

Taller Group by

1. Cargar el CSV phone_data y convertir la columna date de string a tiempo

```
import pandas as pd
import dateutil

data = read_csv('phone_data.csv')
data['date'] = data['date'].apply(dateutil.parser.parse, dayfirst=True)
```

index	date	duration	item	month	network	network_type
0	15/10/14 06:58	34.429	data	2014-11	data	data
1	15/10/14 06:58	13.000	call	2014-11	Vodafone	mobile
2	15/10/14 14:46	23.000	call	2014-11	Meteor	mobile
3	15/10/14 14:48	4.000	call	2014-11	Tesco	mobile
4	15/10/14 17:27	4.000	call	2014-11	Tesco	mobile
5	15/10/14 18:55	4.000	call	2014-11	Tesco	mobile
6	16/10/14 06:58	34.429	data	2014-11	data	data

2. Cuál fue el ítem (llamada/datos) de mayor duración?

```
data['duration'].max()
Out[39]: 10528.0
```

Cuál fue la llamada de mayor duración?

Cuál fue el evento de datos de mayor duración?

3. Cuánto fue el total de segundos de todas las llamadas?

```
data['duration'][data['item'] == 'call'].sum()
Out[40]: 92321.0
```

Cuando se hizo esa llamada?

4. Cuantas entradas hay por cada uno de los meses

```
data['month'].value_counts()
Out[41]:
2014-11    230
2015-01    205
2014-12    157
2015-02    137
2015-03    101
dtype: int64
```

Cuantas entradas de datos hay por cada uno de las redes?

Cuantas entradas de llamada hay por cada uno de las redes?

5. Obtener la suma de la duración por mes

```
data.groupby('month')['duration'].sum()
Out[70]:
month
2014-11    26639.441
2014-12    14641.870
2015-01    18223.299
2015-02    15522.299
2015-03    22750.441
```

Obtener el promedio de llamadas por cada una de las redes

Obtener el promedio de eventos de datos por cada una de las redes

6. Obtener el número de entradas por mes

```
data.groupby('month')['date'].count()
Out[74]:
month
2014-11    230
2014-12    157
2015-01    205
2015-02    137
2015-03    101
```

Obtener el número de entradas tipo llamada por cada una de las redes

Obtener el número de entradas tipo sms por cada una de las redes

7. Cuántos eventos de llamada, datos y sms hubo por mes

```
data.groupby(['month', 'item'])['date'].count()
```

```
Out[76]:
```

month	item	
2014-11	call	107
	data	29
	sms	94
2014-12	call	79
	data	30
	sms	48
2015-01	call	88
	data	31
	sms	86
2015-02	call	67
	data	31
	sms	39
2015-03	call	47
	data	29
	sms	25

Por cada uno de los meses cuántas veces se usaron cada una de las redes (para cualquier evento)

Por cada uno de los meses cuántas veces se usaron cada una de las redes discriminando el tipo de evento