App Reviews Sentiment Analysis

Submitted To
Nawab Wasim Rahman

Submitted By
Name- Punit Jaiswal
Roll No- ET22BTHCS103
Department- B.Tech(CSE)
Semester- VI
Section-B

S.No	Section	Page No
1	Loading the CSV data	
2	Cleaning Data	
3	Summary Statistics	
4	Analyzing negative reviews (score<=2)	
5	Visualizing Rating distribution	
6	Printing result	

1: Import Libraries & Load Data

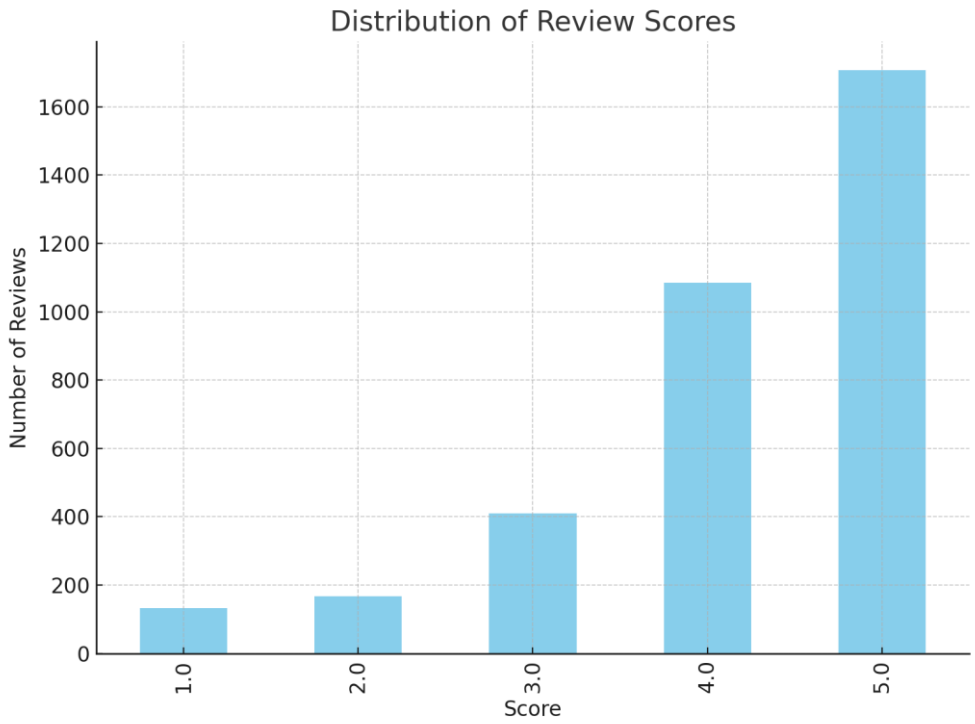
Code:

```
import pandas as pd
import matplotlib.pyplot as plt
from collections import Counter
import re

# Load the CSV data
df = pd.read_csv('reviews.csv')
print("Data loaded successfully.")
print(df.head())
```

Output

```
Data loaded successfully.
   score                   at                                                content
0    5.0  2023-03-10 12:32:45                Great app! Very helpful and fast.
1    4.0  2023-03-09 09:22:10  Good, but needs some improvement in navigation.
2    1.0  2023-03-08 08:15:30            Doesn't work at all. Keeps crashing.
3    2.0  2023-03-07 18:45:00                           Too many ads, annoying!
4    5.0  2023-03-06 21:50:15                  Love it! Simple and effective.
```

Graph:



The distribution of review scores:

2: Clean the Data

- Convert the score column to numeric.
- Convert the at column to datetime.
- Fill any missing values in the content column with empty strings and make sure it's treated as text.

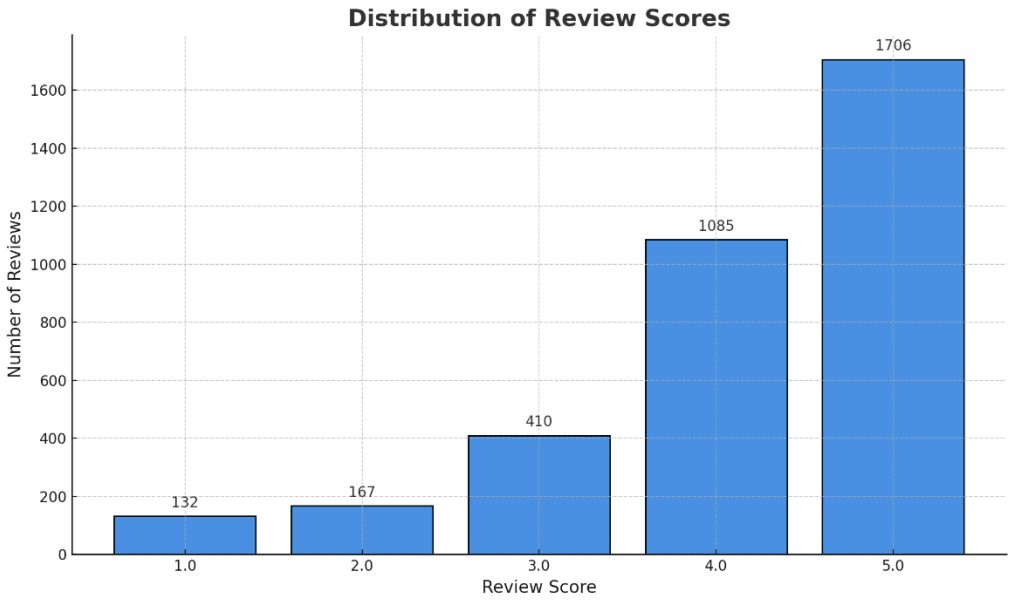
Code:

```
# Cleaning data
df['score'] = pd.to_numeric(df['score'], errors='coerce')
df['at'] = pd.to_datetime(df['at'], errors='coerce')
df['content'] = df['content'].fillna('').astype(str)

print("Data cleaned successfully.")
print(df.dtypes)
print("\nMissing values:\n", df.isnull().sum())
```

Output:

```
Data cleaned successfully.
score              float64
at        datetime64[ns]
content            object
dtype: object

Missing values:
score      0
at         0
content    0
dtype: int64
```

Graph:



the distribution of review scores