C-Innova: Ciencia de Data Taller

Instrucciones de Instalación y más información sobre datos

**\*\* Seguir en orden por favor \*\***

1. Instala **Python3**
   1. **Link:** [**https://www.python.org/downloads/**](https://www.python.org/downloads/)
      1. **Descarga la última versión!**
   2. Que es python?
      1. Nos permite programar -> programar: Proceso de codificación de una regla o proceso en una notación (en este caso: un lenguaje de programación) para ser ejecutado por un ordenador

1. Python: El lenguaje / notación que hemos seleccionado que nos permite comunicarnos y expresar lo que queremos que el ordenador haga

* 1. **Terminal de la línea de comandos** 
     1. ¿Qué es un terminal de línea de comandos?

1. Una ventana / interfaz de usuario para escribir comandos y comunicarse con el sistema operativo de su computadora (para llevar a cabo actividades), como por ejemplo ejecutar archivos python

ii.Si tiene un ordenador portátil Apple MAC, no necesita instalar Git Bash, simplemente utilice el "Terminal del MAC"

**iii. Si no tiene un Apple MAC, instale Git Bash**

1. ¿Qué es Git Bash?

2. Es un terminal de línea de comandos para comandos unix (por lo que puede utilizar los mismos comandos que alguien utiliza en un terminal Apple MAC)

* 1. **Pip3**

¿Qué es pip?

1. Pip es un sistema de gestión de paquetes utilizado para instalar y administrar paquetes de software escritos en Python

¡Las bibliotecas ahora se podrán instalar porque tienes Pip3!

**1. Numpy**

a. Instrucciones:

i. Abra su terminal de línea de comandos (MAC: "Terminal", otro: "Git Bash")

ii.Escriba: pip3 install numpy

b¿Qué es Numpy?

* + 1. Es la biblioteca para el lenguaje de programación de python para soportar tablas y matrices multidimensionales grandes (estructuras de datos) y colección para funciones matemáticas de alto nivel para operar + manipular estas tablas y matrices

c .Nuestro propósito: Es una biblioteca dependiente de Pandas, Matplotlib etc; Fines numéricos y manejo de números / estructuras de datos

**2. Pandas** (depende de Numpy, así que asegúrese de instalar Numpy primero)

a. Instrucciones:

i. Abre tu terminal

Ii. Escriba en: pip3 install pandas

* 1. **¿Qué es Pandas?**
     1. No es un oso.

Ii. Es una biblioteca de software escrita para el lenguaje de programación python para la manipulación / edición y análisis de datos

* + - 1. Las funciones de python escritas para componer la biblioteca ofrecen estructuras de datos y operaciones para manipular tablas numéricas y series temporales (series de puntos de datos indexados en orden de tiempo / cómo cambian las cosas a lo largo del tiempo)
      2. El nombre Pandas se inspiró en "datos de panel" - término de econometría para conjuntos de datos estructurados multidimensionales
  1. Nuestro propósito: Usaremos esto para trazar figuras, recopilar / extraer datos, limpiar datos y analizar datos para su visualización.

**3. Matplotlib**

a. Instrucciones:

i. Abre tu terminal

Ii. Escriba: pip3 install matplotlib

b. ¿Qué es matplotlib?

i.Trazador de la biblioteca para el lenguaje de programación de python y es una extensión numérica de Numpy

c. Nuestro propósito: Usaremos esto para trazar figuras y visualizar datos.

**. Jupyter notebook**

1. Instrucciones:

a. Enlace para lectura: https://jupyter.readthedocs.io/es/latest/install.html

i. Requiere Python! Ya deberías haber instalado esto.

* + - * 1. Abre el terminal
        2. Escriba: pip3 install --upgrade pip
        3. Escriba: pip3 install jupyter
        4. Ejecución del jupyter notebook

i.Enlace para más lectura: https://jupyter.readthedocs.io/es/latest/running.html#running - jupyter notebook

Ii. Como se debe ver tu terminal:

1. $ jupyter notebook [I 08: 58: 24.417 NotebookApp]

atendiendo ordenadores portátiles desde el directorio local: /Users/catherine [I 08:58:24.417 NotebookApp] 0 active kernels [I 08:58:24.417 NotebookApp]

The Jupyter Notebook is running at: http://localhost:8888/ [I 08:58:24.417 NotebookApp]

2. Utilice Control-C para detener este servidor y apagar todos los kernels (dos veces para omitir la confirmación).

a. A continuación, abrirá su navegador web predeterminado a esta URL.

segundo.

b. Cuando el notebook se abra en su navegador, verá el Notebook Dashboard, que mostrará una lista de las carpetas, archivos y subdirectorios en el directorio donde se inició el servidor de notebook. La mayoría de las veces, usted deseará iniciar un servidor de notebook en el directorio de más alto nivel que contiene los notebooks. A menudo este será su directorio de inicio.

**2.¿Qué es Jupyter Notebook?**

a. Aplicación web de código abierto que le permite crear y compartir documentos que contienen código en vivo, ecuaciones, visualizaciones y texto explicativo

3. Nuestro propósito:

a. Proporcionar instrucciones para codificar para que revise

b. Para que usted tome sus propias notas de las actividades que hacemos

* + - 1. Otros visualización de datos / lenguajes de análisis:

i. R (se puede integrar con Tableau, un software de visualización de datos / inteligencia de negocios del que hablaremos más)

Ii. Matlab

**2. Instalar un editor de texto para escribir código**

a.Editor de textos recomendado: **Sublime**

i.Otras opciones que puede explorar: Atom, EMacs, Vim / Vi

Ii. Direcciones de editor de texto **sublimes**:

1. Enlace: http://www.sublimetext.com/3

a. Se requiere OS X (10.7 o posterior)

b.Windows - también disponible como versión portátil

c.Windows 64 bit - también disponible como versión portátil

Linux repos- también disponible como un tarball de 64 bits o 32 bits

**3. Tableu**

a. Software de visualización de datos - funciones de drag ans drop; Fácil de usar para visualizar datos; use server, excel, csv y otras fuentes de datos

b. Otros programas similares: Spotfire, SAS, etc

c. Cargaremos CSV para fines de prueba

i.Puede integrarse con R y Python para afectar las visualizaciones de datos para análisis avanzados - buena conexión / transición