SENTIMENT ANALYSIS FOR MARKETING

Phase-1 Document Submission

Project: Airline Sentiment Analysis

Team Members:

110521106016-R Narmadha

110521106019-Richelle Sofia Stuart

110521106011-R Kaviya

110521106017-V Priyadharshini

OVERVIEW:

Sentiment analysis is a vital component of modern marketing strategies, enabling businesses to gain valuable insights into customer perceptions, preferences, and emotions. This abstract outline a modular approach to sentiment analysis for marketing, breaking down the system into key modules to ensure effectiveness and scalability. By implementing these modules, marketers can better understand and respond to customer sentiment, optimizing their campaigns and strategies for improved results.

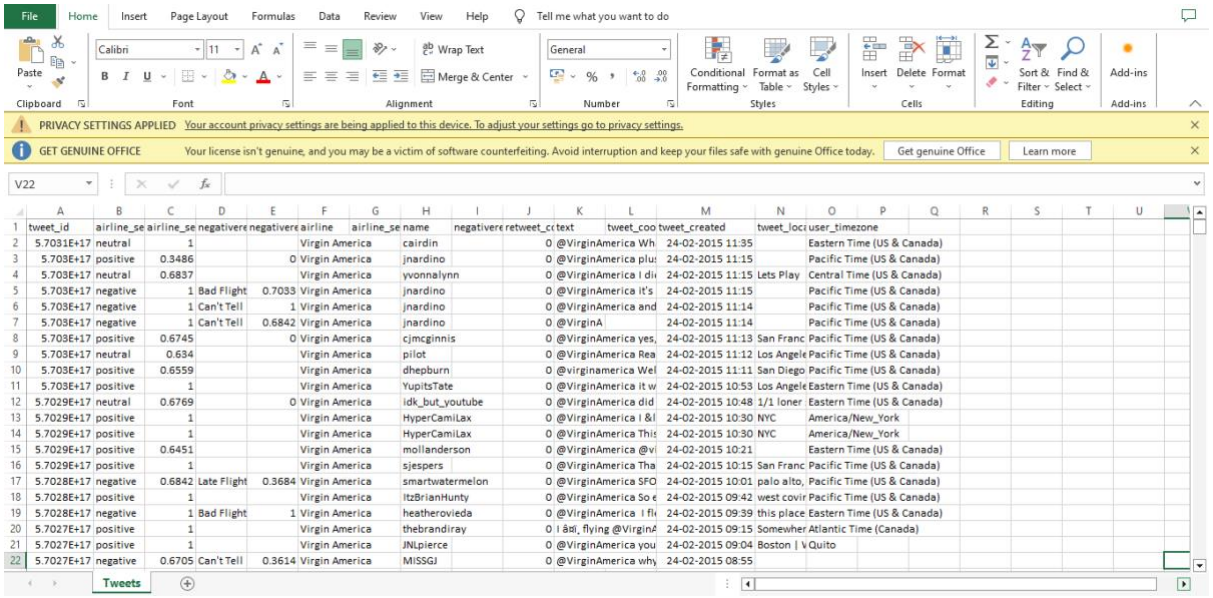
DESIGN THINKING:

DATA COLLECTION MODULE

Data collection for marketing sentiment analysis involves gathering customer feedback, social media posts, reviews, and survey responses. This data, obtained through web scraping, APIs, and user interactions, provides valuable insights into consumer emotions and preferences, enabling businesses to optimize marketing strategies and enhance customer satisfaction.

Dataset Link:

https://www.kaggle.com/datasets/crowdflower/twitter-airline-sentiment



tweet_id	airline_se	airline_se	negativere	negativere	airline	airline_se	name	negativere	retweet_c	text	tweet_coo	tweet_created	tweet_loc	user_timezone
5.7031E+17	neutral	1			Virgin America		cairdin		0	@VirginAmerica Wh		24-02-2015 11:35		Eastern Time (US & Canada)
5.703E+17	positive	0.3486		0	Virgin America		jnardino		0	@VirginAmerica plus		24-02-2015 11:15		Pacific Time (US & Canada)
5.703E+17	neutral	0.6837			Virgin America		yvonnalynn		0	@VirginAmerica I di		24-02-2015 11:15	Lets Play	Central Time (US & Canada)
5.703E+17	negative	1	Bad Flight	0.7033	Virgin America		jnardino		0	@VirginAmerica it's		24-02-2015 11:15		Pacific Time (US & Canada)
5.703E+17	negative	1	Can't Tell	1	Virgin America		jnardino		0	@VirginAmerica and		24-02-2015 11:14		Pacific Time (US & Canada)
5.703E+17	negative	1	Can't Tell	0.6842	Virgin America		jnardino		0	@VirginA		24-02-2015 11:14		Pacific Time (US & Canada)
5.703E+17	positive	0.6745		0	Virgin America		cjmcginnis		0	@VirginAmerica yes,		24-02-2015 11:13	San Franc	Pacific Time (US & Canada)
5.703E+17	neutral	0.634			Virgin America		pilot		0	@VirginAmerica Rea		24-02-2015 11:12	Los Angele	Pacific Time (US & Canada)
5.703E+17	positive	0.6559			Virgin America		dhepburn		0	@virginamerica Wel		24-02-2015 11:11	San Diego	Pacific Time (US & Canada)
5.703E+17	positive	1			Virgin America		YupitsTate		0	@VirginAmerica It w		24-02-2015 10:53	Los Angele	Eastern Time (US & Canada)
5.7029E+17	neutral	0.6769		0	Virgin America		idk_but_youtube		0	@VirginAmerica did		24-02-2015 10:48	1/1 loner	Eastern Time (US & Canada)
5.7029E+17	positive	1			Virgin America		HyperCamiLax		0	@VirginAmerica I &l		24-02-2015 10:30	NYC	America/New_York
5.7029E+17	positive	1			Virgin America		HyperCamiLax		0	@VirginAmerica This		24-02-2015 10:30	NYC	America/New_York
5.7029E+17	positive	0.6451			Virgin America		mollanderson		0	@VirginAmerica @vi		24-02-2015 10:21		Eastern Time (US & Canada)
5.7029E+17	positive	1			Virgin America		sjespers		0	@VirginAmerica Tha		24-02-2015 10:15	San Franc	Pacific Time (US & Canada)
5.7028E+17	negative	0.6842	Late Flight	0.3684	Virgin America		smartwatermelon		0	@VirginAmerica SFO		24-02-2015 10:01	palo alto,	Pacific Time (US & Canada)
5.7028E+17	positive	1			Virgin America		ItzBrianHunty		0	@VirginAmerica So e		24-02-2015 09:42	west covir	Pacific Time (US & Canada)
5.7028E+17	negative	1	Bad Flight	1	Virgin America		heatherovieda		0	@VirginAmerica I fl		24-02-2015 09:39	this place	Eastern Time (US & Canada)
5.7027E+17	positive	1			Virgin America		thebrandiray		0	I âœ¡ flying @VirginA		24-02-2015 09:15	Somewher	Atlantic Time (Canada)
5.7027E+17	positive	1			Virgin America		JNLpierce		0	@VirginAmerica you		24-02-2015 09:04	Boston | V	Quito
5.7027E+17	negative	0.6705	Can't Tell	0.3614	Virgin America		MISSGJ		0	@VirginAmerica why		24-02-2015 08:55		

Data Preprocessing:

Data preprocessing for marketing sentiment analysis is crucial for ensuring accurate results. It involves text cleaning, tokenization, stopword removal, and handling special characters. This process standardizes and prepares raw data from sources like customer reviews and social media for sentiment analysis algorithms, enhancing the quality of insights for marketing decision-makers.

Sentiment Analysis Techniques:

Sentiment analysis techniques for marketing encompass Natural Language Processing (NLP), machine learning, and sentiment lexicons. These methods classify text data into positive, negative, or neutral sentiments, aiding marketers in understanding customer opinions and tailoring strategies.

Feature Extraction:

Feature extraction for sentiment analysis in marketing involves identifying relevant keywords and attributes within text data. Techniques like TF-IDF, word embeddings, and named entity recognition help extract meaningful information, enabling marketers to uncover valuable insights from customer feedback and reviews.

Visualization:

Visualization of sentiment analysis results in marketing helps transform data into actionable insights. It includes creating interactive dashboards, charts, and graphs to display sentiment trends, sentiment distribution, and sentiment sentiment distribution across various marketing channels. These visualizations aid marketers in making data-driven decisions and optimizing campaigns.

Insights Generation:

Sentiment analysis in marketing generates actionable insights by identifying customer sentiment trends, highlighting pain points, and revealing strengths. These insights inform marketing strategies, helping businesses enhance customer satisfaction and engagement.

PYTHON PROGRAMMING:

```
from mpl_toolkits.mplot3d import Axes3D
from sklearn.preprocessing import StandardScaler
import matplotlib.pyplot as plt # plotting
import numpy as np # linear algebra
import os # accessing directory structure
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
```

There is 1 csv file in the current version of the dataset:

```
print(os.listdir('../input'))
```

```
['database.sqlite', 'Tweets.csv']
```

```
nRowsRead = 1000 # specify 'None' if want to read whole file
# Tweets.csv has 14640 rows in reality, but we are only loading/previewing t
he first 1000 rows
df1 = pd.read_csv('../input/Tweets.csv', delimiter=',', nrows = nRowsRead)
df1.dataframeName = 'Tweets.csv'
nRow, nCol = df1.shape
print(f'There are {nRow} rows and {nCol} columns')
```

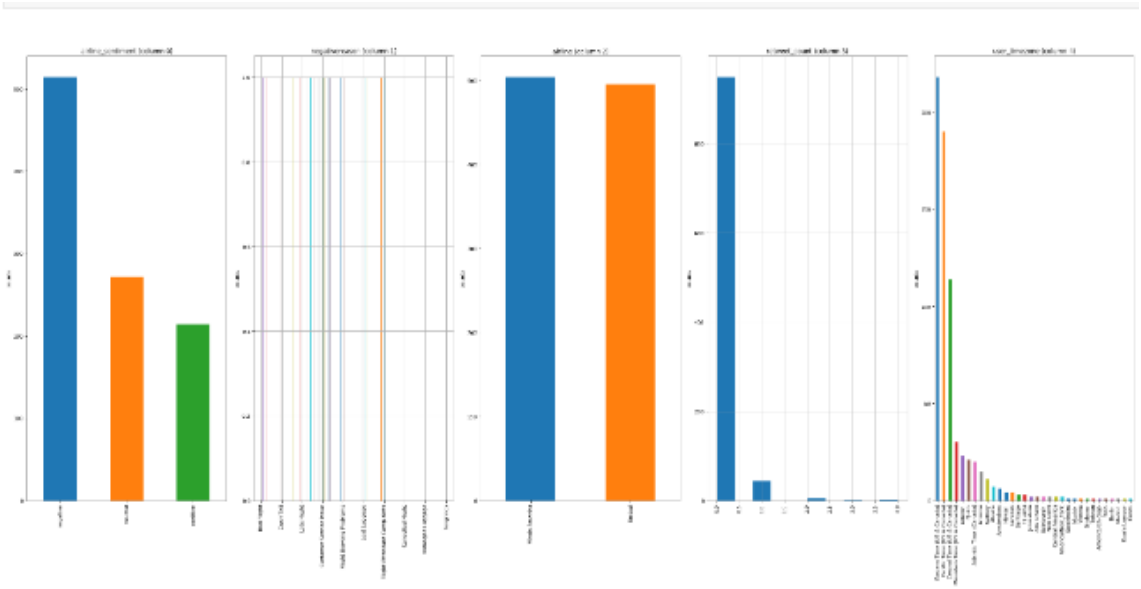
```
There are 1000 rows and 15 columns
```

```
df1.head(5)
```

	tweet_id	airline_sentiment	airline_sentiment_confidence	negativereason	negative
0	570306133677760513	neutral	1.0000	NaN	NaN
1	570301130888122368	positive	0.3486	NaN	0.0000
2	570301083672813571	neutral	0.6837	NaN	NaN
3	570301031407624196	negative	1.0000	Bad Flight	0.7033
4	570300817074462722	negative	1.0000	Can't Tell	1.0000

Distribution graphs (histogram/bar graph) of sampled columns:

```
linkcode
plotPerColumnDistribution(df1, 10, 5)
```



CONCLUSION:

In conclusion, sentiment analysis is a powerful tool for modern marketi ng. By harnessing the capabilities of this technology, businesses can gain deep i nsights into customer emotions, preferences, and perceptions. These insights e nable data-driven decision-making, the optimization of marketing strategies, an d ultimately, the enhancement of customer satisfaction and brand success in a dynamic and competitive marketplace. B