```
from google.colab import drive
drive.mount('/content/drive')
Mounted at /content/drive
#import library
import pandas as pd
import plotly.express as pr
import plotly.graph_objects as go
# Read the data
df = pd.read csv('/content/drive/MyDrive/project dataset.csv.xls')
# To display the data
df
       Unnamed: 0 Year
                          Quarter Month
                                           DayofMonth
                                                        DayOfWeek
FlightDate \
          1295781
                                                                    1998 -
                    1998
                                 2
                                        4
                                                     2
                                                                 4
04-02
          1125375
                    2013
                                 2
                                        5
                                                    13
                                                                   2013-
1
                                                                 1
05 - 13
           118824
                    1993
                                 3
                                        9
                                                    25
                                                                    1993 -
2
                                                                 6
09-25
           634825
                   1994
                                4
                                       11
                                                    12
                                                                 6
                                                                   1994-
3
11-12
          1888125
                    2017
                                 3
                                        8
                                                    17
                                                                 4 2017-
4
08-17
                               . . .
                                      . . .
. . .
               . . .
                     . . .
                                                   . . .
                                 1
                                        1
26995
           821542
                    2017
                                                    24
                                                                 2
                                                                    2017-
01-24
          1910565
                                 2
                                                    27
                                                                   2013-
26996
                    2013
                                        6
                                                                 4
06-27
             9055
                                 3
                                        8
                                                    26
                                                                 5 2016-
26997
                   2016
08-26
26998
            84136
                    2009
                                 3
                                        8
                                                     8
                                                                   2009 -
                                                                 6
08-08
26999
           113029
                    1993
                                 3
                                        7
                                                    17
                                                                    1993 -
                                                                 6
07 - 17
      Reporting_Airline DOT_ID_Reporting_Airline
IATA CODE Reporting Airline \
0
                      AS
                                               19930
AS
                      ΕV
                                               20366
1
ΕV
                      UA
                                               19977
2
UA
```

19991

3

HP

HP

4 UA		UA		19977	
26995		DL		19790	
DL 26996		В6		20409	
B6 26997		AA		19805	
AA 26998		YV		20378	
YV 26999 DL		DL		19790	
0	Div	4WheelsOff NaN	Div4TailNum NaN	Div5Airport NaN	Div5AirportID \
1 2		NaN	NaN	NaN	NaN
2 3		NaN NaN	NaN NaN	NaN NaN	NaN NaN
4		NaN	NaN	NaN	NaN
 26995		 NaN	 NaN	 NaN	 NaN
26996		NaN	NaN	NaN	NaN
26997		NaN	NaN	NaN	NaN
26998		NaN	NaN	NaN	NaN
26999	DivEAir	NaN	NaN	NaN iv5TotalGTime	NaN
Div5Lo	ngestGTi	me \			
0		NaN	NaN	NaN	NaN
1		NaN	NaN	NaN	NaN
2		NaN	NaN	NaN	NaN
3		NaN	NaN	NaN	NaN
4		NaN	NaN	NaN	NaN
26995		NaN	NaN	NaN	NaN
26996		NaN	NaN	NaN	NaN
26997		NaN	NaN	NaN	NaN
26998		NaN	NaN	NaN	NaN

26999 NaN NaN NaN NaN

	Div5WheelsOff	Div5TailNum
0	NaN	NaN
1	NaN	NaN
2	NaN	NaN
3	NaN	NaN
4	NaN	NaN
26995	NaN	NaN
26996	NaN	NaN
26997	NaN	NaN
26998	NaN	NaN
26999	NaN	NaN

[27000 rows x 110 columns]

For find the total Rows and columns
df.shape

(27000, 110)

for copy file fone one file to another file
df1 = df.copy()

To display values from Top to Bottom
dfl.head(12)

F1 -		Year	Quarter	Month	DayofMonth	DayOfWeek	
9 0 02	ghtDate \ 1295781	1998	2	4	2	4	1998-04-
1 13	1125375	2013	2	5	13	1	2013-05-
2	118824	1993	3	9	25	6	1993-09-
25 3 12	634825	1994	4	11	12	6	1994-11-
4	1888125	2017	3	8	17	4	2017-08-
17 5	1133769	2014	2	4	2	3	2014-04-
02 6	1465620	2011	3	8	15	1	2011-08-
15 7	172607	2019	1	1	6	7	2019-01-
96 8	1975839	2002	1	1	25	5	2002-01-
25 9	1912303	2005	3	8	1	1	2005-08-

01 10 17 11 05	525812 1963835	2000 2005	4 1 3	12 17 7 5		900 - 12 - 905 - 07 -
	Reporting_Ai A_CODE_Repor			19930 20366 19977 19991 19977 19805 19393 20366 19805 19805 19877 19393		
0 1 2 3 4 5 6 7 8 9 10 11		NaN NaN NaN NaN NaN NaN NaN NaN NaN	Div4TailNum NaN NaN NaN NaN NaN NaN NaN NaN NaN Na	Div5Airport NaN NaN NaN NaN NaN NaN NaN NaN NaN Na	NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN	ime \ NaN

2	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN
5	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN

Div5WheelsOff Div5TailNum 0 NaN NaN 1 NaN NaN 2 NaN NaN 3 NaN NaN 4 NaN NaN 5 NaN NaN 6 NaN NaN 7 NaN NaN 8 NaN NaN 9 NaN NaN 10 NaN NaN 11 NaN NaN

[12 rows x 110 columns]

To check the dtat type

df1.dtypes

Unnamed: 0 int64 Year int64 **Ouarter** int64 Month int64 DayofMonth int64 . . . Div5Wheels0n float64 Div5TotalGTime float64 Div5LongestGTime float64 Div5WheelsOff float64 Div5TailNum float64

Length: 110, dtype: object

display the information of the dataset dfl.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 27000 entries, 0 to 26999

Columns: 110 entries, Unnamed: 0 to Div5TailNum

dtypes: float64(74), int64(19), object(17)

memory usage: 22.7+ MB

#To display dicription of the dataset dfl.describe()

Unnamed: 0	Year	Quarter	Month		
DayofMonth \ count 2.700000e+04	27000.000000	27000.000000	27000.000000		
27000.000000 mean 9.987625e+05	2004.304926	2.505889	6.518667		
15.714370 std 5.767890e+05	9.189835	1.114589	3.444192		
8.816548 min 1.480000e+02 1.000000	1987.000000	1.000000	1.000000		
25% 5.029945e+05 8.000000	1997.000000	2.000000	4.000000		
50% 9.964900e+05 16.000000	2005.000000	3.000000	7.000000		
75% 1.500778e+06 23.000000	2012.000000	3.000000	9.000000		
max 1.999844e+06 31.000000	2020.000000	4.000000	12.000000		
DayOfWeek count 27000.000000 mean 3.939407 std 1.986853 min 1.0000000 25% 2.000000 50% 4.000000 75% 6.000000 max 7.000000	00 27000.000000 07 19923.175222 53 367.912667 00 19386.000000 00 19704.000000 00 19805.000000 00 20355.000000				
Flight_Number OriginAirportSeqID	_Reporting_Air	line OriginAi	rportID		
count 2.700000e+04	27000.00	0000 27000	.000000		
mean 1.274039e+06	1727.94	7963 12740	.372481		
std 1.543894e+05	1658.88	2759 1543	.896695		
min 1.013501e+06	1.00	0000 10135	.000000		
25% 1.129202e+06	528.00	0000 11292	.000000		
50% 1.289201e+06	1173.00	0000 12892	.000000		
75% 1.410001e+06	2229.00	0000 14100	.000000		
max 1.644001e+06	9576.00	0000 16440	.000000		

```
Div4WheelsOff
                              Div4TailNum
                                             Div5Airport
                                                           Div5AirportID
                         0.0
count
                                       0.0
                                                      0.0
                                                                       0.0
                                       NaN
                        NaN
                                                      NaN
                                                                       NaN
mean
        . . .
                        NaN
                                       NaN
                                                      NaN
                                                                       NaN
std
min
                        NaN
                                       NaN
                                                      NaN
                                                                       NaN
                        NaN
                                       NaN
                                                      NaN
                                                                       NaN
25%
50%
                        NaN
                                       NaN
                                                      NaN
                                                                       NaN
        . . .
75%
                        NaN
                                       NaN
                                                                       NaN
                                                      NaN
        . . .
                        NaN
                                                                       NaN
max
                                       NaN
                                                      NaN
        . . .
                            Div5WheelsOn
                                           Div5TotalGTime
        Div5AirportSeqID
Div5LongestGTime
count
                      0.0
                                      0.0
                                                        0.0
0.0
                      NaN
                                      NaN
                                                        NaN
mean
NaN
std
                      NaN
                                                        NaN
                                      NaN
NaN
min
                      NaN
                                      NaN
                                                        NaN
NaN
25%
                      NaN
                                      NaN
                                                        NaN
NaN
50%
                      NaN
                                      NaN
                                                        NaN
NaN
75%
                      NaN
                                                        NaN
                                      NaN
NaN
                      NaN
                                      NaN
                                                        NaN
max
NaN
        Div5WheelsOff
                        Div5TailNum
                   0.0
                                  0.0
count
mean
                   NaN
                                  NaN
                   NaN
                                 NaN
std
                   NaN
                                 NaN
min
25%
                   NaN
                                 NaN
50%
                   NaN
                                  NaN
75%
                   NaN
                                  NaN
                   NaN
                                  NaN
max
[8 rows x 93 columns]
#To check duplicate values
df1.duplicated()
0
          False
1
          False
2
          False
3
          False
4
          False
```

26995 False 26996 False 26997 False 26998 False 26999 False

Length: 27000, dtype: bool

#To find NULL Values

df1.isna()

Unnamed: 0		Year	Quarter	Month	DayofMonth	DayOfWeek	
Flight 0 False	Jate	\ False	False	False	False	False	False
1 False		False	False	False	False	False	False
2 False		False	False	False	False	False	False
3 False		False	False	False	False	False	False
4 False		False	False	False	False	False	False
						• • •	
26995 False		False	False	False	False	False	False
26996 False		False	False	False	False	False	False
26997 False		False	False	False	False	False	False
26998 False		False	False	False	False	False	False
26999 False		False	False	False	False	False	False
0 1 2 3 4	Repo	rting_A	irline False False False False	DOT_ID_R	eportin	g_Airline \ False False False False False	
26995 26996 26997 26998 26999			False False False False False			False False False False False	

IATA_CODE_Reporting_Airline ... Div4WheelsOff Div4TailNum \

```
0
                                 False
                                                         True
                                                                        True
1
                                 False
                                                                        True
                                                         True
2
                                 False
                                                         True
                                                                        True
3
                                 False
                                                         True
                                                                        True
                                         . . .
4
                                 False
                                                         True
                                                                        True
                                         . . .
                                    . . .
                                                           . . .
                                                                         . . .
. . .
26995
                                 False
                                                                        True
                                                         True
                                         . . .
26996
                                 False
                                                         True
                                                                        True
26997
                                 False
                                                         True
                                                                        True
                                         . . .
26998
                                 False
                                                         True
                                                                        True
                                         . . .
26999
                                 False
                                                         True
                                                                        True
                                         . . .
        Div5Airport Div5AirportID
                                        Div5AirportSeqID
                                                            Div5WheelsOn
0
                True
                                 True
                                                      True
                                                                      True
1
                                 True
                True
                                                      True
                                                                      True
2
                True
                                 True
                                                      True
                                                                      True
3
                                 True
                                                      True
                                                                      True
                True
4
                True
                                 True
                                                      True
                                                                      True
                 . . .
                                  . . .
                                                       . . .
                                                                       . . .
. . .
26995
                True
                                 True
                                                      True
                                                                      True
26996
                                 True
                                                      True
                                                                      True
                True
26997
                True
                                 True
                                                      True
                                                                      True
26998
                True
                                 True
                                                      True
                                                                      True
26999
                                 True
                                                      True
                                                                      True
                True
                                               Div5WheelsOff
                                                                Div5TailNum
        Div5TotalGTime
                          Div5LongestGTime
0
                   True
                                        True
                                                         True
                                                                        True
1
                   True
                                        True
                                                                        True
                                                         True
2
                   True
                                        True
                                                         True
                                                                        True
3
                   True
                                        True
                                                         True
                                                                        True
4
                   True
                                        True
                                                         True
                                                                        True
                    . . .
                                         . . .
                                                          . . .
                                                                         . . .
26995
                   True
                                                         True
                                                                        True
                                        True
26996
                   True
                                        True
                                                         True
                                                                        True
26997
                   True
                                        True
                                                         True
                                                                        True
26998
                   True
                                        True
                                                         True
                                                                        True
26999
                   True
                                        True
                                                         True
                                                                        True
[27000 rows x 110 columns]
#calculate the total NULL values in each column
df1.isna().sum()
Unnamed: 0
                           0
                           0
Year
                           0
Ouarter
Month
                           0
DayofMonth
                           0
```

Div5Wheels0n

27000

27000
27000
27000
27000
int64

df1.isna().sum()/df1.shape[0]

Unnamed: 0	0.0
Year	0.0
Quarter	0.0
Month	0.0
DayofMonth	0.0
Div5WheelsOn	1.0
Div5TotalGTime	1.0
Div5LongestGTime	1.0
Div5WheelsOff	1.0
Div5TailNum	1.0
Length: 110, dtype:	float64

df1.dropna(axis = 1 , inplace=True)

df1

	Unnamed: 0	Year	Quarter	Month	DayofMonth	DayOfWeek	
Flight							
0	1295781	1998	2	4	2	4	1998 -
04-02							
1	1125375	2013	2	5	13	1	2013-
05-13							
2	118824	1993	3	9	25	6	1993 -
09-25							
3	634825	1994	4	11	12	6	1994-
11-12							
4	1888125	2017	3	8	17	4	2017-
08-17							
26005	021542	2017	1	1	2.4	2	2017
26995	821542	2017	1	1	24	2	2017-
01-24	1010565	2012	2	c	27	4	2012
26996	1910565	2013	2	6	27	4	2013-
06-27	0055	2016	2	0	26	-	2016
26997	9055	2016	3	8	26	5	2016-
08-26	0.4126	2000	2	0	0	6	2000
26998	84136	2009	3	8	8	6	2009 -
08-08	112000	1000	_	_		-	1000
26999	113029	1993	3	7	17	6	1993 -
07 - 17							

Reporting_Airline DOT_ID_Reporting_Airline

_	ODE_R	eporting_	Airline \		10020	
0 AS			AS		19930	
1			EV		20366	
EV 2			UA		19977	
UA 3			НР		19991	
HP 4			UA		19977	
ÜA						
			• • •		• • •	
26995 DL			DL		19790	
26996			B6		20409	
B6 26997			AA		19805	
AA 26998			YV		20378	
YV 26999			DL		19790	
DL			DL		19790	
		DestWac	CRSDepTime	DepTimeBlk	CRSArrTime	ArrTimeBlk
Cancel 0	led 	\ 93	1330	1300-1359	1426	1400 - 1459
0.0 1		38	1301	1300-1359	1423	1400-1459
0.0						
2 0.0		41	1650	1600-1659	1730	1700 - 1759
3 0.0		81	1245	1200 - 1259	1457	1400-1459
4		82	755	0700-0759	902	0900-0959
0.0						
 26995		63	910	0900-0959	1028	1000 - 1059
0.0 26996		74	1320	1300 - 1359	1612	1600-1659
0.0 26997		22	920	0900-0959	1405	1400-1459
0.0						
26998 0.0		82	1139	1100-1159	1217	1200 - 1259
26999 0.0		74	615	0600-0659	721	0700-0759
0.0						

Diverted Flights Distance DistanceGroup

0	0.0	1.0	224.0	1
1	0.0	1.0	277.0	2
2	0.0	1.0	130.0	1
3	0.0	1.0	370.0	2
4	0.0	1.0	692.0	3
26995	0.0	1.0	297.0	2
26996	0.0	1.0	1562.0	7
26997	0.0	1.0	1521.0	7
26998	0.0	1.0	73.0	1
26999	0.0	1.0	304.0	2

[27000 rows x 32 columns]

data = df1.sample(n=500 , random_state = 40)

data

	Unnamed: 0	Year	Quarter	Month	DayofMonth	DayOfWeek	
Flight	Date \				-	-	
12781	975274	2003	3	7	11	5	2003 -
07-11							
23751	1471411	1998	2	6	7	7	1998 -
06-07							
2796	1469316	2001	4	12	1	6	2001-
12-01							
18633	1468783	1988	2	5	30	1	1988 -
05-30							
23187	1861806	2019	4	12	1	7	2019-
12-01							
 17100	420E2E	100E	2	6	17	6	100E
17188 06-17	428525	1995	2	O	17	6	1995 -
24952	1300641	2018	1	3	26	1	2018-
03-26	1300041	2010	1	3	20	1	2010-
17089	1410356	2015	1	3	16	1	2015-
03-16	1410330	2013	1	2	10	1	2013-
11144	607318	1998	3	9	10	4	1998-
09-10	007510	1990	3	9	10	4	1990-
19762	301716	2005	3	7	10	7	2005 -
07-10	301/10	2003	3	,	10	,	2003-
07-10							

	g_Airline DOT_ID_	Reporting_Airline
IATA_CODE_Repo	rting_Airline \	
12781	WN	19393
WN		
23751	US	20355
US		
2796	UA	19977
UA		
23751 US 2796	0 5	

18633 AL 23187 DL 17188 US 24952 EV 17089 AA 11144 DL 19762 00			US DL US EV AA DL 00				20355 19790 20355 20366 19805 19790 20304		
Cancel 12781 0.0 23751 0.0 2796 0.0 18633 0.0 23187 0.0 17188 0.0 24952 0.0 17089 0.0 11144 0.0 19762 0.0		91 91 85 23 71 74 72 41	CRSD	epTime 2045 2035 2230 630 1530 1820 1155 1122 1855 1205	2 2 0 1 1 1 1	pTimeBlk 000-2059 000-2059 200-2259 600-0659 500-1559 800-1859 100-1159 800-1859 200-1259	CRSAr	rTime 2140 2203 535 723 1757 1926 1328 1325 1922 1229	ArrTimeBlk 2100-2159 2200-2259 0001-0559 0700-0759 1700-1759 1900-1959 1300-1359 1300-1359 1900-1959
12781 23751 2796 18633 23187	Diverted 0.0 0.0 0.0 0.0 0.0		1.0 1.0 1.0 1.0 1.0 1.0	Distan 197 448 2615 297 867	.0	Distance	eGroup 1 2 11 2 4 		

24952	0.0	1.0	374.0	2
17089	0.0	1.0	985.0	4
11144	0.0	1.0	425.0	2
19762	0.0	1.0	296.0	2

[500 rows x 32 columns]

#To find total number of rows and Columns

data.shape

(500, 32)

data.head

	nethod NDFr			Unnamed	d: 0	Year	Quarter	Month
рауотмоп 12781	ith DayOfW 975274	еек Ft. 2003	igntvate 3	7		11	5	2003-
07-11								
23751	1471411	1998	2	6		7	7	1998 -
06-07	1460216	2001	1	10		1	6	2001
2796 12-01	1469316	2001	4	12		1	6	2001-
18633	1468783	1988	2	5		30	1	1988-
05-30	1100705	2500	_	J			_	1500
23187	1861806	2019	4	12		1	7	2019-
12-01								
						• • •		
 17188	428525	1995	2	6		17	6	1995 -
06-17	120323	1333	_	Ü		Τ,	· ·	1333
24952	1300641	2018	1	3		26	1	2018-
03-26								
17089	1410356	2015	1	3		16	1	2015 -
03-16 11144	607318	1998	3	9		10	4	1998-
09-10	007310	1990	3	9		10	4	1990-
19762	301716	2005	3	7		10	7	2005 -
07-10								

	Reporting_Airline DOT_ID	_Reporting_Airline
IATA (CODE Reporting Airline \	_
$1278\overline{1}$	_ · WN	19393
WN		
23751	US	20355
US		
2796	UA	19977
UA		
18633	US	20355
AL		
23187	DL	19790
DL		

17188 US				US				20	355			
24952 EV				EV				20	366			
17089 AA				AA				19	805			
11144 DL				DL				19	790			
19762 00				00				20	304			
Cancal	 1 a d	Dest	Wac	CRSD	epTime	De	pTimeBl	k C	RSArrTime	Ar	rTimeB	lk
Cancel 12781 0.0		\	33		2045	2	000-2059	9	2140	2	2100-21	59
23751 0.0			23		2035	2	000-2059	9	2203	2	2200-22!	59
2796 0.0			91		2230	2	200-2259	9	535	6	0001-05	59
18633 0.0			91		630	0	600-0659	9	723	(700-07!	59
23187 0.0			85		1530	1	500-1559	9	1757]	L700 - 17!	59
								•			•	
17188 0.0			23		1820	1	800 - 1859	9	1926]	1900 - 19	59
24952 0.0			71		1155	1	100-1159	9	1328]	L300 - 13!	59
17089 0.0			74		1122	1	100-1159	9	1325	1	L300 - 13!	59
11144 0.0			72		1855	1	800 - 1859	9	1922	1	1900 - 19	59
19762 0.0			41		1205	1	200-1259	9	1229	1	L200 - 12!	59
	Dive	erted 0.0	Flig	ghts	Distan		Distand	ceGr	1_			
12781 23751		0.0		1.0	197 448	.0			1 2			
2796 18633 23187		0.0 0.0 0.0		1.0 1.0 1.0	2615 297 867	.0			11 2 4			
17188		0.0		1.0	267				2			
24952 17089		0.0		1.0	374 985	.0			2 4			
11144 19762		0.0 0.0		$1.0 \\ 1.0$	425 296				4 2 2			

[500 rows x 32 columns]>

data.tail

		_	: 0 Yea	r Quarter	Month
2003	gntbate 3	7	11	5	2003-
1998	2	6	7	7	1998-
2001	4	12	1	6	2001-
1988	2	5	30	1	1988 -
2019	4	12	1	7	2019-
1005		6			1995 -
		-			
				_	2018-
2015	1	3	16	1	2015 -
1998	3	9	10	4	1998 -
2005	3	7	10	7	2005 -
	Veek Flig 2003 1998 2001 1988 2019 1995 2018 2015 1998	1998 2 2001 4 1988 2 2019 4 1995 2 2018 1 2015 1 1998 3	Week FlightDate 2003 3 7 1998 2 6 2001 4 12 1988 2 5 2019 4 12 1995 2 6 2018 1 3 2015 1 3 1998 3 9	Week FlightDate 2003 3 7 11 1998 2 6 7 2001 4 12 1 1988 2 5 30 2019 4 12 1 1995 2 6 17 2018 1 3 26 2015 1 3 16 1998 3 9 10	Neek FlightDate 2003 3 7 11 5 1998 2 6 7 7 2001 4 12 1 6 1988 2 5 30 1 2019 4 12 1 7 1995 2 6 17 6 2018 1 3 26 1 2015 1 3 16 1 1998 3 9 10 4

TATA	Reporting_Airline DOT_ID_Repo	rting_Airline
$1278\overline{1}$	CODE_Reporting_Airline \ WN	19393
WN 23751	US	20355
US 2796	UA	19977
UA 18633	US	20355
AL 23187 DL	DL	19790
	• • •	
17188	US	20355
US 24952	EV	20366
EV 17089 AA	AA	19805

11144 DL			DL		19790						
19762 00			00				20304				
Cancel		DestWa	c CRSD	epTime	Dep	TimeBlk	CRSA	rrTime	ArrTimeBlk		
12781 0.0	led \ 	3	3	2045	20	00-2059		2140	2100-2159		
23751 0.0		2	3	2035	20	00-2059		2203	2200-2259		
2796 0.0		9	1	2230	22	200-2259		535	0001-0559		
18633 0.0		9	1	630	06	600-0659		723	0700-0759		
23187 0.0		8	5	1530	15	500 - 1559		1757	1700-1759		
			•								
17188 0.0		2	3	1820	18	800 - 1859		1926	1900-1959		
24952 0.0		7	1	1155	11	.00-1159		1328	1300-1359		
17089 0.0		7	4	1122	11	1159		1325	1300-1359		
11144 0.0		7	2	1855	18	800 - 1859		1922	1900-1959		
19762 0.0		4	1	1205	12	200 - 1259		1229	1200-1259		
12781 23751 2796 18633 23187		ted F 0.0 0.0 0.0 0.0 0.0	1.0 1.0 1.0 1.0 1.0	Distan 197 448 2615 297 867	.0 .0 .0	Distanc	eGroup 1 2 11 2 4	- - -			
17188 24952 17089 11144 19762		0.0 0.0 0.0 0.0 0.0	1.0 1.0 1.0 1.0	267 374 985	. 0 . 0 . 0		 2 2 4 2 2	<u>!</u> !			

[500 rows x 32 columns]>

data.count

07-11 23751	1471411	1998	2	6	7	7	1998-
06-07 2796	1469316	2001	4	12	1	6	2001-
12-01 18633	1468783	1988	2	5	30	1	1988-
05-30 23187 12-01	1861806	2019	4	12	1	7	2019-
17188 06-17	428525	1995	2	6	17	6	1995 -
24952 03-26	1300641	2018	1	3	26	1	2018-
17089 03-16	1410356	2015	1	3	16	1	2015-
11144 09-10	607318	1998	3	9	10	4	1998-
19762 07-10	301716	2005	3	7	10	7	2005 -
	Reporting_Ai	rline	DOT TD	Reporti	ng Airline		
IATA_C	ODE_Reportin	g_Airl: WN	ine \		19393		
WN 23751		US			20355		
US 2796		UA			19977		
UA 18633		US			20355		
AL 23187		DL			19790		
DL 							
 17188		US			20355		
US 24952		EV			20366		
EV 17089		AA			19805		
AA 11144		DL			19790		
DL 19762		00			20304		
00							

... DestWac CRSDepTime DepTimeBlk CRSArrTime ArrTimeBlk Cancelled \
12781 ... 33 2045 2000-2059 2140 2100-2159

23751 23	0 0														
2796 91				23		2	035	20	00-2	2059		2203	3	2200-2	259
18633 91 630 0600-0659 723 0700-0759 0.0	2796			91		2	230	22	00-2	2259		535	5	0001-0	559
23187 85	18633			91			630	06	00-6	9659		723	}	0700-0	759
17188 23	23187			85		1	530	15	00-1	1559		1757	7	1700-1	759
17188 23													ı		
24952 71	17188			23		1	820	18	00-1	1859		1926	6	1900 - 1	959
17089 74	24952			71		1	155	11	00-1	1159		1328	3	1300-1	359
11144 72	17089			74		1	122	11	00-1	1159		1325	5	1300-1	359
19762	11144			72		1	855	18	00-1	1859		1922	<u>)</u>	1900 - 1	959
Diverted Flights Distance DistanceGroup	19762			41		1	205	12	00-1	1259		1229)	1200 - 1	259
12781 2003 3 7 11 5 2003-07-11 23751 1998 2 6 7 7 1998-06-07 2796 2001 4 12 1 6 2001-12-01 18633 1988 2 5 30 1 1988-05-30 23187 2019 4 12 1 7 2019-12-01 17188 1995 2 6 17 6 1995-06-17 24952 2018 1 3 26 1 2018-03-26	12781 23751 2796 18633 23187 17188 24952 17089 11144 19762 [500 red	OWS X	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	olumr	1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0		197. 448. 2615. 297. 867. 267. 374. 985. 425. 296.	0 0 0 0 0 0 0 0 0 0 0 0			 2 11 2 4 2 2 4				
	12781 23751 2796 18633 23187 17188 24952	2003 1998 2001 1988 2019 1995 2018	Qua	3 2 4 2 4 2 1		7 6 12 5 12 6 3	Dayo	fMo	11 7 1 30 1 17 26	Day	5 7 6 1 7	206 199 206 198 201 199 201)3 - ())8 - ())1 - ())8 - () 19 - () 18 - ()	07 - 11 06 - 07 12 - 01 05 - 30 12 - 01 06 - 17 03 - 26	\

11144 19762	1998 2005	3 3	9 7	10 10		4 7	1998-09-10 2005-07-10	
Reporting_Airline DOT_ID_Reporting_Airline IATA CODE Reporting Airline \								
12781 WN	орс_керо	WN	rne (19393		
23751 US		US				20355		
2796 UA		UA				19977		
18633 AL		US				20355		
23187 DL		DL				19790		
17188 US		US				20355		
24952 EV		EV				20366		
17089 AA		AA				19805		
11144 DL		DL				19790		
19762 00		00				20304		
D T '		Number_Repo	orting_	Airline		DestWa	c CRSDepTim	e
12781 2000-2	eBlk \			2613		3:	3 204	5
2000-2 23751 2000-2				563		2:	3 203	5
2796 2200-2				30		9	1 223	0
18633 0600-0				1546		9	1 63	0
23187 1500-1				2793		8	5 153	0
	333							
17188 1800 1	050			470		2:	3 182	0
1800-1 24952 1100-1				4388		7	1 115	5
17089				2371		7.	4 112	2
1100-1 11144 1800-1				351		72	2 185	5

```
1205
19762
                                    6982 ...
                                                     41
1200 - 1259
      CRSArrTime ArrTimeBlk
                               Cancelled
                                          Diverted
                                                     Flights
                                                               Distance \
12781
            2140
                   2100-2159
                                     0.0
                                                0.0
                                                          1.0
                                                                  197.0
23751
            2203
                   2200-2259
                                     0.0
                                                0.0
                                                          1.0
                                                                  448.0
2796
             535
                   0001-0559
                                     0.0
                                                0.0
                                                          1.0
                                                                 2615.0
             723
                   0700-0759
                                                0.0
18633
                                     0.0
                                                          1.0
                                                                  297.0
23187
            1757
                   1700 - 1759
                                     0.0
                                                0.0
                                                          1.0
                                                                  867.0
                                                          . . .
. . .
              . . .
                          . . .
                                     . . .
                                                . . .
                                                                     . . .
            1926
                   1900 - 1959
17188
                                     0.0
                                                0.0
                                                          1.0
                                                                  267.0
            1328
24952
                   1300 - 1359
                                     0.0
                                                0.0
                                                          1.0
                                                                  374.0
17089
            1325
                   1300 - 1359
                                     0.0
                                                0.0
                                                          1.0
                                                                  985.0
            1922
11144
                   1900 - 1959
                                     0.0
                                                0.0
                                                          1.0
                                                                  425.0
19762
            1229
                   1200 - 1259
                                     0.0
                                                0.0
                                                          1.0
                                                                  296.0
      DistanceGroup
12781
                   1
                   2
23751
2796
                  11
                   2
18633
                   4
23187
17188
                   2
                   2
24952
                   4
17089
11144
                   2
19762
                   2
[500 rows x 31 columns]
data.shape
(500, 31)
#To visualize Scatter plot for Distance vs Departure time
fig = go.Figure(data = go.Scatter(x=df['Distance'],y =
df['DepTime'],mode='markers', marker=dict(color='red')))
fig.update layout(title = 'Scatter Plot for Distance vs Departure
time', xaxis title = 'Distance Time', yaxis title = 'Departure Time')
fig.show()
line data = df.groupby('Month')['ArrDelay'].mean().reset index()
line data
    Month ArrDelay
0
        1
           7.097321
1
        2
           6.530025
2
        3
           5.429533
3
        4 4.040055
4
        5
           5.874889
```

```
5
           9.347515
        6
6
        7
           8.851436
7
        8
           7.402630
8
        9
           2.198975
9
       10
           3.242506
10
       11
           4.228961
11
       12
           9.474149
#To display Line plot of Monthly average delay plot for an year
fig = go.Figure(data = go.Scatter(x=line_data['Month'],y =
line_data['ArrDelay'],mode='lines', marker=dict(color='green')))
fig.update_layout(title = 'Monthly average delay plot for an year',
xaxis title = 'Months', yaxis title = 'Delay of flights (in minutes)')
fig.show()
bar data = df.groupby('DestState')['Flights'].sum().reset index()
bar data
   DestState
               Flights
0
          ΑK
                 192.0
1
          ΑL
                 163.0
2
          AR
                 118.0
3
          ΑZ
                 842.0
4
          CA
                3174.0
5
          C0
                 970.0
6
          \mathsf{CT}
                 122.0
7
           FL
                1804.0
8
          GΑ
                1553.0
9
          ΗI
                 289.0
10
           IΑ
                  84.0
                  72.0
11
           ID
12
           ΙL
                1766.0
13
           IN
                 192.0
14
          KS
                  44.0
15
          ΚY
                 463.0
16
          LA
                 265.0
17
          MA
                 484.0
18
          MD
                 421.0
19
          ME
                  34.0
20
          ΜI
                 768.0
21
          MN
                 591.0
22
          MO
                 730.0
23
          MS
                  67.0
24
                  79.0
          ΜT
25
          NC
                 907.0
26
          ND
                  53.0
27
          NE
                 106.0
28
          NH
                  34.0
29
          NJ
                 539.0
                 142.0
30
          NM
31
          NV
                 704.0
```

```
1233.0
32
          NY
33
                 538.0
          0H
34
          0K
                 203.0
35
          0R
                 298.0
36
          PA
                 834.0
37
          PR
                  92.0
38
          RΙ
                  71.0
39
          SC
                 118.0
40
          SD
                  34.0
41
          TN
                 612.0
42
          TT
                   3.0
43
                3023.0
          TX
44
          UT
                 439.0
45
                 882.0
          VA
46
                  16.0
          VI
47
          VT
                  23.0
                 545.0
48
          WA
49
          WΙ
                 210.0
50
          WV
                  20.0
51
          WY
                  27.0
#To dispaly Bar chart of Total number of flights from each destination
states
fig = pr.bar(bar data,x='DestState',y='Flights', title = 'Total number
of flights from each destination states')
fig.show()
bub data = df.groupby('Reporting Airline')
['Flights'].sum().reset index()
bub_data
   Reporting_Airline
                       Flights
0
                          262.0
                   9E
1
                   AA
                        3151.0
2
                   AS
                          710.0
3
                   B6
                         499.0
4
                   C0
                         1208.0
5
                   DH
                           85.0
6
                   DL
                        3530.0
7
                   EΑ
                          120.0
8
                   ΕV
                          895.0
9
                   F9
                          154.0
10
                   FL
                          344.0
11
                   G4
                           36.0
12
                   HA
                          162.0
13
                   HP
                          503.0
14
                   KΗ
                           19.0
15
               ML (1)
                            9.0
16
                   MQ
                         1096.0
17
                   NK
                          121.0
18
                   NW
                         1517.0
```

```
19
                   0H
                         321.0
20
                   00
                        1518.0
21
               PA (1)
                          41.0
22
                   PΙ
                         119.0
23
                   PS
                           9.0
24
                   TW
                         516.0
25
                   ΤZ
                           34.0
26
                   UA
                        2570.0
27
                   US
                        2342.0
28
                   VX
                          59.0
29
                        4186.0
                   WN
30
                   XΕ
                         469.0
31
                   Y۷
                         289.0
32
                   YΧ
                         106.0
```

#To dispaly Bubble chart of Total number of flights per reporting airlines

```
fig = pr.scatter(bub_data,x='Reporting_Airline',y='Flights', title =
'Total number of flights per reporting airlines',size =
'Flights',hover_name='Reporting_Airline',size_max=60)
fig.show()
```