

5. Otras herramientas de transformación

June 21, 2022

1 Transformación de datos

1.0.1 Otras herramientas de transformación

Carga de librerías

```
[2]: import pandas as pd
import os
```

Importado de datos

```
[4]: path = r'C:/Users/Carlos/OneDrive/Formación/Python/Transformación de datos'
file = 'nycflights.csv'
df = pd.read_csv(os.path.join(path,file), sep = ',')
```

```
[5]: df.head()
```

```
[5]:   year  month  day  dep_time  dep_delay  arr_time  arr_delay  carrier  tailnum \
0  2013     6   30     940         15     1216         -4         VX   N626VA
1  2013     5    7    1657         -3     2104          10         DL   N3760C
2  2013    12    8     859         -1     1238          11         DL   N712TW
3  2013     5   14    1841         -4     2122         -34         DL   N914DL
4  2013     7   21    1102         -3     1230          -8         9E   N823AY
```

```
   flight  origin  dest  air_time  distance  hour  minute
0     407    JFK   LAX       313     2475     9      40
1     329    JFK   SJU       216     1598    16      57
2     422    JFK   LAX       376     2475     8      59
3    2391    JFK   TPA       135     1005    18      41
4    3652    LGA   ORF        50       296    11        2
```

Reemplazar datos dentro de una dataframe Reemplazo del texto JFK en la columna 'origin' por ABC

```
[8]: # Verificación del estado inicial
df['origin'].unique().tolist()
```

```
[8]: ['JFK', 'LGA', 'EWR']
```

```
[9]: df1 = df.replace({'origin':{'JFK':'ABC'}})
df1['origin'].unique().tolist()
```

```
[9]: ['ABC', 'LGA', 'EWR']
```

Reemplazo de varios valores en diferentes columnas

```
[27]: # Verificación del estado inicial
x1 = df['origin'].unique().tolist()
x2 = df['carrier'].unique().tolist()
x3 = df['month'].unique().tolist()
print(f' origin:  {x1} \n carrier: {x2} \n month:  {x3}')
```

origin: ['JFK', 'LGA', 'EWR']
carrier: ['VX', 'DL', '9E', 'AA', 'WN', 'B6', 'EV', 'FL', 'UA', 'US', 'MQ', 'F9', 'YV', 'AS', 'HA', 'OO']
month: [6, 5, 12, 7, 1, 8, 9, 4, 11, 3, 10, 2]

```
[29]: # Reemplazo
detail = {"origin"    : {'JFK': 'ABC', 'LGA': 'DEF'},
          "carrier"   : {'VX': 'New value', 'DL':1, '9E':'Nine'},
          "month"     : {6: -6, 12: -12}}
```

df1 = df.replace(detail)

```
x1 = df1['origin'].unique().tolist()
x2 = df1['carrier'].unique().tolist()
x3 = df1['month'].unique().tolist()
print(f' origin:  {x1} \n carrier: {x2} \n month:  {x3}')
```

origin: ['ABC', 'DEF', 'EWR']
carrier: ['New value', 1, 'Nine', 'AA', 'WN', 'B6', 'EV', 'FL', 'UA', 'US', 'MQ', 'F9', 'YV', 'AS', 'HA', 'OO']
month: [-6, 5, -12, 7, 1, 8, 9, 4, 11, 3, 10, 2]

```
[ ]:
```

```
[ ]:
```