```
import warnings
warnings.filterwarnings('ignore')
import pandas as pd
import matplotlib.pyplot as plt
```

```
##importing Bedsure review data
import pandas as pd
import matplotlib.pyplot as plt
from textblob import TextBlob
import warnings
warnings.filterwarnings('ignore')

# Load the data from the CSV file
file_path = "/content/drive/MyDrive/internpulse projects/AmazonReviewsScraper-bedsure.csv"
df = pd.read_csv(file_path, encoding='latin-1')

df.head()
```

	asin	brand	product_name	product_stars	rating_count	review_rating	reviewer_name	review_text	Is_verified	helpful
0	B07YKCZHGK	Bedsure	Bedsure Queen Sheet Set, Cooling Sheets for Qu...	4.4	60591	5	Patty barboza	I can honestly say these are the best Queen si...	True	
1	B07YKCZHGK	Bedsure	Bedsure Queen Sheet Set, Cooling Sheets for Qu...	4.4	60591	5	Samuel Lacey	These sheets have been a game changer for me a...	True	
			Bedsure Queen Sheet					Sheets are soft and		

Next steps: Generate code with df | View recommended plots | New interactive sheet

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 370 entries, 0 to 369
Data columns (total 12 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   asin           370 non-null    object
 1   brand          370 non-null    object
 2   product_name   370 non-null    object
 3   product_stars  370 non-null    float64
 4   rating_count   370 non-null    int64
 5   review_rating  370 non-null    int64
 6   reviewer_name  370 non-null    object
 7   review_text    370 non-null    object
 8   Is_verified    370 non-null    bool
 9   helpful_count  370 non-null    int64
 10  date           370 non-null    object
 11  country        370 non-null    object
dtypes: bool(1), float64(1), int64(3), object(7)
memory usage: 32.3+ KB
```

```
df['date'] = pd.to_datetime(df['date'], format='mixed') ##changing date data type from object to datetime
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 370 entries, 0 to 369
Data columns (total 12 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   asin           370 non-null    object
 1   brand          370 non-null    object
 2   product_name   370 non-null    object
 3   product_stars  370 non-null    float64
 4   rating_count   370 non-null    int64
 5   review_rating  370 non-null    int64
 6   reviewer_name  370 non-null    object
 7   review_text    370 non-null    object
 8   Is_verified    370 non-null    bool
```

```
 9   helpful_count  370 non-null    int64
 10  date           370 non-null    datetime64[ns]
 11  country        370 non-null    object
dtypes: bool(1), datetime64[ns](1), float64(1), int64(3), object(6)
memory usage: 32.3+ KB
```

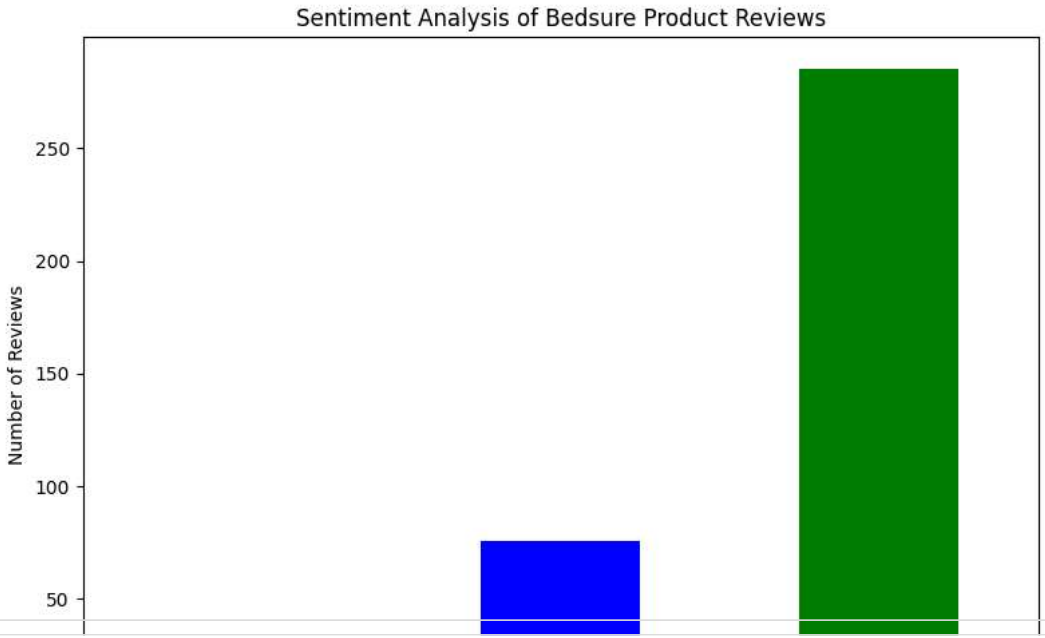
```
df.dropna(inplace=True) ##to remove null values
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 370 entries, 0 to 369
Data columns (total 12 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   asin           370 non-null    object
 1   brand          370 non-null    object
 2   product_name   370 non-null    object
 3   product_stars  370 non-null    float64
 4   rating_count   370 non-null    int64
 5   review_rating  370 non-null    int64
 6   reviewer_name  370 non-null    object
 7   review_text    370 non-null    object
 8   Is_verified    370 non-null    bool
 9   helpful_count  370 non-null    int64
 10  date           370 non-null    datetime64[ns]
 11  country        370 non-null    object
dtypes: bool(1), datetime64[ns](1), float64(1), int64(3), object(6)
memory usage: 32.3+ KB
```

```
import pandas as pd
import matplotlib.pyplot as plt
from textblob import TextBlob
import warnings
warnings.filterwarnings('ignore')

# Load the data from the CSV file
file_name = "/content/drive/MyDrive/internpulse projects/AmazonReviewsScraper-bedsure.csv"
df = pd.read_csv(file_name, encoding='latin-1')

# Perform sentiment analysis on the 'review_text' column
df['Sentiment_Polarity'] = df['review_text'].apply(lambda x: TextBlob(str(x)).sentiment.polarity)

# Categorize the sentiment based on the polarity score
def categorize_sentiment(score):
    if score > 0.1:
        return 'Positive'
    elif score < -0.1:
        return 'Negative'
    else:
        return 'Neutral'

df['Sentiment'] = df['Sentiment_Polarity'].apply(categorize_sentiment)

# Get the value counts for each sentiment category
sentiment_counts = df['Sentiment'].value_counts()

# Plot the sentiment distribution
plt.figure(figsize=(8, 6))
sentiment_counts.sort_index().plot(kind='bar', color=['red', 'blue', 'green'])
plt.title('Sentiment Analysis of Bedsure Product Reviews')
plt.xlabel('Sentiment Category')
plt.ylabel('Number of Reviews')
plt.xticks(rotation=0)
plt.tight_layout()
plt.savefig('sentiment_analysis_bar_chart.png')

# Print the final sentiment counts
print("\nSentiment Analysis Counts:")
print(sentiment_counts)
```

```
Sentiment Analysis Counts:
Sentiment
Positive    285
Neutral      76
Negative      9
Name: count, dtype: int64
```




```
## importing breescrape review data
# Load the data from the CSV file
file_path = "/content/drive/MyDrive/internpulse projects/AmazonReviewsScraper -breescrape.csv"
df_2 = pd.read_csv(file_path, encoding='latin-1')

df_2.head()
```

	ASIN	Brand	Product_name	Product_stars	Rating_count	Review_rating	Reviewer_name	Review_title	Review_content	
0	B0DSPZKCCN	Breescape	Breescape Cooling Sheet Set Queen Size - Blend...	4.3	279.0	5.0	Brooklyn	They really are cooling, worth the price	These sheets are SO worth the price. I was ske...	h
1	B0DSPZKCCN	Breescape	Breescape Cooling Sheet Set Queen Size - Blend...	4.3	279.0	5.0	Angela Cherry	Gate keep I can not!	Have you heard of these cooling sheets? They□r...	h
2	B0DSPZKCCN	Breescape	Breescape Cooling Sheet Set Queen Size - Blend...	4.3	279.0	5.0	Jessica H	Comfy and Cooling	These sheets are amazing. Very smooth and cool...	h
3	B0DSPZKCCN	Breescape	Breescape Cooling Sheet Set Queen Size - Blend...	4.3	279.0	4.0	D	Soft nice but not cooling so much	Very nice sheets well made but like everyone s...	h
4	B0DSPZKCCN	Breescape	Breescape Cooling Sheet Set Queen Size - Blend...	4.3	279.0	5.0	Aubrey HalmanAubrey Halman	Hot sleeper approved, cold fiancée approved, p...	I am in love with these sheets. They are soft ...	h

Next steps: Generate code with df_2 | View recommended plots | New interactive sheet

```
df_2.info
```

```
pandas.core.frame.DataFrame.info
def info(verbose: bool | None=None, buf: WriteBuffer[str] | None=None, max_cols: int | None=None,
memory_usage: bool | str | None=None, show_counts: bool | None=None) -> None
```

```
df_2.dropna(inplace=True)
```

```
df_2.info
```

```
pandas.core.frame.DataFrame.info
def info(verbose: bool | None=None, buf: WriteBuffer[str] | None=None, max_cols: int | None=None,
memory_usage: bool | str | None=None, show_counts: bool | None=None) -> None
```

```
/usr/local/lib/python3.12/dist-packages/pandas/core/frame.py
Print a concise summary of a DataFrame.

This method prints information about a DataFrame including
the index dtype and columns, non-null values and memory usage.
```

```
# Perform sentiment analysis on the 'Review_title' column
df_2['Sentiment_Polarity'] = df_2['Review_title'].apply(lambda x: TextBlob(str(x)).sentiment.polarity)

# Categorize the sentiment based on the polarity score
def categorize_sentiment(score):
    if score > 0.1:
        return 'Positive'
    elif score < -0.1:
        return 'Negative'
    else:
        return 'Neutral'

df_2['Sentiment'] = df_2['Sentiment_Polarity'].apply(categorize_sentiment)

# Get the value counts for each sentiment category
sentiment_counts = df_2['Sentiment'].value_counts()

# Plot the sentiment distribution
plt.figure(figsize=(8, 6))
sentiment_counts.sort_index().plot(kind='bar', color=['red', 'blue', 'green'])
plt.title('Sentiment Analysis of Breescrape Product Reviews')
plt.xlabel('Sentiment Category')
plt.ylabel('Number of Reviews')
plt.xticks(rotation=0)
plt.tight_layout()
plt.savefig('sentiment_analysis_bar_chart.png')

# Print the final sentiment counts
print("\nSentiment Analysis Counts:")
print(sentiment_counts)
```

```
Sentiment Analysis Counts:
Sentiment
Positive    6
Negative    2
Name: count, dtype: int64
```