



Iniciándose en Python

Diana Iveth López Atienzo
Carlos Lizárraga Celaya

20 septiembre del 2016

1 Introducción

En muchas ocasiones, al momento de hacer una investigación el análisis de los resultados se complican un poco, ya que son demasiadas variables o la cantidad de datos es muy extensa. Cuando estamos en esta situación, revisar, ordenar, graficar e interpretar todos y cada uno de estos datos se convierte en una tarea muy tardada y en muchas ocasiones, es algo imposible como por ejemplo, en documentos que tienen millones y millones de datos.

Este tipo de situaciones han provocado que el ser humano busque la manera de poder interpretar gigantescas cantidades de información de manera rápida y precisa, creando así distintos lenguajes de programación. En esta actividad utilizaremos Python, que es un lenguaje de programación interpretado, es decir, es un programa informático capaz de analizar y ejecutar otros programas. soporta orientación a objetos, programación imperativa y, en menor medida, programación funcional. Usa tipado dinámico y es multiplataforma.

2 Paso I

Con la información que tenemos en el documento csv de la actividad 2, trabajaremos en esta actividad.

3 Paso II

Abre una sesión en ipython notebook. Para ello, abrimos la terminal. Una vez que estemos en la carpeta donde guardaremos nuestros avances escribimos ipython notebook, lo que nos abrirá una ventana de internet a una plataforma online donde podremos correr nuestros datos.

4 Paso III

Abrimos un nuevo cuaderno y definimos los siguientes datos.

```
import pandas as pd

import numpy as np

import matplotlib as plt

df=pd.read_csv("/home/dilopez/Escritorio/Computacional/Actividad2/datosCAPE12.csv")

y lo corremos
```

5 Paso IV

En la siguiente celda escribimos

```
df.head(25)
```

Lo corremos y nos aparecera una tabla con los primeros 25 datos de nuestro documento csv.

6 Paso V

En la siguiente celda escribimos

```
df.describe()
```

Lo corremos y nos muestra una una tabla con la descripcion del documento, los cuartiles, maximos y minimos, la mediana, etc.

7 Paso VI

En la siguiente celda escribimos

```
df.apply(lambda x: sum(x.isnull()),axis=0)
```

Lo corremos y nos muestra la siguiente tabla.

8 Paso VII

Después corremos el siguiente código

```
df_complete=df.dropna()
```

y después

```
df_complete
```

Nos mostrará una tabla con los primeros 30 datos y los últimos 30 datos de nuestro documento.

9 Paso VIII

Para saber cuantos datos tenemos en total utilizamos el siguiente código

```
len(df)
```

En el caso de mi documento tenía 316 datos.
Utilizamos el comando

```
df.columns
```

Para saber cuál es el nombre que utilizaremos para cada columna a la hora de graficar o buscar datos específicos.

10 Paso IX

Corriendo el siguiente código

```
df[u' CAPE'].head(20)
```

Podemos ver los primeros 20 datos del CAPE

11 Paso XII

Utilizando la siguiente función, podemos graficar los datos de la columna CAPE.

```
df[u' CAPE'].hist(bins=100)
```

12 Paso XIII

Utilizando el siguiente comando

```
df_complete.boxplot(column= u' CAPE')
```

Se nos muestra un diagrama donde se muestran la mediana, los cuartiles, etc.

13 Paso XIV

Después utilizamos el siguiente código

```
df_complete.boxplot(column=u' CAPE',sym='')
```

y aparece un diagrama de caja.

14 Paso XV

Por último colocamos el siguiente comando

```
df_complete['month'] = pd.DatetimeIndex(df_complete[u'Date']).month
```

Lo corremos y después escribimos el siguiente:

```
df_complete.boxplot(column=u' CAPE', by = 'month', sym='')
```

El cual nos hará n diagrama de caja para cada mes del año de nuestra base de datos.

15 Conclusión

Python es una herramienta muy útil a la hora de editar documentos que contienen una gran cantidad de información ya que analiza los datos y los tabula, grafica y sintetiza de una manera muy fácil y rápida.

16 Bibliografía

University of Wyoming/ College of Engineering/ Department of Atmospheric Science : [http : //weather.uwyo.edu/upperair/sounding.html](http://weather.uwyo.edu/upperair/sounding.html)

Pydata/ Python Data Analysis Library : [http : //pandas.pydata.org/pandas-docs/stable/tutorials.html](http://pandas.pydata.org/pandas-docs/stable/tutorials.html)