

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

a) Programación funcional

```
In [ ]: ruta = 'data/ecobici/'
viajes = [x for x in os.listdir(ruta)]
viajes.remove('estaciones-de-ecobici.csv')
viajes
```

```
Out[ ]: ['2018-10.csv',
'2019-05.csv',
'2018-07.csv',
'2019-01.csv',
'2017-12.csv',
'2018-05.csv',
'2018-12.csv',
'2018-11.csv',
'2018-09.csv',
'2019-06.csv',
'2019-10.csv',
'2019-09.csv',
'2019-07.csv',
'2017-11.csv',
'2018-03.csv',
'2018-08.csv',
'2018-02.csv',
'2019-04.csv',
'2018-06.csv',
'2019-03.csv',
'2018-04.csv',
'2019-02.csv',
'2019-08.csv',
'2018-01.csv']
```

```
In [ ]: merged_data = pd.concat([pd.read_csv(ruta + x) for x in viajes])
```

```

/tmp/ipykernel_5017/4103827754.py:1: DtypeWarning: Columns (2) have mixed type
s. Specify dtype option on import or set low_memory=False.
merged_data = pd.concat([pd.read_csv(ruta + x) for x in viajes])
/tmp/ipykernel_5017/4103827754.py:1: DtypeWarning: Columns (2) have mixed type
s. Specify dtype option on import or set low_memory=False.
merged_data = pd.concat([pd.read_csv(ruta + x) for x in viajes])
/tmp/ipykernel_5017/4103827754.py:1: DtypeWarning: Columns (2) have mixed type
s. Specify dtype option on import or set low_memory=False.
merged_data = pd.concat([pd.read_csv(ruta + x) for x in viajes])
/tmp/ipykernel_5017/4103827754.py:1: DtypeWarning: Columns (2) have mixed type
s. Specify dtype option on import or set low_memory=False.
merged_data = pd.concat([pd.read_csv(ruta + x) for x in viajes])
/tmp/ipykernel_5017/4103827754.py:1: DtypeWarning: Columns (2) have mixed type
s. Specify dtype option on import or set low_memory=False.
merged_data = pd.concat([pd.read_csv(ruta + x) for x in viajes])
/tmp/ipykernel_5017/4103827754.py:1: DtypeWarning: Columns (2) have mixed type
s. Specify dtype option on import or set low_memory=False.
merged_data = pd.concat([pd.read_csv(ruta + x) for x in viajes])
/tmp/ipykernel_5017/4103827754.py:1: DtypeWarning: Columns (2,6,9) have mixed
types. Specify dtype option on import or set low_memory=False.
merged_data = pd.concat([pd.read_csv(ruta + x) for x in viajes])
/tmp/ipykernel_5017/4103827754.py:1: DtypeWarning: Columns (2) have mixed type
s. Specify dtype option on import or set low_memory=False.
merged_data = pd.concat([pd.read_csv(ruta + x) for x in viajes])
/tmp/ipykernel_5017/4103827754.py:1: DtypeWarning: Columns (2) have mixed type
s. Specify dtype option on import or set low_memory=False.
merged_data = pd.concat([pd.read_csv(ruta + x) for x in viajes])
/tmp/ipykernel_5017/4103827754.py:1: DtypeWarning: Columns (2) have mixed type
s. Specify dtype option on import or set low_memory=False.
merged_data = pd.concat([pd.read_csv(ruta + x) for x in viajes])

```

b) Propuesta para la limpieza de los datos

In []: merged_data.head()

Out[]:

	Genero_Usuario	Edad_Usuario	Bici	Ciclo_Estacion_Retiro	Fecha_Retiro	Hora_Retiro	Ciclo_f
0	M	58	2176	299	01/10/2018	0:00:02	
1	F	42	9816	43	01/10/2018	0:00:23	
2	M	30	11219	37	01/10/2018	0:00:34	
3	M	26	3910	123	01/10/2018	0:01:36	
4	M	29	11487	41	01/10/2018	0:02:39	

In []: merged_data.info()

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 17104573 entries, 0 to 686291
Data columns (total 10 columns):
#   Column                Dtype
---  -
0   Genero_Usuario        object
1   Edad_Usuario          int64
2   Bici                  object
3   Ciclo_Estacion_Retiro int64
4   Fecha_Retiro          object
5   Hora_Retiro           object
6   Ciclo_Estacion_Arribo object
7   Fecha_Arribo          object
8   Hora_Arribo           object
9   Unnamed: 9            object
dtypes: int64(2), object(8)
memory usage: 1.4+ GB
```

- Genero_Usuario: Confirmación que los únicos valores sean 'M' y 'F'.
- Bici: Pasar todos los salores a string para no eliminar las letras que se tienen.
- Fecha_Retiro: Pasar los datos a formato datetime junto con la Hora_Retiro
- Fecha_Arribo: Pasar los datos a formato datetime junto con la Hora_Arribo