

# Analisis

February 2, 2017

## 1 Encuesta sobre lenguajes de programación y género

### 2 # Introducción

una breve introducción de la idea con link a la encuesta. También cargamos los módulos que usaremos

```
In [34]: # -*- coding: utf-8 -*-
import requests
import io
import csv
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.cm as cm
import plotly.plotly as py
import cufflinks as cf
cf.set_config_file(offline=False, world_readable=True, theme='ggplot')
%matplotlib inline
```

### 3 # Descripción de los datos obtenidos

Aquí hablamos de cada uno de los campos de la encuesta y para cada uno mostraremos un gráfico con los obtenidos. Primero cargamos los datos de la encuesta

```
In [40]: headers={}
headers["User-Agent"] = "Mozilla/5.0 (Windows NT 6.2; WOW64; rv:22.0) Gecko"
headers["DNT"] = "1"
headers["Accept"] = "text/html,application/xhtml+xml,application/xml;q=0.9"
headers["Accept-Encoding"] = "deflate"
headers["Accept-Language"] = "es-ES,en;q=0.5"
headers["Content-Type"] = "application/x-www-form-urlencoded; charset=UTF-8"
lines = []

file_id="1xFpri9AF6N23vo6J5UUA_d5rY_oeVwpQ-EUBani007M"
url = "https://docs.google.com/spreadsheets/d/{0}/export?format=csv".format
```

```

r = requests.get(url)
r.encoding = 'utf-8'

data = {}
cols = []
genero_lenguaje = {}
residencia_lenguaje = {}
lenguaje = {}
sio = io.StringIO( r.text, newline=None)
reader = csv.reader(sio, dialect=csv.excel)
rownum = 0

for row in reader:
    if rownum == 0:
        for col in row:
            col=col.upper();
            col=col.strip();
            data[col] = ''
            cols.append(col)

    else:
        orden_columnas=[0,1,3,2]
        for i in orden_columnas:
            col=row[i];
            col=col.upper()
            col=col.strip();
            data[cols[i]] = col
            if (i==1):
                if (not col in genero_lenguaje):
                    genero_lenguaje[col]={}
            elif (i==2):
                genero=data[cols[i-1]]
                residencia=data[cols[i+1]]
                if (col in genero_lenguaje[genero]):
                    genero_lenguaje[genero][col]+=1
                else:
                    genero_lenguaje[genero][col]=1;
                if (col in residencia_lenguaje[residencia]):
                    residencia_lenguaje[residencia][col]+=1
                else:
                    residencia_lenguaje[residencia][col]=1;
                if (not col in lenguaje):
                    lenguaje[col]=1;
                else:
                    lenguaje[col]+=1;

```

```

        elif (i==3):
            if (not col in residencia_lenguaje):
                residencia_lenguaje[col]={}

    rownum = rownum + 1

    generos=genero_lenguaje.keys();
    residencias=residencia_lenguaje.keys();
    lenguajes=lenguaje.keys();
    #ordenamos alfabéticamente
    generos=sorted(generos)
    residencias=sorted(residencias)
    lenguajes=sorted(lenguajes)

```

- Género

```

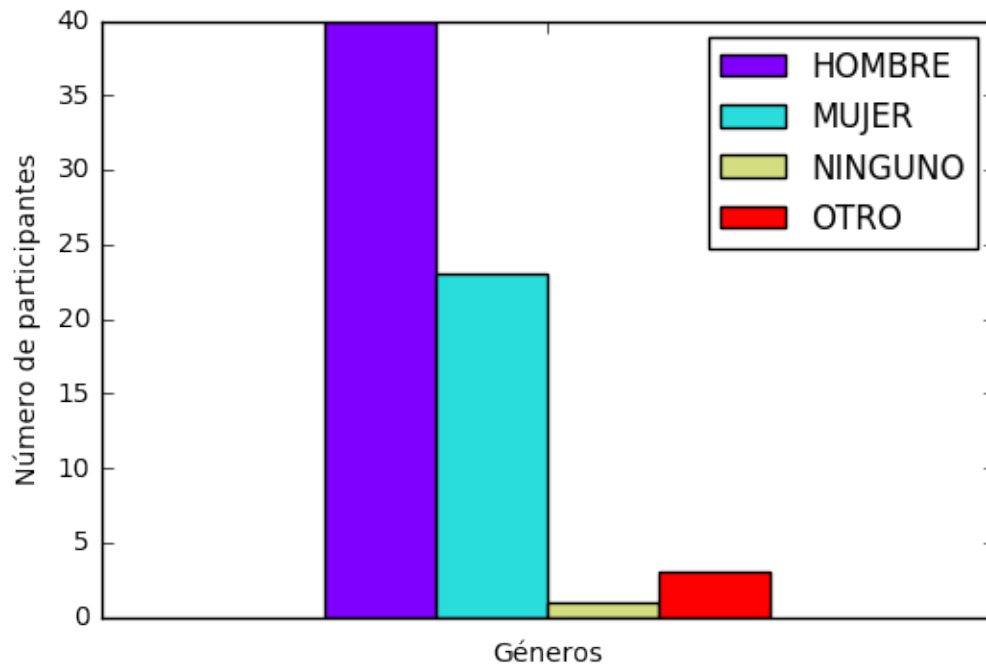
In [37]: #gráfico de barras de cada genero
cantidad_genero={};
colors = cm.rainbow(np.linspace(0, 1, len(generos)))
for g in generos:
    datas=genero_lenguaje[g];
    total=sum(datas.values());
    cantidad_genero[g]=total;
df = pd.DataFrame(data=cantidad_genero,index=[""],columns=generos)
df.head()
genero_plot = df.plot(kind='bar',label='Gráfico de barras de cada genero',
genero_plot.set_ylabel("Número de participantes")
genero_plot.set_xlabel("Géneros")

```

```

Out[37]: <matplotlib.text.Text at 0x7fc00eb9ebe0>

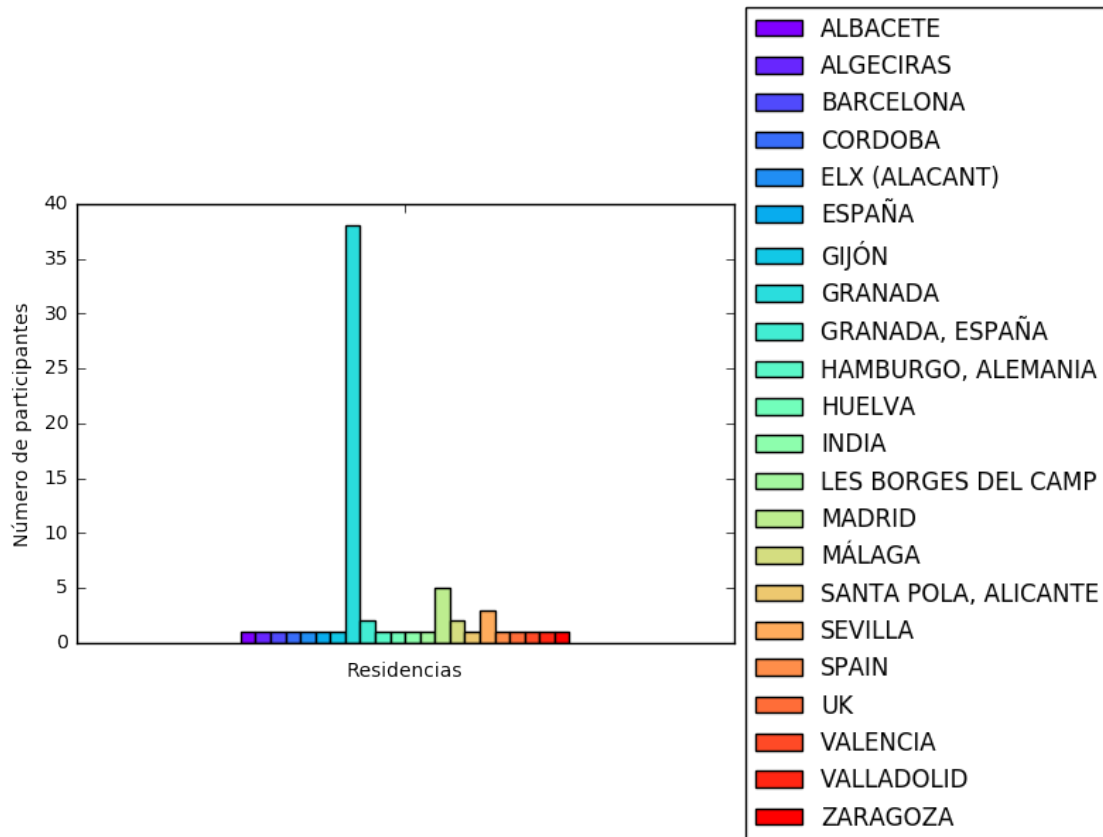
```



- Residencia

```
In [38]: #gráfico de barras de cada residencia
cantidad_residencia={};
colors = cm.rainbow(np.linspace(0, 1, len(residencias)))
for g in residencias:
    datas=residencia_lenguaje[g];
    total=sum(datas.values());
    cantidad_residencia[g]=total;
df = pd.DataFrame(data=cantidad_residencia,index=[""],columns=residencias)
df.head()
residencia_plot = df.plot(kind='bar',label='Gráfico de barras de cada resi
residencia_plot.set_ylabel("Número de participantes")
residencia_plot.set_xlabel("Residencias")
residencia_plot.legend(loc='center left', bbox_to_anchor=(1.0, 0.5))
```

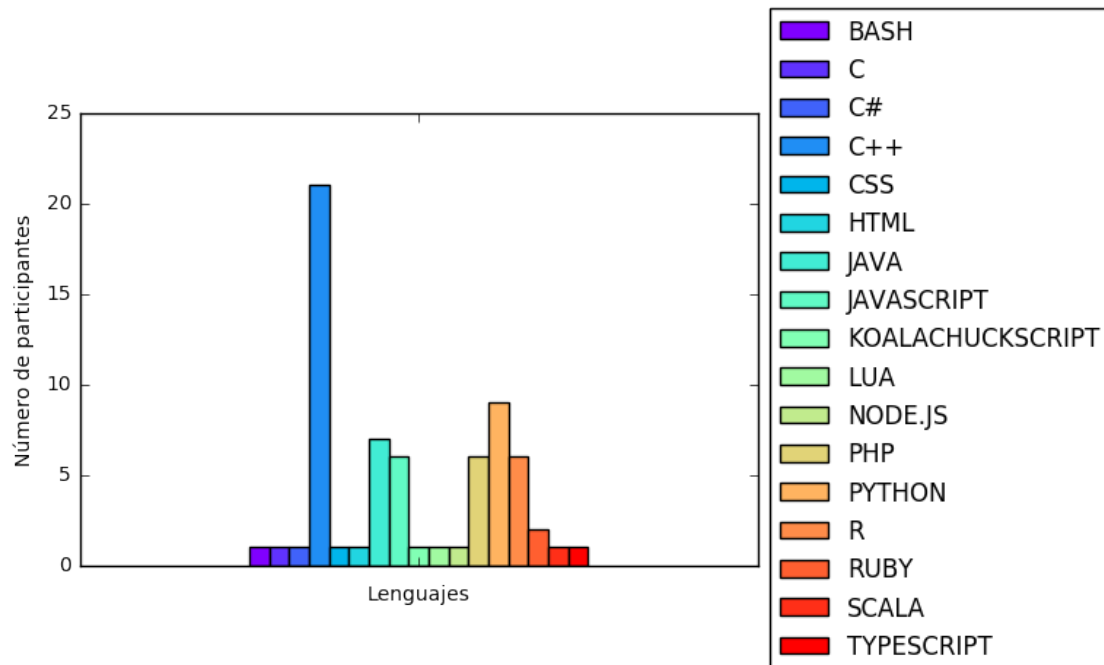
```
Out[38]: <matplotlib.legend.Legend at 0x7fc00eb193c8>
```



- Lenguaje de programación

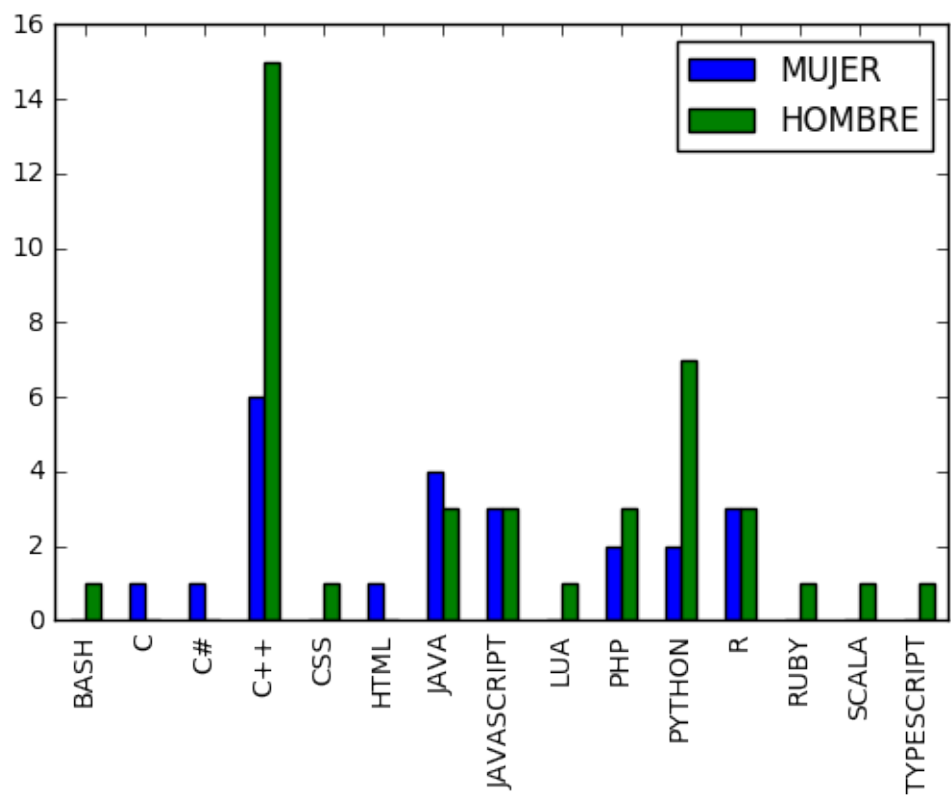
```
In [41]: #gráfico de barras de lenguajes de programación
colors = cm.rainbow(np.linspace(0, 1, len(lenguajes)))
df = pd.DataFrame(data=lenguaje, index=[""], columns=lenguajes)
df.head()
lengua_plot = df.plot(kind='bar', label='Gráfico de barras de cada lenguaje')
lengua_plot.set_ylabel("Número de participantes")
lengua_plot.set_xlabel("Lenguajes")
lengua_plot.legend(loc='center left', bbox_to_anchor=(1.0, 0.5))

Out[41]: <matplotlib.legend.Legend at 0x7fc00ea98898>
```



```
In [42]: d=[{'MUJER':genero_lenguaje['MUJER']['C++'],'HOMBRE':genero_lenguaje['HOMBRE']

df = pd.DataFrame(data=genero_lenguaje,columns=['MUJER','HOMBRE'])
df.head()
my_plot = df.plot(kind='bar')
```



In [ ]: