```
import pandas as pd
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt
data = pd.read excel(r"C:\Users\saisn\Downloads\Onyx Data -DataDNA
Dataset Challenge - Email Analysis Dataset - July 2024\Onyx Data -
DataDNA Dataset Challenge - Email Analysis Dataset - July 2024.xlsx")
data.head()
   Email id
                From Name From seniority
                                               From Department \
0
            Leona McAree
                                 C-level
                                          Executive Management
          1
1
          2
            Leona McAree
                                 C-level Executive Management
2
          3
            Leona McAree
                                 C-level Executive Management
3
          4 Leona McAree
                                 C-level
                                          Executive Management
4
          5 Leona McAree
                                 C-level Executive Management
              To Name
                                  To seniority
                                                  To Department
  Elsinore Waterland Middle level management Human Resources
0
1
  Elsinore Waterland Middle level management
                                                Human Resources
   Elsinore Waterland Middle level management
                                                Human Resources
3
   Elsinore Waterland Middle level management
                                                Human Resources
4 Elsinore Waterland Middle level management
                                                Human Resources
          Email topic
                            Date Sentiment Is opened?
                                                        Device \
   Operational Issues 2024-03-14
0
                                   neutral
                                               opened
                                                       desktop
1
  HR related topics 2024-03-22
                                                        mobile
                                  positive
                                               opened
2
  Meeting Scheduling 2024-03-26
                                 neutral
                                             unopened
                                                       desktop
3
   HR related topics 2024-03-11
                                   neutral
                                               opened
                                                        mobile
   HR related topics 2024-03-11 positive
                                               opened
                                                        mobile
 Within work hours Within workdays
0
                yes
1
                no
                                yes
2
                yes
                                yes
3
                yes
                                yes
4
                yes
                                ves
data.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1132 entries, 0 to 1131
Data columns (total 14 columns):
                        Non-Null Count
#
     Column
                                        Dtype
     -----
0
     Email id
                        1132 non-null
                                        int64
1
     From Name
                        1132 non-null
                                        object
 2
     From seniority
                        1132 non-null
                                        object
 3
     From Department
                        1132 non-null
                                        object
 4
     To Name
                        1132 non-null
                                        object
```

```
5
     To seniority
                        1132 non-null
                                        object
     To Department
                        1132 non-null
                                        object
 6
 7
     Email topic
                        1132 non-null
                                        object
 8
     Date
                        1132 non-null
                                        datetime64[ns]
 9
     Sentiment
                        1132 non-null
                                        object
 10
    Is opened?
                        1132 non-null
                                        object
     Device
                        1132 non-null
                                        object
 11
12
    Within work hours
                        1132 non-null
                                        object
    Within workdays
                        1132 non-null
13
                                        object
dtypes: datetime64[ns](1), int64(1), object(12)
memory usage: 123.9+ KB
data.describe()
          Email id
                                             Date
count
       1132.000000
                                             1132
        624.666961 2024-03-16 11:07:50.671378176
mean
min
          1.000000
                              2024-03-01 00:00:00
                              2024-03-11 00:00:00
25%
        348.750000
        631.500000
                              2024-03-15 00:00:00
50%
75%
        914.250000
                              2024-03-25 00:00:00
       1197.000000
                              2024-03-31 00:00:00
max
std
       337.942989
                                              NaN
data.tail()
      Email id
                     From Name
                                         From seniority From
Department \
                Emmey Matoshin Middle level management
1127
          1193
Sales
1128
          1194
                Emmey Matoshin Middle level management
Sales
1129
          1195
                Emmey Matoshin Middle level management
Sales
1130
          1196 Emmey Matoshin Middle level management
Sales
1131
          1197
                Emmey Matoshin Middle level management
Sales
              To Name
                                                         To Department
                                  To seniority
     Blithe Arnaudet Middle level management Finance and Accounting
1127
1128
      Blithe Arnaudet Middle level management Finance and Accounting
1129
     Blithe Arnaudet Middle level management Finance and Accounting
1130
      Blithe Arnaudet Middle level management Finance and Accounting
1131 Blithe Arnaudet Middle level management Finance and Accounting
```

		Email t	opic	Date	Sentiment	t Is (pened?	Device	<u>.</u>
\ 1127	Finance a	nd Accour	iting 202	4-03-28	positive	9	opened	mobile	2
1128	Finance a	nd Accour	iting 202	4-03-07	negative	e	opened	desktop	כ
1129	0pera	tional Is	sues 202	4-03-14	neutra	l	opened	desktop)
1130	Meeti	ng Schedu	ıling 202	4-03-08	positive	e	opened	desktop	כ
1131	Finance a	nd Accour	iting 202	4-03-23	neutra ⁻	l	opened	mobile	ē
1127 1128 1129 1130 1131	√ithin wor	k hours w yes yes no yes no	/ithin wo	rkdays yes yes yes yes no					
data.isnull()									
0 1 2 3 4	Email id False False False False False	From Nam Fals Fals Fals Fals	se se se	seniorit Fals Fals Fals Fals	se se se se	· F F F	ment T Talse Talse Talse Talse	o Name False False False False False	\
1127 1128 1129 1130 1131	False False False False	Fals Fals Fals Fals	e e e e	Fals Fals Fals Fals	se se se	F F	alse alse alse alse alse	False False False False	
opene	To senior	ity To D	epartmen	t Email	l topic	Date	Sentim	ent Is	
0 False		lse	Fals	е	False I	False	Fa	lse	
1 False	Fa	lse	Fals	е	False I	False	Fa	lse	
2 False	False False		Fals	е	False I	False False		False False	
3 False			Fals	е	False I				
4	Fa	lse	Fals	е	False I	False	Fa	lse	
False									

1127 False	False	False	False	False	False					
1128	False	False	False	False	False					
False 1129	False	False	False	False	False					
False 1130	False	False	False	False	False					
False 1131 False	False	False	False	False	False					
1127 Fals 1128 Fals 1129 Fals 1130 Fals	se se se se se se se	hours Within False	n workdays. False False False False False False False False False							
		ratse	ratse	=						
_	x 14 columns]									
data.descr	<pre>data.describe().T</pre>									
min \	count		mean							
Email id 1132.0 624.666961 1.0										
Date 1132 2024-03-16 11:07:50.671378176 2024-03-01 00:00:00										
7E0. \		25%	5	50%						
75% \ Email id	34	8.75	631	1.5						
914.25 Date 00:00:00	2024-03-11 00:0	0:00 2024-03	3-15 00:00:	:00 2024-0	3-25					
Email id Date :	11 2024-03-31 00:0	max 97.0 337.942 0:00	std 1989 NaN							
data										
Emai Department	\	Name	From se	eniority	From					
0	1 Leona M	caree		C-level E	xecutive					

Managemen						_			
1 Managemen	2	Leona	McAree			C -	level	Executi	ve
Managemen 2	3	Leona	McAree			C -	level	Executi	ve
Managemen 3	4	Leona	McAree			C -	level	Executi	ve
Managemen 4	5	Leona	McAree			C -	level	Executi	ve
Managemen	t					_			
1127	1102	Emmov M	latochin	Middlo	lovol	manag	omont		
1127 Sales	1193	cilliley i	latoshin	Middle	tevet	manay	ement		
1128	1194	Emmey M	latoshin	Middle	level	manag	ement		
Sales	1131	Lilling	ia cosiiiii	HILUUCC		mariag	Cilicite		
1129	1195	Emmey M	latoshin	Middle	level	manag	ement		
Sales		,				J			
1130	1196	Emmey M	latoshin	Middle	level	manag	ement		
Sales									
1131	1197	Emmey M	latoshin	Middle	level	manag	ement		
Sales									
		To Nam	ie	To	senio	ritv		To	
Departmen	t \	10 Han		. ,	5 50111	J. 123		. 0	
		Waterlan	d Middl	e level	manage	ement		Human	
Resources					J				
		Waterlan	d Middl	e level	manage	ement		Human	
Resources				, ,					
2 Els Resources	inore v	Waterlan	a Mlaat	e level	manage	ement		Human	
	inore V	Waterlan	lbbiM b	e level	manage	ement		Human	
Resources	11101 6 1	va cer can	d Hiidac	c cevec	mariage	Silicite		Haman	
	inore V	Waterlan	d Middl	e level	manage	ement		Human	
Resources					J				
	Blithe	Arnaude	t Middl	e level	manage	ement	Finan	ce and	
Accountin	g				J				
_	_	Arnaude	t Middl	e level	manage	ement	Finan	ce and	
Accountin	_								
		Arnaude	t Middl	e level	manage	ement	Finan	ce and	
Accountin	_	A	- M'-1-17	. 11			F:		
1130 Accountin		Arnaude	t MIOOL	e level	manage	ement	rinan	ce and	
	_	Arnaude	t Middl	e level	manag	amant	Finan	ce and	
Accountin		ATTIAUUE	t HILUUC	c tevet	mariage	Sillett	THAIR	Le and	
ACCOUNTERIN	9								
		Email	topic	Date	e Sent:	iment	Is ope	ned? D	evice
\									

```
0
          Operational Issues 2024-03-14
                                            neutral
                                                                 desktop
                                                         opened
1
           HR related topics 2024-03-22
                                           positive
                                                         opened
                                                                  mobile
2
          Meeting Scheduling 2024-03-26
                                            neutral
                                                       unopened
                                                                 desktop
3
           HR related topics 2024-03-11
                                            neutral
                                                         opened
                                                                  mobile
           HR related topics 2024-03-11
                                           positive
                                                         opened
                                                                  mobile
1127 Finance and Accounting 2024-03-28
                                           positive
                                                         opened
                                                                  mobile
1128
     Finance and Accounting 2024-03-07
                                           negative
                                                         opened
                                                                 desktop
          Operational Issues 2024-03-14
1129
                                            neutral
                                                         opened
                                                                 desktop
1130
          Meeting Scheduling 2024-03-08
                                           positive
                                                         opened
                                                                 desktop
1131 Finance and Accounting 2024-03-23
                                            neutral
                                                         opened
                                                                  mobile
     Within work hours Within workdays
0
                    yes
                                     yes
1
                     no
                                     yes
2
                    yes
                                     yes
3
                                     yes
                    yes
4
                    yes
                                     yes
                    . . .
                                     . . .
1127
                    yes
                                     yes
1128
                    yes
                                     yes
1129
                     no
                                     yes
1130
                    yes
                                     yes
1131
                     no
                                      no
[1132 rows x 14 columns]
data.head(5)
   Email id
                 From Name From seniority
                                                 From Department \
0
             Leona McAree
                                  C-level
                                            Executive Management
          1
1
          2
             Leona McAree
                                  C-level
                                            Executive Management
2
          3
             Leona McAree
                                   C-level
                                            Executive Management
3
             Leona McAree
                                   C-level
                                            Executive Management
          5
             Leona McAree
4
                                   C-level
                                            Executive Management
```

To seniority

Middle level management

Middle level management

Middle level management

To Department

Human Resources

Human Resources

Human Resources

To Name

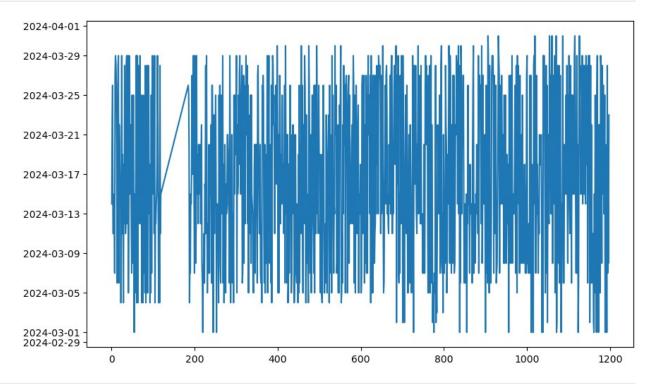
Elsinore Waterland

Elsinore Waterland

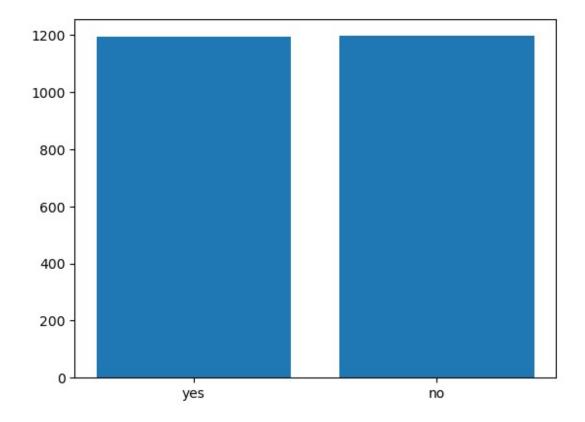
Elsinore Waterland

0

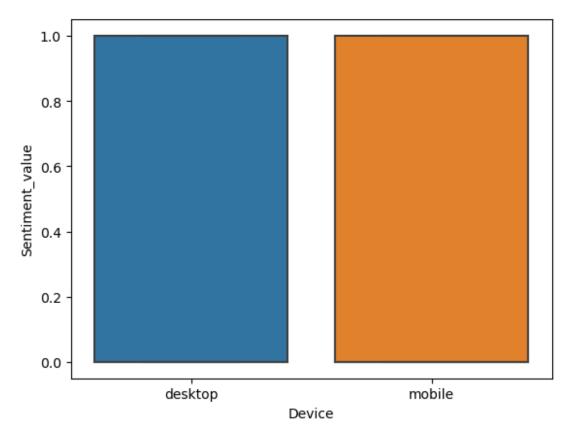
```
3
   Elsinore Waterland
                       Middle level management
                                                  Human Resources
4
   Elsinore Waterland
                       Middle level management
                                                  Human Resources
          Email topic
                             Date Sentiment Is opened?
                                                          Device \
   Operational Issues 2024-03-14
0
                                    neutral
                                                 opened
                                                         desktop
1
    HR related topics 2024-03-22
                                   positive
                                                 opened
                                                          mobile
2
   Meeting Scheduling 2024-03-26
                                               unopened
                                                         desktop
                                    neutral
3
                                                          mobile
    HR related topics 2024-03-11
                                    neutral
                                                 opened
    HR related topics 2024-03-11
                                                          mobile
4
                                                 opened
                                   positive
  Within work hours Within workdays
0
                yes
                                 yes
1
                 no
                                 yes
2
                yes
                                 yes
3
                yes
                                 yes
4
                yes
                                 yes
plt.figure(figsize=(10, 6))
plt.plot(data['Email id'],data['Date'])
plt.show()
```



plt.bar(data['Within work hours'],data['Email id'])
<BarContainer object of 1132 artists>

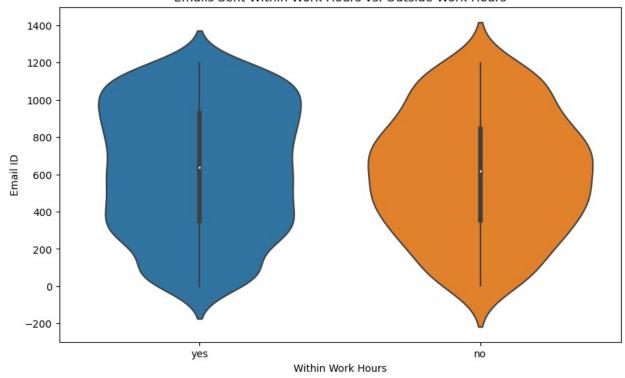


import seaborn as sns



```
# Violin Plot: Emails Sent Within Work Hours
plt.figure(figsize=(10, 6))
sns.violinplot(x='Within work hours', y='Email id', data=df)
plt.title('Emails Sent Within Work Hours vs. Outside Work Hours')
plt.xlabel('Within Work Hours')
plt.ylabel('Email ID')
plt.show()
```

Emails Sent Within Work Hours vs. Outside Work Hours

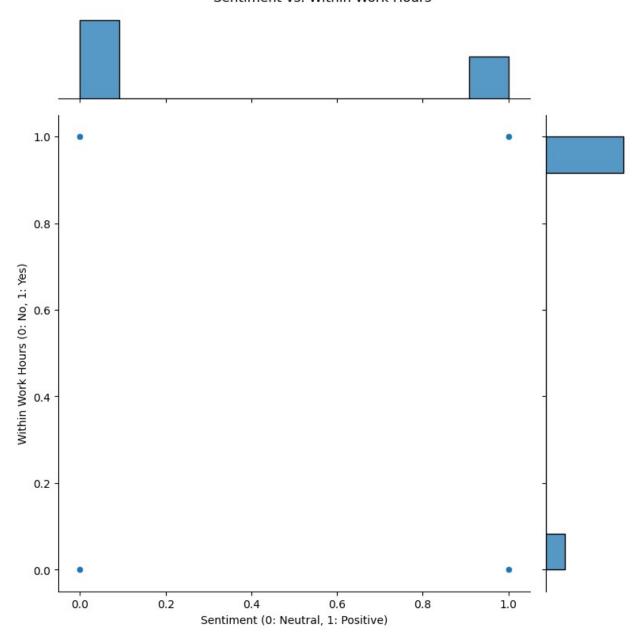


```
# Mapping sentiment to numerical values for plotting
sentiment_map = {'neutral': 0, 'positive': 1}
df['Sentiment_value'] = df['Sentiment'].map(sentiment_map)
plt.figure(figsize=(8, 8))
# Convert 'Within work hours' to numerical values for plotting
work_hours_map = {'yes': 1, 'no': 0}
df['Within_work_hours_value'] = df['Within work
hours'].map(work_hours_map)

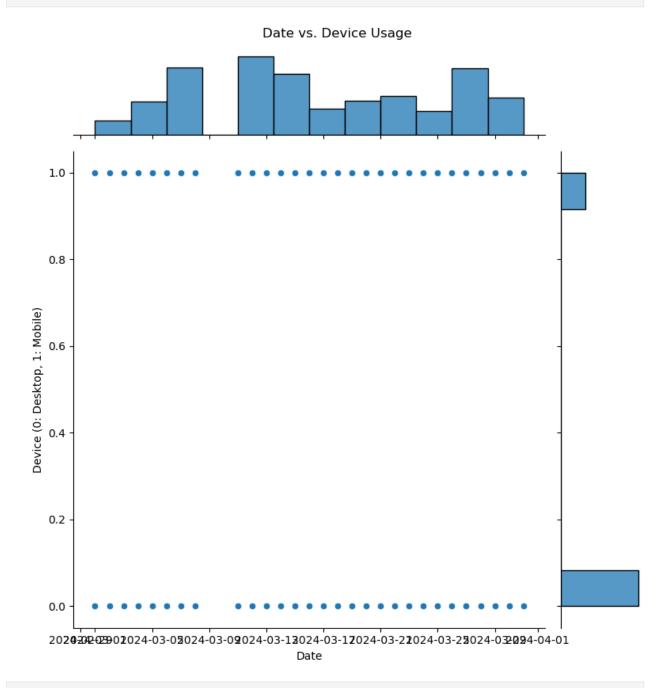
sns.jointplot(x='Sentiment_value', y='Within_work_hours_value',
data=df, kind='scatter', height=8)
plt.suptitle('Sentiment vs. Within Work Hours', y=1.02)
plt.xlabel('Sentiment (0: Neutral, 1: Positive)')
plt.ylabel('Within Work Hours (0: No, 1: Yes)')
plt.show()

<Figure size 800x800 with 0 Axes>
```

Sentiment vs. Within Work Hours



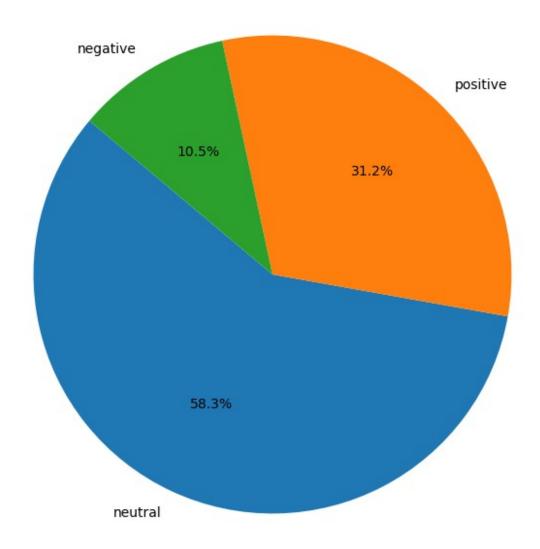
```
# Joint Plot: Date vs. Device
# Convert 'Device' to numerical values for plotting
device_map = {'desktop': 0, 'mobile': 1}
df['Device_value'] = df['Device'].map(device_map)
plt.figure(figsize=(8, 8))
sns.jointplot(x='Date', y='Device_value', data=df, kind='scatter',
height=8)
plt.suptitle('Date vs. Device Usage', y=1.02)
plt.xlabel('Date')
plt.ylabel('Date')
plt.ylabel('Device (0: Desktop, 1: Mobile)')
plt.show()
```



```
df = pd.DataFrame(data)
# Count the occurrences of each sentiment
sentiment_counts = df['Sentiment'].value_counts()
# Plotting the pie chart
plt.figure(figsize=(8, 8))
```

```
plt.pie(sentiment_counts, labels=sentiment_counts.index,
autopct='%1.1f%%', startangle=140)
plt.title('Distribution of Email Sentiments')
plt.show()
```

Distribution of Email Sentiments



```
df = pd.DataFrame(data)
# Filter to only include numerical columns
numerical_df = df.select_dtypes(include=['number'])
# Generate descriptive statistics for numerical data
```

```
desc = numerical_df.describe()

# Plot the heatmap
plt.figure(figsize=(10, 6))
sns.heatmap(desc, annot=True, linewidths=0.5, cmap='coolwarm')
plt.title('Summary Statistics Heatmap')
plt.show()
```

Summary Statistics Heatmap

