Kamaleshs-MacBook-Pro.local::/Users/kamalesh\_das/Desktop/Python/AcadGild$python DataClean\_Assignment\_12\_1.py

Original DF:

Airline FlightNumber From\_To RecentDelays

0 KLM(!) 10045.0 LoNDon\_paris [23, 47]

1 <Air France> (12) NaN MAdrid\_miLAN []

2 (British Airways. ) 10065.0 londON\_StockhOlm [24, 43, 87]

3 12. Air France NaN Budapest\_PaRis [13]

4 "Swiss Air" 10085.0 Brussels\_londOn [67, 32]

After filling missing value of FlightNumber:

Airline FlightNumber From\_To RecentDelays

0 KLM(!) 10045 LoNDon\_paris [23, 47]

1 <Air France> (12) 10055 MAdrid\_miLAN []

2 (British Airways. ) 10065 londON\_StockhOlm [24, 43, 87]

3 12. Air France 10075 Budapest\_PaRis [13]

4 "Swiss Air" 10085 Brussels\_londOn [67, 32]

Temp DataFrame after splitting From\_To column:

From To

0 LoNDon paris

1 MAdrid miLAN

2 londON StockhOlm

3 Budapest PaRis

4 Brussels londOn

Temp DataFrame after Standardising the city names:

From To

0 London Paris

1 Madrid Milan

2 London Stockholm

3 Budapest Paris

4 Brussels London

Deleted From\_To column from df and added temp in df:

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

New DataFrame after spliting RecentDelays:

Airline FlightNumber From To delays\_1 delays\_2 delays\_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