In [28]:

```
import pandas as pd
import numpy as np
```

In [29]:

```
df = pd.DataFrame({'From_To':['LoNDon_paris','MAdrid_miLAN',
                              'londON_StockhOlm',
                              'Budapest_PaRis','Brussels_londOn'],
                   'FlightNumber':[10045,np.nan,10065,np.nan,10085],
                   'RecentDelays':[[23,47],[],[24,43,87],[13],[67,32]],
                   'Airline':['KLM(!)','<Air France>(12)','(British Airways.)',
                            '12.Air France',"'Swiss Air'"]})
```

In [30]:

```
df
```

Out[30]:

	From_To	FlightNumber	RecentDelays	Airline
0	LoNDon_paris	10045.0	[23, 47]	KLM(!)
1	MAdrid_miLAN	NaN	[]	<Air France>(12)
2	londON_StockhOlm	10065.0	[24, 43, 87]	(British Airways.)
3	Budapest_PaRis	NaN	[13]	12.Air France
4	Brussels_londOn	10085.0	[67, 32]	'Swiss Air'

In [31]:

```
df['FlightNumber'].isnull().sum()
```

Out[31]:

2

In [32]:

```
new_index =np.arange(1,df.From_To.count()+1)
new_index
df.set_index(new_index,inplace=True)
df
```

Out[32]:

	From_To	FlightNumber	RecentDelays	Airline
1	LoNDon_paris	10045.0	[23, 47]	KLM(!)
2	MAdrid_miLAN	NaN	[]	<Air France>(12)
3	londON_StockhOlm	10065.0	[24, 43, 87]	(British Airways.)
4	Budapest_PaRis	NaN	[13]	12.Air France
5	Brussels_londOn	10085.0	[67, 32]	'Swiss Air'

In [33]:

```
for i in np.arange(1,df.From_To.count()+1):
    if pd.isnull(df.FlightNumber.loc[i,]):
        df.loc[i,'FlightNumber'] = df.FlightNumber.loc[i-1] +10
df['FlightNumber']
df
```

Out[33]:

	From_To	FlightNumber	RecentDelays	Airline
1	LoNDon_paris	10045.0	[23, 47]	KLM(!)
2	MAdrid_miLAN	10055.0	[]	<Air France>(12)
3	londON_StockhOlm	10065.0	[24, 43, 87]	(British Airways.)
4	Budapest_PaRis	10075.0	[13]	12.Air France
5	Brussels_londOn	10085.0	[67, 32]	'Swiss Air'

In [34]:

```
df['FlightNumber'].astype(int)
```

Out[34]:

```
1    10045
2    10055
3    10065
4    10075
5    10085
Name: FlightNumber, dtype: int32
```

In [35]:

```
df['From_To']
```

Out[35]:

```
1        LoNDon_paris
2        MAdrid_miLAN
3    londON_StockhOlm
4      Budapest_PaRis
5     Brussels_londOn
Name: From_To, dtype: object
```

In [36]:

```
df1 = df.copy()
```

In [37]:

```
df1[['From','To']] = df1.From_To.str.split("_",expand = True)
df1
```

Out[37]:

	From_To	FlightNumber	RecentDelays	Airline	From	To
1	LoNDon_paris	10045.0	[23, 47]	KLM(!)	LoNDon	paris
2	MAdrid_miLAN	10055.0	[]	<Air France>(12)	MAdrid	miLAN
3	londON_StockhOlm	10065.0	[24, 43, 87]	(British Airways.)	londON	StockhOlm
4	Budapest_PaRis	10075.0	[13]	12.Air France	Budapest	PaRis
5	Brussels_londOn	10085.0	[67, 32]	'Swiss Air'	Brussels	londOn

Notice how the Capitalisation of the city names is all mixed up in this temporary DataFrame .Standarise the strings so that only the first letter is uppercase(eg "londOn" should become "London")

In [38]:

```
df1.From = df1.From.str.capitalize()
df1.To= df1.To.str.capitalize()
df1.head()
```

Out[38]:

	From_To	FlightNumber	RecentDelays	Airline	From	To
1	LoNDon_paris	10045.0	[23, 47]	KLM(!)	London	Paris
2	MAdrid_miLAN	10055.0	[]	<Air France>(12)	Madrid	Milan
3	londON_StockhOlm	10065.0	[24, 43, 87]	(British Airways.)	London	Stockholm
4	Budapest_PaRis	10075.0	[13]	12.Air France	Budapest	Paris
5	Brussels_londOn	10085.0	[67, 32]	'Swiss Air'	Brussels	London

In [39]:

```
df
```

Out[39]:

	From_To	FlightNumber	RecentDelays	Airline
1	LoNDon_paris	10045.0	[23, 47]	KLM(!)
2	MAdrid_miLAN	10055.0	[]	<Air France>(12)
3	londON_StockhOlm	10065.0	[24, 43, 87]	(British Airways.)
4	Budapest_PaRis	10075.0	[13]	12.Air France
5	Brussels_londOn	10085.0	[67, 32]	'Swiss Air'

In [40]:

```
df.drop('From_To',axis=1,inplace = True)
df
```

Out[40]:

	FlightNumber	RecentDelays	Airline
1	10045.0	[23, 47]	KLM(!)
2	10055.0	[]	<Air France>(12)
3	10065.0	[24, 43, 87]	(British Airways.)
4	10075.0	[13]	12.Air France
5	10085.0	[67, 32]	'Swiss Air'

In [41]:

```
df['From'] = df1['From']
df['To'] = df1['To']
```

In [42]:

```
df
```

Out[42]:

	FlightNumber	RecentDelays	Airline	From	To
1	10045.0	[23, 47]	KLM(!)	London	Paris
2	10055.0	[]	<Air France>(12)	Madrid	Milan
3	10065.0	[24, 43, 87]	(British Airways.)	London	Stockholm
4	10075.0	[13]	12.Air France	Budapest	Paris
5	10085.0	[67, 32]	'Swiss Air'	Brussels	London

In [44]:

```
```

In [45]:

```
```

Out[45]:

```
pandas.core.series.Series
```

In [51]:

```
df
```

Out[51]:

	FlightNumber	RecentDelays	Airline	From	To
1	10045.0	[23, 47]	KLM(!)	London	Paris
2	10055.0	[]	<Air France>(12)	Madrid	Milan
3	10065.0	[24, 43, 87]	(British Airways.)	London	Stockholm
4	10075.0	[13]	12.Air France	Budapest	Paris
5	10085.0	[67, 32]	'Swiss Air'	Brussels	London

In [52]:

```
df[['Delay1','Delay2','Delay3']] = pd.DataFrame(df.RecentDelays.tolist(),index =df.index)
df
```

Out[52]:

	FlightNumber	RecentDelays	Airline	From	To	Delay1	Delay2	Delay3
1	10045.0	[23, 47]	KLM(!)	London	Paris	23.0	47.0	NaN
2	10055.0	[]	<Air France>(12)	Madrid	Milan	NaN	NaN	NaN
3	10065.0	[24, 43, 87]	(British Airways.)	London	Stockholm	24.0	43.0	87.0
4	10075.0	[13]	12.Air France	Budapest	Paris	13.0	NaN	NaN
5	10085.0	[67, 32]	'Swiss Air'	Brussels	London	67.0	32.0	NaN

In [53]:

```
df.drop('RecentDelays',axis=1,inplace=True)
```

In [54]:

```
df
```

Out[54]:

	FlightNumber	Airline	From	To	Delay1	Delay2	Delay3
1	10045.0	KLM(!)	London	Paris	23.0	47.0	NaN
2	10055.0	<Air France>(12)	Madrid	Milan	NaN	NaN	NaN
3	10065.0	(British Airways.)	London	Stockholm	24.0	43.0	87.0
4	10075.0	12.Air France	Budapest	Paris	13.0	NaN	NaN
5	10085.0	'Swiss Air'	Brussels	London	67.0	32.0	NaN

In [ ]:

```

```