

Import Libraries

```
import pandas as pd
```

```
import numpy as np
```

```
pip install plotly
```

```
Requirement already satisfied: plotly in c:\users\shivdas\anaconda3\lib\site-packages (5.8.0)
```

```
Requirement already satisfied: tenacity>=6.2.0 in c:\users\shivdas\anaconda3\lib\site-packages (from plotly) (8.0.1)
```

```
Note: you may need to restart the kernel to use updated packages.
```

```
train_data = pd.read_csv('train.csv')
```

Reading the data from CSV file

```
import matplotlib.pyplot as plt
```

```
import matplotlib.patches as patches
```

```
import seaborn as sns
```

```
import plotly.express as px
```

```
import plotly.graph_objs as go
```

```
from plotly.offline import iplot
```

```
train_data.head()
```

	Unnamed: 0	id	Gender	Customer Type	Age	Type of Travel
0	0	70172	Male	Loyal Customer	13	Personal Travel
1	1	5047	Male	disloyal Customer	25	Business travel
2	2	110028	Female	Loyal Customer	26	Business travel
3	3	24026	Female	Loyal Customer	25	Business travel
4	4	119299	Male	Loyal Customer	61	Business travel

	Class	Flight Distance	Inflight wifi service
0	Eco Plus	460	3
1	Business	235	3
2	Business	1142	2
3	Business	562	2
4	Business	214	3

	Departure/Arrival time convenient	...	Inflight entertainment
0	4	...	5
1	2	...	1
2	2	...	5
3	5	...	2

4 3 ... 3

	On-board service	Leg room service	Baggage handling	Checkin service \
0	4	3	4	
4				
1	1	5	3	
1				
2	4	3	4	
4				
3	2	5	3	
1				
4	3	4	4	
3				

	Inflight service	Cleanliness	Departure Delay in Minutes \
0	5	5	25
1	4	1	1
2	4	5	0
3	4	2	11
4	3	3	0

	Arrival Delay in Minutes	satisfaction
0	18.0	neutral or dissatisfied
1	6.0	neutral or dissatisfied
2	0.0	satisfied
3	9.0	neutral or dissatisfied
4	0.0	satisfied

[5 rows x 25 columns]

```
test_data = pd.read_csv('test.csv')
```

```
test_data.head()
```

	Unnamed: 0	id	Gender	Customer Type	Age	Type of Travel
0	0	19556	Female	Loyal Customer	52	Business travel
1	1	90035	Female	Loyal Customer	36	Business travel
2	2	12360	Male	disloyal Customer	20	Business travel
3	3	77959	Male	Loyal Customer	44	Business travel
4	4	36875	Female	Loyal Customer	49	Business travel

	Class	Flight Distance	Inflight wifi service \
0	Eco	160	5

1	Business	2863	1
2	Eco	192	2
3	Business	3377	0
4	Eco	1182	2

	Departure/Arrival time convenient	...	Inflight entertainment	\
0	4	...	5	
1	1	...	4	
2	0	...	2	
3	0	...	1	
4	3	...	2	

	On-board service	Leg room service	Baggage handling	Checkin service \
0	5	5	5	
2				
1	4	4	4	
3				
2	4	1	3	
2				
3	1	1	1	
3				
4	2	2	2	
4				

	Inflight service	Cleanliness	Departure Delay in Minutes	\
0	5	5	50	
1	4	5	0	
2	2	2	0	
3	1	4	0	
4	2	4	0	

	Arrival Delay in Minutes	satisfaction
0	44.0	satisfied
1	0.0	satisfied
2	0.0	neutral or dissatisfied
3	6.0	satisfied
4	20.0	satisfied

[5 rows x 25 columns]

Understand Data Variables

#need to remove first column

```
train_data = train_data.iloc[:, 1:]
test_data = test_data.iloc[:, 1:]

train_data.head()
```

Class	id	Gender	Customer Type	Age	Type of Travel	
0	70172	Male	Loyal Customer	13	Personal Travel	Eco Plus
1	5047	Male	disloyal Customer	25	Business travel	Business
2	110028	Female	Loyal Customer	26	Business travel	Business
3	24026	Female	Loyal Customer	25	Business travel	Business
4	119299	Male	Loyal Customer	61	Business travel	Business

	Flight Distance	Inflight wifi service	Departure/Arrival time
0	460	3	
1	235	3	
2	1142	2	
3	562	2	
4	214	3	

	Ease of Online booking	Inflight entertainment	On-board
0	3	5	
1	3	1	
2	2	5	
3	5	2	
4	3	3	

	Leg room service	Baggage handling	Checkin service	Inflight
0	3	4	4	
1	5	3	1	
2	3	4	4	
3	5	3	1	

4		4		4		3
3						
	Cleanliness	Departure Delay in Minutes		Arrival Delay in		
	Minutes \					
0	5		25			18.0
1	1		1			6.0
2	5		0			0.0
3	2		11			9.0
4	3		0			0.0

	satisfaction
0	neutral or dissatisfied
1	neutral or dissatisfied
2	satisfied
3	neutral or dissatisfied
4	satisfied

[5 rows x 24 columns]

test_data.head()

	id	Gender	Customer Type	Age	Type of Travel	Class \
0	19556	Female	Loyal Customer	52	Business travel	Eco
1	90035	Female	Loyal Customer	36	Business travel	Business
2	12360	Male	disloyal Customer	20	Business travel	Eco
3	77959	Male	Loyal Customer	44	Business travel	Business
4	36875	Female	Loyal Customer	49	Business travel	Eco

	Flight Distance	Inflight wifi service	Departure/Arrival time
	convenient \		
0	160		5
4			
1	2863		1
1			
2	192		2
0			
3	3377		0
0			
4	1182		2
3			

	Ease of Online booking	...	Inflight entertainment	On-board
	service \			
0		3	...	5

```

5
1          3 ...          4
4
2          2 ...          2
4
3          0 ...          1
1
4          4 ...          2
2

```

```

    Leg room service  Baggage handling  Checkin service  Inflight
service \
0          5          5          2
5
1          4          4          3
4
2          1          3          2
2
3          1          1          3
1
4          2          2          4
2

```

```

    Cleanliness  Departure Delay in Minutes  Arrival Delay in
Minutes \
0          5          50          44.0
1          5          0          0.0
2          2          0          0.0
3          4          0          6.0
4          4          0          20.0

```

```

    satisfaction
0          satisfied
1          satisfied
2  neutral or dissatisfied
3          satisfied
4          satisfied

```

[5 rows x 24 columns]

```
train_data.describe()
```

```

            id          Age  Flight Distance  Inflight wifi
service \
count  103904.000000  103904.000000    103904.000000

```

103904.000000			
mean	64924.210502	39.379706	1189.448375
2.729683			
std	37463.812252	15.114964	997.147281
1.327829			
min	1.000000	7.000000	31.000000
0.000000			
25%	32533.750000	27.000000	414.000000
2.000000			
50%	64856.500000	40.000000	843.000000
3.000000			
75%	97368.250000	51.000000	1743.000000
4.000000			
max	129880.000000	85.000000	4983.000000
5.000000			

	Departure/Arrival time convenient	Ease of Online booking \
count	103904.000000	103904.000000
mean	3.060296	2.756901
std	1.525075	1.398929
min	0.000000	0.000000
25%	2.000000	2.000000
50%	3.000000	3.000000
75%	4.000000	4.000000
max	5.000000	5.000000

	Gate location	Food and drink	Online boarding	Seat
count	103904.000000	103904.000000	103904.000000	103904.000000
mean	2.976883	3.202129	3.250375	3.439396
std	1.277621	1.329533	1.349509	1.319088
min	0.000000	0.000000	0.000000	0.000000
25%	2.000000	2.000000	2.000000	2.000000
50%	3.000000	3.000000	3.000000	4.000000
75%	4.000000	4.000000	4.000000	5.000000
max	5.000000	5.000000	5.000000	5.000000

	Inflight entertainment	On-board service	Leg room service \
count	103904.000000	103904.000000	103904.000000
mean	3.358158	3.382363	3.351055
std	1.332991	1.288354	1.315605

min	0.000000	0.000000	0.000000
25%	2.000000	2.000000	2.000000
50%	4.000000	4.000000	4.000000
75%	4.000000	4.000000	4.000000
max	5.000000	5.000000	5.000000

	Baggage handling	Checkin service	Inflight service
Cleanliness \			
count	103904.000000	103904.000000	103904.000000
103904.000000			
mean	3.631833	3.304290	3.640428
3.286351			
std	1.180903	1.265396	1.175663
1.312273			
min	1.000000	0.000000	0.000000
0.000000			
25%	3.000000	3.000000	3.000000
2.000000			
50%	4.000000	3.000000	4.000000
3.000000			
75%	5.000000	4.000000	5.000000
4.000000			
max	5.000000	5.000000	5.000000
5.000000			

	Departure Delay in Minutes	Arrival Delay in Minutes
count	103904.000000	103594.000000
mean	14.815618	15.178678
std	38.230901	38.698682
min	0.000000	0.000000
25%	0.000000	0.000000
50%	0.000000	0.000000
75%	12.000000	13.000000
max	1592.000000	1584.000000

```
train_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 103904 entries, 0 to 103903
Data columns (total 24 columns):
```

#	Column	Non-Null Count	Dtype
0	id	103904 non-null	int64
1	Gender	103904 non-null	object
2	Customer Type	103904 non-null	object
3	Age	103904 non-null	int64
4	Type of Travel	103904 non-null	object
5	Class	103904 non-null	object
6	Flight Distance	103904 non-null	int64
7	Inflight wifi service	103904 non-null	int64


```

8   Departure/Arrival time convenient 103904 non-null int64
9   Ease of Online booking            103904 non-null int64
10  Gate location                     103904 non-null int64
11  Food and drink                    103904 non-null int64
12  Online boarding                   103904 non-null int64
13  Seat comfort                      103904 non-null int64
14  Inflight entertainment            103904 non-null int64
15  On-board service                  103904 non-null int64
16  Leg room service                  103904 non-null int64
17  Baggage handling                  103904 non-null int64
18  Checkin service                   103904 non-null int64
19  Inflight service                  103904 non-null int64
20  Cleanliness                       103904 non-null int64
21  Departure Delay in Minutes        103904 non-null int64
22  Arrival Delay in Minutes          103594 non-null float64
23  satisfaction                      103904 non-null object

dtypes: float64(1), int64(18), object(5)
memory usage: 19.0+ MB

```

```
test_data.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 25976 entries, 0 to 25975
Data columns (total 24 columns):

```

#	Column	Non-Null Count	Dtype
0	id	25976 non-null	int64
1	Gender	25976 non-null	object
2	Customer Type	25976 non-null	object
3	Age	25976 non-null	int64
4	Type of Travel	25976 non-null	object
5	Class	25976 non-null	object
6	Flight Distance	25976 non-null	int64
7	Inflight wifi service	25976 non-null	int64
8	Departure/Arrival time convenient	25976 non-null	int64
9	Ease of Online booking	25976 non-null	int64
10	Gate location	25976 non-null	int64
11	Food and drink	25976 non-null	int64
12	Online boarding	25976 non-null	int64
13	Seat comfort	25976 non-null	int64
14	Inflight entertainment	25976 non-null	int64
15	On-board service	25976 non-null	int64
16	Leg room service	25976 non-null	int64
17	Baggage handling	25976 non-null	int64
18	Checkin service	25976 non-null	int64
19	Inflight service	25976 non-null	int64
20	Cleanliness	25976 non-null	int64
21	Departure Delay in Minutes	25976 non-null	int64
22	Arrival Delay in Minutes	25893 non-null	float64
23	satisfaction	25976 non-null	object

```
dtypes: float64(1), int64(18), object(5)
memory usage: 4.8+ MB
```

```
len(train_data.describe().columns) # num of numeric columns in data set
```

```
19
```

```
train_data.columns
```

```
Index(['id', 'Gender', 'Customer Type', 'Age', 'Type of Travel', 'Class',
      'Flight Distance', 'Inflight wifi service',
      'Departure/Arrival time convenient', 'Ease of Online booking',
      'Gate location', 'Food and drink', 'Online boarding', 'Seat comfort',
      'Inflight entertainment', 'On-board service', 'Leg room service',
      'Baggage handling', 'Checkin service', 'Inflight service',
      'Cleanliness', 'Departure Delay in Minutes', 'Arrival Delay in Minutes',
      'satisfaction'],
      dtype='object')
```

```
train_data.shape
```

```
(103904, 24)
```

```
train_data.isnull().head()
```

	id	Gender	Customer Type	Age	Type of Travel	Class	\
0	False	False	False	False	False	False	
1	False	False	False	False	False	False	
2	False	False	False	False	False	False	
3	False	False	False	False	False	False	
4	False	False	False	False	False	False	

	Flight Distance	Inflight wifi service	Departure/Arrival time convenient	\
0	False	False		
1	False	False		
2	False	False		
3	False	False		
4	False	False		

	Ease of Online booking	...	Inflight entertainment	On-board service	\
--	------------------------	-----	------------------------	------------------	---

0	False	...	False
False			
1	False	...	False
False			
2	False	...	False
False			
3	False	...	False
False			
4	False	...	False
False			

	Leg room service	Baggage handling	Checkin service	Inflight
service \				
0	False	False	False	
False				
1	False	False	False	
False				
2	False	False	False	
False				
3	False	False	False	
False				
4	False	False	False	
False				

	Cleanliness	Departure Delay in Minutes	Arrival Delay in
Minutes \			
0	False	False	False
1	False	False	False
2	False	False	False
3	False	False	False
4	False	False	False

	satisfaction
0	False
1	False
2	False
3	False
4	False

[5 rows x 24 columns]

`train_data.isnull().sum()` *# checking for any null values*

id	0
Gender	0

Customer Type	0
Age	0
Type of Travel	0
Class	0
Flight Distance	0
Inflight wifi service	0
Departure/Arrival time convenient	0
Ease of Online booking	0
Gate location	0
Food and drink	0
Online boarding	0
Seat comfort	0
Inflight entertainment	0
On-board service	0
Leg room service	0
Baggage handling	0
Checkin service	0
Inflight service	0
Cleanliness	0
Departure Delay in Minutes	0
Arrival Delay in Minutes	310
satisfaction	0
dtype: int64	

Correlations !!!

```
labels = train_data["Customer Type"].value_counts().index
```

```
labels
```

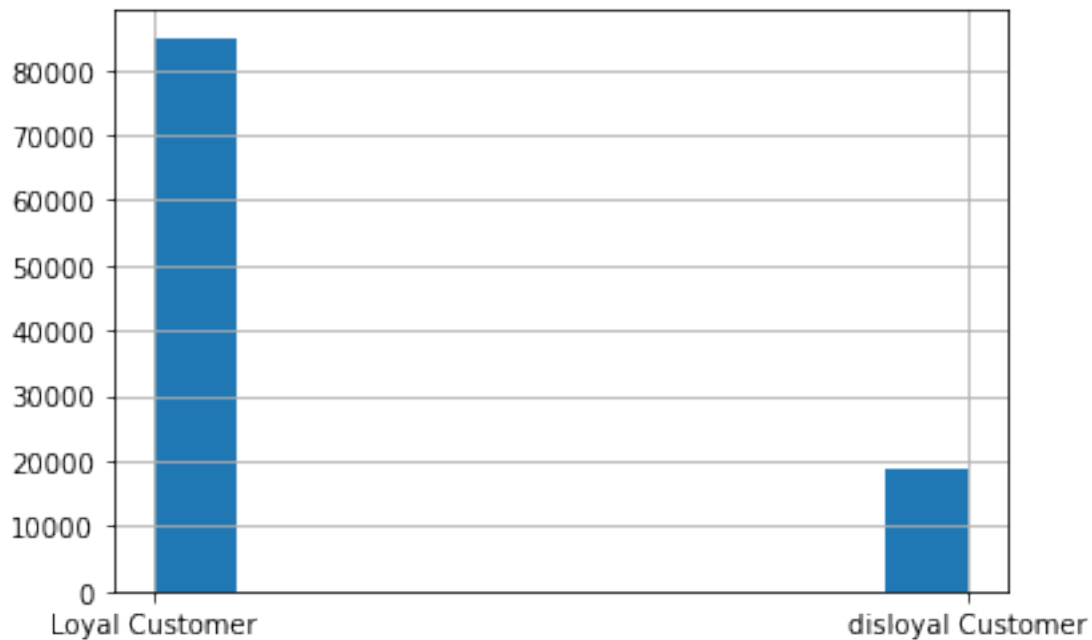
```
Index(['Loyal Customer', 'disloyal Customer'], dtype='object')
```

```
train_data["Customer Type"].value_counts().values
```

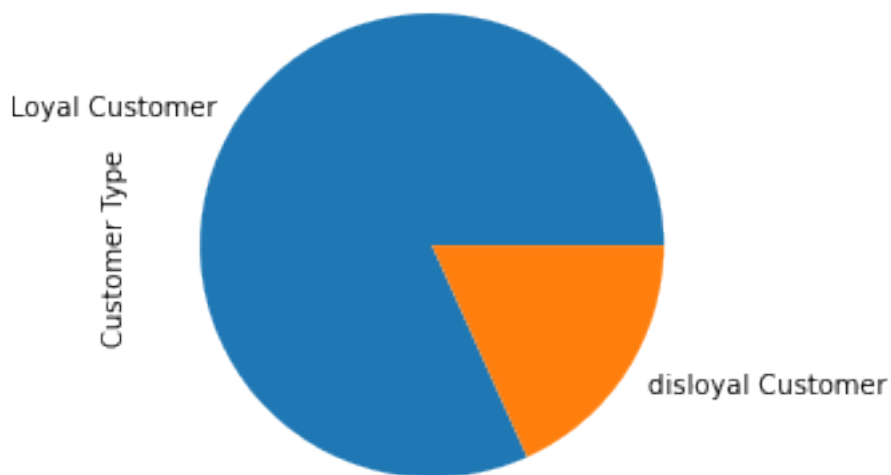
```
array([84923, 18981], dtype=int64)
```

```
train_data["Customer Type"].hist()
```

```
<AxesSubplot:>
```



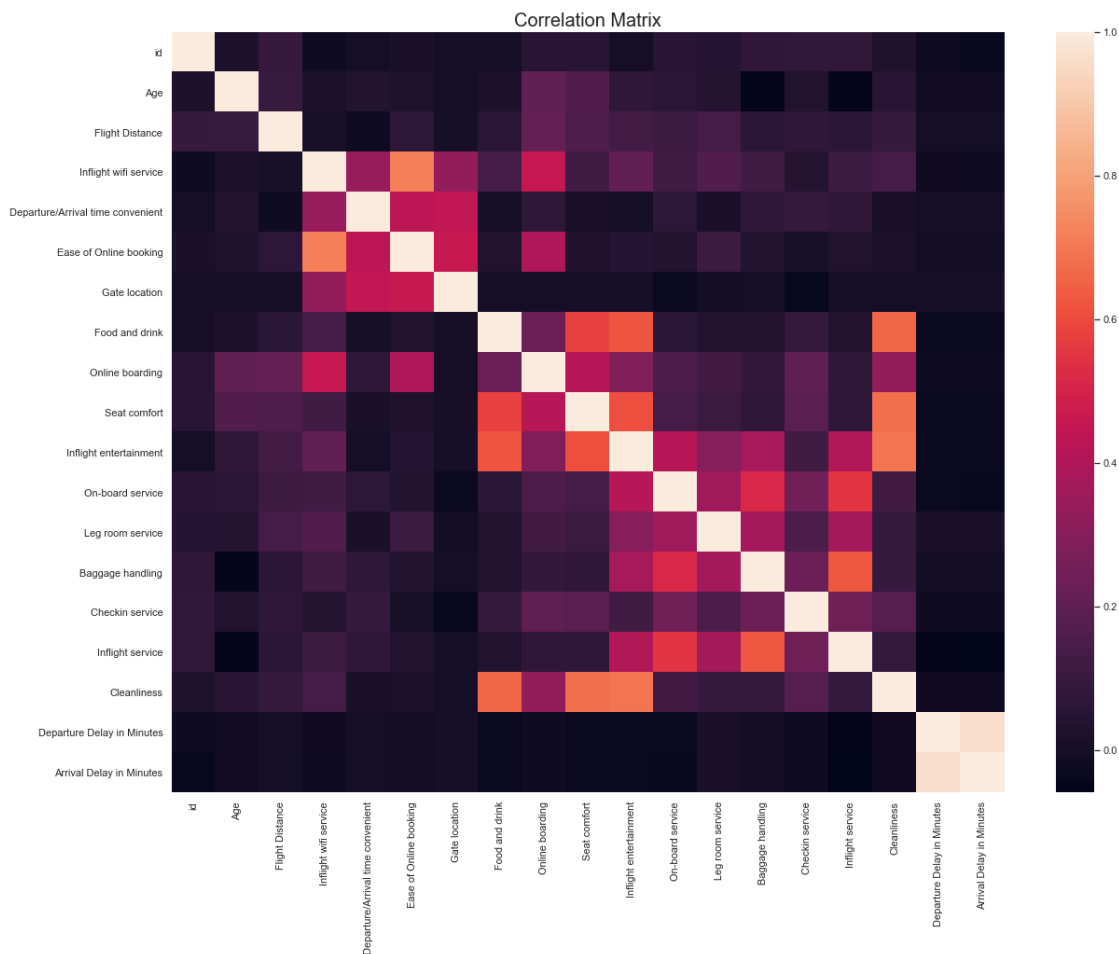
```
train_data["Customer Type"].value_counts().plot.pie()
<AxesSubplot:ylabel='Customer Type'>
```



```
plt.figure(figsize=(12,12))
sns.set(style = "white")
corr = train_data.corr()
mask = np.triu(np.ones_like(corr))
f, ax = plt.subplots(figsize=(20, 15))
cmap = sns.diverging_palette(100, 15, as_cmap=True)
plt.title('Correlation Matrix', fontsize=20)
sns.heatmap(corr)
```

```
<AxesSubplot:title={'center':'Correlation Matrix'}>
```

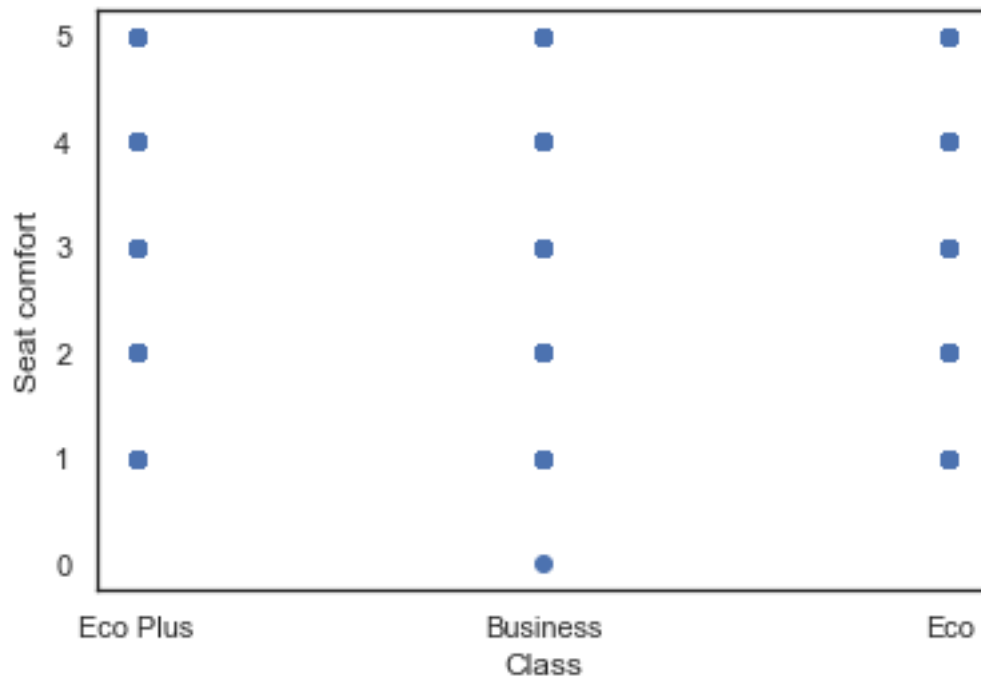
```
<Figure size 864x864 with 0 Axes>
```



Pattern Between Different Feature And Their Business Impact

```
def scatter(train_data, column1, column2):  
    plt.figure()  
    plt.scatter(train_data[column1] , train_data[column2])  
    plt.ylabel(column2)  
    plt.xlabel(column1)
```

```
# scatter(train_data, 'Age',) # with respect to Fuel_Price  
scatter(train_data, 'Class', 'Seat comfort') # with respect to Size  
# scatter(train_data, 'Seat comfort') # with respect to CPI  
# scatter(train_data, 'Type') # with respect to Type  
# scatter(train_data, 'IsHoliday') # with respect to IsHoliday  
# scatter(train_data, 'Unemployment') # with respect to Unemployment  
# scatter(train_data, 'Temperature') # with respect to Temperature  
# scatter(train_data, 'Store') # with respect to Store  
# scatter(train_data, 'Dept') # with respect to Dept
```



```
train_data.head()
```

	id	Gender	Customer Type	Age	Type of Travel	
Class \						
0	70172	Male	Loyal Customer	13	Personal Travel	Eco Plus
1	5047	Male	disloyal Customer	25	Business travel	Business
2	110028	Female	Loyal Customer	26	Business travel	Business
3	24026	Female	Loyal Customer	25	Business travel	Business
4	119299	Male	Loyal Customer	61	Business travel	Business

	Flight Distance	Inflight wifi service	Departure/Arrival time
convenient \			
0	460	3	
4			
1	235	3	
2			
2	1142	2	
2			
3	562	2	
5			
4	214	3	
3			

	Ease of service	Online booking	...	Inflight entertainment	On-board
0		3	...		5
4					
1		3	...		1
1					
2		2	...		5
4					
3		5	...		2
2					
4		3	...		3
3					

	Leg room service	Baggage handling	Checkin service	Inflight
0	3	4	4	
5				
1	5	3	1	
4				
2	3	4	4	
4				
3	5	3	1	
4				
4	4	4	3	
3				

	Cleanliness	Departure Delay in Minutes	Arrival Delay in
Minutes \			
0	5	25	18.0
1	1	1	6.0
2	5	0	0.0
3	2	11	9.0
4	3	0	0.0

	satisfaction
0	neutral or dissatisfied
1	neutral or dissatisfied
2	satisfied
3	neutral or dissatisfied
4	satisfied

[5 rows x 24 columns]

`train_data.corr()[train_data.corr().>0.5]['Seat comfort']`

id	NaN
Age	NaN
Flight Distance	NaN
Inflight wifi service	NaN
Departure/Arrival time convenient	NaN
Ease of Online booking	NaN
Gate location	NaN
Food and drink	0.574556
Online boarding	NaN
Seat comfort	1.000000
Inflight entertainment	0.610590
On-board service	NaN
Leg room service	NaN
Baggage handling	NaN
Checkin service	NaN
Inflight service	NaN
Cleanliness	0.678534
Departure Delay in Minutes	NaN
Arrival Delay in Minutes	NaN

Name: Seat comfort, dtype: float64

-As per above analysis, Seat comfort has high correlation with Food and drink, Inflight entertainment and Cleanliness

`train_data.corr()[train_data.corr()>0.5]`

	id	Age	Flight Distance	\
id	1.0	NaN	NaN	
Age	NaN	1.0	NaN	
Flight Distance	NaN	NaN	1.0	
Inflight wifi service	NaN	NaN	NaN	
Departure/Arrival time convenient	NaN	NaN	NaN	
Ease of Online booking	NaN	NaN	NaN	
Gate location	NaN	NaN	NaN	
Food and drink	NaN	NaN	NaN	
Online boarding	NaN	NaN	NaN	
Seat comfort	NaN	NaN	NaN	
Inflight entertainment	NaN	NaN	NaN	
On-board service	NaN	NaN	NaN	
Leg room service	NaN	NaN	NaN	
Baggage handling	NaN	NaN	NaN	
Checkin service	NaN	NaN	NaN	
Inflight service	NaN	NaN	NaN	
Cleanliness	NaN	NaN	NaN	
Departure Delay in Minutes	NaN	NaN	NaN	
Arrival Delay in Minutes	NaN	NaN	NaN	

	Inflight wifi service	\
id	NaN	
Age	NaN	
Flight Distance	NaN	

Inflight wifi service	1.000000
Departure/Arrival time convenient	NaN
Ease of Online booking	0.715856
Gate location	NaN
Food and drink	NaN
Online boarding	NaN
Seat comfort	NaN
Inflight entertainment	NaN
On-board service	NaN
Leg room service	NaN
Baggage handling	NaN
Checkin service	NaN
Inflight service	NaN
Cleanliness	NaN
Departure Delay in Minutes	NaN
Arrival Delay in Minutes	NaN

	Departure/Arrival time	
convenient \		
id		NaN
Age		NaN
Flight Distance		NaN
Inflight wifi service		NaN
Departure/Arrival time convenient		1.0
Ease of Online booking		NaN
Gate location		NaN
Food and drink		NaN
Online boarding		NaN
Seat comfort		NaN
Inflight entertainment		NaN
On-board service		NaN
Leg room service		NaN
Baggage handling		NaN
Checkin service		NaN

Inflight service	NaN
Cleanliness	NaN
Departure Delay in Minutes	NaN
Arrival Delay in Minutes	NaN

	Ease of Online booking	Gate
location \		
id	NaN	
NaN		
Age	NaN	
NaN		
Flight Distance	NaN	
NaN		
Inflight wifi service	0.715856	
NaN		
Departure/Arrival time convenient	NaN	
NaN		
Ease of Online booking	1.000000	
NaN		
Gate location	NaN	
1.0		
Food and drink	NaN	
NaN		
Online boarding	NaN	
NaN		
Seat comfort	NaN	
NaN		
Inflight entertainment	NaN	
NaN		
On-board service	NaN	
NaN		
Leg room service	NaN	
NaN		
Baggage handling	NaN	
NaN		
Checkin service	NaN	
NaN		
Inflight service	NaN	
NaN		
Cleanliness	NaN	
NaN		
Departure Delay in Minutes	NaN	
NaN		
Arrival Delay in Minutes	NaN	
NaN		

	Food and drink	Online boarding \
id	NaN	NaN
Age	NaN	NaN
Flight Distance	NaN	NaN
Inflight wifi service	NaN	NaN
Departure/Arrival time convenient	NaN	NaN
Ease of Online booking	NaN	NaN
Gate location	NaN	NaN
Food and drink	1.000000	NaN
Online boarding	NaN	1.0
Seat comfort	0.574556	NaN
Inflight entertainment	0.622512	NaN
On-board service	NaN	NaN
Leg room service	NaN	NaN
Baggage handling	NaN	NaN
Checkin service	NaN	NaN
Inflight service	NaN	NaN
Cleanliness	0.657760	NaN
Departure Delay in Minutes	NaN	NaN
Arrival Delay in Minutes	NaN	NaN

	Seat comfort	Inflight
entertainment \		
id	NaN	
NaN		
Age	NaN	
NaN		
Flight Distance	NaN	
NaN		
Inflight wifi service	NaN	
NaN		
Departure/Arrival time convenient	NaN	
NaN		
Ease of Online booking	NaN	
NaN		
Gate location	NaN	
NaN		
Food and drink	0.574556	
0.622512		
Online boarding	NaN	
NaN		
Seat comfort	1.000000	
0.610590		
Inflight entertainment	0.610590	
1.000000		
On-board service	NaN	
NaN		
Leg room service	NaN	
NaN		
Baggage handling	NaN	

NaN	
Checkin service	NaN
NaN	
Inflight service	NaN
NaN	
Cleanliness	0.678534
0.691815	
Departure Delay in Minutes	NaN
NaN	
Arrival Delay in Minutes	NaN
NaN	

	On-board service	Leg room service
\		
id	NaN	NaN
Age	NaN	NaN
Flight Distance	NaN	NaN
Inflight wifi service	NaN	NaN
Departure/Arrival time convenient	NaN	NaN
Ease of Online booking	NaN	NaN
Gate location	NaN	NaN
Food and drink	NaN	NaN
Online boarding	NaN	NaN
Seat comfort	NaN	NaN
Inflight entertainment	NaN	NaN
On-board service	1.000000	NaN
Leg room service	NaN	1.0
Baggage handling	0.519134	NaN
Checkin service	NaN	NaN
Inflight service	0.550782	NaN
Cleanliness	NaN	NaN

Departure Delay in Minutes	NaN	NaN
Arrival Delay in Minutes	NaN	NaN

	Baggage handling	Checkin
service \ id	NaN	NaN
Age	NaN	NaN
Flight Distance	NaN	NaN
Inflight wifi service	NaN	NaN
Departure/Arrival time convenient	NaN	NaN
Ease of Online booking	NaN	NaN
Gate location	NaN	NaN
Food and drink	NaN	NaN
Online boarding	NaN	NaN
Seat comfort	NaN	NaN
Inflight entertainment	NaN	NaN
On-board service	0.519134	NaN
Leg room service	NaN	NaN
Baggage handling	1.000000	NaN
Checkin service	NaN	1.0
Inflight service	0.628561	NaN
Cleanliness	NaN	NaN
Departure Delay in Minutes	NaN	NaN
Arrival Delay in Minutes	NaN	NaN

	Inflight service	Cleanliness \
id	NaN	NaN

Age	NaN	NaN
Flight Distance	NaN	NaN
Inflight wifi service	NaN	NaN
Departure/Arrival time convenient	NaN	NaN
Ease of Online booking	NaN	NaN
Gate location	NaN	NaN
Food and drink	NaN	0.657760
Online boarding	NaN	NaN
Seat comfort	NaN	0.678534
Inflight entertainment	NaN	0.691815
On-board service	0.550782	NaN
Leg room service	NaN	NaN
Baggage handling	0.628561	NaN
Checkin service	NaN	NaN
Inflight service	1.000000	NaN
Cleanliness	NaN	1.000000
Departure Delay in Minutes	NaN	NaN
Arrival Delay in Minutes	NaN	NaN

	Departure Delay in Minutes	\
id		NaN
Age		NaN
Flight Distance		NaN
Inflight wifi service		NaN
Departure/Arrival time convenient		NaN
Ease of Online booking		NaN
Gate location		NaN
Food and drink		NaN
Online boarding		NaN
Seat comfort		NaN
Inflight entertainment		NaN
On-board service		NaN
Leg room service		NaN
Baggage handling		NaN
Checkin service		NaN
Inflight service		NaN
Cleanliness		NaN
Departure Delay in Minutes		1.000000
Arrival Delay in Minutes		0.965481

	Arrival Delay in Minutes
id	NaN
Age	NaN
Flight Distance	NaN
Inflight wifi service	NaN
Departure/Arrival time convenient	NaN
Ease of Online booking	NaN
Gate location	NaN
Food and drink	NaN
Online boarding	NaN

Seat comfort	NaN
Inflight entertainment	NaN
On-board service	NaN
Leg room service	NaN
Baggage handling	NaN
Checkin service	NaN
Inflight service	NaN
Cleanliness	NaN
Departure Delay in Minutes	0.965481
Arrival Delay in Minutes	1.000000

-Arrival delay in minutes has high correlation to departure delay in minutes which is expected.

-Inflight WIFI service has high correlation with Ease of Online booking which means flight that has WIFI service provide easy booking for their customers

```
train_data.corr()[train_data.corr()<0]
```

	id	Age	Flight Distance
\			
id	NaN	NaN	NaN
Age	NaN	NaN	NaN
Flight Distance	NaN	NaN	NaN
Inflight wifi service	-0.021276	NaN	NaN
Departure/Arrival time convenient	-0.002110	NaN	-0.020043
Ease of Online booking	NaN	NaN	NaN
Gate location	-0.000606	-0.001330	NaN
Food and drink	NaN	NaN	NaN
Online boarding	NaN	NaN	NaN
Seat comfort	NaN	NaN	NaN
Inflight entertainment	NaN	NaN	NaN
On-board service	NaN	NaN	NaN
Leg room service	NaN	NaN	NaN
Baggage handling	NaN	-0.047529	NaN
Checkin service	NaN	NaN	NaN

Inflight service	NaN	-0.049427	NaN
Cleanliness	NaN	NaN	NaN
Departure Delay in Minutes	-0.019546	-0.010152	NaN
Arrival Delay in Minutes	-0.037254	-0.012147	-0.002426

	Inflight wifi service \
id	-0.021276
Age	NaN
Flight Distance	NaN
Inflight wifi service	NaN
Departure/Arrival time convenient	NaN
Ease of Online booking	NaN
Gate location	NaN
Food and drink	NaN
Online boarding	NaN
Seat comfort	NaN
Inflight entertainment	NaN
On-board service	NaN
Leg room service	NaN
Baggage handling	NaN
Checkin service	NaN
Inflight service	NaN
Cleanliness	NaN
Departure Delay in Minutes	-0.017402
Arrival Delay in Minutes	-0.019095

	Departure/Arrival time
convenient \	
id	-0.002110
Age	NaN
Flight Distance	-0.020043
Inflight wifi service	NaN
Departure/Arrival time convenient	NaN
Ease of Online booking	NaN
Gate location	NaN
Food and drink	NaN

Online boarding	NaN
Seat comfort	NaN
Inflight entertainment	-0.004861
On-board service	NaN
Leg room service	NaN
Baggage handling	NaN
Checkin service	NaN
Inflight service	NaN
Cleanliness	NaN
Departure Delay in Minutes	NaN
Arrival Delay in Minutes	-0.000864

	Ease of Online booking	Gate
location \		
id	NaN	-
0.000606		
Age	NaN	-
0.001330		
Flight Distance	NaN	
NaN		
Inflight wifi service	NaN	
NaN		
Departure/Arrival time convenient	NaN	
NaN		
Ease of Online booking	NaN	
NaN		
Gate location	NaN	
NaN		
Food and drink	NaN	-
0.001159		
Online boarding	NaN	
NaN		
Seat comfort	NaN	
NaN		
Inflight entertainment	NaN	
NaN		
On-board service	NaN	-
0.028373		

Leg room service	NaN	-
0.005873		
Baggage handling	NaN	
NaN		
Checkin service	NaN	-
0.035427		
Inflight service	NaN	
NaN		
Cleanliness	NaN	-
0.003830		
Departure Delay in Minutes	-0.006371	
NaN		
Arrival Delay in Minutes	-0.007984	
NaN		

	Food and drink	Online boarding	\
id	NaN	NaN	
Age	NaN	NaN	
Flight Distance	NaN	NaN	
Inflight wifi service	NaN	NaN	
Departure/Arrival time convenient	NaN	NaN	
Ease of Online booking	NaN	NaN	
Gate location	-0.001159	NaN	
Food and drink	NaN	NaN	
Online boarding	NaN	NaN	
Seat comfort	NaN	NaN	
Inflight entertainment	NaN	NaN	
On-board service	NaN	NaN	
Leg room service	NaN	NaN	
Baggage handling	NaN	NaN	
Checkin service	NaN	NaN	
Inflight service	NaN	NaN	
Cleanliness	NaN	NaN	
Departure Delay in Minutes	-0.029926	-0.018982	
Arrival Delay in Minutes	-0.032524	-0.021949	

	Seat comfort	Inflight
entertainment \		
id	NaN	
NaN		
Age	NaN	
NaN		
Flight Distance	NaN	
NaN		
Inflight wifi service	NaN	
NaN		
Departure/Arrival time convenient	NaN	-
0.004861		
Ease of Online booking	NaN	
NaN		

Gate location	NaN	
NaN		
Food and drink	NaN	
NaN		
Online boarding	NaN	
NaN		
Seat comfort	NaN	
NaN		
Inflight entertainment	NaN	
NaN		
On-board service	NaN	
NaN		
Leg room service	NaN	
NaN		
Baggage handling	NaN	
NaN		
Checkin service	NaN	
NaN		
Inflight service	NaN	
NaN		
Cleanliness	NaN	
NaN		
Departure Delay in Minutes	-0.027898	-
0.027489		
Arrival Delay in Minutes	-0.029900	-
0.030703		

	On-board service	Leg room service
\		
id	NaN	NaN
Age	NaN	NaN
Flight Distance	NaN	NaN
Inflight wifi service	NaN	NaN
Departure/Arrival time convenient	NaN	NaN
Ease of Online booking	NaN	NaN
Gate location	-0.028373	-0.005873
Food and drink	NaN	NaN
Online boarding	NaN	NaN
Seat comfort	NaN	NaN

Inflight entertainment	NaN	NaN
On-board service	NaN	NaN
Leg room service	NaN	NaN
Baggage handling	NaN	NaN
Checkin service	NaN	NaN
Inflight service	NaN	NaN
Cleanliness	NaN	NaN
Departure Delay in Minutes	-0.031569	NaN
Arrival Delay in Minutes	-0.035227	NaN

	Baggage handling	Checkin
service \ id	NaN	NaN
Age	-0.047529	NaN
Flight Distance	NaN	NaN
Inflight wifi service	NaN	NaN
Departure/Arrival time convenient	NaN	NaN
Ease of Online booking	NaN	NaN
Gate location	NaN	-0.035427
Food and drink	NaN	NaN
Online boarding	NaN	NaN
Seat comfort	NaN	NaN
Inflight entertainment	NaN	NaN
On-board service	NaN	NaN
Leg room service	NaN	NaN
Baggage handling	NaN	NaN

Checkin service	NaN	NaN
Inflight service	NaN	NaN
Cleanliness	NaN	NaN
Departure Delay in Minutes	-0.005573	-0.018453
Arrival Delay in Minutes	-0.008542	-0.020369

	Inflight service	Cleanliness \
id	NaN	NaN
Age	-0.049427	NaN
Flight Distance	NaN	NaN
Inflight wifi service	NaN	NaN
Departure/Arrival time convenient	NaN	NaN
Ease of Online booking	NaN	NaN
Gate location	NaN	-0.003830
Food and drink	NaN	NaN
Online boarding	NaN	NaN
Seat comfort	NaN	NaN
Inflight entertainment	NaN	NaN
On-board service	NaN	NaN
Leg room service	NaN	NaN
Baggage handling	NaN	NaN
Checkin service	NaN	NaN
Inflight service	NaN	NaN
Cleanliness	NaN	NaN
Departure Delay in Minutes	-0.054813	-0.014093
Arrival Delay in Minutes	-0.059196	-0.015774

	Departure Delay in Minutes \
id	-0.019546
Age	-0.010152
Flight Distance	NaN
Inflight wifi service	-0.017402
Departure/Arrival time convenient	NaN
Ease of Online booking	-0.006371
Gate location	NaN
Food and drink	-0.029926
Online boarding	-0.018982
Seat comfort	-0.027898
Inflight entertainment	-0.027489
On-board service	-0.031569
Leg room service	NaN
Baggage handling	-0.005573
Checkin service	-0.018453

Inflight service	-0.054813
Cleanliness	-0.014093
Departure Delay in Minutes	NaN
Arrival Delay in Minutes	NaN

	Arrival Delay in Minutes
id	-0.037254
Age	-0.012147
Flight Distance	-0.002426
Inflight wifi service	-0.019095
Departure/Arrival time convenient	-0.000864
Ease of Online booking	-0.007984
Gate location	NaN
Food and drink	-0.032524
Online boarding	-0.021949
Seat comfort	-0.029900
Inflight entertainment	-0.030703
On-board service	-0.035227
Leg room service	NaN
Baggage handling	-0.008542
Checkin service	-0.020369
Inflight service	-0.059196
Cleanliness	-0.015774
Departure Delay in Minutes	NaN
Arrival Delay in Minutes	NaN

Finding of people with higher class has more satisfaction

train_data.head()

	id	Gender	Customer Type	Age	Type of Travel	
Class \						
0	70172	Male	Loyal Customer	13	Personal Travel	Eco Plus
1	5047	Male	disloyal Customer	25	Business travel	Business
2	110028	Female	Loyal Customer	26	Business travel	Business
3	24026	Female	Loyal Customer	25	Business travel	Business
4	119299	Male	Loyal Customer	61	Business travel	Business

	Flight Distance	Inflight wifi service	Departure/Arrival time convenient \
0	460	3	
4			
1	235	3	
2			
2	1142	2	
2			

3	562	2
5		
4	214	3
3		

	Ease of Online booking service \	...	Inflight entertainment	On-board
0	3	...	5	
4				
1	3	...	1	
1				
2	2	...	5	
4				
3	5	...	2	
2				
4	3	...	3	
3				

	Leg room service	Baggage handling	Checkin service	Inflight
0	3	4	4	
5				
1	5	3	1	
4				
2	3	4	4	
4				
3	5	3	1	
4				
4	4	4	3	
3				

	Cleanliness Minutes \	Departure Delay in Minutes	Arrival Delay in
0	5	25	18.0
1	1	1	6.0
2	5	0	0.0
3	2	11	9.0
4	3	0	0.0

	satisfaction
0	neutral or dissatisfied
1	neutral or dissatisfied
2	satisfied
3	neutral or dissatisfied

4 satisfied

[5 rows x 24 columns]

```
class_satisfaction_df =  
train_data[['Class', 'satisfaction']].value_counts().sort_index()
```

```
class_satisfaction_df
```

```
Class      satisfaction  
Business  neutral or dissatisfied    15185  
          satisfied                 34480  
Eco        neutral or dissatisfied    38044  
          satisfied                 8701  
Eco Plus   neutral or dissatisfied    5650  
          satisfied                 1844  
dtype: int64
```

- Percentage of people from Business Class that were satisfied:

```
class_satisfaction_df['Business']['satisfied']*100/  
(class_satisfaction_df['Business']['satisfied']  
+class_satisfaction_df['Business']['neutral or dissatisfied'])  
  
69.42514849491593
```

- Percentage of people from Economy class that were satisfied:

```
class_satisfaction_df['Eco']['satisfied']*100/  
(class_satisfaction_df['Eco']['satisfied']  
+class_satisfaction_df['Eco']['neutral or dissatisfied'])  
  
18.61375548186972
```

- Percentage of people from Economy Plus class that were satisfied:

```
class_satisfaction_df['Eco Plus']  
['satisfied']*100/(class_satisfaction_df['Eco Plus']['satisfied']  
+class_satisfaction_df['Eco Plus']['neutral or dissatisfied'])  
  
24.60635174806512
```

From above results as per the class, Business class people are more satisfied than others

```
eco_class_data = train_data[train_data['Class']=='Eco']
```

```
eco_class_data
```

	id	Gender	Customer Type	Age	Type of Travel	Class
5	111157	Female	Loyal Customer	26	Personal Travel	Eco
6	82113	Male	Loyal Customer	47	Personal Travel	Eco

9	65725	Male	disloyal Customer	20	Business travel	Eco
10	34991	Female	disloyal Customer	24	Business travel	Eco
12	98628	Male	Loyal Customer	53	Business travel	Eco
...
103895	66030	Female	disloyal Customer	24	Business travel	Eco
103896	71445	Male	Loyal Customer	57	Business travel	Eco
103898	60666	Male	Loyal Customer	50	Personal Travel	Eco
103899	94171	Female	disloyal Customer	23	Business travel	Eco
103902	54173	Female	disloyal Customer	22	Business travel	Eco

	Flight Distance	Inflight wifi service	\
5	1180	3	
6	1276	2	
9	1061	3	
10	1182	4	
12	834	1	
...	
103895	1055	1	
103896	867	4	
103898	1620	3	
103899	192	2	
103902	1000	1	

	Departure/Arrival time convenient	Ease of Online booking	...
\			
5	4	2	...
6	4	2	...
9	3	3	...
10	5	5	...
12	4	4	...
...
103895	1	1	...

103896	5	5	...
103898	1	3	...
103899	1	2	...
103902	1	1	...

	Inflight entertainment	On-board service	Leg room service	\
5	1	3	4	
6	2	3	3	
9	2	2	3	
10	2	3	3	
12	1	1	1	
...	
103895	1	3	3	
103896	4	3	4	
103898	2	4	3	
103899	2	3	1	
103902	1	4	5	

	Baggage handling	Checkin service	Inflight service	
Cleanliness \				
5	4	4	4	
1				
6	4	3	5	
2				
9	4	4	3	
2				
10	5	3	5	
2				
12	3	4	4	
1				
...
...				
103895	5	5	4	
1				
103896	3	1	3	
4				
103898	4	2	4	
2				
103899	4	2	3	
2				
103902	1	5	4	
1				

	Departure Delay in Minutes	Arrival Delay in Minutes	\
5	0	0.0	

6	9	23.0
9	0	0.0
10	0	0.0
12	28	8.0
...
103895	13	10.0
103896	0	0.0
103898	0	0.0
103899	3	0.0
103902	0	0.0

	satisfaction
5	neutral or dissatisfied
6	neutral or dissatisfied
9	neutral or dissatisfied
10	neutral or dissatisfied
12	neutral or dissatisfied
...	...
103895	neutral or dissatisfied
103896	neutral or dissatisfied
103898	neutral or dissatisfied
103899	neutral or dissatisfied
103902	neutral or dissatisfied

[46745 rows x 24 columns]

eco_class_data.corr()[eco_class_data.corr()<0]

	id	Age	Flight Distance
\			
id	NaN	-0.007525	NaN
Age	-0.007525	NaN	-0.022706
Flight Distance	NaN	-0.022706	NaN
Inflight wifi service	-0.070385	NaN	-0.023015
Departure/Arrival time convenient	NaN	NaN	NaN
Ease of Online booking	-0.004869	NaN	NaN
Gate location	NaN	-0.009461	-0.005444
Food and drink	-0.034884	NaN	-0.031528
Online boarding	-0.036725	NaN	NaN
Seat comfort	-0.033220	NaN	-0.014670

Inflight entertainment	-0.052020	NaN	-0.029777
On-board service	-0.006808	-0.022584	NaN
Leg room service	-0.001907	-0.069000	NaN
Baggage handling	NaN	-0.108044	NaN
Checkin service	NaN	-0.009091	NaN
Inflight service	NaN	-0.100225	-0.006603
Cleanliness	-0.040218	NaN	-0.014125
Departure Delay in Minutes	-0.022996	-0.009110	NaN
Arrival Delay in Minutes	-0.038271	-0.009990	NaN

	Inflight wifi service \
id	-0.070385
Age	NaN
Flight Distance	-0.023015
Inflight wifi service	NaN
Departure/Arrival time convenient	NaN
Ease of Online booking	NaN
Gate location	-0.002613
Food and drink	NaN
Online boarding	NaN
Seat comfort	NaN
Inflight entertainment	NaN
On-board service	NaN
Leg room service	NaN
Baggage handling	NaN
Checkin service	NaN
Inflight service	NaN
Cleanliness	NaN
Departure Delay in Minutes	-0.027428
Arrival Delay in Minutes	-0.029354

	Departure/Arrival time
convenient \	
id	NaN
Age	NaN
Flight Distance	NaN

Inflight wifi service	NaN
Departure/Arrival time convenient	NaN
Ease of Online booking	NaN
Gate location	NaN
Food and drink	NaN
Online boarding	NaN
Seat comfort	NaN
Inflight entertainment	NaN
On-board service	NaN
Leg room service	NaN
Baggage handling	NaN
Checkin service	NaN
Inflight service	NaN
Cleanliness	NaN
Departure Delay in Minutes	-0.002111
Arrival Delay in Minutes	-0.005674

	Ease of Online booking	Gate
location \		
id	-0.004869	
NaN		
Age	NaN	-
0.009461		
Flight Distance	NaN	-
0.005444		
Inflight wifi service	NaN	-
0.002613		
Departure/Arrival time convenient	NaN	
NaN		
Ease of Online booking	NaN	
NaN		
Gate location	NaN	

NaN		
Food and drink	NaN	-
0.007587		
Online boarding	NaN	-
0.006598		
Seat comfort	NaN	-
0.003072		
Inflight entertainment	NaN	-
0.002842		
On-board service	-0.002885	-
0.051601		
Leg room service	NaN	-
0.016577		
Baggage handling	-0.007487	
NaN		
Checkin service	-0.010811	-
0.061150		
Inflight service	-0.009421	
NaN		
Cleanliness	NaN	-
0.013192		
Departure Delay in Minutes	-0.021561	
NaN		
Arrival Delay in Minutes	-0.023021	
NaN		

	Food and drink	Online boarding \
id	-0.034884	-0.036725
Age	NaN	NaN
Flight Distance	-0.031528	NaN
Inflight wifi service	NaN	NaN
Departure/Arrival time convenient	NaN	NaN
Ease of Online booking	NaN	NaN
Gate location	-0.007587	-0.006598
Food and drink	NaN	NaN
Online boarding	NaN	NaN
Seat comfort	NaN	NaN
Inflight entertainment	NaN	NaN
On-board service	NaN	NaN
Leg room service	-0.023098	-0.050992
Baggage handling	-0.023224	-0.063666
Checkin service	NaN	NaN
Inflight service	-0.021133	-0.074496
Cleanliness	NaN	NaN
Departure Delay in Minutes	-0.055879	NaN
Arrival Delay in Minutes	-0.058440	NaN

	Seat comfort	Inflight
entertainment \		
id	-0.033220	-

0.052020		
Age	NaN	
NaN		
Flight Distance	-0.014670	-
0.029777		
Inflight wifi service	NaN	
NaN		
Departure/Arrival time convenient	NaN	
NaN		
Ease of Online booking	NaN	
NaN		
Gate location	-0.003072	-
0.002842		
Food and drink	NaN	
NaN		
Online boarding	NaN	
NaN		
Seat comfort	NaN	
NaN		
Inflight entertainment	NaN	
NaN		
On-board service	-0.003510	
NaN		
Leg room service	-0.040585	
NaN		
Baggage handling	-0.046765	
NaN		
Checkin service	NaN	
NaN		
Inflight service	-0.049472	
NaN		
Cleanliness	NaN	
NaN		
Departure Delay in Minutes	-0.016166	-
0.015353		
Arrival Delay in Minutes	-0.015412	-
0.015678		

	On-board service	Leg room service
\		
id	-0.006808	-0.001907
Age	-0.022584	-0.069000
Flight Distance	NaN	NaN
Inflight wifi service	NaN	NaN
Departure/Arrival time convenient	NaN	NaN

Ease of Online booking	-0.002885	NaN
Gate location	-0.051601	-0.016577
Food and drink	NaN	-0.023098
Online boarding	NaN	-0.050992
Seat comfort	-0.003510	-0.040585
Inflight entertainment	NaN	NaN
On-board service	NaN	NaN
Leg room service	NaN	NaN
Baggage handling	NaN	NaN
Checkin service	NaN	NaN
Inflight service	NaN	NaN
Cleanliness	NaN	-0.018669
Departure Delay in Minutes	-0.009943	NaN
Arrival Delay in Minutes	-0.009736	NaN

	Baggage handling	Checkin
service \ id	NaN	NaN
Age	-0.108044	-0.009091
Flight Distance	NaN	NaN
Inflight wifi service	NaN	NaN
Departure/Arrival time convenient	NaN	NaN
Ease of Online booking	-0.007487	-0.010811
Gate location	NaN	-0.061150
Food and drink	-0.023224	NaN
Online boarding	-0.063666	NaN

Seat comfort	-0.046765	NaN
Inflight entertainment	NaN	NaN
On-board service	NaN	NaN
Leg room service	NaN	NaN
Baggage handling	NaN	NaN
Checkin service	NaN	NaN
Inflight service	NaN	NaN
Cleanliness	-0.011486	NaN
Departure Delay in Minutes	NaN	-0.019120
Arrival Delay in Minutes	NaN	-0.019002

	Inflight service	Cleanliness \
id	NaN	-0.040218
Age	-0.100225	NaN
Flight Distance	-0.006603	-0.014125
Inflight wifi service	NaN	NaN
Departure/Arrival time convenient	NaN	NaN
Ease of Online booking	-0.009421	NaN
Gate location	NaN	-0.013192
Food and drink	-0.021133	NaN
Online boarding	-0.074496	NaN
Seat comfort	-0.049472	NaN
Inflight entertainment	NaN	NaN
On-board service	NaN	NaN
Leg room service	NaN	-0.018669
Baggage handling	NaN	-0.011486
Checkin service	NaN	NaN
Inflight service	NaN	-0.019246
Cleanliness	-0.019246	NaN
Departure Delay in Minutes	-0.047880	-0.011394
Arrival Delay in Minutes	-0.048572	-0.010818

	Departure Delay in Minutes \
id	-0.022996
Age	-0.009110
Flight Distance	NaN
Inflight wifi service	-0.027428
Departure/Arrival time convenient	-0.002111

Ease of Online booking	-0.021561
Gate location	NaN
Food and drink	-0.055879
Online boarding	NaN
Seat comfort	-0.016166
Inflight entertainment	-0.015353
On-board service	-0.009943
Leg room service	NaN
Baggage handling	NaN
Checkin service	-0.019120
Inflight service	-0.047880
Cleanliness	-0.011394
Departure Delay in Minutes	NaN
Arrival Delay in Minutes	NaN

	Arrival Delay in Minutes
id	-0.038271
Age	-0.009990
Flight Distance	NaN
Inflight wifi service	-0.029354
Departure/Arrival time convenient	-0.005674
Ease of Online booking	-0.023021
Gate location	NaN
Food and drink	-0.058440
Online boarding	NaN
Seat comfort	-0.015412
Inflight entertainment	-0.015678
On-board service	-0.009736
Leg room service	NaN
Baggage handling	NaN
Checkin service	-0.019002
Inflight service	-0.048572
Cleanliness	-0.010818
Departure Delay in Minutes	NaN
Arrival Delay in Minutes	NaN

For Economy class, Leg room service and Seat comfort has negative correlation and should be improved.