algumas em menos de uma semana, outras em 1 mês, dois..mas sempre são resolvidas!
Translated Review: It is worth the super useful, easy to navigate .. of all occurrences I reported 90% have already been resolved, some in less than one week, others in 1 month, two..but are always resolved!
RoBERTa Positive Score: 0.9110659956932068

```python
# Function to display top 10 discrepant reviews based on the given query and sor
def display_top_discrepant_reviews(query, sort_by, ascending, description):
```

```python
        top_discrepancy_reviews = reviews.query(query).sort_values(sort_by, ascendin
        # Print the description with larger headline
        print(f"\033[1;34mTop 10 reviews for {description}:\033[0m\n")
        for index, row in top_discrepancy_reviews.iterrows():
            print(f"Original Review: {row['review']}")
            print(f"Translated Review: {row['review_english']}")
            print(f"{sort_by.capitalize()}: {row[sort_by]}\n")


# 1. Reviews with a rating of 5, sorted by RoBERTa negative score in descending
display_top_discrepant_reviews('rating == 5', 'roberta_neg', False, 'rating of 5

# 2. Reviews with a rating of 5, sorted by RoBERTa positive score in descending
display_top_discrepant_reviews('rating == 5', 'roberta_pos', False, 'rating of 5

# 3. Reviews with a rating of 1, sorted by RoBERTa negative score in descending
display_top_discrepant_reviews('rating == 1', 'roberta_neg', False, 'rating of 1

# 4. Reviews with a rating of 1, sorted by RoBERTa positive score in descending
display_top_discrepant_reviews('rating == 1', 'roberta_pos', False, 'rating of 1
```

Top 10 reviews for rating of 5 sorted by RoBERTa negative score:

Original Review: Na minha rua anda um pit bul enorme e atira se as pessoas e aos animais e o dono ameaça as pessoas Boa da para resolver porque o animal ataca já liguei para a PSP e nada fazem até que algo grave aconteca
Translated Review: In my street a huge pit bul is and shoots if people and animals and the owner threatens the good people to solve because the animal attacks I already called PSP and do nothing until something serious happens
Roberta_neg: 0.9545252919197083

Original Review: 1) Na Praça José Queróz estão montados entradas de passageiros em dezenas de carreiras várias da carris outras muitas que vão para o exterior da cidade. tanto uma da carris como exteriores têm aqui o seu terminal. São centenas de pessoas que no passeio Este da Praça, junto das paragens passam. Porque é que neste local não é instalado serviços sanitários, convenientes se há espaço? Até onde existe ainda uma casa fechada, com ainda publicidade que foi de venda de jornais e revistas, não é alargada e podia ser aí. Quiosque São todos os dias dezenas de pessoas (até os motoristas das carreiras com terminal aqui) a urinar e defecar, junto das paredes dos prédios. É vergonhoso para quem aqui reside, (centenas de famílias) e turismo em geral (muito má imagem) na entrada duma das Rua principais do Parque das Nações, (R. João Pinto Ribeiro) 2) As limpezas diárias do local são raras, existe erva alta de aspeto vergonhoso, para quem passa por falta de higiene. e os caixotes do lixo são do antigos com tantos utilizadores, também se torna uma imagem muito desagradável. E porque se há espaço não são feitos os lixos subterrâneos. Cumprimentos um residente no 115, 4ºC Fernando Beato Marques
Translated Review: 1) José Queroz Square are set up passenger entries in dozens of careers several from Carris many others that go outside the city.Both from Carris and Outdoors have their terminal here.There are hundreds of people who on the ride this from the square, next to the stops.Why is not in this place sanitary services, convenient if there is space?As far as there is still a closed house, with still advertising that was the sale of newspapers and magazines, is not extended and could be there.Kiosk are dozens of people every day (even career drivers with terminal here) to urinate and defecate, next to the walls of the buildings.It is shameful for those who live here, (hundreds of families) and general tourism (very bad image) at the entrance of the main street of Parque das Nações, (R. João Pinto Ribeiro) 2) The daily cleaning of the place are rare, there isHigh herb of shameful appearance, for those who undergo lack of hygiene.And the rubbish crates are from the old ones with so many users, it also becomes a very unpleasant image.And because if there is space, underground waste is not made.Greetings a resident in 115, 4ºC Fernando Beato Marques
Roberta_neg: 0.8218367099761963

Original Review: Vandalizar propriedade-O bloco esquerda estava a pintar na entrada principal do instituto superior técnico o seu cartaz nos únicos muros limpos na entrada do Ist (frente para o ist lado direito).
Acho inadmissível que um partido político possa promover assim as suas ideologias estragando propriedade pública . Deviam ser obrigados a pintar a fachada !
(Tenho fotos que provam está ação)
Translated Review: Vandalizing Property The left block was painting at the main entrance of the Instituto Superior Technical, his poster on the only clean walls at the entrance of the IST (front to the right side).
I find it inadmissible for a political party to promote its ideologies spoiling public property.They should be required to paint the facade!
(I have photos that prove this action)
Roberta_neg: 0.8142930865287781

Original Review: Sou administrador do condominio da travessa da conceição da glória ,7 que tem cerca de 90 apartamentos e desde á 2 dias não é feita a recolha do lixo na travessa. Será que estão em greve ? ou então o que se passa ? por chamada telefónica para os serviços de recolha de lixo da cml ninguém atende

Translated Review: I am an administrator of the condominium of Travessa da Concei
ção da Gloria, 7 which has about 90 apartments and since 2 days is not made the g
arbage collection on the cross.Are they on strike?Or what's going on?by telephone
call for CML's garbage collection services no one answers
Roberta_neg: 0.7171677947044373

Original Review: Funcional e útil no entanto sem seguimento practice-De início a
app estava inoperacional, actualmente e após contacto pela equipa de programação
da app (surpreendentemente agradável e atenta ás minhas questões) conseguiram ap
rimorar a App até resolverem as questões por mim apresentadas. Infelizmente a nív
el prático não funciona , as ocorrências quando são encaminhavas para as juntas d
e freguesia, morrem lá ..... perpetua-se o problema. Tornando a APP um mero instr
umento eleitoral... o que é pena pois até tinha potencial.
Translated Review: Functional and Useful However without follow-up on the start-u
p the app was inoperative, currently and after contact with the app programming t
eam (surprisingly pleasant and aware of my questions) managed to improve the app
until the questions presented by me were resolved.Unfortunately at the practical
level it does not work, the occurrences when they are sent to the parish council
s, die there ..... the problem is perpetuated.Making the app a mere electoral ins
trument ... which is pity because it even had potential.
Roberta_neg: 0.6943075656890869

Original Review: Não obtive resposta eficaz além de "iremos analisar a sua queixa
".
Translated Review: I got no effective response beyond "we will analyze your compl
aint."
Roberta_neg: 0.6881999373435974

Original Review: Muito difícil no registo: >>Email já registado!!! >>Palavra Pass
e incorrecta.... >>etc... Parece que a Ideia é desincentivar o Munícipe a partici
par as ocorrências na via Pública, que é de todos e uma questão de Cidadania. Por
que não um Email específico, como era no passado? Só a complicar??? Obrigado.
Translated Review: Very difficult in the record: >> Email already registered !!!>
> Word Incorrects .... >> etc ... It seems that the idea is to discourage the cit
izen to participate in the public road, which is everyone and a matter of citizen
ship.Why not a specific email, as it was in the past?Just complicating ???Thanks.
Roberta_neg: 0.6138259172439575

Original Review: Desde há algumas semanas que não é possível nem na app nem no si
te participar ocorrências. Antes funcionava muito bem mas desde um pouco antes da
s eleições que deixou de funcionar. Update 28 de Novembro 2021: O problema só aco
ntece quando se faz o login com as credenciais do facebook. Se entrarmos no site
ou app com email e password já funciona.
Translated Review: Since a few weeks ago it is not possible in the app or on the
site to participate in occurrences.It before it worked very well but since just b
efore the elections it stopped working.UPDATE November 28, 2021: The problem only
happens when you log in to Facebook credentials.If we enter the website or app wi
th email and password already works.
Roberta_neg: 0.5339022278785706

Original Review: Prático eficaz.
Translated Review: Effective practical.
Roberta_neg: 0.2423071265220642

Original Review: Oh até parece que não há aplicações de location spoofer, em vez
de estarem aí a reclamar com isso. Aposto que se fosse o jogo do Pokémon tinham s
e logo lembrado disso mas pronto... Finalmente que a câmara de Lisboa descobre os
smartphones, já não era sem tempo
Translated Review: Oh it seems that there are no location spoofer applications, i
nstead of being there to complain about it.I bet if it was the Pokémon game they

had soon remembered this but ready ... Finally, the Lisbon Chamber discovers the smartphones, it was no longer timeless
Roberta_neg: 0.24164289236068726

Top 10 reviews for rating of 5 sorted by RoBERTa positive score:

Original Review: Este Serviço da CM Lisboa, tem sido extraordinário é muito eficiente. Todas as ocorrências por mim endereçadas, são sempre executadas atempadamente. Parabéns a toda a equipa que faz deste Serviço "excelente". Teresa Sousa
Translated Review: This service of CM Lisbon has been extraordinary is very efficient.All occurrences addressed by me are always performed in time.Congratulations to the whole team that makes this service "excellent".Teresa Sousa
Roberta_pos: 0.9843660593032837

Original Review: Esta app está agora muito mais fácil de utilizar. Acessível e intuitiva. Parabéns. Esperemos que os serviços que os serviços disponibilizados possam passar a ser tão eficientes quanto esta APP.
Translated Review: This app is now much easier to use.Accessible and intuitive.Congratulations.We hope that the services that the services available can become as efficient as this app.
Roberta_pos: 0.9830625057220459

Original Review: Fantástica app. boa resposta às ocorrências colocadas. Parabéns.
Translated Review: Fantastic App.good response to the occurrences placed.Congratulations.
Roberta_pos: 0.9796355962753296

Original Review: Bastante útil, esperemos que os serviços da CML correspondam
Translated Review: Very useful, we hope that CML services will correspond
Roberta_pos: 0.9578760266304016

Original Review: Muito bom
Translated Review: Very good
Roberta_pos: 0.9518744945526123

Original Review: Muito positiva.O que reportei foi rápidamente atendido.
Translated Review: Very positive. I reported quickly.
Roberta_pos: 0.9516088366508484

Original Review: Muito boa....
Translated Review: Very good ....
Roberta_pos: 0.95113205909729

Original Review: Ferramenta super útil para os lisboetas. Parabéns-Simples e muito útil para podermos informar a câmara sobre as ocorrências com que nos deparamos na rua.
Translated Review: Super useful tool for Lisbon.Congratulations-Simple and very useful so that we can inform the House about the occurrences with which we come across the street.
Roberta_pos: 0.9449506998062134

Original Review: Muito útil.
Translated Review: Very useful.
Roberta_pos: 0.9302197098731995

Original Review: VALE A PENA-App super útil , fácil de navegar.. de todas as ocorrências q reportei 90% já foram resolvidas, algumas em menos de uma semana, outras em 1 mês, dois..mas sempre são resolvidas!
Translated Review: It is worth the super useful, easy to navigate .. of all occurrences I reported 90% have already been resolved, some in less than one week, oth

ers in 1 month, two..but are always resolved!
Roberta_pos: 0.9110659956932068

Top 10 reviews for rating of 1 sorted by RoBERTa negative score:

Original Review: Absolutamente inacreditável Nao consigo identificar o numero de pirta
correctamente ESTA APP É LIXO
Translated Review: Absolutely unbelievable I cannot identify the door number correctly this app is garbage
Roberta_neg: 0.975937008857727

Original Review: Não funciona-Ideia óptima, mas a app não funciona
Translated Review: Does not work great, but the app does not work
Roberta_neg: 0.9728612303733826

Original Review: Kafkiana-A coisa mais inútil que já vi
Translated Review: Kafkiana-The most useless thing I've ever seen
Roberta_neg: 0.9713993072509766

Original Review: Inútil-A aplicação está sem funcionar há meses
Translated Review: Useless-the application has been without working for months
Roberta_neg: 0.9666875600814819

Original Review: Funciona mal-Funciona mal, não se conseguindo reportar problemas por não encontrar a localização
Translated Review: It works badly badly, unable to report problems for not finding the location
Roberta_neg: 0.9660051465034485

Original Review: Funciona mal-Não permite colocar fotografias e, no geral, funciona muito mal.
Translated Review: It works badly not allows you to put photographs and, in general, works very badly.
Roberta_neg: 0.9643632173538208

Original Review: Não funciona. Testei em dois telemóveis diferentes e não consigo criar ocorrências. Carrego no botão e nada acontece. A incompetência é surreal!
Translated Review: Does not work.I tested two different phones and I can't create occurrences.I load the button and nothing happens.Incompetence is surreal!
Roberta_neg: 0.9604612588882446

Original Review: Péssima App que em nada prestigia a cidade de Lisboa-Para além de funcionar pessimamente, estou há vários dias para relatar algumas situações nesta App, porém depois de inserir fotografias, inserir a localização e descrever a situação, quando clico no último botão "avançar" tudo desaparece!!!! Também lamentavelmente esta App parece não acolher por parte da Câmara Municipal de Lisboa grande simpatia, sendo reprovável e lamentável o desrespeito relativamente aos seus concidadãos pois a maior parte das vezes não é dada resposta às sugestões! Eu próprio tenho situações que aqui coloquei há mais de dois anos e que ainda se encontram em análise!!! Vergonha!
Translated Review: Bad app that does not honor the city of Lisbon-Para in addition to working in badly, I am several days ago to report some situations in this app, but after inserting photographs, inserting the location and describing the situation, when I click the last button "Next"Everything disappears !!!!Also unfortunately this app does not seem to welcome from the City Council of Lisbon Grande sympathy, being reprehensible and unfortunate the disrespect for its fellow citizens because most of the time is not given response to the suggestions!I myself have situations that I have put here for over two years and are still under analysis !!!Shame!

Roberta_neg: 0.9583747386932373

Original Review: Aplicação não funciona-Após ter sido descrita a ocorrência, a ap
licação bloqueia no final e a mesma não é criada.
Aplicação inútil!
Translated Review: Application does not work to have the occurrence described, th
e application blocks in the end and it is not created.
Useless application!
Roberta_neg: 0.9553852081298828

Original Review: Nao funciona-A app está sempre off
Translated Review: Does not work it app is always off
Roberta_neg: 0.9552300572395325

Top 10 reviews for rating of 1 sorted by RoBERTa positive score:

Original Review: Uma bela m*-O título diz tudo nada funciona
Viva o legado do merdina e do PS
Translated Review: A beautiful m*-The title says all nothing works
Long live the legacy of Mercdina and PS
Roberta_pos: 0.7601968050003052

Original Review: Entrar após registo-Nem consegui entrar na aplicação após regist
o. Dá mensagem "Credenciais inválidas". Incrível …
Translated Review: Entering after registration-nam I was able to enter the applic
ation after registration.Gives "invalid credentials".Incredible …
Roberta_pos: 0.7237253189086914

Original Review: Uma óptima aplicação para receber e-mails a dizer que a ocorrênc
ia foi resolvida quando de facto não foi...
Translated Review: A great application to receive emails to say that the occurren
ce was resolved when it was not ...
Roberta_pos: 0.44900140166282654

Original Review: A app tem muitos bugs-A app tem muitos bugs. Era bom que funcion
asse pois seria muito útil.
Translated Review: The app has many bugs-a app has many bugs.It was good to work
because it would be very useful.
Roberta_pos: 0.3196195662021637

Original Review: Assuntos "resolvidos"-Para além do design da app ser um pouco co
nfuso, por exemplo nos processos de reportar com 40 menus e opções diferentes que
nem sempre são obvias, após várias queixas semelhantes resolvem os assuntos com m
eias soluções que, dado que os assuntos continuam a ser reportados, deviam ser ob
viamente reavaliadas. Uma pena: a ideia é boa, a app medíocre, mas a resolução do
s assuntos péssima.
Translated Review: "Resolved" subjects-besides the app design is a bit confusing,
for example in the processes of reporting with 40 menus and different options tha
t are not always obvious, after several similar complaints solve the subjects wit
h half solutions that, given that the subjectsThey continue to be reported, they
should obviously be reevaluated.A pity: the idea is good, the mediocre app, but t
he resolution of poor matters.
Roberta_pos: 0.2881270796060562

Original Review: No updates for the last 2 years-The app is pretty clunky, not to
mention the terrible internal workflows from the CML, who simply closes the issue
s saying they have already been resolved, but without any prove nor way of contes
ting the assertion. If the app could be updated to support a better conversation
between the city and citizens, that would be a major improvement, though that see
ms unlikely.

Translated Review: In Updates for the Last 2 Years-The The Terrible Internal Work
flows from the CML, Who Simply Closses The Issues Say Already Been Resolve, Butt
Whit Prové Nor Way Of Contesting The Assertion.If the app could be updated to sup
port a better conversation beteen the city and citizens, that woul be a major imp
rovement, that that sems unlikely.
Roberta_pos: 0.2704237252473831

Original Review: Muito pouco funcional-No iphone8S não consegui descobrir como cr
iar uma ocorrência. Consegui ver apena o que já tinha reportado no portal.
Nas faq não explicam como se cria uma ocorrência.
Translated Review: Very little functional-in iPhone8s could not find out how to c
reate an occurrence.I managed to see what he had already reported on the portal.
In FAQ do not explain how an occurrence is created.
Roberta_pos: 0.2660457342863083

Original Review: Já tentei por diversas vezes reportar várias situações, mas a ap
p bloqueia ou não deixa avançar com o pedido. (Este medina é brilhante)
Translated Review: I've tried several times to report several situations, but the
app blocks or does not let it advance with the order.(This Medina is brilliant)
Roberta_pos: 0.2642149478197098

Original Review: Inconsequente-Nao basta criar uma app para dar ideia que se está
na proximidade dos municipes. Tem de haver alguém que dê seguimento às ocorrência
s lá no gabinete. 1 semana para resolver falhas de iluminaçao publica quando ha v
arias ocorrencias reportadas referentes ao mesmo incidente é demasiado inconseque
nte.
Translated Review: Inconsequently, just create an app to give an idea that you ar
e in the proximity of the municipalities.There has to be someone who follows the
occurrences there in the cabinet.1 week to resolve public light failures when the
re are several reported occurrences regarding the same incident is too inconseque
ntial.
Roberta_pos: 0.2612108588218689

Original Review: Nem login da-Fiz registo no pc com Google account, saco a app e
não dá para fazer login pelo Google, só por email ou Facebook ... tenho que criar
outra conta??
Translated Review: Enem Login Da-Foz Registration on PC with Google Account, bag
the app and you can not login through Google, only by email or Facebook ... do I
have to create another account ??
Roberta_pos: 0.26039405167102814

Code to Create rating_by_roberta Column:

```
In [131…    # Function to map RoBERTa sentiment to rating
            def map_sentiment_to_rating(sentiment):
                if sentiment == 'Negative':
                    return 1
                elif sentiment == 'Neutral':
                    return 3
                elif sentiment == 'Positive':
                    return 5

            # Apply the mapping function to create the 'rating_by_roberta' column
            reviews['rating_by_roberta'] = reviews['roberta_sentiment'].apply(map_sentiment_
```

```python
# Display a sample of the DataFrame to check the new column
print(reviews[['review_english', 'roberta_sentiment', 'rating_by_roberta']].head
```

```
                                                          review_engli
sh  \
0                                     Very practical and effecti
ve
1                             It is not possible to submit, invalid emai
l!
2                                         Impossible to access app
...
3  Application presents the error, "can not present information" is not functiona
l.
4                                         An excellent idea badly implemente
d.

   roberta_sentiment  rating_by_roberta
0           Positive                  5
1           Negative                  1
2           Negative                  1
3           Negative                  1
4           Positive                  5
```

Final Ratings

```python
# Calculate the average user rating
final_user_rating = reviews['rating'].mean()

# Calculate the average user rating for Google Store app users
final_user_rating_google = reviews[reviews['app'] == 'google']['rating'].mean()

# Calculate the average user rating for Apple Store app users
final_user_rating_apple = reviews[reviews['app'] == 'apple']['rating'].mean()

# Calculate the average rating based on the `rating_by_roberta` column
final_roberta_rating = reviews['rating_by_roberta'].mean()

# Calculate the average rating based on the `rating_by_roberta` column for Googl
final_roberta_rating_google = reviews[reviews['app'] == 'google']['rating_by_rob

# Calculate the average rating based on the `rating_by_roberta` column for Apple
final_roberta_rating_apple = reviews[reviews['app'] == 'apple']['rating_by_rober

# Print the final ratings
print('1 - True rating by Apple users: 1,4')
print('2 - True rating by users of the Google Store app: 2,5')
print(f"3 - Final rating by users: {final_user_rating:.2f}")
print(f"4 - Final rating by users of the Google Store app: {final_user_rating_go
print(f"5 - Final rating by Apple users: {final_user_rating_apple:.2f}")
print(f"6 - Final rating by users (RoBERTa): {final_roberta_rating:.2f}")
print(f"7 - Final rating by Google Store app users (RoBERTa): {final_roberta_rat
print(f"8 - Final rating by Apple users (RoBERTa): {final_roberta_rating_apple:.
```

```
1 - True rating by Apple users: 1,4
2 - True rating by users of the Google Store app: 2,5
3 - Final rating by users: 2.04
4 - Final rating by users of the Google Store app: 2.38
5 - Final rating by Apple users: 1.61
6 - Final rating by users (RoBERTa): 1.93
7 - Final rating by Google Store app users (RoBERTa): 2.10
8 - Final rating by Apple users (RoBERTa): 1.71
```

Final Technical Analysis of the *Na Minha Rua* Application Reviews

Note:This analysis was conducted using a selection of review samples extracted from user feedback on the *Na Minha Rua* application (**positive_reviews.csv & negativee_reviews.csv**). These samples were processed and analyzed using Claude, an advanced AI tool, to derive the following insights. The analysis focuses on identifying key problem areas, recurring themes, and actionable recommendations to improve the application.

Key Problems Identified

Access and Login Issues

- **Description**: Many users experience significant difficulties accessing the application. Reported problems include:
  - Invalid credentials (e.g., João Araujo, Aze Teka, Diana David).
  - Server errors (e.g., João Mm, Nazare Tojal).
  - Inability to recover passwords (e.g., Bruno, Ângela Braga).
- **Impact**: These issues prevent users from using the application, resulting in a frustrating and unsatisfactory experience.

Difficulties in Reporting Issues

- **Description**: Users frequently cannot report issues due to:
  - Failures in the photo attachment functionality (e.g., Ricardo Carvalho, Cátia Barroso).
  - Problems with manual location settings (e.g., Tons & Volumes Cabeleireiro, Joana Pereira).
  - Dependence on GPS for creating reports and location inaccuracies (e.g., Célia Neves, a Google user).
- **Impact**: These challenges compromise the application's effectiveness in addressing civic complaints.

Resolution of Reported Issues

- **Description**: Users express frustration with the lack of actual resolution for reported issues. Many cases are marked as resolved without effective intervention (e.g., Rui

Machado, Nuno Pereira, Mário Macheta).

- **Impact**: The application's credibility is diminished, leading to perceptions of inefficacy and distrust among users.

Usability and Interface

- **Description**: The application interface is criticized for being unintuitive and technically problematic, including:
  - Non-functional buttons (e.g., Rita De Almeida, Pedro Dom).
  - Absence of creation dates for reported issues (e.g., F C).
  - Problems with the search field (e.g., Célia Neves).
- **Impact**: These issues negatively affect the user experience, hindering efficient navigation and use.

Feedback and Customer Support

- **Description**: Recurring lack of effective customer support, including:
  - Insufficient or absent feedback (e.g., João Ribeiro, Rute Calheiros).
  - Difficulty in reopening or tracking reported issues (e.g., Rui Machado, Maria João Pires).
- **Impact**: User dissatisfaction grows, leading to a negative perception of the application's service.

Recurring Themes

Need for Technical and Functional Improvements

- **Description**: Urgent fixes are required for login systems, GPS functionality, and photo attachment processes. Overall application stability and usability need significant improvement to avoid frequent failures (e.g., Artur Dias, Rodolfo Costa).

Efficiency in Issue Resolution

- **Description**: A more effective system for tracking and resolving issues is critical. Cases should not be marked as resolved without actual intervention (e.g., Rui Machado, Maria João Pires).

Enhancement of Interface and Usability

- **Description**: The interface needs to be simplified and made more intuitive. Improvements are required in search fields, location settings, and overall functionality (e.g., Rita De Almeida, Pedro Dom).

Better Customer Support

- **Description**: Strengthening customer support is essential, with clear communication channels and prompt responses to reported issues (e.g., João Ribeiro, Rute

Calheiros).

Recommendations

Immediate Technical Fixes

- **Actions**:
  - Address login problems, particularly with Facebook and Google credentials (e.g., João Araujo, Aze Teka).
  - Improve GPS functionality and enable more efficient manual location selection (e.g., Célia Neves, a Google user).

Enhance User Interface

- **Actions**:
  - Make the interface more intuitive and functional, improving search fields, photo attachment features, and general navigation (e.g., Rita De Almeida, Pedro Dom).
  - Include essential information, such as creation dates for reported issues visible on the map (e.g., F C).

Improve Issue Resolution Processes

- **Actions**:
  - Implement a more effective system for tracking and resolving issues, ensuring cases are not closed without real intervention (e.g., Rui Machado, Maria João Pires).
  - Provide clear and consistent feedback to users about the status of their issues (e.g., João Ribeiro, Rute Calheiros).

Strengthen Customer Support

- **Actions**:
  - Bolster customer support with clear communication channels and quick responses to reported problems (e.g., João Ribeiro, Rute Calheiros).
  - Train the support team to handle technical issues and provide efficient solutions.