**INNOVATION DOCUMENT**

To convert the outlined design thinking approach into innovation and solve the problem effectively

**ADVANCED SENTIMENT ANALYSIS MODELS** -Explore the latest developments in NLP, including state-of-the-art transformer models like GPT-4 or similar successors to enhance the accuracy of sentiment analysis.

**CUSTOM DATASET CREATION-** Instead of relying solely on existing datasets, consider creating a custom dataset specific to competitor products in the airline industry. This can involve scraping social media, customer service interactions, and surveys to gather a more targeted dataset.

**REAL -TIME ANALYSIS -**Develop a real-time sentiment analysis system that can process and analyze customer feedback as it comes in. This allows companies to respond to issues or positive feedback promptly.

**MULTIMODAL ANALYSIS** -Incorporate multimodal analysis by analyzing not only text but also images and audio in customer feedback. This can provide a more comprehensive understanding of sentiment.

**EMOTION DETECTION -**Go beyond simple positive, negative, or neutral sentiment classification. Implement emotion detection to understand the specific emotions expressed in customer feedback, such as anger, joy, or frustration.

**PERSONALIZATION**-Customize sentiment analysis for each airline based on its specific products and services. This personalization can lead to more actionable insights for each company.

**INTERACTIVE DASHBOARDS**-Create interactive dashboards that allow stakeholders to explore sentiment data visually. Use tools like Tableau or Power BI to make data-driven decisions easier.

**AUTOMATED RESPONSE SYSTEM**- Develop an automated response system that suggests appropriate responses to customer feedback, especially for negative sentiments. This can improve customer service and satisfaction.

**PREDICTIVE ANALYSIS -**Implement predictive analytics to forecast future sentiment trends based on historical data. This can help airlines proactively address potential issues.

**FEEDBACK LOOP-** Establish a feedback loop where insights gained from sentiment analysis are directly integrated into the product development and marketing strategies. Continuously iterate and improve based on customer feedback.

**ETHICAL CONSIDERATIONS** -nsure ethical handling of customer data and feedback. Anonymize data when necessary and adhere to privacy regulations like GDPR or CCPA.

**USER -FRIENDLY INTERFACES**-Create user-friendly interfaces for non-technical stakeholders to access and understand sentiment analysis results easily.

**MACHINE LEARNING EXPLAINABILITY**-ncorporate machine learning explainability techniques to make the sentiment analysis model’s decisions more transparent and understandable.

**COLLOBRATIVE DECISION MAKING** -Foster collaboration between departments (e.g., marketing, customer service, product development) based on the insights generated from sentiment analysis.

**MONITORING AND ADAPTATION -**Continuously monitor the performance of the sentiment analysis system and adapt it to evolving customer feedback patterns and changes in the competitive landscape.

**INSTALLING AND IMPORTING DEPENDENCIES**

Import numpy as np # linear algebra

Import pandas as pd # data processing

Import os

#import tweepy as tw #for accessing Twitter API

#For Preprocessing

Import re # RegEx for removing non-letter characters

Import nltk #natural language processing

Nltk.download(“stopwords”)

From nltk.corpus import stopwords

From nltk.stem.porter import \*

# For Building the model

From sklearn.model\_selection import train\_test\_split

Import tensorflow as tf

Import seaborn as sns

#For data visualization

Import matplotlib.pyplot as plt

Import matplotlib.patches as mpatches

%matplotlib inline

Pd.options.plotting.backend = “plotly”

[nltk\_data] Error loading stopwords: <urlopen error [Errno -3]

[nltk\_data] Temporary failure in name resolution>

**EXPLORATORY DATA ANALYSIS**

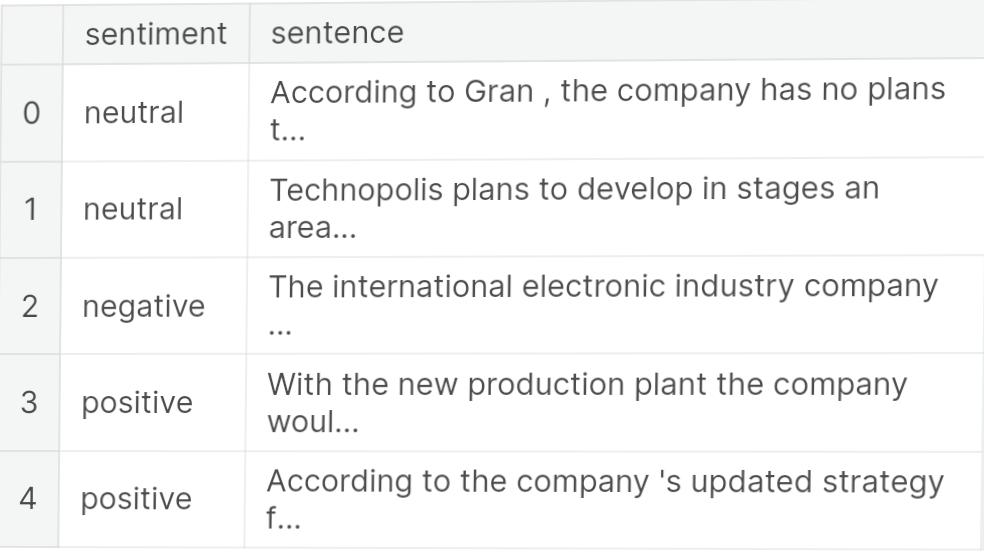
CLEANING AND PREPPING DATA SET

# Load training dataset

Df = pd.read\_csv(‘../input/sentiment-analysis-for-financial-news/all-data.csv’, names=[‘sentiment’,’sentence’])

# Output first five rows

Df.head()

****

*# Check for missing data*

df.isnull().sum()

Out[3]:

sentiment 0

sentence 0

dtype: int64

In [4]:

*# drop missing rows*

df.dropna(axis=0, inplace=True)

In [5]:

*# dimensionality of the data*

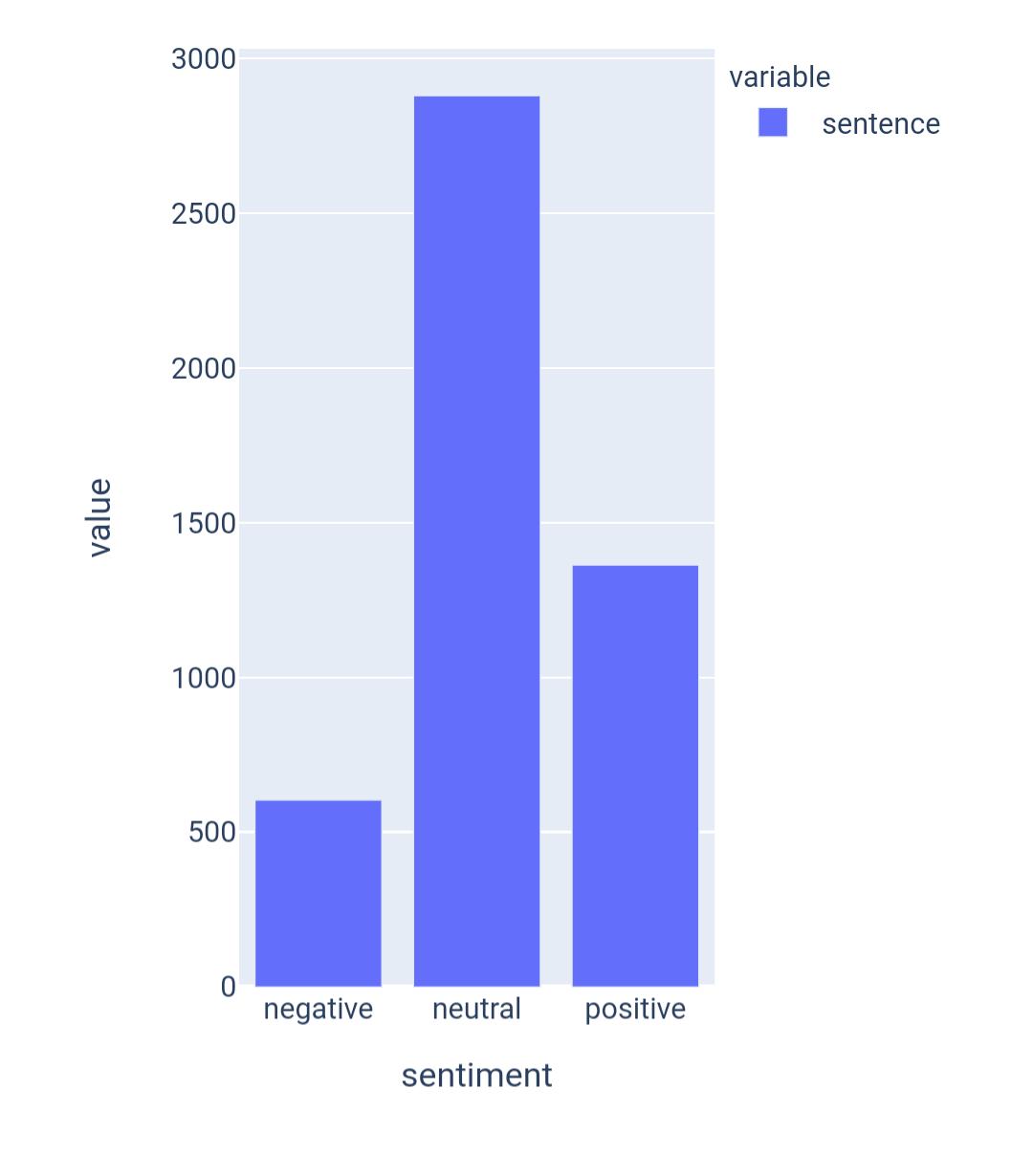
*# df1.shape*

## **Exploratory Data Analysis**

Data Visualisation

*# The distribution of sentiments*

df.groupby('sentiment').count().plot(kind='bar')



Plotting the distribution of headline lengths

df = df.rename(columns={‘sentence’: ‘clean\_text’, ‘sentiment’:’category’})

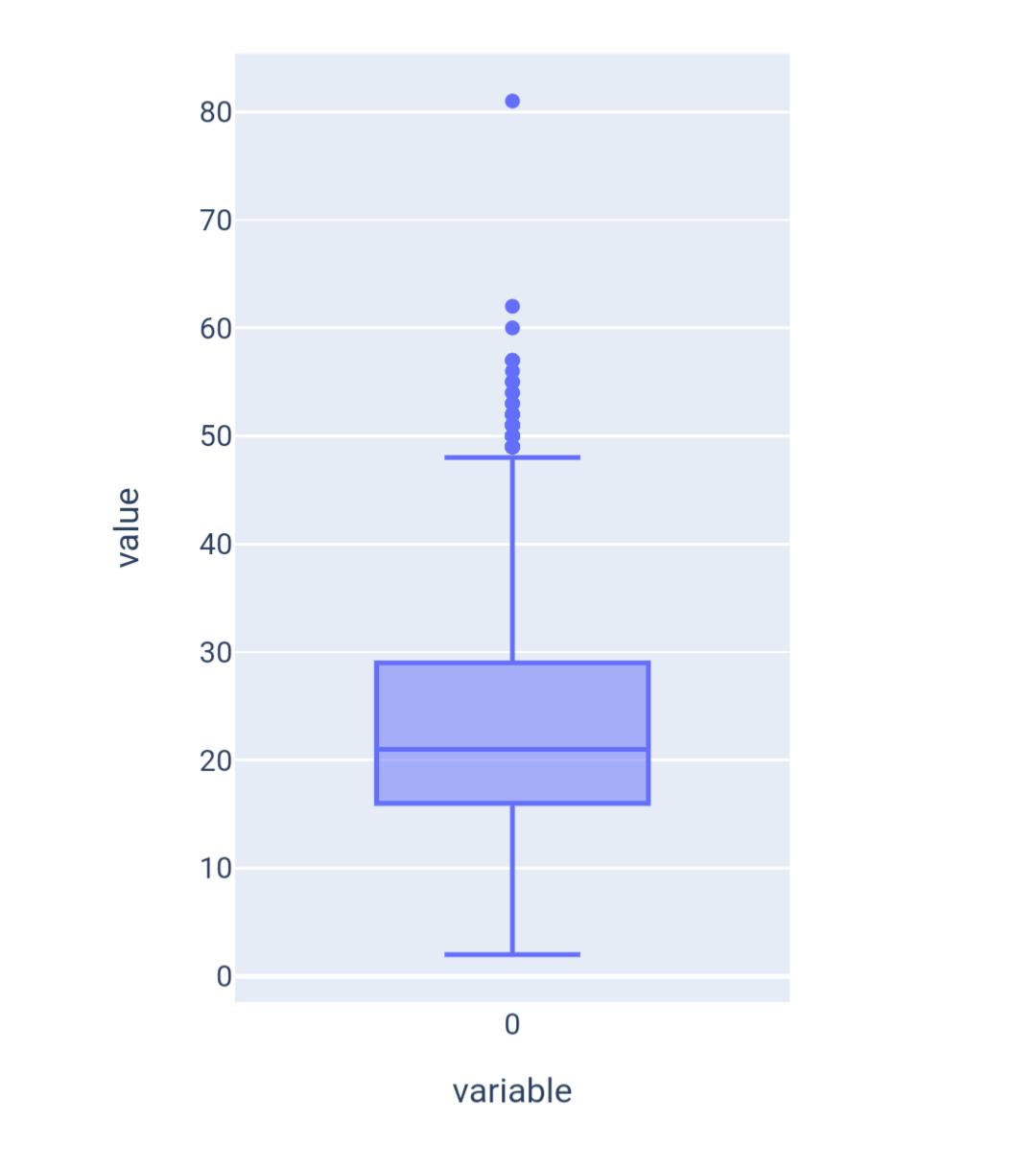
df['category'] = df['category'].map({'neutral': 'Neutral', 'positive': 'Positive', 'negative':'Negative'})

In [9]:

*# Calculate headline lengths*

headline\_len = pd.Series([len(headline.split()) for headline **in** df['clean\_text']])

*# The distribution of headline text lengths*

headline\_len.plot(kind='box')

Plotting the distribution of text length for positive sentiment headlines

In [10]:

fig = plt.figure(figsize=(14,7))

df['length'] = df.clean\_text.str.split().apply(len)

ax1 = fig.add\_subplot(122)

sns.histplot(df[df['category']=='Positive']['length'], ax=ax1,color='green')

describe = df.length[df.category=='Positive'].describe().to\_frame().round(2)

ax2 = fig.add\_subplot(121)

ax2.axis('off')

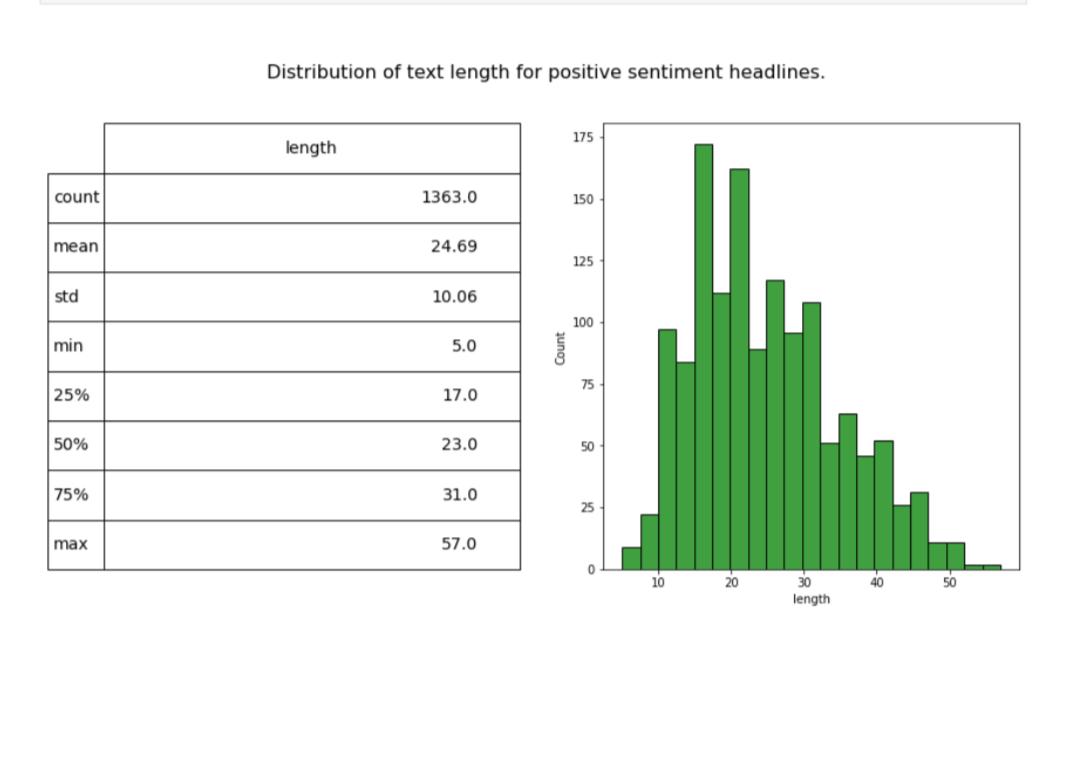
font\_size = 14

bbox = [0, 0, 1, 1]

table = ax2.table(cellText = describe.values, rowLabels = describe.index, bbox=bbox, colLabels=describe.columns)

table.set\_fontsize(font\_size)

fig.suptitle('Distribution of text length for positive sentiment headlines.', fontsize=16)

plt.show()

Plotting the Distribution of text length for Negative sentiment headlines.

In [11]:

fig = plt.figure(figsize=(14,7))

df['length'] = df.clean\_text.str.split().apply(len)

ax1 = fig.add\_subplot(122)

sns.histplot(df[df['category']=='Negative']['length'], ax=ax1,color='red')

describe = df.length[df.category=='Negative'].describe().to\_frame().round(2)

ax2 = fig.add\_subplot(121)

ax2.axis('off')

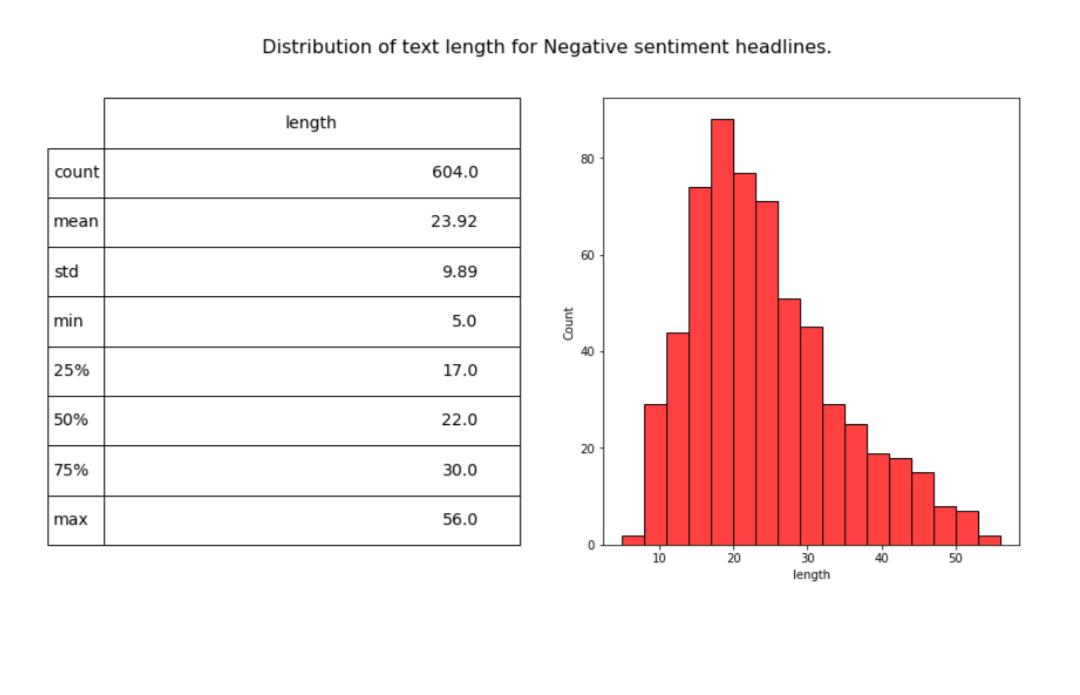
font\_size = 14

bbox = [0, 0, 1, 1]

table = ax2.table(cellText = describe.values, rowLabels = describe.index, bbox=bbox, colLabels=describe.columns)

table.set\_fontsize(font\_size)

fig.suptitle('Distribution of text length for Negative sentiment headlines.', fontsize=16)

plt.show()

*# #### Visualizing data into wordclouds*

*# from wordcloud import WordCloud, STOPWORDS*

*# def wordcount\_gen(df, category):*

*# '''*

*# Generating Word Cloud*

*# inputs:*

*# - df: tweets dataset*

*# - category: Positive/Negative/Neutral*

*# '''*

*# # Combine all tweets*

*# combined\_tweets = " ".join([tweet for tweet in df[df.category==category]['clean\_text']])*

*# # Initialize wordcloud object*

*# wc = WordCloud(background\_color='white',*

*# max\_words=50,*

*# stopwords = STOPWORDS)*

*# # Generate and plot wordcloud*

*# plt.figure(figsize=(10,10))*

*# plt.imshow(wc.generate(combined\_tweets))*

*# plt.title('{} Sentiment Words'.format(category), fontsize=20)*

*# plt.axis('off')*

*# plt.show()*

*# # Positive tweet words*

*# wordcount\_gen(df, 'Positive')*

*# # Negative tweet words*

*# wordcount\_gen(df, 'Negative')*

*# # Neutral tweet words*

*# wordcount\_gen(df, 'Neutral')*

By infusing innovation into design thinking approach, can create a powerful solution for performing sentiment analysis on customer in the airline industry, ultimately leading to improved products and services, enhanced customer satisfaction, and a competitive edge In the market.