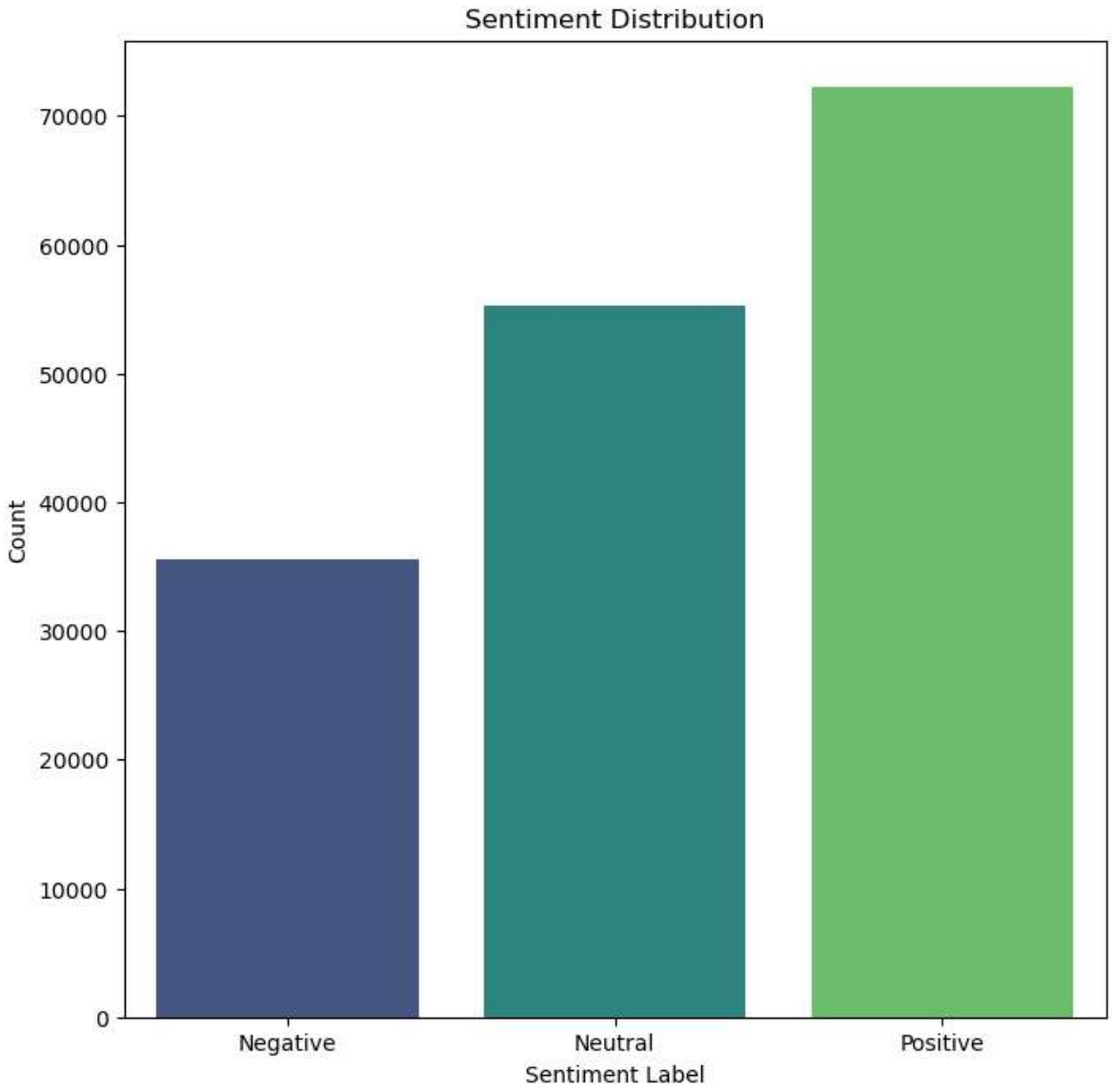
In [2]:
```
# Import necessary libraries
import pandas as pd
from textblob import TextBlob
import matplotlib.pyplot as plt
import seaborn as sns

# Load the dataset
dataset_path = "C:\\Users\\ARYAN PARIKH\\Downloads\\archive (2)\\Twitter_Data.
df = pd.read_csv(dataset_path)

# Display basic information about the dataset
print("Original Dataset Info:")
print(df.info())

# Sentiment Analysis using TextBlob
df['Sentiment'] = df['clean_text'].apply(lambda x: TextBlob(str(x)).sentiment.

# Convert sentiment polarity to sentiment labels (positive, negative, neutral)
df['Sentiment_Label'] = df['Sentiment'].apply(lambda x: 'Positive' if x > 0 el

# Data Visualization: Plotting a pie chart for sentiment distribution
plt.figure(figsize=(8, 8))
sns.countplot(x='Sentiment_Label', data=df, palette='viridis')
plt.title('Sentiment Distribution')
plt.xlabel('Sentiment Label')
plt.ylabel('Count')
plt.show()

# Save the dataset with sentiment analysis results (if needed)
# df.to_csv("twitter_data_with_sentiment.csv", index=False)
```

```
Original Dataset Info:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 162980 entries, 0 to 162979
Data columns (total 2 columns):
 #   Column      Non-Null Count   Dtype
---  ------      --------------   -----
 0   clean_text  162976 non-null  object
 1   category    162973 non-null  float64
dtypes: float64(1), object(1)
memory usage: 2.5+ MB
None
```


Sentiment Distribution
70000
60000
50000
40000
30000
20000
10000
0
Count
Negative
Neutral
Positive
Sentiment Label

In [ ]: