



NYC-Airbnb-Recommendation-Engine-NLP / notebooks / Sentiment_Analysis.ipynb

 **kamalova** Created using Colaboratory

History

 1 contributor

1.63 MB

...



Airbnb Recommendation Engine for NYC through Sentiment Analysis

Author: Nurgul Kurbanali kyzy

Table of Contentss

- Business Case
- What is Sentiment Analysis
- Aim of the Notebook
- Data Understanding
- Exploratory Data Analysis(EDA)
- Preprocessing Comments
- Estimate Sentiment Polarity
- Explore Polarity Variable and Correlation with Score Features

1. Business Case

About Airbnb: *You can host anything, anywhere, so guests can enjoy everything, everywhere.*

Nowadays the demand for short and long-term temporary accommodation is increasing thanks to easing travel conditions. This demand positively affects the number of online platforms that allow you to make reservations before traveling. **Airbnb** is one such platform, which allows travelers to make accommodation reservations based on the fact that the host leases all or part of his or her home to the traveler.

Customer reviews play an important role in the customer's decision to purchase a product or use a service. Customer preferences and opinions are affected by other customers' reviews online, on blogs or over social networking platforms

The main goal of this work is to combine both recommendation system and sentiment analysis in order to recommend the most accurate listings for users based on their preferences in **New York City**. Since both domains suffer from the lack of labeled data, to overcome that, this project detects the opinions polarity score using **NLTK VADER** (Valence Aware Dictionary and Sentiment Reasoner) Lexicon.

We'll therefore split our approaches into following sections:

- Exploring available AirBnb listings in NYC
- Measuring polarity/sentiment scores along with vader_lexicon. This polarity

measurement adapts to *pos*, *neu*, *neg*, and compound. By simply taking the compound from these values, a new feature was created on the data.

- Building a recommendation engine with Collaborative Filtering to predict sentiment score for all reviewer-listing pairs and making personalised recommendations for each user based on their ranked preferences.


2. What is Sentiment Analysis(opinion mining)?

Sentiment analysis, also referred to as opinion mining, is an approach to natural language processing (NLP) that identifies the emotional tone behind a body of text. This is a popular way for organizations to determine and categorize opinions about a product, service, or idea. It involves the use of data mining, machine learning (ML) and artificial intelligence (AI) to mine text for sentiment and subjective information.

In addition to identifying sentiment, opinion mining can extract the polarity (or the amount of positivity and negativity), subject and opinion holder within the text. [more](#)

3. Aim of the Notebook

The goal of this Notebook is to perform necessary text pre-processing steps for each reviewer's comments (including language detection). Further we will calculate the compound polarity of each comment by applying **NLTK's Vader lexicon**

 Project Notebook was run in **Google Colab**

4. Data Understanding

The dataset is obtained from [Inside Airbnb](#). It is a mission driven project that provides data and advocacy about Airbnb's impact on residential communities. For the purpose of this project we downloaded the most recent quarterly datasets between December, 2021 - September, 2022 which includes Airbnb **summary review** data and Listing ID in New York City. For the sake of the memory we had to load 20000 dataset from the each quarter which made 80000 in total. Let's discover further in detail.

4.1. Data Load

```
In [1]: # Loading Necessary Libraries
import numpy as np
import pandas as pd
pd.set_option('display.max_colwidth', None)
import string
string.punctuation
import re
from sklearn.feature_extraction.text import TfidfVectorizer

# Data visualization
import seaborn as sns
from sklearn.model_selection import train_test_split
import matplotlib.pyplot as plt
%matplotlib inline
# Seaborn's beautiful styling
import seaborn as sns
sns.set_style('whitegrid')
import plotly.express as px
# to get rid of the warnings
import warnings
warnings.filterwarnings("ignore")
from nltk.sentiment.vader import SentimentIntensityAnalyzer
from nltk.tokenize import TweetTokenizer
from wordcloud import WordCloud
from wordcloud.wordcloud import STOPWORDS
```

```
In [ ]: %rm -rf sample_data/
```

```
In [ ]: # Loading Datasets into Colab
reviews_march = pd.read_csv('/content/reviews_march.csv', compression='gzip', on_bad_lines='skip',
                             low_memory=False, nrows=20000);
reviews_jun = pd.read_csv('/content/reviews_jun.csv', compression='gzip', on_bad_lines='skip',
                           low_memory=False, nrows=20000);
reviews_sep = pd.read_csv('/content/reviews_sep.csv', compression='gzip', on_bad_lines='skip',
                           low_memory=False, nrows=20000);
reviews_dec = pd.read_csv('/content/reviews_dec21.csv', compression='gzip', on_bad_lines='skip',
                           low_memory=False, nrows=20000);
```

low_memory=True, nrows=20000,

```
In [ ]: # Print dimensionality of DataFrame
print(reviews_march.shape, reviews_jun.shape, reviews_sep.shape, reviews_dec.shape)

(20000, 6) (20000, 6) (20000, 6) (20000, 6)
```

```
In [ ]: # Concatinate Reviews from each quarter
df_reviews = pd.concat([reviews_march, reviews_jun, reviews_sep, reviews_dec])
```

```
In [ ]: # Preview DataFrame samples
df_reviews
```

Out[]:

	listing_id	id	date	reviewer_id	reviewer_name	comments
						Notre séjour de trois nuits.\r Nous avons apprécié L'appartement qui est très bien situé. Agréable,\r propre et bien soigné. C'est idéal pour une famille de 3 ou 4\r personnes.\r Petits soucis en arrivant il y avait personne pour nous recevoir, et il\r manquait le savon pour la douche, le liquide vaisselle, nous les avons\r reçu de surlendemain. Il y a aussi le bruit du Métro de NY, donc une\r première nuit difficile si on est pas habitué. Jennifer est correcte le\r remboursement de la caution était très rapide. A part ces petits\r détails notre court séjour c'est bien passé. \r Si j'ai la possibilité de revenir sur NY pour les vacances, je reprendrai\r à "The Midtown Castle"\r Jean\r Possession - Ile de La Réunion\r
0	2595	17857	2009-11-21	50679	Jean	
1	2595	19176	2009-12-05	53267	Cate	Great experience.
2	2595	19760	2009-12-10	38960	Anita	I've stayed with my friend at the Midtown Castle for six days and it\r was a lovely place to be. A big spacious room with a pointy roof,\r which really makes you feel like staying in a castle. The location is\r perfect. It is just a few steps from Macy's Time Square and Theatre\r District. Everything worked just perfect with the keys etc. Thank you\r so much Jennifer, we had a great time in New York.\r Attention: it's on the 4th floor without a lift :-) but definately worth it!
3	2595	24320	2010-	71130	Kai-I lwe	We've been staying here for about 9 nights, enjoying to be in the\r center of the city, that never sleeps...short ways to everywhere in\r Manhattan, by subway or by walk. Midtown castle is a beautiful and

...	04-09	tastful place, Jennifer and Tori relaxed and friendly hosts - thats why we - the three Berliners - recommand that place! Good to have WiFi and a little kitchen too!
4	2595	46312	2010-05-25	117113	Alicia	We had a wonderful stay at Jennifer's charming apartment! They were very organized and helpful; I would definitely recommend staying at the Midtown Castle!
...
19995	72265	161050979	2017-06-16	109542482	John	Vanessa was very pleasant and communication was very good.
19996	72265	163401732	2017-06-24	1282541	Sofia	Great location! Close to G train.
						Highly recommend. Cannot beat this value.
19997	72265	252657179	2018-04-11	8936723	Yo	Great location - 4 minute walk to subway and 20 sec to bus which connects you easily and quickly to various parts of Manhattan and Brooklyn. Organic (as well as regular) grocery stores and lots of awesome restaurants and stores near by. Very safe neighborhoods Nice room - not big but it's plenty enough and everything works well. It's nice & warm in the winter. Even though the bedroom is separated by a curtain to the kitchen, because the host is mainly in the other section of the apartment, you have a lot of privacy. Vanessa is a very friendly, interesting and helpful host.
19998	72265	277084426	2018-06-15	17160406	Ioannis	Vanessa is a great and very polite host and gives you as much privacy as you want. The room can be seen in the photos and has everything you need. The location is amazing as well with plenty of bars, restaurants and stores around and literally half a block away from G train.
19999	72265	294169497	2018-07-21	165490874	Elsa	Vanessa is a very hospitable and friendly person, I was able to interact with her a lot. The apartment is ideally located, the room very convenient. I have nothing to say except that I had a great time, and it was a great experience for me. Thank you Vanessa, I would definitely go back!

80000 rows × 6 columns

```
In [ ]: # Drop unnecessary columns
df_reviews = df_reviews.drop(columns=['id', 'reviewer_name'])
```

```
# Check for null values
df_reviews.isnull().sum()
```

```
Out[ ]: listing_id    0
        date         0
        reviewer_id  0
        comments     6
        dtype: int64
```

5. Exploratory Data Analysis

```
In [ ]: # Seperate date column into day month and year
df_reviews['year'],df_reviews['month'],df_reviews['days'] = df_reviews['date'].str.split('-',2).str
df_reviews.head()
```

```
Out[ ]: listing_id  date  reviewer_id  comments  year  month  days
```

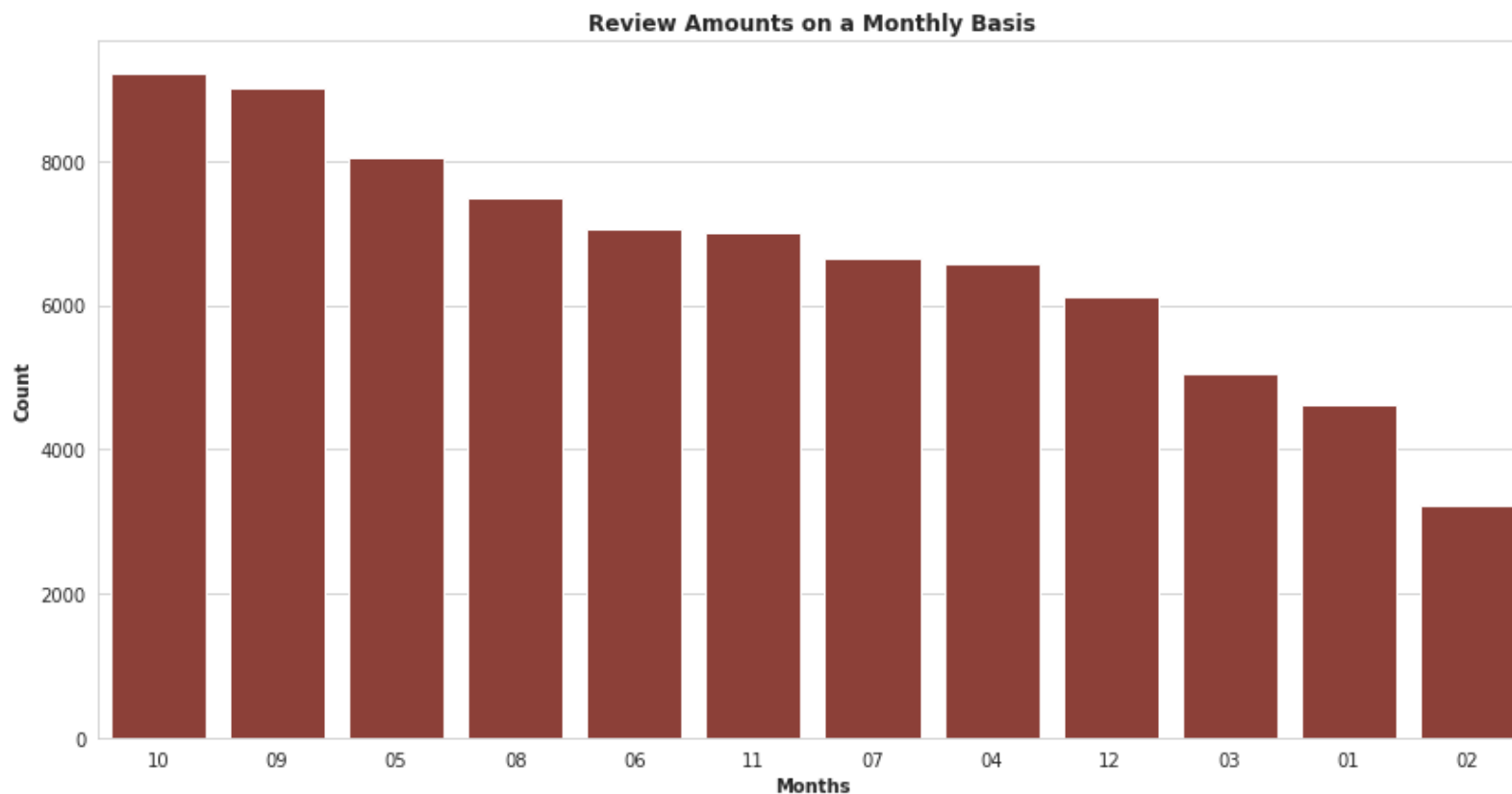
	listing_id	date	reviewer_id	comments	year	month	days
0	2595	2009-11-21	50679	Notre séjour de trois nuits.\r Nous avons apprécié L'appartement qui est très bien situé. Agréable, propre et bien soigné. C'est idéal pour une famille de 3 ou 4 personnes.\r Petits soucis en arrivant il y avait personne pour nous recevoir, et il manquait le savon pour la douche, le liquide vaisselle, nous les avons reçu de surlendemain. Il y a aussi le bruit du Métro de NY, donc une première nuit difficile si on est pas habitué. Jennifer est correcte le remboursement de la caution était très rapide. A part ces petits détails notre court séjour c'est bien passé. \r Si j'ai la possibilité de revenir sur NY pour les vacances, je reprendrai à "The Midtown Castle"\r Jean\r Possession - Ile de La Réunion\r	2009	11	21
1	2595	2009-12-05	53267	Great experiance.	2009	12	05
2	2595	2009-12-10	38960	I've stayed with my friend at the Midtown Castle for six days and it was a lovely place to be. A big spacious room with a pointy roof, which really makes you feel like staying in a castle. The location is perfect. It is just a few steps from Macy's Time Square and Theatre District. Everything worked just perfect with the keys etc. Thank you so much Jennifer, we had a great time in New York.\r Attention: it's on the 4th floor without a lift :-)) but definetely worth it!	2009	12	10
3	2595	2010-04-09	71130	We've been staying here for about 9 nights, enjoying to be in the center of the city, that never sleeps...short ways to everywhere in Manhattan, by subway or by walk. Midtown castle is a beauftiful and tastful place, Jennifer and Tori relaxed and friendly hosts - thats why we - the three Berliners - recommand that place! Good	2010	04	09

to have WiFi and a little kitchen too!

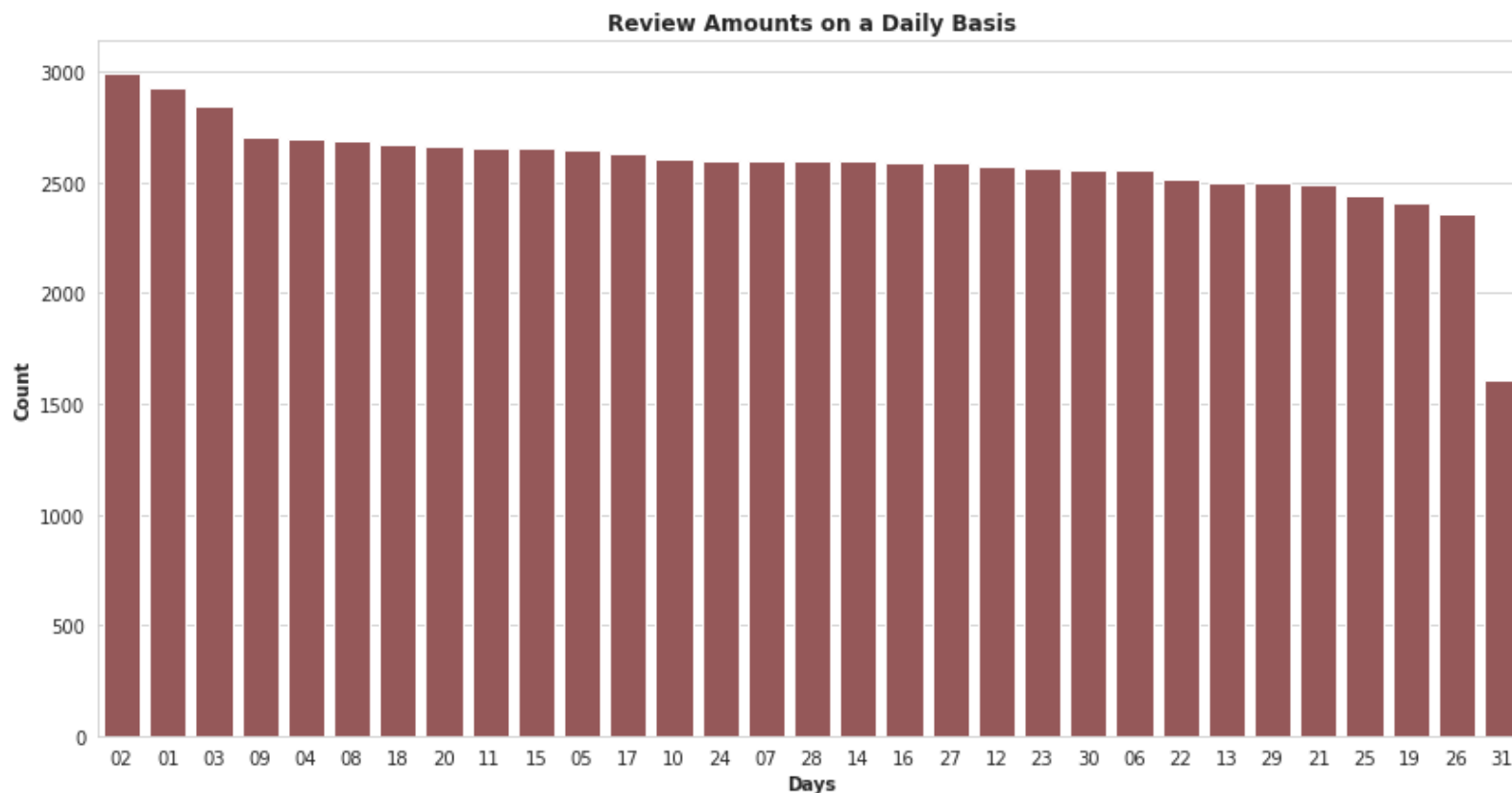
4 2595 2010-05-25 117113 We had a wonderful stay at Jennifer's charming apartment! They were very organized and helpful; I would definitely recommend staying at the Midtown Castle! 2010 05 25

```
In [ ]: # Change data types into datetime format and get weekdays
df_reviews['date'] = pd.to_datetime(df_reviews.date)
df_reviews['weekday'] = df_reviews['date'].dt.day_name()
```

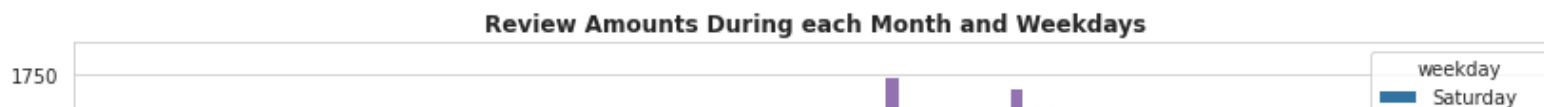
```
In [ ]: # Plot monthly based review amounts
plt.figure(figsize=(14,7))
ax = sns.countplot(x='month', data=df_reviews, order=df_reviews['month'].value_counts().index, color='#9a342a')
plt.title('Review Amounts on a Monthly Basis',fontweight="bold")
ax.set_xlabel('Months', fontweight="bold")
ax.set_ylabel('Count', fontweight="bold");
```

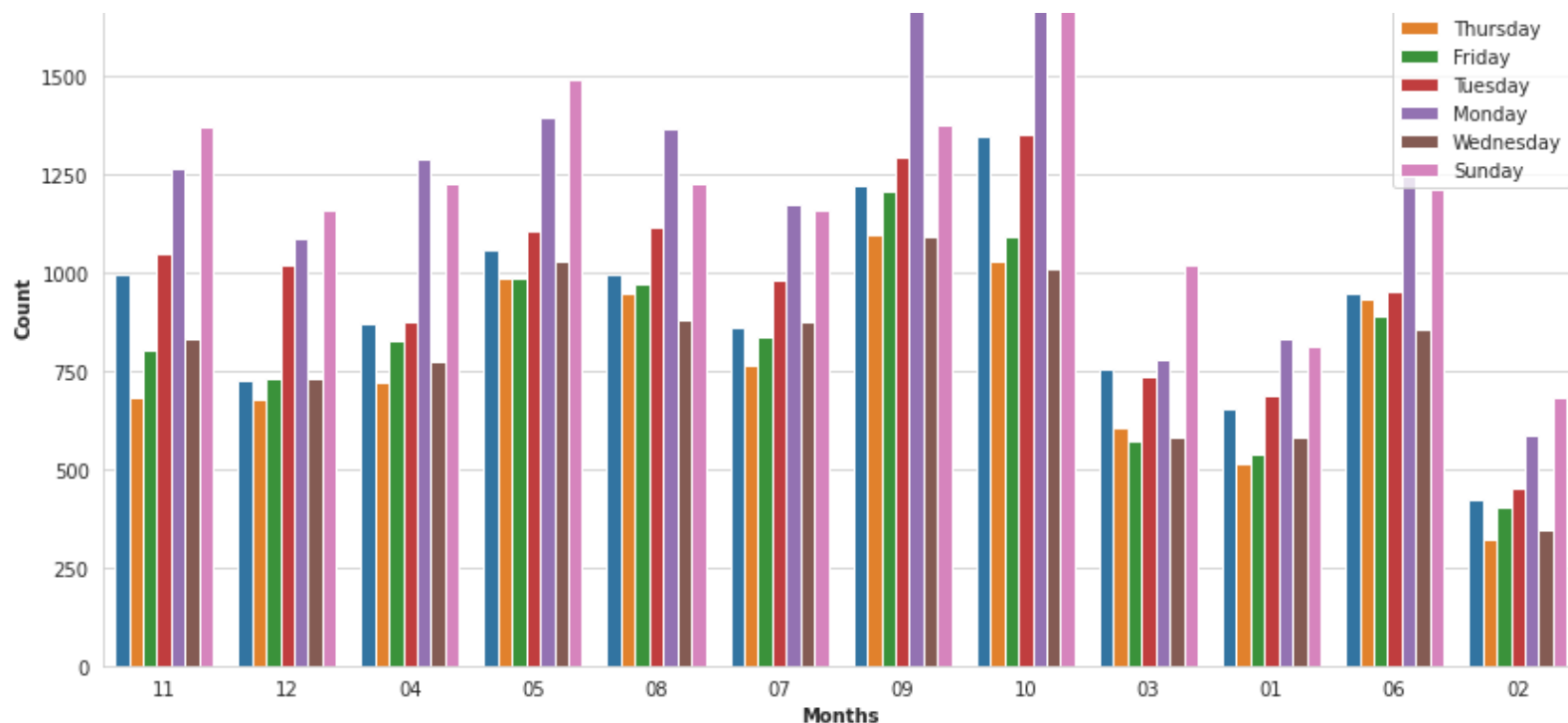



```
In [ ]: # Plot daily based review amounts
plt.figure(figsize=(14,7))
ax = sns.countplot(x='days', data=df_reviews, order=df_reviews['days'].value_counts().index, color='#9f4e4f')
plt.title('Review Amounts on a Daily Basis',fontweight="bold")
ax.set_xlabel('Days', fontweight="bold")
ax.set_ylabel('Count', fontweight="bold");
```



```
In [ ]: plt.figure(figsize=(14,7))
ax = sns.countplot(x='month', data=df_reviews, hue='weekday')
plt.title('Review Amounts During each Month and Weekdays',fontweight="bold")
ax.set_xlabel('Months', fontweight="bold")
ax.set_ylabel('Count', fontweight="bold");
```





```
In [ ]: df_reviews['reviewer_id'].nunique()
```

```
Out[ ]: 28943
```

There are 28943 different ID for reviews in this data.

```
In [ ]: # Print DataFrame columns
df_reviews.columns
```

```
Out[ ]: Index(['listing_id', 'date', 'reviewer_id', 'comments', 'year', 'month',
            'days', 'weekday'],
            dtype='object')
```

```
In [ ]: # Drop unnecessary columns
df_reviews.drop(columns=['date', 'year', 'days'], inplace=True)
```

6. Preprocessing comments for Sentiment Analysis

```
In [ ]: # Drop null values
df_reviews.dropna(subset =['comments'],how='any',axis=0,inplace=True)
```

```
In [ ]: ! pip install langdetect
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Collecting langdetect
  Downloading langdetect-1.0.9.tar.gz (981 kB)
    |████████████████████████████████████████| 981 kB 28.3 MB/s
Requirement already satisfied: six in /usr/local/lib/python3.8/dist-packages (from langdetect) (1.15.0)
Building wheels for collected packages: langdetect
  Building wheel for langdetect (setup.py) ... done
  Created wheel for langdetect: filename=langdetect-1.0.9-py3-none-any.whl size=993242 sha256=6451138fbd43914e971a287f10c61210458192783fd15d883af5cc156a08de5a
  Stored in directory: /root/.cache/pip/wheels/13/c7/b0/79f66658626032e78fc1a83103690ef6797d551cb22e56e734
Successfully built langdetect
Installing collected packages: langdetect
Successfully installed langdetect-1.0.9
```

```
In [ ]: # Import Language detect
from langdetect import detect
```

```
In [ ]: # Create function to detect language of the comments
def detect_lang(x):
    try:
        lang = detect(x)
    except:
        lang = 'other'
    return lang

df_reviews['language'] = df_reviews['comments'].apply(detect_lang)
```

```
In [ ]: # Preview first 5 rows of dataframe
df_reviews
```

```
Out[ ]: listing_id      id  reviewer_id  reviewer_name      comments  month  weekday  language
```

```

                Notre séjour de trois nuits.\r
Nous avons apprécié L'appartement qui
est très bien situé. Agréable, propre et
```

					est très bien situé. Agréable, propre et bien soigné. C'est idéal pour une famille de 3 ou 4 personnes.\r			
0	2595	17857	50679	Jean	Petits soucis en arrivant il y avait personne pour nous recevoir, et il manquait le savon pour la douche, le liquide vaisselle, nous les avons reçu de surlendemain. Il y a aussi le bruit du Métro de NY, donc une première nuit difficile si on est pas habitué. Jennifer est correcte le remboursement de la caution était très rapide. A part ces petits détails notre court séjour c'est bien passé. \r	11	Saturday	fr
					Si j'ai la possibilité de revenir sur NY pour les vacances, je reprendrai à "The Midtown Castle"\r			
					Jean\r			
					Possession - Ile de La Réunion\r			
1	2595	19176	53267	Cate	Great experience.	12	Saturday	ro
					I've stayed with my friend at the Midtown Castle for six days and it was a lovely place to be. A big spacious room with a pointy roof, which really makes you feel like staying in a castle. The location is perfect.			
2	2595	19760	38960	Anita	It is just a few steps from Macy's Time Square and Theatre District. Everything worked just perfect with the keys etc. Thank you so much Jennifer, we had a great time in New York.\r	12	Thursday	en
					Attention: it's on the 4th floor without a lift :-) but definitely worth it!			
					We've been staying here for about 9 nights, enjoying to be in the center of the city, that never sleeps...short ways to everywhere in Manhattan, by subway or by walk. Midtown castle is a beautiful and tastful place, Jennifer and Tori relaxed and friendly hosts - thats why we - the three Berliners - recommand that place! Good to have WiFi and a little kitchen too!			
3	2595	34320	71130	Kai-Uwe		04	Friday	en
					We had a wonderful stay at Jennifer's charming apartment! They were very organized and helpful; I would definitely recommend staying at the Midtown			
4	2595	46312	117113	Alicia		05	Tuesday	en

Castle!									
...
19995	72265	161050979	109542482	John	Vanessa was very pleasant and communication was very good.	06	Friday	en	
19996	72265	163401732	1282541	Sofia	Great location! Close to G train.	06	Saturday	en	
					Highly recommend. Cannot beat this value.				
					Great location - 4 minute walk to subway and 20 sec to bus which connects you easily and quickly to various parts of Manhattan and Brooklyn. Organic (as well as regular) grocery stores and lots of awesome restaurants and stores near by.				
19997	72265	252657179	8936723	Yo	Very safe neighborhoods Nice room - not big but it's plenty enough and everything works well. It's nice & warm in the winter. Even though the bedroom is separated by a curtain to the kitchen, because the host is mainly in the other section of the apartment, you have a lot of privacy. Vanessa is a very friendly, interesting and helpful host.	04	Wednesday	en	
					Vanessa is a great and very polite host and gives you as much privacy as you want. The room can be seen in the photos and has everything you need. The location is amazing as well with plenty of bars, restaurants and stores around and literally half a block away from G train.				
19998	72265	277084426	17160406	Ioannis		06	Friday	en	
					Vanessa is a very hospitable and friendly person, I was able to interact with her a lot. The apartment is ideally located, the room very convenient. I have nothing to say except that I had a great time, and it was a great experience for me. Thank you Vanessa, I would definitely go back!				
19999	72265	294169497	165490874	Elsa		07	Saturday	en	

79994 rows × 8 columns

```
In [ ]: # Count total language of all DataFrame
df_reviews.language.value_counts()
```

```
Out[ ]: en      71630
fr       3059
es       2071
de       1009
it        498
pt        316
nl        254
zh-cn     150
ja        141
ko        108
other       90
ro         83
sv         82
ru         80
af         51
no         51
da         47
so         40
ca         34
zh-tw     29
pl         26
tl         24
fi         23
cs         20
cy         18
id         11
sw         10
tr          9
hu          8
he          6
el          5
th          3
hr          3
bg          2
et          1
sl          1
vi          1
Name: language, dtype: int64
```

```
In [ ]: # Plot language column
df_reviews['language'].value_counts().plot.bar()
```



```
In [ ]: # Sample of French Language
df_reviews.loc[df_reviews['language']=='fr']
```

	listing_id	reviewer_id	comments	month	weekday	language
			Notre séjour de trois nuits.\r\nNous avons apprécié L'appartement qui est très bien situé. Agréable, propre et bien soigné. C'est idéal pour une famille de 3 ou 4 personnes.\r\nPetits soucis en arrivant il y avait personne pour nous recevoir, et il manquait le savon pour la douche, le liquide vaisselle, nous les avons reçus de surlendemain. Il y a aussi le bruit du Métro de NY, donc une première nuit difficile si on est pas habitué. Jennifer est correcte le remboursement de la caution était très rapide. A part ces petits détails notre court séjour c'est bien passé. \r\nSi j'ai la possibilité de revenir sur NY pour les vacances, je reprendrai à "The Midtown Castle"\r\nJean\r\nPossession - Ile de La Réunion\r	11	Saturday	fr
123	5178	118953	chambre tres bien située, du bruit mais c'est normal on est en plein Manhatan. Le propriétaire est tres discret je vous le recommande tres bonne adresse	08	Wednesday	fr
			Première fois à New York et très bon séjour ! L'appartement de Shun est vraiment bien situé, à proximité de tout et la chambre est plutôt grande et confortable. De plus le métro est à côté donc c'est également très			

149	5178	311594	confortable. De plus le metro est a cote donc c'est egalement tres pratique pour aller dans les lieux a visiter qui n'etaient pas proche. Times Square et Central park a cote ! Et enfin Shun est vraiment tres gentil et discret. Il etait la si nous avions des questions. Encore merci a lui pour ce tres bon sejour, nous recommandons cet hebergement, si nous devons retourner a New York nous choisirons encore son appartement !	03	Tuesday	fr
200	5178	1337014	Nous avons séjourné une semaine chez Shun. Grande chambre très propre et bien équipée (frigo, micro-onde, bureau, ordinateur et wifi, étagère, grand placard...).\r Shun a été très attentif à ce que nous manquions de rien tout en étant très discret. Nous l'avons très peu vu car nous rentrions souvent tard. Nous recommandons cette chambre.	04	Monday	fr
229	5178	3768630	La chambre était confortable bien que dépouillé de toutes décorations, surtout l'emplacement est idéal, proche des stations de métro et de broadway. \r Shun est un hote discret mais disponible	01	Saturday	fr
...
19636	68974	71309926	Idéalement placé, au carrefour de Soho, Little Italie, Chinatown. Boutiques pour les filles, resto pour le soir et à 100 mètres du métro! Logement très agréable, bien équipé (cuisine, ordinateur a dispo etc....) et bien chauffé dans lequel on se sent bien. Pas particulièrement bruyant car au 4eme étage (et je suis de la campagne donc je suis sensible au bruit). Alicia très arrangeante quant aux horaires d'arrivée (22h) et de départ(19h), très réactive par sms ou mails si on a des questions. Nous étions 7 (2 couples avec enfants) et avons été, grands et petits, super contents de retrouver notre loft le soir après une journée de visites. Très agréable, pas seulement un endroit pour dormir. On y a passé de bonnes soirées et les enfants ont pu dormir tard le matin sans être réveillé par le bruit. Ce qu'il faut savoir: 1) 4 étages d'escalier à monter. Personnellement, je ne suis pas sportive mais ça ne m'a pas posé de problème. 2) 1 seul salle de bain donc il faut s'organiser quand on est 7 Merci Alicia. J'y reviendrais!	01	Monday	fr
19661	68974	28711110	Super appartement dans un quartier animé et central. La décoration est sympa. La cuisine très bien équipée. Salon cosy. Bien chauffé en hiver. Très recommandable. Nous avons adoré ce loft! Nous étions une famille de 5 et c'était idéal pour notre séjour de 6 jours, bien situé pour explorer à pied le sud de Manhattan (puis en métro depuis Grand Street a 2 pas du loft, pour le centre et nord). L'appart est tres propre, bien aménagé et organisé, très bien climatisé, confortable et très spacieux! Exercices cardio gratuits en	12	Tuesday	fr

			prime pour monter 4 (dernier) etages tous les jours (pour nous ce n'était pas un problème mais il faut le savoir). De très beaux livres et une deco design avec certains meubles vintage, exactement comme sur les photo.			
19675	68974	190923463	Vue sur l'avenue et la one world tower en arrière plan. Machine à laver et sèche linge neufs, parfait avec une famille. Le bâtiment en brique rouge, typique NYC, est relativement calme, le salon donne sur l'avenue mais étant au quatrième étage, c'est tout à fait acceptable. Nous avons été accueilli par l'housekeeper d'Alicia mais Alicia réagit très rapidement par messages, elle communique aussi par téléphone si nécessaire. Elle nous a donné de bonnes adresses 'take away' pour dîner à l'appart, et s'est montré très flexible pour l'horaire de départ. Je recommande vivement ce loft!	09	Sunday	fr
			Bel appartement, clair, bien placé, conforme aux photos même si la propreté laissait un peu à désirer (ménage fait de manière superficielle et rideau de douche très sale)...			
19823	71248	17152030	Bien pour 2 adultes et un enfant (matelas futon dans la pièce commune). La climatisation un peu bruyante, mais très utile! De bonnes adresses et de bons conseils.	08	Sunday	fr
19980	72190	45444395	Très bon séjour chez Hudson, qui est très arrangeant et réactif. Excellent emplacement. Appartement de caractère, caractéristique de New York.	01	Saturday	fr

3059 rows × 6 columns

```
In [ ]: df_reviews.loc[df_reviews['language']=='sl']
```

```
Out [ ]:
```

	listing_id	date	reviewer_id	comments	language
6458	105790	2014-11-24	10169217	Privet vsem! Ya po russki. Arman ochen' gostepriimniy, obchitel'niy chozyain. Raspologenie kvartiri horoshee, Komnata super. Eche poedu obyazatel'no ostanovlus' u Armana.\r	sl
13841	56859	2017-05-14	116271063	Skvělá lokalita+bezpečna, blízkost obchodu, metro do centra 5 min., dobrá komunikace s hostitelem, koupelna sdílena s hostitelem primerene cista, pokoj studenejsi s hlukem z ulice. (Spatne uzavreni okna)	sl

```
In [ ]: df_reviews.loc[df_reviews['language']=='et']
```

```
Out [ ]: listing_id date reviewer_id comments language
```

```
In [ ]: df_reviews.loc[df_reviews['language']=='pl']
```

Out[]:

	listing_id	date	reviewer_id		comments	language	
	1129	6990	2009-10-31	7691		nice	pl
	7910	27644	2014-06-27	15813035	Emma była bardzo miły gospodarz. Spędziliśmy wspaniały czas w Nowym Jorku.Apartament znajduje się w bardzo wygodnym miejscu. W pobliżu stacji metra. Jest duży, czysty i dobrze wyposażone, ale meble są dość stare.	kr	pl
	11590	45542	2018-09-11	214030038		nice guy nice place	pl
	11848	45556	2019-11-29	131269496		Very nice!	pl
	12450	53469	2013-08-15	7462716	Bezpieczna, spokojna i cicha okolica. Sprawna komunikacja z kartą MTA autobus i metro. Wygodna łazienka i ogarnięta otwarta kuchnia. Gospodarz bardzo pomocny i otwarty.		pl
	19235	74680	2014-06-19	9845175		Super miejsce. Gorąco polecam	pl
	1134	45542	2018-09-11	214030038		nice guy nice place	pl
	1730	45556	2019-11-29	131269496		Very nice!	pl
	3495	89621	2017-09-17	143337753		Was nice	pl
	3496	89621	2017-09-21	150733803	Pokój bardzo czysty i przyjemny. Właściciele bardzo mili, nie było problemu z moim późniejszym przyjazdem. Lokalizacja bardzo wygodna, blisko do komunikacji miejskiej. Jak najbardziej mogę polecić to miejsce :)		pl
	3707	6990	2009-10-31	7691		nice	pl
	3840	53469	2013-08-15	7462716	Bezpieczna, spokojna i cicha okolica. Sprawna komunikacja z kartą MTA autobus i metro. Wygodna łazienka i ogarnięta otwarta kuchnia. Gospodarz bardzo pomocny i otwarty.		pl
	15824	141154	2019-09-30	28863242		nice	pl
			2015		Wszystko było dokładnie zgodne z opisem. Lokalizacja jest znakomita blisko centralnego parku i ulicy columbus. Niedaleko od stacji metra. Bezpieczna okolica, sklepy		

2015_

16396	148825	2019-05-04	31344052	całodobowe, czego więcej chcieć :). Przyjechaliśmy po 22:00 ale kolega właściciela czekał na nas z kluczami. Polecamy.	pl
1182	6990	2009-10-31	7691	nice	pl
7906	27644	2014-06-27	15813035	Emma była bardzo miły gospodarz. Spędziliśmy wspaniały czas w Nowym Jorku. Apartament znajduje się w bardzo wygodnym miejscu. W pobliżu stacji metra. Jest duży, czysty i dobrze wyposażony, ale meble są dość stare. Krzysztof	pl
8527	29628	2022-07-16	4126638	To był naprawdę świetny pobyt. Warunki w domu u Chrisa były bardzo dobre: 1. Czysto 2. Duży pokój i łazienka 3. Cicha lokalizacja ale bardzo blisko metra z szybkim dojazdem na Manhattan 4. Skorzystaliśmy z bezpłatnych kart wstępu do muzeów 5. Bardzo przyjazny pies Jitsu Zdecydowanie polecamy!	pl
11383	45542	2018-09-11	214030038	nice guy nice place	pl
11641	45556	2019-11-29	131269496	Very nice!	pl
12255	53469	2013-08-15	7462716	Bezpieczna, spokojna i cicha okolica. Sprawna komunikacja z kartą MTA autobus i metro. Wygodna łazienka i ogarnięta otwarta kuchnia. Gospodarz bardzo pomocny i otwarty.	pl
19464	74680	2014-06-19	9845175	Super miejsce. Gorąco polecam	pl
1520	6990	2009-10-31	7691	nice	pl
8296	27644	2014-06-27	15813035	Emma była bardzo miły gospodarz. Spędziliśmy wspaniały czas w Nowym Jorku. Apartament znajduje się w bardzo wygodnym miejscu. W pobliżu stacji metra. Jest duży, czysty i dobrze wyposażony, ale meble są dość stare. Krzysztof	pl
12081	45542	2018-09-11	214030038	nice guy nice place	pl
13012	53469	2013-08-15	7462716	Bezpieczna, spokojna i cicha okolica. Sprawna komunikacja z kartą MTA autobus i metro. Wygodna łazienka i ogarnięta otwarta kuchnia. Gospodarz bardzo pomocny i otwarty.	pl

```
In [ ]: # Filter out only english comments
df_reviews = df_reviews.loc[(df_reviews['language'] == 'en')]
```

```
In [ ]: # Count total english comments
df_reviews.language.value_counts()
```

```
Out[ ]: en    71636
Name: language, dtype: int64
```

```
In [ ]: # Create a function to preprocess text of comments
def clean_text(text):
    """
    input- 'text' to be preprocessed
    output- converts input 'text' to lowercase,remove square brackets,links,punctuation
    and words containing numbers. Removes common accent characters and returns clean text.
    """
    text = text.lower()
    text = re.sub('|', "", text)
    text = re.sub(r'^a-zA-Z0-9', ' ', text) # removes punctuations
    text = re.sub('
                    .*?
                ', '', text)
    text = re.sub('\w*\d\w*', ' ', text)
    text = re.sub('[%s]' % re.escape(string.punctuation), '', text)
    return text
```

```
In [ ]: # apply preprocess function in comment column
df_reviews['comments'] = df_reviews['comments'].apply(lambda x: clean_text(x))
df_reviews.head()
```

```
Out[ ]:
```

	listing_id	date	reviewer_id	comments	year	month	days	weekday	language
2	2595	2009-12-10	38960	i ve stayed with my friend at the midtown castle for six days and it was a lovely place to be a big spacious room with a pointy roof which really makes you feel like staying in a castle the location is perfect it is just a few steps from macy s time square and theatre district everything worked just perfect with the keys etc thank you so much jennifer we had a great time in new york attention it s on the floor without a lift but definately worth it	2009	12	10	Thursday	en

3	2595	2010-04-09	71130	we ve been staying here for about nights enjoying to be in the center of the city that never sleeps short ways to everywhere in manhattan by subway or by walk midtown castle is a beautiful and tastful place jennifer and tori relaxed and friendly hosts thats why we the three berliners recommand that place good to have wifi and a little kitchen too	2010	04	09	Friday	en
4	2595	2010-05-25	117113	we had a wonderful stay at jennifer s charming apartment they were very organized and helpful i would definitely recommend staying at the midtown castle	2010	05	25	Tuesday	en
5	2595	2012-05-07	1783688	hi to everyone would say our greatest compliments to jennifer the host of midtown castle we spent in this lovely apartment in the heart of manhattan one month april and will remember this time as ours best the apartment is pretty spacious and great located the th ave right around the corner there is everything you can need during your short or long stay jennifer is very friendly vigorous and very responsible host thanks her and highly recomend this apartment for everyone who are looking for a quiet place right in the center of the boiling midtown	2012	05	07	Monday	en
6	2595	2012-05-17	1870771	jennifer was very friendly and helpful and her place is exactly as advertised the location is very convenient and it was a pleasure to stay at the midtown castle i definitely recommend it thanks	2012	05	17	Thursday	en

```
In [ ]: # Reset Index
df_reviews.reset_index(inplace=True, drop=True)
```

```
In [ ]: # Calculate the length of the comments
df_reviews['text_length'] = df_reviews['comments'].apply(lambda x : len(x))
df_reviews.head(4)
```

```
Out[ ]: listing_id  date  reviewer_id  comments  year  month  days  weekday  language  text_length
```

i ve stayed with my friend at the midtown castle for six days and it was a lovely place to be a big spacious room with a pointy roof which really makes you feel like staying in a castle the location is perfect it is just a few

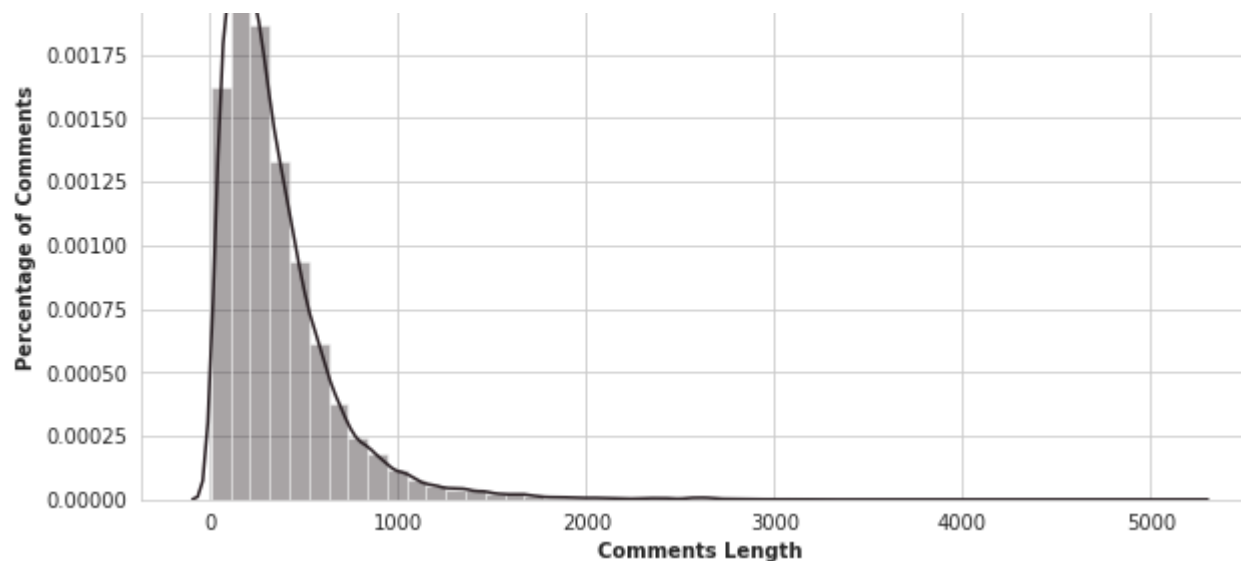
anna_

2	2595	2009-12-10	38960	steps from macy s time square and theatre district everything worked just perfect with the keys etc thank you so much jennifer we had a great time in new york attention it s on the floor without a lift but definately worth it	2009	12	10	Thursday	en	468
3	2595	2010-04-09	71130	we ve been staying here for about nights enjoying to be in the center of the city that never sleeps short ways to everywhere in manhattan by subway or by walk midtown castle is a beautiful and tastful place jennifer and tori relaxed and friendly hosts thats why we the three berliners recommand that place good to have wifi and a little kitchen too	2010	04	09	Friday	en	366
4	2595	2010-05-25	117113	we had a wonderful stay at jennifer s charming apartment they were very organized and helpful i would definitely recommend staying at the midtown castle	2010	05	25	Tuesday	en	155
5	2595	2012-05-07	1783688	hi to everyone would say our greatest compliments to jennifer the host of midtown castle we spent in this lovely apartment in the heart of manhattan one month april and will remember this time as ours best the apartment is pretty spacious and great located the th ave right around the corner there is everything you can need during your short or long stay jennifer is very friendly vigorous and very responsible host thanks her and highly recomend this apartment for everyone who are looking for a quiet place right in the center of the boiling midtown	2012	05	07	Monday	en	570

In []:

```
# Plot distribution of the comment length
plt.figure(figsize=(10, 5))
sns.distplot(df_reviews.text_length, kde=True, bins=50, color='#261c20')
plt.title('NYC Airbnb Review Lengths Distributin', fontweight="bold")
plt.xlabel('Comments Length', fontweight="bold")
plt.ylabel('Percentage of Comments', fontweight="bold");
```





```
In [ ]: import nltk
        nltk.download('wordnet')
        nltk.download('omw-1.4')
```

[nltk_data] Downloading package wordnet to /root/nltk_data...
[nltk_data] Downloading package omw-1.4 to /root/nltk_data...

Out[]: True

```
In [ ]: # Print DataFrame dimentionality
        df_reviews.shape
```

Out[]: (71648, 6)

```
In [ ]: df_reviews.info()
```

```
RangeIndex: 71648 entries, 0 to 71647
Data columns (total 6 columns):
#   Column      Non-Null Count  Dtype
---  -
0   listing_id  71648 non-null  int64
1   date        71648 non-null  object
2   reviewer_id 71648 non-null  int64
3   comments    71648 non-null  object
```

```
4 language      71648 non-null object
5 text_length   71648 non-null int64
dtypes: int64(3), object(3)
memory usage: 3.3+ MB
```

```
In [ ]: df_reviews
```

	listing_id	date	reviewer_id	comments	language	text_length
0	2595	2009-12-10	38960	i ve stayed with my friend at the midtown castle for six days and it was a lovely place to be a big spacious room with a pointy roof which really makes you feel like staying in a castle the location is perfect it is just a few steps from macy s time square and theatre district everything worked just perfect with the keys etc thank you so much jennifer we had a great time in new york attention it s on the floor without a lift but definetely worth it	en	468
1	2595	2010-04-09	71130	we ve been staying here for about nights enjoying to be in the center of the city that never sleeps short ways to everywhere in manhattan by subway or by walk midtown castle is a beautiful and tastful place jennifer and tori relaxed and friendly hosts thats why we the three berliners recommand that place good to have wifi and a little kitchen too	en	366
2	2595	2010-05-25	117113	we had a wonderful stay at jennifer s charming apartment they were very organized and helpful i would definitely recommend staying at the midtown castle	en	155
3	2595	2012-05-07	1783688	hi to everyone would say our greatest compliments to jennifer the host of midtown castle we spent in this lovely apartment in the heart of manhattan one month april and will remember this time as ours best the apartment is pretty spacious and great located the th ave right around the corner there is everything you can need during your short or long stay jennifer is very friendly vigorous and very responsible host thanks her and highly recomend this apartment for everyone who are looking for a quiet place right in the center of the boiling midtown	en	570
4	2595	2012-05-17	1870771	jennifer was very friendly and helpful and her place is exactly as advertised the location is very convenient and it was a pleasure to stay at the midtown castle i definitely recommend it thanks	en	204
...
71636	72265	2017-06-16	109542482	vanessa was very pleasant and communication was very good	en	58
71637	72265	2017-06-24	1282541	great location close to g train	en	34

				highly recommend cannot beat this value great location minute walk to subway and sec to bus which connects you easily and quickly to various parts of manhattan and brooklyn organic as well as regular grocery stores and lots of awesome restaurants and stores near by very safe neighborhoods nice room not big but it s plenty enough and everything works well it s nice warm in the winter even though the bedroom is separated by a curtain to the kitchen because the host is mainly in the other section of the apartment you have a lot of privacy vanessa is a very friendly interesting and helpful host	en	626
71638	72265	2018-04-11	8936723			
				vanessa is a great and very polite host and gives you as much privacy as you want the room can be seen in the photos and has everything you need the location is amazing as well with plenty of bars restaurants and stores around and literally half a block away from g train	en	275
71639	72265	2018-06-15	17160406			
				vanessa is a very hospitable and friendly person i was able to interact with her a lot the apartment is ideally located the room very convenient i have nothing to say except that i had a great time and it was a great experience for me thank you vanessa i would definitely go back	en	287
71640	72265	2018-07-21	165490874			

71641 rows × 6 columns

The process of transforming text into a numerical feature is called text **vectorization**.

TF-IDF(Term Frequency-Inverse Document Frequency) of records - is a text vectorizer that *transforms the text into a usable vector*. It can be defined as the calculation of how relevant a word in a series or corpus is to a text.

- *TF* - is the number of occurrences(importance) of a specific term in a document.
- *DF* - is the number of documents containing a specific term by indicating how common the term is.
- *IDF* - is a weight indicating how commonly(relevant) a word is used.

```
In [ ]: vectorizer = TfidfVectorizer(min_df=53, max_features=50, stop_words='english', ngram_range=(1,2), use_idf=True)
matrix = vectorizer.fit_transform(df_reviews['comments'])
```

```
In [ ]: df_reviews['comments'][0]
```

```
Out[ ]: 'Great host '
```

```
In [ ]: # Display first vector, IDF values for first comment
nd DataFrame(matrix[0].T.todense().index=vectorizer.get_feature_names()
```

```
pd.DataFrame(matrix[0].t.coef_, index=vectorizer.get_feature_names(),  
             columns=["tfidf"]).sort_values(['tfidf'], ascending=False).style.bar(subset=['tfidf'], color=['#'
```

Out[]:

	tfidf
just	0.431949
perfect	0.402302
time	0.393531
lovely	0.244261
new york	0.241695
york	0.241684
staying	0.233586
like	0.233292
new	0.227994
really	0.200927
room	0.183414
br	0.172514
location	0.149560
place	0.135511
great	0.121419
recommend	0.000000
nice	0.000000
nyc	0.000000
bed	0.000000
quiet	0.000000
city	0.000000
neighborhood	0.000000
space	0.000000
stay	0.000000

away	0.000000
subway	0.000000
area	0.000000
walk	0.000000
wonderful	0.000000
restaurants	0.000000
need	0.000000
clean	0.000000
great location	0.000000
close	0.000000
comfortable	0.000000
definitely	0.000000
easy	0.000000
experience	0.000000
friendly	0.000000
good	0.000000
helpful	0.000000
manhattan	0.000000
highly	0.000000
home	0.000000
host	0.000000
house	0.000000
brooklyn	0.000000
apartment	0.000000
br br	0.000000
amazing	0.000000

6. Estimate Sentiment Polarity

VADER (Valence Aware Dictionary and Sentiment Reasoner)

It is a lexical database and rule-based sentiment analysis tool that is optimized for social media sentiments. It makes use of a variety of techniques. A sentiment lexicon is a collection of lexical features (e.g., words) that are classified as positive or negative depending on their sentiment polarity. It not only exhibits the Positivity and Negativity scores, but also the degree to which a sentiment is positive or negative

The **compound score** is the sum of positive, negative & neutral scores which is then normalized between -1 (*most extreme negative*) and $+1$ (*most extreme positive*)

```
In [ ]: import nltk
        nltk.download('vader_lexicon')
```

[nltk_data] Downloading package vader_lexicon to /root/nltk_data...

```
Out[ ]: True
```

```
In [ ]: # Instantiate SentimentAnalyzer
        analyzer = SentimentIntensityAnalyzer()
```

```
In [ ]: # Test SentimentAnalyzer on first comment
        def sentiment_analyzer(comment):
            polarity_score = analyzer.polarity_scores(comment)
            return(pd.DataFrame.from_dict(polarity_score, orient='index'))

        sentiment_analyzer(df_reviews['comments'][0])
```

```
Out[ ]:      0
        neg  0.0000
        neu  0.8060
        pos  0.1940
        compound  0.9274
```

```
In [ ]: # Take 'compound' score of the sentiments score
        compound_polarity= lambda s: (analyzer.polarity_scores(s))['compound']
```

```
In [ ]: # Get compound polarity score of all comments
        df_reviews['polarity_score'] = df_reviews.comments.map(compound_polarity)
```

```
In [ ]: df_reviews[:5].style.bar(subset=['polarity_score'], align='mid')
```

```
Out[ ]: listing_id  reviewer_id  comments  month  weekday  language  text_length  polarity_score
```

2	2595	38960	I've stayed with my friend at the Midtown Castle for six days and it was a lovely place to be. A big spacious room with a pointy roof, which really makes you feel like staying in a castle. The location is perfect. It is just a few steps from Macy's Time Square and Theatre District. Everything worked just perfect with the keys etc. Thank you so much Jennifer, we had a great time in New York. Attention: it's on the 4th floor without a lift :-)	12	Thursday	en	475	0.924800
3	2595	71130	We've been staying here for about 9 nights, enjoying to be in the center of the city, that never sleeps...short ways to everywhere in Manhattan, by subway or by walk. Midtown castle is a beautiful and tastful place, Jennifer and Tori relaxed and friendly hosts - thats why we - the three Berliners - recommand that place! Good to have WiFi and a little kitchen too!	04	Friday	en	366	0.922900
4	2595	117113	We had a wonderful stay at Jennifer's charming apartment! They were very organized and helpful; I would definitely recommend staying at the Midtown Castle!	05	Tuesday	en	155	0.946400
5	2595	1783688	Hi to everyone! Would say our greatest compliments to Jennifer, the host of Midtown Castle. We spent in this lovely apartment in the heart of Manhattan one month (April, 2012) and will remember this time as ours best. The apartment is pretty spacious and great located - the 5-th Ave right around the corner.	05	Monday	en	583	0.986900

There is everything you can need during your short or long stay. Jennifer is very friendly, vigorous and very responsible host. Thanks her and highly recomend this apartment for everyone who are looking for a quiet place right in the center of the boiling Midtown!

Jennifer was very friendly and helpful, and her place is exactly as advertised. The location is very convenient, and it was a pleasure to stay at the Midtown Castle. I definitely recommend it :)

Thanks !

6	2595	1870771		05	Thursday	en	214	0.956800
---	------	---------	--	----	----------	----	-----	----------

```
In [ ]: df_reviews[df_reviews['polarity_score']<0][10:15].style.bar(subset=['polarity_score'], align='mid')
```

	listing_id	reviewer_id	comments	month	weekday	language	text_length	polarity_score
297	5178	4337113	This place is horrible and not what is represented in the photos. Not to mention very dirty. I had a horrible experience!	10	Friday	en	121	-0.888100
			He is very quiet but gave no idea about the area					
323	5178	14767392	The room was a bit small and the walls and windows are dirty as was the microwave. We could not use the kitchen sink as it was blocked . Good location Dave	05	Friday	en	216	-0.665200
			Our apartment building in Philadelphia caught fire, and much of our possessions were ruined and the whole building rendered uninhabitable by water damage the day before I had planned to stay with Shunichi. So I feel asleep within an hour after arriving at his apartment and didn't wake up until the next morning. Which goes to show that the room faces a quiet street and that the bed is comfortable, even if a metal frame. The room is as it looks in the photo, with the additions of a table cloth, cups, mugs and a small mirror, without the plant, stool and dark full length curtain, now replaced by a shorter one that lets the light through it. Shunichi was very communicative about my move in and	08	Wednesday	en	875	-0.851900

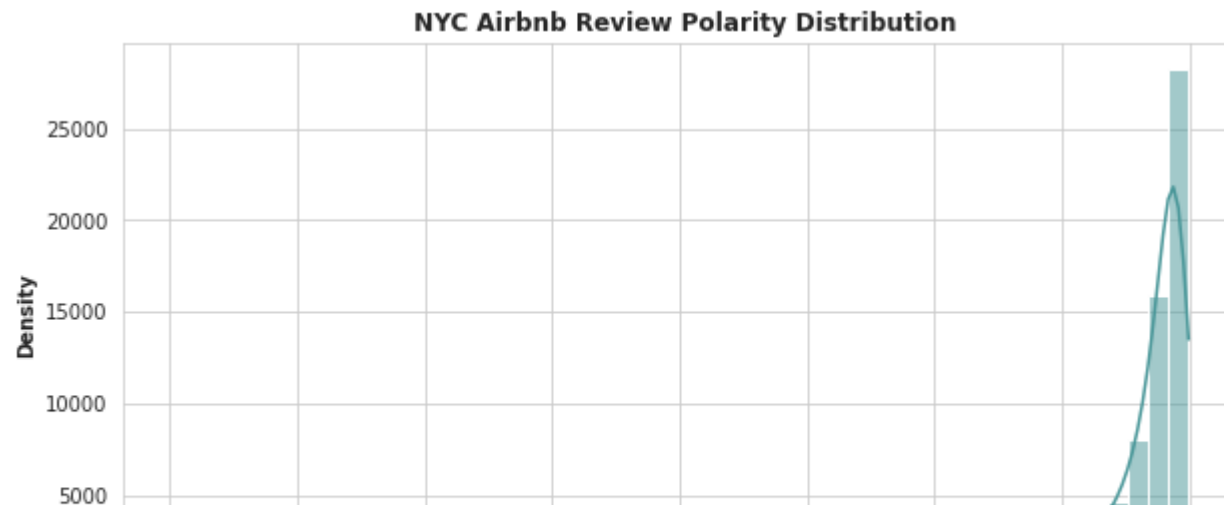
very comminative about my move in and had no problem with me coming earlier than I initially indicated,nor with me leaving my luggage at his place until I left the city the following day. Then I crashed out.

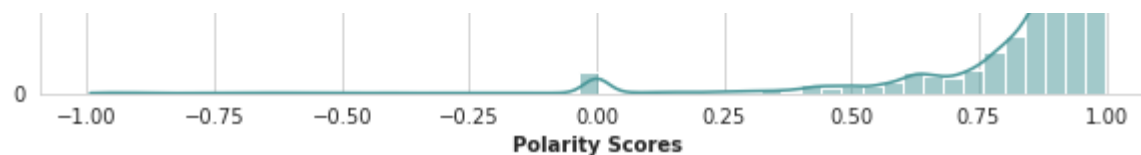
The apartment was in a great location! It was four blocks from time square! We found a really cool burger place to eat at! The air conditioner was the disappointing part of it all. We got there before the sun went down and had to leave because it was just that hot in the room. No fan was available. It was hot in the city and very uncomfortable to sleep. This was my first airbnb experience.

437	5178	133592606	all. We got there before the sun went down and had to leave because it was just that hot in the room. No fan was available. It was hot in the city and very uncomfortable to sleep. This was my first airbnb experience.	06	Wednesday	en	392	-0.312900
529	5178	11934009	the bathroom is dirty, the kitchen could not be used because it was full and dirty, the photos correspond but really very messy everything I would not recommend this place is not the complete apt just a room and bathroom shared	06	Monday	en	227	-0.763000

In []:

```
# Plot distribution of the comment length
plt.figure(figsize=(10, 5))
sns.histplot(df_reviews['polarity_score'], kde=True, bins=50, color='#469597')
plt.title(' NYC Airbnb Review Polarity Distribution', fontweight="bold")
plt.xlabel('Polarity Scores', fontweight="bold")
plt.ylabel('Density', fontweight="bold");
```





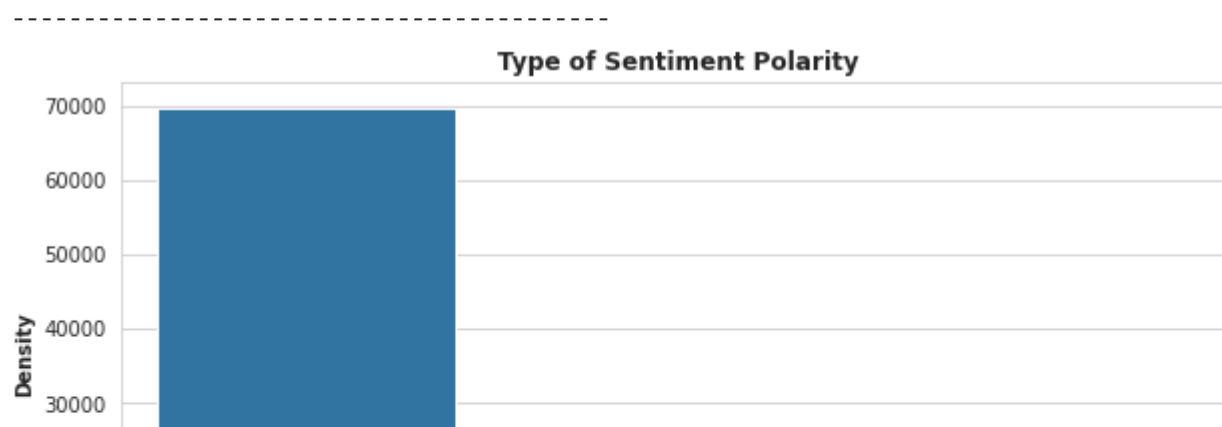
In []:

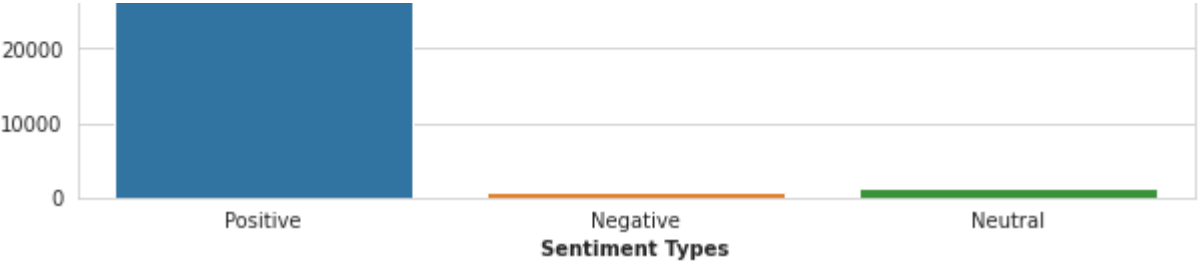
```
# Decide sentiment as positive, negative or neutral;
def sentiment_type(df_reviews):
    if df_reviews['polarity_score'] >= 0.05:
        val = "Positive"
    elif df_reviews['polarity_score'] <= - 0.05:
        val = "Negative"
    else:
        val = "Neutral"
    return val

df_reviews['sentiment_type'] = df_reviews.apply(sentiment_type, axis=1)

print(df_reviews.sentiment_type.value_counts())
print('-----')
plt.figure(figsize=(10,5))
sns.set_style("whitegrid")
ax = sns.countplot(x="sentiment_type", data=df_reviews)
plt.title('Type of Sentiment Polarity', fontweight="bold")
plt.xlabel('Sentiment Types', fontweight="bold")
plt.ylabel('Density', fontweight="bold");
```

```
Positive    69674
Neutral      1254
Negative      708
Name: sentiment_type, dtype: int64
```





```
In [ ]: df_reviews[df_reviews['sentiment_type']=='Neutral'][:10].style.bar(subset=['polarity_score'])
```

Out[]:

	listing_id	reviewer_id	comments	month	weekday	language	text_length	polarity_score	sentiment_type	
	43	2595	168460710	10 / 10 would stay again	05	Tuesday	en	24	0.000000	Neutral
	170	5178	404983	The host canceled my reservation 6 days before arrival.	09	Thursday	en	55	0.000000	Neutral
	232	5178	5029546	The reservation was canceled 107 days before arrival. This is an automated posting.	02	Tuesday	en	83	0.000000	Neutral
	290	5178	11031194	The reservation was canceled 4 days before arrival. This is an automated posting.	08	Friday	en	81	0.000000	Neutral
	457	5178	18110000	It is 2 floor , there is not an elevador.	10	Wednesday	en	42	0.000000	Neutral
	531	5178	269542502	Location and price is great, but hygiene is not acceptable.	07	Tuesday	en	59	0.027600	Neutral
	585	5178	414552453	Exactly as advertised.	07	Tuesday	en	22	0.000000	Neutral
	668	5203	1544337	The host canceled my reservation 92 days before arrival.	02	Thursday	en	56	0.000000	Neutral
	1026	6848	16767305	It has a grat location. Metro just 5 minutes away and you are in 15 minutes in Manhatten.	12	Wednesday	en	89	0.000000	Neutral

listing_id	reviewer_id	comments	month	weekday	language	text_length	polarity_score	sentiment_type
1067	6848	117326652	02	Sunday	en	53	0.000000	Neutral

Polite hosts and a cozy apartment with all amneties.

Neutral reviews mostly includes prototypicals *the host canceled this reservation X days before arrival this is an automated posting*. We can drop this rows

```
In [ ]: # Drop rows which includes canceled reservations
df_reviews = df_reviews[df_reviews["comments"].str.contains("automated")==False]
df_reviews = df_reviews[df_reviews["comments"].str.contains("canceled")==False]
```

```
In [ ]: # Check for unique values
df_reviews.sentiment_type.value_counts()
```

```
Out[ ]: Positive    69567
Negative     749
Neutral      489
Name: sentiment_type, dtype: int64
```

```
In [ ]: df_reviews[df_reviews['sentiment_type']=='Neutral'][:10].style.bar(subset=['polarity_score'],align='mid')
```

```
Out[ ]:
```

	listing_id	reviewer_id	comments	month	weekday	language	text_length	polarity_score	sentiment_type
43	2595	168460710	10 / 10 would stay again	05	Tuesday	en	24	0.000000	Neutral
457	5178	18110000	It is 2 floor , there is not an elevador.	10	Wednesday	en	42	0.000000	Neutral
531	5178	269542502	Location and price is great, but hygiene is not acceptable.	07	Tuesday	en	59	0.027600	Neutral
585	5178	414552453	Exactly as advertised.	07	Tuesday	en	22	0.000000	Neutral
1026	6848	16767305	It has a grat location. Metro just 5 minutes away and you are in 15 minutes in Manhattan.	12	Wednesday	en	89	0.000000	Neutral
1067	6848	117326652	Polite hosts and a cozy apartment with all amneties.	02	Sunday	en	53	0.000000	Neutral

Eric apartment laden

			epic apartment laden with character! Stellar backyard, fantastically located near Fort Greene park! Why would we ever consider a hotel ever again?!	04	Sunday	en	149	0.000000	Neutral
1416	7097	1316201							
1547	7097	64651723	Not at this time.	11	Saturday	en	17	0.000000	Neutral
			Laurie's place is warm and charming, a lovely home. Laurie is not just the nicest person I met on my trip to New York, but one of the nicest people I've met. Period.	12	Friday	en	165	-0.001500	Neutral
2208	10962	228894527							
2346	12192	5611145	Small room in big city.	11	Thursday	en	24	0.000000	Neutral

```
In [ ]: # Create seperate DataFrame of positive reviews
pos_reviews = df_reviews[df_reviews['sentiment_type']=='Positive']
```

```
In [ ]: pos_reviews.head()
```

```
Out[ ]: listing_id    id  reviewer_id  reviewer_name  comments  month  weekday  language  text_length  polarity_score  sentiment_
```

				i ve stayed with my friend at the midtown castle for six days and it was a lovely place to be a big spacious room with a pointy roof which really makes you feel like staying in a castle the location is perfect it is						
--	--	--	--	--	--	--	--	--	--	--

0	2595	19760	38960	Anita	just a few steps from macy s time square and theatre district everything worked just perfect with the keys etc thank you so much jennifer we had a great time in new york attention it s on the floor without a lift but definetely worth it	12	Thursday	en	468	0.9274	Po
1	2595	34320	71130	Kai-Uwe	we ve been staying here for about nights enjoying to be in the center of the city that never sleeps short ways to everywhere in manhattan by subway or by walk midtown castle is a beauftiful and tastful place jennifer and tori relaxed and friendly	04	Friday	en	366	0.9136	Po

hi to
everyone
would say
our greatest
compliments
to jennifer
the host of
midtown
castle we
spent in this
lovely
apartment in
the heart of
manhattan
one month
april and will
remember
this time as

2	2595	46312	117113	Alicia	Very organized	05	Tuesday	en	155	0.9409	Po
---	------	-------	--------	--------	-------------------	----	---------	----	-----	--------	----

					ours best						
					the						
					apartment is						
					pretty						
					spacious						
					and great						
					located the						
3	2595	1238204	1783688	Sergey	th ave right	05	Monday	en	570	0.9863	Po
					around the						
					corner there						
					is everything						
					you can						
					need during						
					your short						
					or long stay						
					jennifer is						
					very friendly						
					vigorous						
					and very						
					responsible						
					host thanks						
					her and						
					highly						
					recomend						
					this						
					apartment						
					for everyone						
					who are						
					looking for a						
					quiet place						
					right in the						
					center of the						
					boiling						
					midtown						
					jennifer was						
					very friendly						
					and helpful						
					and her						
					place is						
					exactly as						
					advertised						
					the location						
4	2595	1293632	1870771	Loïc	is very	05	Thursday	en	204	0.9542	Po
					convenient						
					and it was a						
					pleasure to						

reviewer to
stay at the
midtown
castle i
definitely
recommend
it thanks



```
In [ ]: # Create seperate dataframe of negative reviews
neg_reviews = df_reviews[df_reviews['sentiment_type']=='Negative']
```

```
In [ ]: neg_reviews.head()
```

listing_id	id	reviewer_id	reviewer_name	comments	month	weekday	language	text_length	polarity_score	sen
				this is a nice apartment and it looks like in the photo it is well located and jennifer was very nice and super responsive she also let me have a late check out at no extra cost which was very kind of her stars for that on the negative side it was rather cold in the room i stayed there for nights end of november beginning of						

					december so					
					the central					
					heating					
					should be on					
					but it only					
					worked					
					occasionally					
					this is a					
					building					
					from phone					
					number					
					hidden by					
					airbnb and					
					the					
					amenities					
					are a bit					
					worn and					
					probably					
					never have					
					been					
					updated					
21	2595	216355494	2888785	Claudia	there is an	12	Sunday	en	1121	-0.4455
					extra heater					
					but it is not					
					sufficient to					
					heat up the					
					room cause					
					the ceiling is					
					high and the					
					glas roof is					
					not a good					
					insulation					
					consequently					
					for the					
					majority of					
					the time the					
					temperature					
					inside was					
					between					
					phone					
					number					
					hidden by					
					airbnb c					
					which was					
					disagreeable					
					second point					

					was the					
					noise from					
					construction					
					sites in the					
					street one					
					just opposite					
					the building					
					they started					
					work early					
					and the					
					clearing of					
					rubble was					
					sometimes					
					done past					
					midnight					
					and very					
					loud once					
					the heating					
					problem is					
					sorted out					
					and the					
					construction					
					is done it will					
					be a nice					
					place to stay					
					the location					
					made					
					getting to					
					time square					
					easily					
					accessible					
30	2595	331029784	174975601	Julie	apartment	10	Monday	en	122	-0.2960
					was quiet					
					and no					
					disturbances					
					from					
					neighbors at					
					all					
					first of all i					
					have to say					
					that jennifer					
					is not					
					allowed to					
					sublet her					
					apartment					

by her
landlord i
was
instructed by
her to not
mention
airbnb to
anybody
beforehand
which was
fine but one
evening i
was stopped
by her
landlord in
the hallway
and
questioned
about where
i am staying
how long
where
jennifer is etc
afterwards
although i
didn t say
anything
wrong
jennifer
freaked out
about it and
i really
thought for a
few hours
that i would
have to
move out
and find
another
place to stay
right away
from that
point on i
felt really
unsafe there
and always

					tried to hide					
					from the					
					landlord for					
					my stay the					
					apartment					
					was also not					
					ideal there is					
					no couch					
41	2595	514642186	33952280	Laura	only	08	Wednesday	en	1359	-0.2495
					cushions on					
					the floor and					
					there is no					
					real table					
					and chairs i					
					had to work					
					from home					
					sometimes					
					and could					
					not really do					
					it anywhere					
					the place					
					was really					
					clean in					
					general only					
					the					
					bathroom					
					and the					
					shower were					
					really not					
					okay					
					especially					
					the shower					
					curtain was					
					mildewed					
					the kitchen					
					was clean					
					and well					
					equipped					
					apartment					
					has an ac					
					which is					
					definitely a					
					pro also					
					what should					
					be					
					mentioned					
					the hallway					

the railway
is really
narrow and
the stairs are
quite steep
so with a
huge
suitcase not
ideal as the
apartment is
walk up to
the floor the
location of
the
apartment is
alright in my
opinion it is
a bit busy in
that area and
there are
nicer places
to live bit
that is just
my personal
opinion

the
apartment is
not the place
shown in the
pictures it
was not
clean never
met akiko
only her
friend
website
hidden by
airbnb
unfortunetly
i have
nothing
positive to
say about
my stay sorry
but i feel
cheated

104	5178	29344	35310	Serdar	03	Sunday	en	219	-0.7262
-----	------	-------	-------	--------	----	--------	----	-----	---------

113	5178	69674	106633	Flore	the bad	07	Tuesday	en	537	-0.4911
-----	------	-------	--------	-------	---------	----	---------	----	-----	---------

```
In [ ]: # Create seperate DataFrame of neutral reviews
neut_reviews = df_reviews[df_reviews['sentiment_type']=='Neutral']
```

```
In [ ]: neut_reviews.head()
```

Out []:

	listing_id	id	reviewer_id	reviewer_name	comments	month	weekday	language	text_length	polarity_score
38	2595	449815873	168460710	Christopher	would stay again	05	Tuesday	en	22	0.00
					the location of this apartment is perfect we have privacy in our room the room has a small refrigerator microwave and some kitchen stuff the room was clean					
345	5178	122582806	24406166	Jael	although everything including the bed was very old the only problem	12	Saturday	en	315	0.00