In [1]:

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

In [2]:

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

1. Some values in the the FlightNumber column are missing. These numbers are meant to increase by 10 with each row so 10055 and 10075 need to be put in place. Fill in these missing numbers and make the column an integer column (instead of a float column).

In [3]:

```
df.isnull().sum()
```

Out[3]:

```
Airline         0
FlightNumber    2
From_To         0
RecentDelays    0
dtype: int64
```

In [4]:

```
df["FlightNumber"].iloc[1] = 10055
df["FlightNumber"].iloc[3] = 10075
```

```
C:\ProgramData\Anaconda3\lib\site-packages\pandas\core\indexing.py:194: SettingWithCopyWarnin
g:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.ht
ml#indexing-view-versus-copy
  self._setitem_with_indexer(indexer, value)
```

In [5]:

```
df["FlightNumber"]
```

Out[5]:

```
0    10045.0
1    10055.0
2    10065.0
3    10075.0
4    10085.0
Name: FlightNumber, dtype: float64
```

In [6]:

```
df["FlightNumber"].dtype
```

Out[6]:

```
dtype('float64')
```

In [7]:

```
df["FlightNumber"] = df["FlightNumber"].astype(int)
```

In [8]:

```
df["FlightNumber"].dtype
```

Out[8]:

```
dtype('int32')
```

2.The FromTo column would be better as two separate columns! Split each string on the underscore delimiter to give a new temporary DataFrame with the correct values. Assign the correct column names to this temporary DataFrame.

In [9]:

```
dftemp = df['From_To'].str.split('_',expand=True)
```

In [10]:

```
dftemp.columns = ['From', 'To']
dftemp
```

Out[10]:

	From	To
0	LoNDon	paris
1	MAdrid	miLAN
2	londON	StockhOlm
3	Budapest	PaRis
4	Brussels	londOn

3.Notice how the capitalisation of the city names is all mixed up in this temporary DataFrame. Standardise the strings so that only the first letter is uppercase (e.g. "londON" should become "London".)

In [11]:

```
dftemp['From'] = dftemp['From'].str.capitalize()
dftemp['To'] = dftemp['To'].str.capitalize()
```

In [12]:

```
dftemp
```

Out[12]:

	From	To
0	London	Paris
1	Madrid	Milan
2	London	Stockholm
3	Budapest	Paris
4	Brussels	London

4.Delete the From_To column from df and attach the temporary DataFrame from the previous questions.

In [13]:

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

In [14]:

```
df
```

Out[14]:

	Airline	FlightNumber	RecentDelays
0	KLM(!)	10045	[23, 47]
1	<Air France> (12)	10055	[]
2	(British Airways.)	10065	[24, 43, 87]
3	12. Air France	10075	[13]
4	"Swiss Air"	10085	[67, 32]

In [15]:

```
df_cat1 = pd.concat([df,dftemp], axis=1)
df_cat1
```

Out[15]:

	Airline	FlightNumber	RecentDelays	From	To
0	KLM(!)	10045	[23, 47]	London	Paris
1	<Air France> (12)	10055	[]	Madrid	Milan
2	(British Airways.)	10065	[24, 43, 87]	London	Stockholm
3	12. Air France	10075	[13]	Budapest	Paris
4	"Swiss Air"	10085	[67, 32]	Brussels	London

4.In the RecentDelays column, the values have been entered into the DataFrame as a list. We would like each first value in its own column, each second value in its own column, and so on. If there isn't an Nth value, the value should be NaN.

In [19]:

```
delays = pd.DataFrame(df_cat1['RecentDelays'].tolist())
```

In [20]:

```
delays.columns = ['delay_1', 'delay_2', 'delay_3']
delays
```

Out[20]:

	delay_1	delay_2	delay_3
0	23.0	47.0	NaN
1	NaN	NaN	NaN
2	24.0	43.0	87.0
3	13.0	NaN	NaN
4	67.0	32.0	NaN

In [21]:

```
del df_cat1['RecentDelays']
```

In [22]:

```
df2 = pd.concat([df_cat1,delays], axis=1)
df2
```

Out[22]:

	Airline	FlightNumber	From	To	delay_1	delay_2	delay_3
0	KLM(!)	10045	London	Paris	23.0	47.0	NaN
1	<Air France> (12)	10055	Madrid	Milan	NaN	NaN	NaN
2	(British Airways.)	10065	London	Stockholm	24.0	43.0	87.0
3	12. Air France	10075	Budapest	Paris	13.0	NaN	NaN
4	"Swiss Air"	10085	Brussels	London	67.0	32.0	NaN