

Taller Open Source: Análisis y visualización de datos con Python en Google Colab

Gonzalo Uribarri

•



@gonzauri

Agenda

Presnetación

(20 min) ¿Por qué Python en Colab?

Hands-on

(60 min) Actividad en grupos sobre Notebook.

Puesta en común y Conclusiones

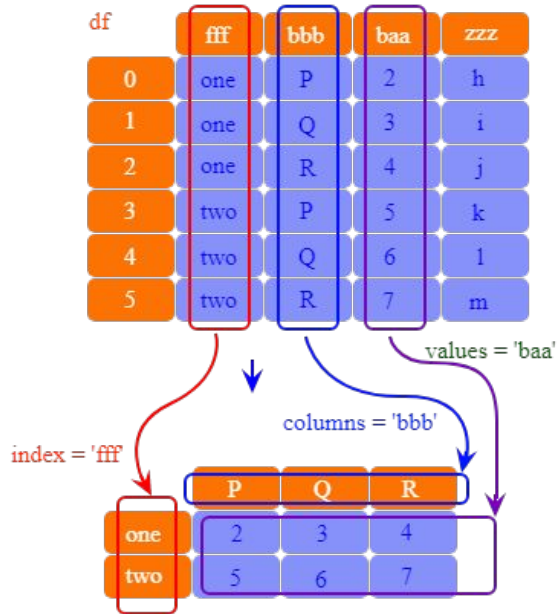
(20 minutos) Comentarios sobre el trabajo realizado y recursos para seguir.



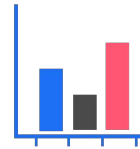
Python para Análisis de Datos

Análisis de Datos

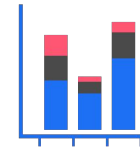
Procesamiento



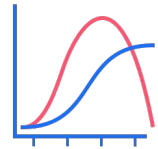
Visualización



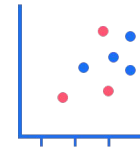
Bar chart



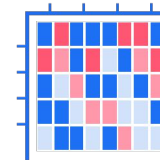
Stacked bar chart



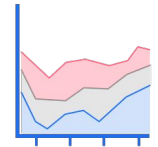
Line graph



Scatter plot



Calendar heatmap



Stacked area chart

Análisis de Datos

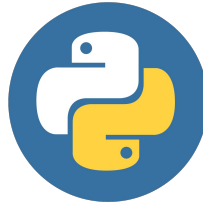
¿Que lenguaje uso?



Alto
Nivel



ORIGIN®



Bajo
Nivel



Herramientas para análisis de Datos

Existe una gran variedad de programas y lenguajes disponibles para trabajar con datos.

Alto
Nivel



ORIGIN®

orange

DATA MINING
FRUITFUL&FUN



Herramientas para análisis de Datos

Existe una gran variedad de programas y lenguajes disponibles para trabajar con datos.



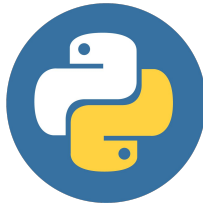
MATLAB



Bajo
Nivel



Alto
Nivel



Bajo
Nivel

Herramientas para análisis de Datos

Existe una gran variedad de programas y lenguajes disponibles para trabajar con datos.

Pero muchas no son Open Source ...



Alto
Nivel



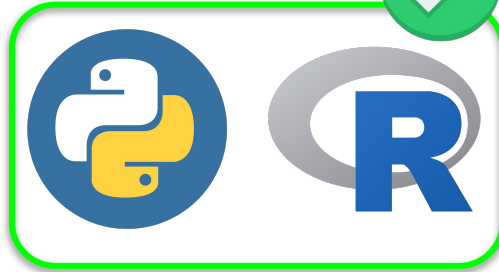
orange
DATA MINING
FRUITFUL&FUN



Herramientas para análisis de Datos

Existe una gran variedad de programas y lenguajes disponibles para trabajar con datos.

Pero muchas no son Open Source ...



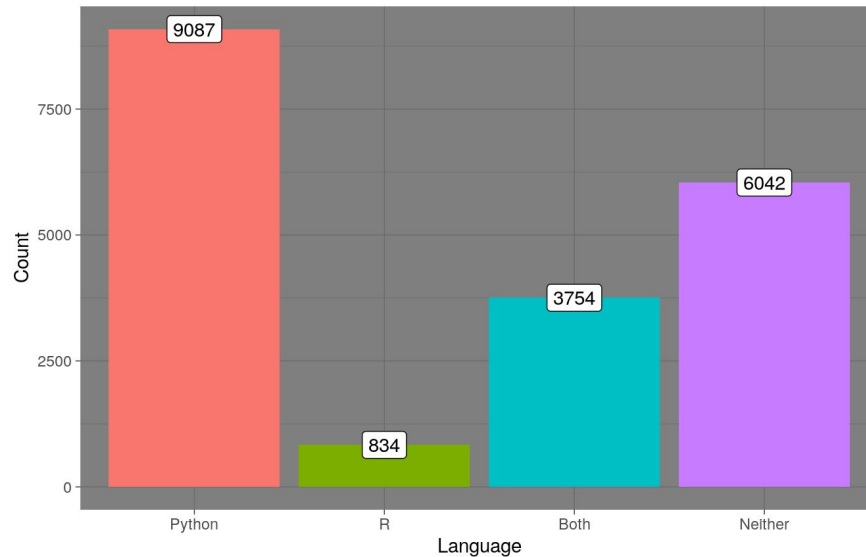
Bajo
Nivel



Python vs R

Which Programming Language Do You Use on a Regular Basis?

Python still dominates among Kaggle users, and the majority of R users also use Python.

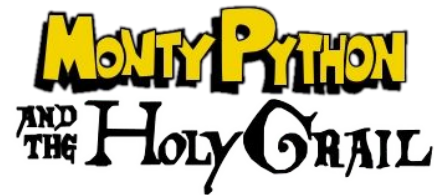
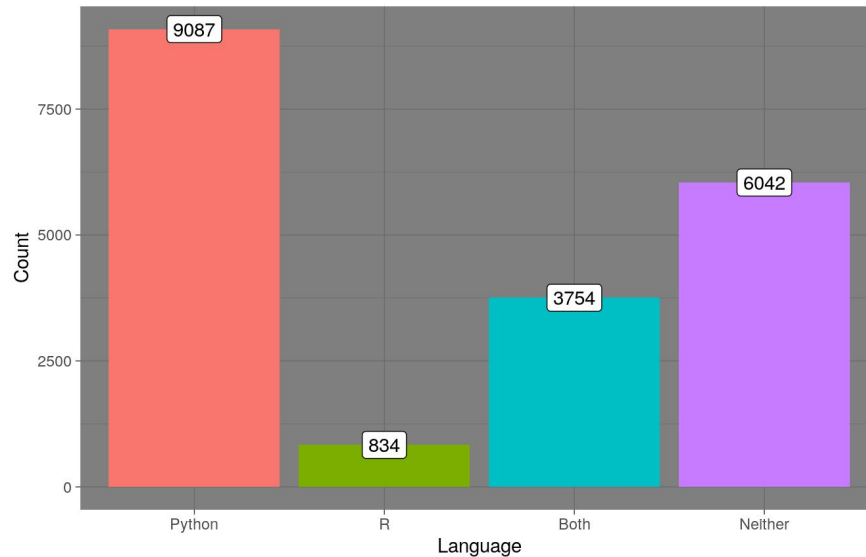


<https://www.kaggle.com/etsc9287/python-vs-r-the-data-science-rivalry>

Python vs R

Which Programming Language Do You Use on a Regular Basis?

Python still dominates among Kaggle users, and the majority of R users also use Python.



<https://www.kaggle.com/etsc9287/python-vs-r-the-data-science-rivalry>

Usando Python en Google Colab



Google Colab

Entorno de Desarrollo Gratuito

- Web - based (Corre en el navegador)
 - Interactivo
 - Usa IPython
 - Basado en Jupyter Notebook (.ipynb)
-



Google Colab

Entorno de Desarrollo Gratuito

- Web - based (Corre en el navegador)
- Interactivo
- Usa IPython
- Basado en Jupyter Notebook (.ipynb)
- Comenzó siendo Open Source...



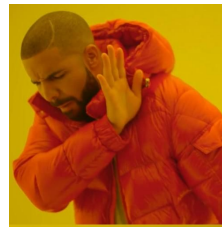


Google Colab

Entorno de Desarrollo Gratuito

- Web - based (Corre en el navegador)
- Interactivo
- Usa IPython
- Basado en Jupyter Notebook (.ipynb)
- Comnzó siendo Open Source...

pero ya no lo es





Nuevo_proyeto ☆

Archivo Editar Ver Insertar Entorno de ejecución Herramientas Ayuda



Comentar



Compartir



+ Código + Texto



RAM

Disco



Editar



Bienvenidos al Taller Open Source

Primera Edición

```
[1] print('Hello World')
```

```
↳ Hello World
```

```
[3] listita = [0,1,2,3,4]
    listita
```

```
↳ [0, 1, 2, 3, 4]
```




Google Colab - Tipos de celdas

Celdas de Código

- Código ejecutable
- Corre en un Entorno (IPython)

```
[3] listita = [0,1,2,3,4]
listita
```

↪ [0, 1, 2, 3, 4]

Celdas de texto

- Texto enriquecido
- Formato con Markdown



Bienvenidos al **Taller Open Source**

Primera Edicion



Google Colab - Tipos de Kernel

Entorno Local

- Corre en nuestra computadora
- Control total del entorno de ejecución (ej. Conda)
- Usa RAM y procesadores de nuestra PC

Entorno remoto

- Corre en un servidor de Google
- Hardware poderoso disponible (variedad)
- Límite de tiempo
- ~~Code~~ Plug and Play



Google Colab - Almacenamiento

Integracion con Drive

- Se guardan automaticamente los notebooks.
- Se puede leer y guardar archivos en carpetas de Google Drive

Integración con Github

- Conexión práctica hacia y desde repositorios
- Mas Info: **Próximo Taller Open Source**



Dataset y objetivos

Datos Experimentales

Movimiento de un roedor en un Open Field

- Adquisición de imágenes a 20 Hz (Video)
 - Movimiento en una caja de 50cm x 50cm
 - Coordenadas en el plano X - Y en función del tiempo
-

Datos Experimentales

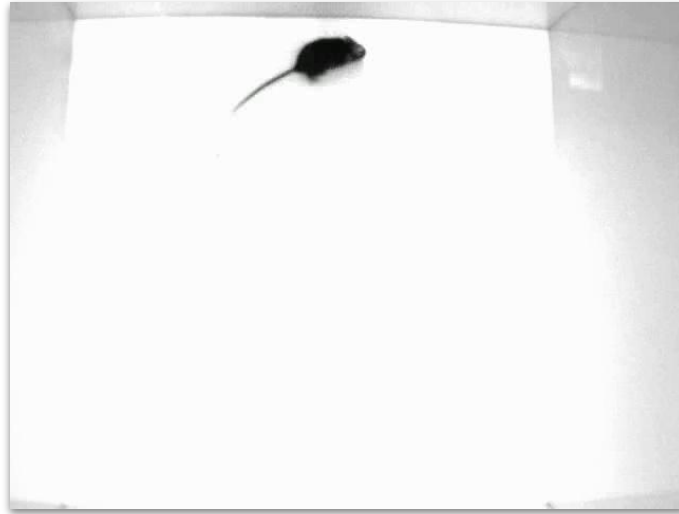
Movimiento de un roedor en un Open Field

- Adquisición de imágenes a 20 Hz (Video)
- Movimiento en una caja de 50cm x 50cm
- Coordenadas en el plano X - Y en función del tiempo



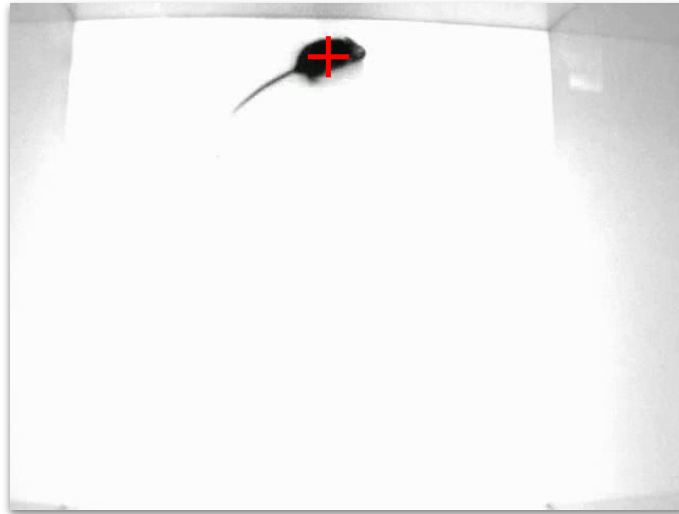
Datos Experimentales

Movimiento de un roedor en un Open Field



Datos Experimentales

Movimiento de un roedor en un Open Field



11900 lines (11900 sloc) | 83.4 KB

1	190,88
2	195,81
3	200,74
4	206,66
5	213,59
6	221,52
7	230,45
8	238,40
9	244,37
10	250,34

Datos Experimentales

Preguntas a responder:

- A) ¿Qué cantidad del tiempo se encuentra el ratón en el centro del recinto (total y porcentaje)?
- B) ¿Se mueve más rápido cuando está en el centro del recinto?
- C) ¿Se mueve más rápido en la primer, segunda o tercer parte del experimento?
- D) ¿Qué lugares del recinto habitó más tiempo el ratón?

Algunas no son usuales en el campo.

Hands-on



Grupos aleatorios de 4 personas

Uno comparte pantalla y trabajan juntos.

Discutir explicaciones y código

No solo avanzar sino también ir entendiendo.

Completar ejercicios y probar

Que sea lo más interactivo posible, prueben modificar partes del código.

Leer explicaciones

Si se traban y no pueden avanzar con algo, me llaman a la sala.



Antes de empezar

Crear la carpeta de trabajo en su Google Drive!

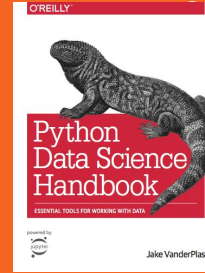
Todos los archivos están en:

https://drive.google.com/drive/folders/1oGPkJRmQQRcqcHMrOh_oimK1TvUceg_C?usp=sharing

Copiar el Notebook y mouse_data.txt a su carpeta de trabajo en su drive

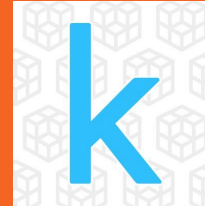
Recursos (Como seguir)

Material en la web



Libro

<https://jakevdp.github.io/PythonDataScienceHandbook/>



Kaggle!

<https://www.kaggle.com/>



Towards Data Science

<https://towardsdatascience.com/>

Próximo Talleres Open Source

- **Git en Github - 7/9**

Control de versionado y
desarrollo colaborativo

- **Arduino - 15/9**

Placa controladora

- **Diseño CAD e Impresión - 22/9**

Fabricación 3D



Muchas Gracias!

Gonzalo Uribarri •



@gonzauri
