

EDS

THEORY

ACTIVITY NO.I

Name of Dataset :-Twitter US Airline Sentiment

Name : Mangesh Zate

Roll No.: CS4-63 Batch : C43

PRN : 202401040110

2304102 ALL PR: Theory Activity | Search | Kaggle | Twitter US Airline Sentiment | Pandas Commands for Twitter | Pandas.ipynb - Colab

colab.research.google.com/drive/1a1uPGEyiOAD6lADYSMfTmOj3R62ZYFFN#scrollTo=EJkQ8yyUd0VU

Pandas.ipynb | Share | Gemini

File Edit View Insert Runtime Tools Help

Commands + Code + Text RAM Disk

path = '/kaggle/Tweets.csv'
df = pd.read_csv(path)
df.head(5)

| | tweet_id | airline_sentiment | airline_sentiment_confidence | negativereason | negativereason_confidence | airline | airline_sentiment_gold | name | negativereason_gold | retw |
|---|--------------------|-------------------|------------------------------|----------------|---------------------------|----------------|------------------------|------------|---------------------|------|
| 0 | 570306133677760513 | neutral | 1.0000 | NaN | NaN | Virgin America | NaN | cairdin | NaN | NaN |
| 1 | 570301130888122368 | positive | 0.3486 | NaN | 0.0000 | Virgin America | NaN | jnardino | NaN | NaN |
| 2 | 570301083672813571 | neutral | 0.6837 | NaN | NaN | Virgin America | NaN | yvonnalynn | NaN | NaN |
| 3 | 570301031407624196 | negative | 1.0000 | Bad Flight | 0.7033 | Virgin America | NaN | jnardino | NaN | NaN |
| 4 | 570300817074462722 | negative | 1.0000 | Can't Tell | 1.0000 | Virgin America | NaN | jnardino | NaN | NaN |

Next steps: Generate code with df | View recommended plots | New interactive sheet | Connected to Python 3 Google Compute Engine backend

USD/INR -0.46% | Search | ⌂ | ENG IN | 18:35 | 28-04-2025

2304102 ALL PR: Theory Activity | Search | Kaggle | Twitter US Airline Sentiment | Pandas Commands for Twitter | Pandas.ipynb - Colab

colab.research.google.com/drive/1a1uPGEyiOAD6lADYSMFtMoj3R62ZYFFN#scrollTo=EJkQ8yyUd0VU

Pandas.ipynb

File Edit View Insert Runtime Tools Help

Commands + Code + Text RAM Disk

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

Q2:-which are last 15 data

df.tail(15)

| | tweet_id | airline_sentiment | airline_sentiment_confidence | negativereason | negativereason_confidence | airline | airline_sentiment_gold | name | negative |
|-------|--------------------|-------------------|------------------------------|------------------------|---------------------------|----------|------------------------|----------------|----------|
| 14625 | 569589643487928321 | positive | 1.0000 | NaN | NaN | American | NaN | DrCaseyJRudkin | 0s |
| 14626 | 569589460226183168 | negative | 1.0000 | Late Flight | 1.0000 | American | NaN | cataattack | |
| 14627 | 569588816438169600 | negative | 1.0000 | Cancelled Flight | 1.0000 | American | NaN | KristinaMeyer7 | |
| 14628 | 569588651925098496 | positive | 1.0000 | NaN | NaN | American | NaN | jhalldc | |
| 14629 | 569588591602458624 | negative | 1.0000 | Customer Service Issue | 1.0000 | American | NaN | jontgreen89 | |
| 14630 | 569588473050611712 | positive | 1.0000 | NaN | NaN | American | NaN | Laurelinesblog | |

Connected to Python 3 Google Compute Engine backend

Hot weather Now

Search

ENG IN

2304102 ALL PR: Theory Activity | Search | Kaggle | Twitter US Airline Sentiment | Pandas Commands for Twitter | Pandas.ipynb - Colab

colab.research.google.com/drive/1a1uPGEyiOAD6IADYSMfTmOj3R62ZYFFN#scrollTo=EJkQ8yyUd0VU

Pandas.ipynb

File Edit View Insert Runtime Tools Help

Commands + Code + Text ✓

14 user_timezone 9820 non-null object
dtypes: float64(2), int64(2), object(11)
memory usage: 1.7+ MB

df.info()

Q4:- how many number of column and there name

```
0s <class 'pandas.core.frame.DataFrame'>
RangeIndex: 14640 entries, 0 to 14639
Data columns (total 15 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   tweet_id         14640 non-null   int64  
 1   airline_sentiment 14640 non-null   object  
 2   airline_sentiment_confidence 14640 non-null   float64 
 3   negativereason    9178 non-null   object  
 4   negativereason_confidence 10522 non-null   float64 
 5   airline           14640 non-null   object  
 6   airline_sentiment_gold 40 non-null    object  
 7   name              14640 non-null   object  
 8   negativereason_gold 32 non-null    object  
 9   retweet_count     14640 non-null   int64  
 10  text              14640 non-null   object  
 11  tweet_coord       1019 non-null   object  
 12  tweet_created     14640 non-null   object  
 13  tweet_location    9907 non-null   object  
 14  user_timezone     9820 non-null   object  
dtypes: float64(2), int64(2), object(11)
memory usage: 1.7+ MB
```

[17] df.describe()



CO Pandas.ipynb ⭐ ⚙

File Edit View Insert Runtime Tools Help

Commands + Code + Text



✓ [17] df.describe()

Q5.what of retweet_count in dataset



| | tweet_id | airline_sentiment_confidence | negativereason_confidence | retweet_count |
|-------|--------------|------------------------------|---------------------------|---------------|
| count | 1.464000e+04 | 14640.000000 | 10522.000000 | 14640.000000 |
| mean | 5.692184e+17 | 0.900169 | 0.638298 | 0.082650 |
| std | 7.791112e+14 | 0.162830 | 0.330440 | 0.745778 |
| min | 5.675883e+17 | 0.335000 | 0.000000 | 0.000000 |
| 25% | 5.685592e+17 | 0.692300 | 0.360600 | 0.000000 |
| 50% | 5.694779e+17 | 1.000000 | 0.670600 | 0.000000 |
| 75% | 5.698905e+17 | 1.000000 | 1.000000 | 0.000000 |
| max | 5.703106e+17 | 1.000000 | 1.000000 | 44.000000 |



✓ ⏴ df.shape

Q6.what is size of dataset



➡ (14640, 15)



✓ [19] df.columns

Commands

+ Code

+ Text

✓ RAM Disk

| | | | | | |
|-----|-----|--------------|----------|----------|-----------|
| ☰ | std | 7.791112e+14 | 0.162830 | 0.330440 | 0.745778 |
| 🔗 | min | 5.675883e+17 | 0.335000 | 0.000000 | 0.000000 |
| 🔍 | 25% | 5.685592e+17 | 0.692300 | 0.360600 | 0.000000 |
| ↔ | 50% | 5.694779e+17 | 1.000000 | 0.670600 | 0.000000 |
| {x} | 75% | 5.698905e+17 | 1.000000 | 1.000000 | 0.000000 |
| | max | 5.703106e+17 | 1.000000 | 1.000000 | 44.000000 |

[18] df.shape

(14640, 15)

[19] df.columns

```
Index(['tweet_id', 'airline_sentiment', 'airline_sentiment_confidence',
       'negativereason', 'negativereason_confidence', 'airline',
       'airline_sentiment_gold', 'name', 'negativereason_gold',
       'retweet_count', 'text', 'tweet_coord', 'tweet_created',
       'tweet_location', 'user_timezone'],
      dtype='object')
```

[20] df.index
df.dtypes

```
0
```

| tweet_id | int64 |
|----------|-------|
|----------|-------|

✓ Connected to Python 3 Google Compute Engine backend

colab.research.google.com/drive/1a1uPGEyiOAD6IADYSMfTmOj3R62ZYFFN#scrollTo=EJkQ8yyUd0VU**Pandas.ipynb**

File Edit View Insert Runtime Tools Help

Commands

+ Code + Text

0s

df.index
df.dtypes

| | |
|------------------------------|---------|
| tweet_id | int64 |
| airline_sentiment | object |
| airline_sentiment_confidence | float64 |
| negativereason | object |
| negativereason_confidence | float64 |
| airline | object |
| airline_sentiment_gold | object |
| name | object |
| negativereason_gold | object |
| retweet_count | int64 |
| text | object |
| tweet_coord | object |
| tweet_created | object |
| tweet_location | object |
| user_timezone | object |

dtype: object

Q7. The retweet_count is a string or categorical data type

Q8. what type of data of tweet_id





Commands + Code + Text

RA
Di

```
US Airways    2913
American     2759
Southwest    2420
Delta        2222
Virgin America 504
dtype: int64
```

Q9. how many airline_sentiment give me value

```
[22] df['airline_sentiment'].value_counts()
df['airline'].unique()

array(['Virgin America', 'United', 'Southwest', 'Delta', 'US Airways',
       'American'], dtype=object)
```

Q10.how many unique airline

```
df['airline_sentiment'].unique()

array(['neutral', 'positive', 'negative'], dtype=object)
```

Q11.which are unique airline_sentiment

Missing Values

```
[24] df.isnull().sum()

tweet_id      0
```

Connected to Python 3 Google Compute Engine backend



Commands

+ Code + Text



Missing Values

0s



df.isnull().sum()

Q12.Which columns may need data cleaning or imputation

0

| | |
|------------------------------|-------|
| tweet_id | 0 |
| airline_sentiment | 0 |
| airline_sentiment_confidence | 0 |
| negativereason | 5462 |
| negativereason_confidence | 4118 |
| airline | 0 |
| airline_sentiment_gold | 14600 |
| name | 0 |
| negativereason_gold | 14608 |
| retweet_count | 0 |
| text | 0 |
| tweet_coord | 13621 |
| tweet_created | 0 |
| tweet_location | 4733 |
| user_timezone | 4820 |

Connected to Python 3 G



2304102 ALL PR: Theory Activity | Search | Kaggle | Twitter US Airline Sentiment | Pandas Commands for Twitter | Pandas.ipynb - Colab

colab.research.google.com/drive/1a1uPGEyiOAD6lADYSMfTmOj3R62ZYFFN#scrollTo=EJkQ8yyUd0VU

Pandas.ipynb | Share | Gemini

File Edit View Insert Runtime Tools Help

Commands + Code + Text RAM Disk

[25] df.dropna()

| | tweet_id | airline_sentiment | airline_sentiment_confidence | negativereason | negativereason_confidence | airline | airline_sentiment_gold | name | negativereason_ |
|------|--------------------|-------------------|------------------------------|------------------|---------------------------|------------|------------------------|-----------------|-----------------|
| 4206 | 567778009013178368 | negative | 1.0000 | Cancelled Flight | 1.0000 | United | negative | realmikesmith | Cancelled |
| 9536 | 569887533267611648 | negative | 0.8563 | Late Flight | 0.5938 | US Airways | negative | ConstanceSCHERE | Late |

[26] df.fillna('Unknown')

| | tweet_id | airline_sentiment | airline_sentiment_confidence | negativereason | negativereason_confidence | airline | airline_sentiment_gold | name | negativereason_ |
|---|--------------------|-------------------|------------------------------|----------------|---------------------------|----------------|------------------------|----------|-----------------|
| 0 | 570306133677760513 | neutral | 1.0000 | Unknown | Unknown | Virgin America | Unknown | cairdin | Unkn |
| 1 | 570301130888122368 | positive | 0.3486 | Unknown | 0.0 | Virgin America | Unknown | jnardino | Unkn |

Connected to Python 3 Google Compute Engine backend

Sports headline Carlo Ancelotti... | Search | ENG IN | 18:40 | 28-04-2025

2304102 ALL PR: Theory Activity | Search | Kaggle | Twitter US Airline Sentiment | Pandas Commands for Twitter | Pandas.ipynb - Colab

colab.research.google.com/drive/1a1uPGEyiOAD6lADYSMfTmOj3R62ZYFFN#scrollTo=EJkQ8yyUd0VU

Pandas.ipynb

File Edit View Insert Runtime Tools Help

Commands + Code + Text RAM Disk

Filtering Data

0s df[df['airline'] == 'Delta']

| | tweet_id | airline_sentiment | airline_sentiment_confidence | negativereason | negativereason_confidence | airline | airline_sentiment_gold | name | negativereason |
|------|--------------------|-------------------|------------------------------|-----------------------------|---------------------------|---------|------------------------|-----------------|----------------|
| 6746 | 570309345281486848 | positive | 0.6469 | NaN | NaN | Delta | NaN | jaxbra | |
| 6747 | 570309308937842688 | neutral | 0.6869 | NaN | NaN | Delta | NaN | Oneladyyouadore | |
| 6748 | 570308513181904901 | neutral | 1.0000 | NaN | NaN | Delta | NaN | Oneladyyouadore | |
| 6749 | 570307820962373632 | negative | 1.0000 | Flight Attendant Complaints | 0.6522 | Delta | NaN | Oneladyyouadore | |

Connected to Python 3 Google Compute Engine backend

Sports headline Carlo Ancelotti... | Search | File Explorer | Microsoft Edge | Microsoft Word | Microsoft Excel | Google Sheets | Google Slides | Powerpoint | ENG IN | 18:40 | 28-04-2025

Pandas.ipynb

File Edit View Insert Runtime Tools Help

Commands + Code + Text

RAM Disk



Share

Gemini



0s df[df['airline_sentiment'] == 'negative']



tweet_id airline_sentiment airline_sentiment_confidence negativereson negativereson_confidence airline airline_sentiment_gold name negativereson_

3 570301031407624196 negative 1.0000 Bad Flight 0.7033 Virgin America NaN jnardino

4 570300817074462722 negative 1.0000 Can't Tell 1.0000 Virgin America NaN jnardino

5 570300767074181121 negative 1.0000 Can't Tell 0.6842 Virgin America NaN jnardino

15 570282469121007616 negative 0.6842 Late Flight 0.3684 Virgin America NaN smartwatermelon

17 570276917301137409 negative 1.0000 Bad Flight 1.0000 Virgin America NaN heatherovieda

...

14631 569588464896876545 negative 1.0000 Bad Flight 1.0000 American NaN MDDavis7

✓ Connected to Python 3 Google Compute Engine backend

9 Sports headline
Carlo Ancelotti...



Search



18:40
28-04-2025

colab.research.google.com/drive/1a1uPGEyiOAD6lADYSMfTmOj3R62ZYFFN#scrollTo=EJkQ8yyUd0VU

Pandas.ipynb Share Gemini

File Edit View Insert Runtime Tools Help

Commands + Code + Text RAM Disk

df[df['airline_sentiment'] == 'negative']

| | tweet_id | airline_sentiment | airline_sentiment_confidence | negativereason | negativereason_confidence | airline | airline_sentiment_gold | name | negativereason_ |
|-------|--------------------|-------------------|------------------------------|----------------|---------------------------|----------------|------------------------|-----------------|-----------------|
| 3 | 570301031407624196 | negative | 1.0000 | Bad Flight | 0.7033 | Virgin America | NaN | jnardino | |
| 4 | 570300817074462722 | negative | 1.0000 | Can't Tell | 1.0000 | Virgin America | NaN | jnardino | |
| 5 | 570300767074181121 | negative | 1.0000 | Can't Tell | 0.6842 | Virgin America | NaN | jnardino | |
| 15 | 570282469121007616 | negative | 0.6842 | Late Flight | 0.3684 | Virgin America | NaN | smartwatermelon | |
| 17 | 570276917301137409 | negative | 1.0000 | Bad Flight | 1.0000 | Virgin America | NaN | heatherovieda | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 14631 | 569588464896876545 | negative | 1.0000 | Bad Flight | 1.0000 | American | NaN | MDDavis7 | |

Connected to Python 3 Google Compute Engine backend

colab.research.google.com/drive/1a1uPGEyiOAD6IADYSMFtMoj3R62ZYFFN#scrollTo=EJkQ8yyUd0VU

Pandas.ipynb

File Edit View Insert Runtime Tools Help

Commands + Code + Text

0s df.loc[0]

Q13.give me data of 1st row

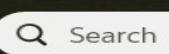
| | |
|------------------------------|-------------------------------------|
| tweet_id | 570306133677760513 |
| airline_sentiment | neutral |
| airline_sentiment_confidence | 1.0 |
| negativereson | NaN |
| negativereson_confidence | NaN |
| airline | Virgin America |
| airline_sentiment_gold | NaN |
| name | cairdin |
| negativereson_gold | NaN |
| retweet_count | 0 |
| text | @VirginAmerica What @dhepburn said. |
| tweet_coord | NaN |
| tweet_created | 2015-02-24 11:35:52 -0800 |
| tweet_location | NaN |
| user_timezone | Eastern Time (US & Canada) |

dtype: object

Connected to Python 3 Google Compute Engine backend



Sunset
6:55 pm



2304102 ALL PR: Theory Activity | Search | Kaggle | Twitter US Airline Sentiment

colab.research.google.com/drive/1a1uPGEyiOAD6IADYSMfTmOj3R62ZYFFN#scrollTo=EJkQ8yyUdC

Pandas.ipynb

File Edit View Insert Runtime Tools Help

Commands + Code + Text

df.iloc[0]

| | |
|------------------------------|-------------------------------------|
| tweet_id | 570306133677760513 |
| airline_sentiment | neutral |
| airline_sentiment_confidence | 1.0 |
| negativereason | NaN |
| negativereason_confidence | NaN |
| airline | Virgin America |
| airline_sentiment_gold | NaN |
| name | cairdin |
| negativereason_gold | NaN |
| retweet_count | 0 |
| text | @VirginAmerica What @dhepburn said. |
| tweet_coord | NaN |
| tweet_created | 2015-02-24 11:35:52 -0800 |
| tweet_location | NaN |
| user_timezone | Eastern Time (US & Canada) |

Connected to Python 3 Google Colab

Sunset 6:55 pm

Search

colab.research.google.com/drive/1a1uPGEyiOAD6IADYSMFtMoj3R62ZYFFN#scrollTo=EJkQ8yyUd0VU

Pandas.ipynb

File Edit View Insert Runtime Tools Help

Commands

+ Code + Text

tweet_location



user_timezone

Eastern Time (US & Canada)



dtype: object



[31] df[['airline', 'text']].head()



airline

text



| | | |
|---|----------------|---|
| 0 | Virgin America | @VirginAmerica What @dhepburn said. |
| 1 | Virgin America | @VirginAmerica plus you've added commercials t... |
| 2 | Virgin America | @VirginAmerica I didn't today... Must mean I n... |
| 3 | Virgin America | @VirginAmerica it's really aggressive to blast... |
| 4 | Virgin America | @VirginAmerica and it's a really big bad thing... |

Data Manipulation

0s ➔ df.rename(columns={'text': 'tweet_text'}, inplace=True)
print(df)

➔ tweet_id airline_sentiment airline_sentiment_confidence
0 570306133677760513 neutral 1.0000
1 570301130888122368 positive 0.3486
2 570301083672813571 neutral 0.6837
3 570301031407624196 negative 1.0000
4 570300817074462722 negative 1.0000
...

✓ Connected to Python 3 Google Compute Engine backend



Commands + Code + Text



- 0 Virgin America @VirginAmerica What @dhepburn said.
- 1 Virgin America @VirginAmerica plus you've added commercials t...
- 2 Virgin America @VirginAmerica I didn't today... Must mean I n...
- 3 Virgin America @VirginAmerica it's really aggressive to blast...
- 4 Virgin America @VirginAmerica and it's a really big bad thing...



Data Manipulation



0s df.rename(columns={'text': 'tweet_text'}, inplace=True)
print(df)

```
0s  df.rename(columns={'text': 'tweet_text'}, inplace=True)  
print(df)  
  
          tweet_id airline_sentiment  airline_sentiment_confidence \\\n0      570306133677760513        neutral                  1.0000  
1      570301130888122368      positive                  0.3486  
2      570301083672813571        neutral                  0.6837  
3      570301031407624196     negative                  1.0000  
4      570300817074462722     negative                  1.0000  
...           ...           ...           ...  
14635   569587686496825344      positive                  0.3487  
14636   569587371693355008     negative                  1.0000  
14637   569587242672398336        neutral                  1.0000  
14638   569587188687634433     negative                  1.0000  
14639   569587140490866689        neutral                  0.6771
```



```
          negativereason  negativereason_confidence  airline \\\n0                 NaN                      NaN  Virgin America  
1                 NaN                      0.0000  Virgin America  
2                 NaN                      NaN  Virgin America
```

✓ Connected to Python 3 Google Compute Engine backend

9 21 min delay
NH60 / Pune Na...



Search



ENG
IN



2304102 ALL PR: Theory Activity | Search | Kaggle | Twitter US Airline Sentiment | Pandas Commands for Twitter | Pandas.ipynb - Colab

colab.research.google.com/drive/1a1uPGEyiOAD6lADYSMfTmOj3R62ZYFFN#scrollTo=EJkQ8yyUd0VU

Pandas.ipynb

File Edit View Insert Runtime Tools Help

Commands + Code + Text RAM Disk

[14640 rows x 15 columns]

df.drop('new_column', axis=1, inplace=True, errors='ignore')
print(df)

| | Customer Service Issue | airline_sentiment_gold | name | negative reason | retweet_count | tweet_text |
|-------|------------------------|------------------------|-----------------|-----------------|---------------|---|
| 14636 | Customer Service Issue | 1.0000 | American | cairdin | NaN | @VirginAmerica What @dhepburn said. |
| 14637 | | NaN | American | jnardino | NaN | @VirginAmerica plus you've added commercials t... |
| 14638 | Customer Service Issue | 0.6659 | American | yvonnalynn | NaN | @VirginAmerica I didn't today... Must mean I n... |
| 14639 | | 0.0000 | American | jnardino | NaN | @VirginAmerica it's really aggressive to blast... |
| | | | | jnardino | NaN | @VirginAmerica and it's a really big bad thing... |
| | | | | ... | ... | ... |
| 14635 | | NaN | KristenReenders | | NaN | @AmericanAir thank you we got on a different f... |
| 14636 | | NaN | itsropes | | NaN | @AmericanAir leaving over 20 minutes Late Flig... |
| 14637 | | NaN | sanyabun | | NaN | @AmericanAir Please bring American Airlines to... |
| 14638 | | NaN | SraJackson | | NaN | |
| 14639 | | NaN | daviddtwu | | NaN | |

Connected to Python 3 Google Compute Engine backend

21 min delay NH60 / Pune Na... | Search | ENG IN | 18:43 | 28-04-2025

Q Commands + Code + Text

✓ RAM Disk

GroupBy & Aggregation

[53] df.groupby('airline_sentiment').size()

0s

| airline_sentiment | 0 |
|-------------------|------|
| negative | 9178 |
| neutral | 3099 |
| positive | 2363 |

dtype: int64

[54] df.groupby('airline')[['airline_sentiment']].value_counts()

0s

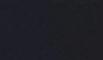
Q14.what is count of american positive airline.

| airline | airline_sentiment | count |
|----------|-------------------|-------|
| American | negative | 1960 |
| | neutral | 463 |
| | positive | 336 |
| Delta | negative | 955 |
| | neutral | 723 |
| | positive | 544 |

1 9 21 min delay
NH60 / Pune Na...



Search



ENG

IN



18:43
28-04-2025

✓ Connected to Python 3 Google Compute Engine backend



Commands

+ Code

+ Text



| | | | |
|--|----------------|----------|-----|
| | | positive | 492 |
| | Virgin America | negative | 181 |
| | | neutral | 171 |
| | | positive | 152 |

dtype: int64

{x}



0s ⏴ df.groupby(['airline', 'airline_sentiment']).size()



| | | airline | airline_sentiment | 0 |
|------------|--|------------|-------------------|------|
| American | | American | negative | 1960 |
| | | | neutral | 463 |
| | | | positive | 336 |
| Delta | | Delta | negative | 955 |
| | | | neutral | 723 |
| | | | positive | 544 |
| Southwest | | Southwest | negative | 1186 |
| | | | neutral | 664 |
| | | | positive | 570 |
| US Airways | | US Airways | negative | 2263 |

✓ Connected to Python 3 Google Compute Engine backend

36°C
Sunny

Search

ENG
IN

 Pandas.ipynb  

File Edit View Insert Runtime Tools Help

   Share

Commands + Code + Text

RAM
Disk

dtype: int64

[56] df.groupby(['airline', 'airline_sentiment']).size().unstack()

airline_sentiment negative neutral positive

airline

| | | | |
|----------------|------|-----|-----|
| American | 1960 | 463 | 336 |
| Delta | 955 | 723 | 544 |
| Southwest | 1186 | 664 | 570 |
| US Airways | 2263 | 381 | 269 |
| United | 2633 | 697 | 492 |
| Virgin America | 181 | 171 | 152 |

df['airline_sentiment'].value_counts(normalize=True) * 100

proportion

airline_sentiment

| | |
|----------|-----------|
| negative | 62.691257 |
| neutral | 21.168033 |
| positive | 16.140710 |

dtype: float64

Connected to Python 3 Google Compute Engine backend

USD/INR
-0.46%

Search

ENG
IN

??

```
positive      16.140710
```

dtype: float64

```
s df.groupby('airline').size()
```

```
0
```

airline

| | |
|-----------------------|------|
| American | 2759 |
| Delta | 2222 |
| Southwest | 2420 |
| US Airways | 2913 |
| United | 3822 |
| Virgin America | 504 |

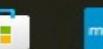
dtype: int64

✓ Connected to Python 3 Google Compute Engine backend

USD/INR
-0.46%



Search



=====

Thank you

=====