

importing the Libraries and dataset loading..

```
In [7]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
#import afinn as Afinn

df=pd.read_csv(r"C:\Users\KARTIK\Downloads\news_data.csv")
df.head()
```

```
Out[7]:
```

	headline	description	date
0	Gaza's journalists 'Targets' or 'casual...	Gaza's journalists are being killed at an u...	23-12-2023
1	Israel orders 'death corridor' evacuat...	Hundreds of Palestinians were forced to eva...	23-12-2023
2	Israel's war on Gaza: List of key events, day 78	UN Security Council passes resolution ...	23-12-2023
3	US accuses Iran of being 'deeply involved'...	White House says Tehran is providing 'Yemeni ...	23-12-2023
4	Israel forces continue attacks amid cr...	The watered-down resolution boosting aid ...	23-12-2023

cleaning and processing

```
In [8]: df.shape

Out[8]: (3338, 3)
```

```
In [9]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3338 entries, 0 to 3337
Data columns (total 3 columns):
 #   Column      Non-Null Count  Dtype  
---  --
 0   headline    3338 non-null   object 
 1   description  3338 non-null   object 
 2   date        3338 non-null   object 
dtypes: object(3)
memory usage: 78.4+ KB
```

```
In [10]: df["date"]=pd.to_datetime(df["date"])

C:\Users\KARTIK\AppData\Local\Temp\ipykernel_10536\1303819999.py:1: UserWarning: Parsing dates in '%d-%m-%Y' format when dayfirst=False (the default) was specified. Pass 'dayfirst=True' or specify a format to silence this warning.
df["date"]=pd.to_datetime(df["date"])
```

```
In [11]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3338 entries, 0 to 3337
Data columns (total 3 columns):
 #   Column      Non-Null Count  Dtype  
---  --
 0   headline    3338 non-null   object 
 1   description  3338 non-null   object 
 2   date        3338 non-null   datetime64[ns]
dtypes: datetime64[ns](1), object(2)
memory usage: 78.4+ KB
```

checking Null values

```
In [12]: df.isnull().sum()
```

```
Out[12]:
```

	headline	description	date
headline	0	0	0
description	0	0	0
date	0	0	0
dtype:	int64	int64	int64

```
In [13]: df.head()
```

```
Out[13]:
```

	headline	description	date
0	Gaza's journalists 'Targets' or 'casual...	Gaza's journalists are being killed at an u...	2023-12-23
1	Israel orders 'death corridor' evacuat...	Hundreds of Palestinians were forced to eva...	2023-12-23
2	Israel's war on Gaza: List of key events, day 78	UN Security Council passes resolution ...	2023-12-23
3	US accuses Iran of being 'deeply involved'...	White House says Tehran is providing 'Yemeni ...	2023-12-23
4	Israel forces continue attacks amid cr...	The watered-down resolution boosting aid ...	2023-12-23

installing afinn class libraries for sentiments analysis

```
In [16]: !pip install afinn

Collecting afinn
  Using cached afinn-0.1.tar.gz (52 kB)
  Preparing metadata (setup.py): started
  Preparing metadata (setup.py): finished with status 'done'
Building wheels for collected packages: afinn
  Building wheel for afinn (setup.py): started
  Building wheel for afinn (setup.py): finished with status 'done'
  Created wheel for afinn: filename=afinn-0.1-py3-none-any.whl size=53427 sha256=7a808931f78b0b2be955658abaa7da480c255db47c376b477307b0ef67585
  Stored in directory: c:\users\kartik\appdata\local\pip\Cache\wheels\5ee\93\ab\19c255ebac29886ac31c2805303782f36399677f606be379725
Successfully built afinn
Installing collected packages: afinn
Successfully installed afinn-0.1
```

```
In [17]: import afinn
afin=afinn.Afinn()
```

```
In [18]: def senti(x):
    return af.scores(x)
df["headline_score"]=df["headline"].apply(senti)
df["description_score"]=df["description"].apply(senti)
```

sentiments score of headlines and description

```
In [19]: df.head()
```

```
Out[19]:
```

	headline	description	date	headline_score	description_score
0	Gaza's journalists 'Targets' or 'casual...	Gaza's journalists are being killed at an u...	2023-12-23	-2.0	-5.0
1	Israel orders 'death corridor' evacuat...	Hundreds of Palestinians were forced to eva...	2023-12-23	-2.0	-1.0
2	Israel's war on Gaza: List of key events, day 78	UN Security Council passes resolution ...	2023-12-23	-2.0	1.0
3	US accuses Iran of being 'deeply involved'...	White House says Tehran is providing 'Yemeni ...	2023-12-23	0.0	-2.0
4	Israel forces continue attacks amid cr...	The watered-down resolution boosting aid ...	2023-12-23	0.0	1.0

```
In [20]: df["headline_score"].value_counts()
```

```
Out[20]:
```

headline_score	count
0.0	1389
-2.0	681
-3.0	487
-1.0	232
-4.0	157
-5.0	118
1.0	112
2.0	89
-6.0	28
3.0	26
-7.0	18
-8.0	6
4.0	6
5.0	3
-9.0	2
-10.0	2
6.0	1
-11.0	1

```
Name: count, dtype: int64
```

```
In [21]: df.head()
```

```
Out[21]:
```

	headline	description	date	headline_score	description_score
0	Gaza's journalists 'Targets' or 'casual...	Gaza's journalists are being killed at an u...	2023-12-23	-2.0	-5.0
1	Israel orders 'death corridor' evacuat...	Hundreds of Palestinians were forced to eva...	2023-12-23	-2.0	-1.0
2	Israel's war on Gaza: List of key events, day 78	UN Security Council passes resolution ...	2023-12-23	-2.0	1.0
3	US accuses Iran of being 'deeply involved'...	White House says Tehran is providing 'Yemeni ...	2023-12-23	0.0	-2.0
4	Israel forces continue attacks amid cr...	The watered-down resolution boosting aid ...	2023-12-23	0.0	1.0

```
In [22]: df.shape

Out[22]: (3338, 5)
```

```
In [23]: df.loc[df["headline_score"]<0].count()
```

```
Out[23]:
```

	headline	description	date	headline_score	description_score
headline	1712	1712			
description	1712	1712			
date	1712	1712			
headline_score	1712				
description_score	1712				
dtype:	int64	int64	int64	int64	int64

```
In [24]: df.loc[df["headline_score"]<0].count()
```

```
Out[24]:
```

	headline	description	date	headline_score	description_score
headline	237	237			
description	237	237			
date	237	237			
headline_score	237				
description_score	237				
dtype:	int64	int64	int64	int64	int64

```
In [25]: df.loc[df["headline_score"]==0].count()
```

```
Out[25]:
```

	headline	description	date	headline_score	description_score
headline	1389	1389			
description	1389	1389			
date	1389	1389			
headline_score	1389				
description_score	1389				
dtype:	int64	int64	int64	int64	int64

creating new columns for year, month, day

```
In [26]: df["year"]=df["date"].dt.year

In [27]: df["month"]=df["date"].dt.month_name()

In [28]: df["day"]=df["date"].dt.day

In [29]: df.head()
```

```
Out[29]:
```

	headline	description	date	headline_score	description_score	year	month	day
0	Gaza's journalists 'Targets' or 'casual...	Gaza's journalists are being killed at an u...	2023-12-23	-2.0	-5.0	2023	December	23
1	Israel orders 'death corridor' evacuat...	Hundreds of Palestinians were forced to eva...	2023-12-23	-2.0	-1.0	2023	December	23
2	Israel's war on Gaza: List of key events, day 78	UN Security Council passes resolution ...	2023-12-23	-2.0	1.0	2023	December	23
3	US accuses Iran of being 'deeply involved'...	White House says Tehran is providing 'Yemeni ...	2023-12-23	0.0	-2.0	2023	December	23
4	Israel forces continue attacks amid cr...	The watered-down resolution boosting aid ...	2023-12-23	0.0	1.0	2023	December	23

```
In [30]: #Negative News headlines
neg=df.loc[(df["year"]==2023)&(df["headline_score"]<0)].count()
neg
```

```
Out[30]:
```

	headline	description	date	headline_score	description_score	year	month	day
headline	1288	1288						
description	1288	1288						
date	1288	1288						
headline_score	1288							
description_score	1288							
year	1288							
month	1288							
day	1288							
dtype:	int64	int64	int64	int64	int64	int64	int64	int64

```
In [31]: #Positive News headlines
df.loc[(df["year"]==2023)&(df["headline_score"]==0)].count()
```

```
Out[31]:
```

	headline	description	date	headline_score	description_score	year	month	day
headline	950	950						
description	950	950						
date	950	950						
headline_score	950							
description_score	950							
year	950							
month	950							
day	950							
dtype:	int64	int64	int64	int64	int64	int64	int64	int64

```
In [32]: #Neutral News Headlines
df.loc[(df["year"]==2023)&(df["headline_score"]>0)].count()
```

```
Out[32]:
```

	headline	description	date	headline_score	description_score	year	month	day
headline	168	168						
description	168	168						
date	168	168						
headline_score	168							
description_score	168							
year	168							
month	168							
day	168							
dtype:	int64	int64	int64	int64	int64	int64	int64	int64

```
In [33]: neg=df.loc[(df["headline_score"]<0)]
```

analysing the negative impact headlines and descriptions sentiments

```
In [34]: neg
```

```
Out[34]:
```

	headline	description	date	headline_score	description_score	year	month	day
0	Gaza's journalists 'Targets' or 'casual...	Gaza's journalists are being killed at an u...	2023-12-23	-2.0	-5.0	2023	December	23
1	Israel orders 'death corridor' evacuat...	Hundreds of Palestinians were forced to eva...	2023-12-23	-2.0	-1.0	2023	December	23
2	Israel's war on Gaza: List of key events, day 78	UN Security Council passes resolution ...	2023-12-23	-2.0	1.0	2023	December	23
3	US accuses Iran of being 'deeply involved'...	White House says Tehran is providing 'Yemeni ...	2023-12-22	-8.0	-1.0	2023	December	22
5	RSF files second war crimes complaint with I...	Latest complaint by Paris-based press freed...	2023-12-22	-2.0	0.0	2023	December	22
10	Israel-Gaza war: Why are culture and socie...	Fatima Bhutto talks with Marc Lamont Hill...	2023-12-22	-2.0	0.0	2023	December	22
...
3328	In Pictures: Global protests in solidarity...	Large demonstrations held in chief cities ar...	2021-05-12	-2.0	0.0	2021	May	12
3329	Jerusalem, the unfolding tragedy	For Netanyahu, Jerusalem and Gaza are the g...	2021-05-11	-2.0	2.0	2021	May	11
3330	Break the fear barrier and speak up for Pale...	Saying the Palestinian issue is 'complex'...	2021-05-11	-2.0	-6.0	2021	May	11
3332	In Pictures: Israel attacks on Gaza kill 24...	Israel continues to pound Gaza with air rail...	2021-05-11	-3.0	0.0	2021	May	11
3336	The War in June 1967	The June 1967 Arab-Israeli war lasted only s...	2017-06-02	-2.0	-2.0	2017	June	2

1712 rows x 8 columns

```
In [35]: neg.groupby("year").describe(include="object").count
```

```
Out[35]:
```

	headline	description	date	headline_score	description_score	year	month	day
year	count	unique	top	freq				
2017	1	1	The War in June 1967	1				
2021	187	186	Israeli forces kill Palestinian man in occ...	2				
2022	236	233	Israeli forces kill Palestinian teenager i...	4				
2023	1288	1288	Gaza's journalists: 'Targets' or 'casual...	1				
description	count	unique	top	freq				
year	1	1	The June 1967 Arab-Israeli war lasted only s...	1				
2021	187	187	Al Jazeera talks to young Palestinians inju...					
2022	236	236	Palestinians fear a continuing Israeli p...					
2023	1288	1287	As the conflict between Israel and Gaza en...					
month	freq	count	unique	top	freq			
year	1	1	1	June	1			
2021	1	187	8	May	86			
2022	1	236	12	May	54			
2023	2	1288	12	October	401			

```
In [36]: neg.shape

Out[36]: (1712, 8)
```

finding the month and year were most news and attack happend

```
In [37]: neg.groupby("year").count()
```

```
Out[37]:
```

	headline	description	date	headline_score	description_score	month	day
year	count	unique	top	freq			
2017	1	1	1	1		1	1
2021	187	187	187	187		187	187
2022	236	236	236	236		236	236
2023	1288	1288	1288	1288		1288	1288

```
In [38]: neg["month"].value_counts().plot(kind="bar")

Out[38]:
```

```
In [40]: neg.describe()
```

```
Out[40]:
```

	date	headline_score	description_score	year	month	day
count	1712	1712.000000	1712.000000	1712.000000	1712.000000	
mean	2023-05-07 23:40:39.252336128	-2.689252	-1.738992	2022.640187	15.303738	
min	2017-06-02 00:00:00	-11.000000	-10.000000	2017.000000	1.000000	
25%	2023-01-04 00:00:00	-3.000000	-3.000000	2023.000000	8.000000	
50%	2023-10-18 00:00:00	-2.000000	-2.000000	2023.000000	15.000000	
75%	2023-11-14 00:00:00	-2.000000	0.000000	2023.000000	22.000000	
max	2023-12-23 00:00:00	-1.000000	8.000000	2023.000000	31.000000	
std		NaN	1.323413	2.070399	0.683082	8.420497

```
In [43]: neg["year"].value_counts().plot(kind="bar")

Out[43]:
```

<Axes: xlabel='year'>

```
In [46]: neg.groupby(["month","year"]).describe().reset_index().sort_values(by=["year","count"],ascending=False)
```

```
Out[46]:
```

	month	year	count	mean	std	min	25%	50%	75%	max
29	October	2023	401.0	19.558603	6.892308	5.0	13.00	19.0	25.00	31.0
26	November	2023	379.0	13.831135	8.622782	1.0	7.00	13.0	21.00	30.0
7	December	2023	260.0	10.834615	6.356794	1.0	6.00	11.0	16.00	23.0
14	July	2023	40.0	12.500000	8.673567	2.0	4.75	10.0	20.00	28.0
23	May	2023	38.0	12.078947	9.059508	2.0	3.25	10.0	17.75	30.0
11	January	2023	35.0	17.971429	10.279865	2.0	6.00	19.0	27.50	31.0
18	June	2023	31.0	16.741935	8.670069	2.0	13.50	19.0	21.50	29.0
9	February	2023	28.0	17.071429	8.764104	2.0	9.75	20.5	24.00	28.0
4	August	2023	26.0	19.153846	9.640300	1.0	11.00	21.0	28.00	31.0
1	April	2023	20.0	8.950000	6.581273	1.0	6.75	7.0	10.25	26.0
30	March	2023	19.0	13.578947	8.883917	2.0	7.50	12.0	19.50	31.0
32	September	2023	11.0	18.727273	10.100405	1.0	12.50	22.0	26.50	30.0
22	May	2022	54.0	16.111111	6.302640	5.0	11.00	15.0	21.00	29.0
3	August	2022	41.0	13.853659	7.853537	2.0	7.00	11.0	18.00	31.0
6	December	2022	22.0	13.363636	10.710678	1.0	4.00	10.0		

