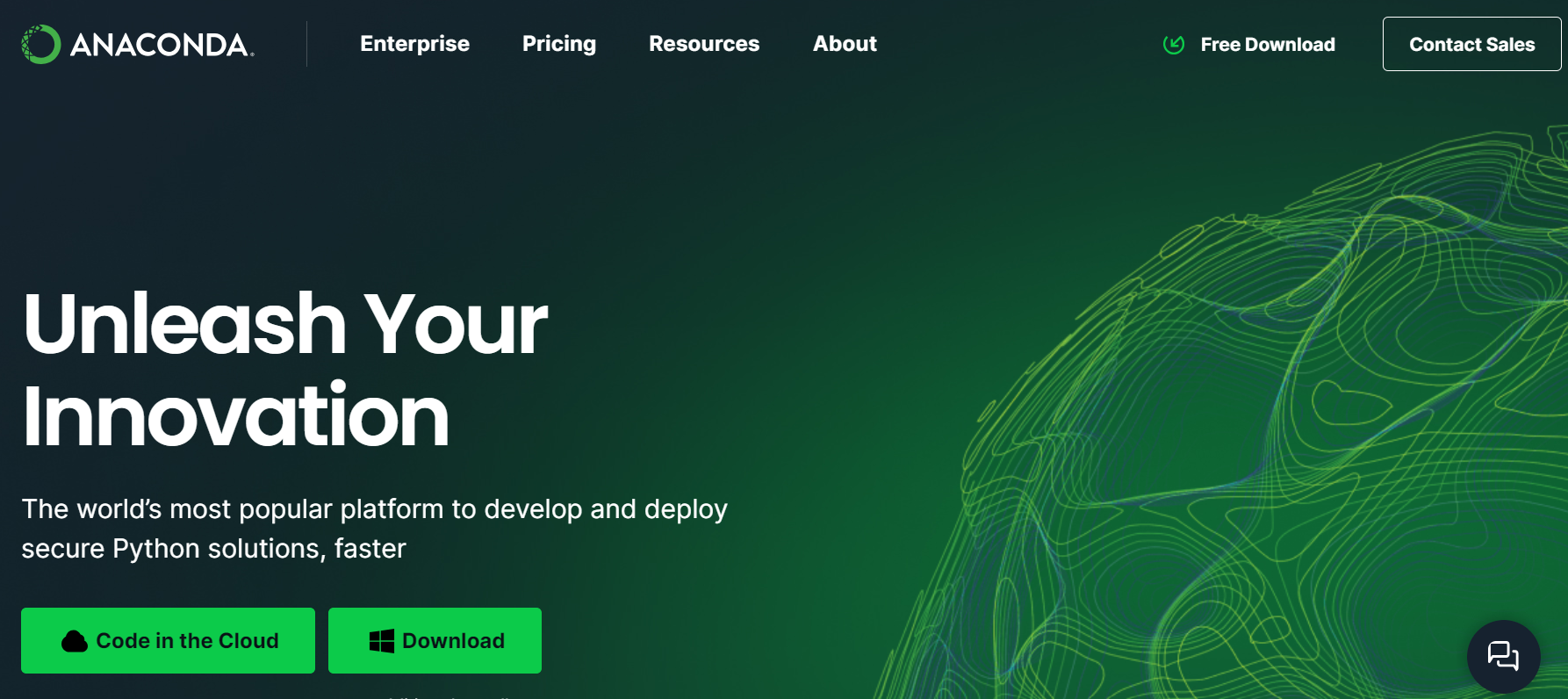
# Instalación de Anaconda

Vamos a trabajar con la suite de Anaconda. **¿Qué es Anaconda?**

La Suite de Anaconda es una plataforma de código abierto para el procesamiento y análisis de datos multi-lenguaje. Es una herramienta altamente popular en el ámbito de la ciencia de datos debido a su amplia gama de paquetes pre-instalados y sus capacidades para crear entornos virtuales de programación. Esta plataforma incluye *software* como Python, R, Jupyter Notebook, JupyterLab y otros paquetes comúnmente utilizados en la programación científica.

El primer paso para empezar a utilizar Anaconda es descargar e instalar la suite de Anaconda en el computador, para lo cual deben seguirse los siguientes pasos:

1. **Descargar Anaconda:** Vaya al sitio web oficial de Anaconda (<https://www.anaconda.com>) e identifique el instalador adecuado para su sistema operativo (Windows, Mac o Linux). Descárguelo e instálelo en su computador.





1. **Instalar Anaconda:** Una vez que haya descargado el instalador correspondiente, ejecútelo y siga las instrucciones del proceso de instalación.
2. **Configurar Anaconda:** Verifique que todo esté correctamente instalado. Para ello, abra una terminal en su sistema operativo y escriba “conda” en el comando. Si todo está correctamente instalado, se mostrarán en la pantalla las opciones disponibles del comando.

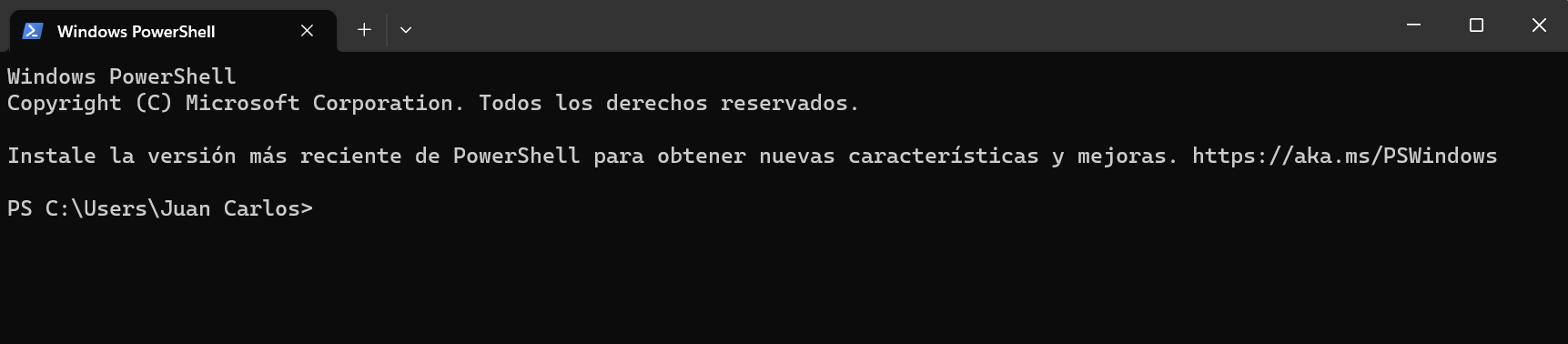
Una vez que Anaconda ha sido instalado correctamente, puede comenzar a trabajar con él. La interfaz de Anaconda incluye una serie de herramientas y bibliotecas, incluyendo el gestor de paquetes “conda”, Jupyter Notebook, JupyterLab, IPython, Spyder, entre otros.

Para empezar a trabajar deben buscar la aplicación **JupyterLab**, puesto que sobre ésta es que se correrán los programas en Python con los que vamos a trabajar.

# Instalando las Librerías en Python

Existen algunos módulos dentro de Python, con los cuales vamos a trabajar durante este curso, siendo ellos principalmente **pandas** y **numpy**. **Pandas** es una Librería de código abierto dentro de los desarrolladores de **Python** y sobre todo dentro del ámbito de Data Science y **Machine Learning**, ya que ofrece unas estructuras muy poderosas y flexibles que facilitan la manipulación y tratamiento de datos. **Pandas** surgió ante la necesidad de contar con una única librería que abarcara todo lo necesario para que un analista de datos pudiera tener en una misma herramienta todas las funcionalidades que necesitaba en su día a día, como son: **cargar datos, modelar, analizar, manipular y prepararlos. Por su parte numpy es una biblioteca de Python** que da soporte para crear vectores y matrices grandes multidimensionales, junto con una gran colección de funciones matemáticas de alto nivel para operar con ellas.

Para instalar estas librerías, abrimos una terminal en el computador. En el buscador de aplicaciones simplemente digitan la palabra terminal y luego, cuando aparezca, le dan *click* ahí. La terminal lucirá de esta manera:



Primero digitan **pip install pandas**

Una vez haya terminado la instalación, hacen lo mismo para numpy: **pip install numpy**

Cada vez que necesiten instalar una librería, siguen este mismo procedimiento con la librería que deseen instalar.

Otra librería que vamos a querer tener instalada es **Matplotlib**, la cual se usa para generar una gran variedad de gráficos de calidad para publicar online o en papel y sin emplear muchas líneas de código. Entonces, para instalar digitamos: **pip install matplotlib**

Una más es **seaborn**, la cual se usa para visualizar datos estadísticospara entender la data. Para instalarla digitamos: **pip install seaborn**

**Recuerden:** siempre que tengan una duda sobre Python, casi con toda seguridad la respuesta está en internet. Pueden digitar en su buscador la duda y alguna respuesta les aparecerá. Dentro de las páginas más útiles a la hora de encontrar respuestas a inquietudes está <https://stackoverflow.com>

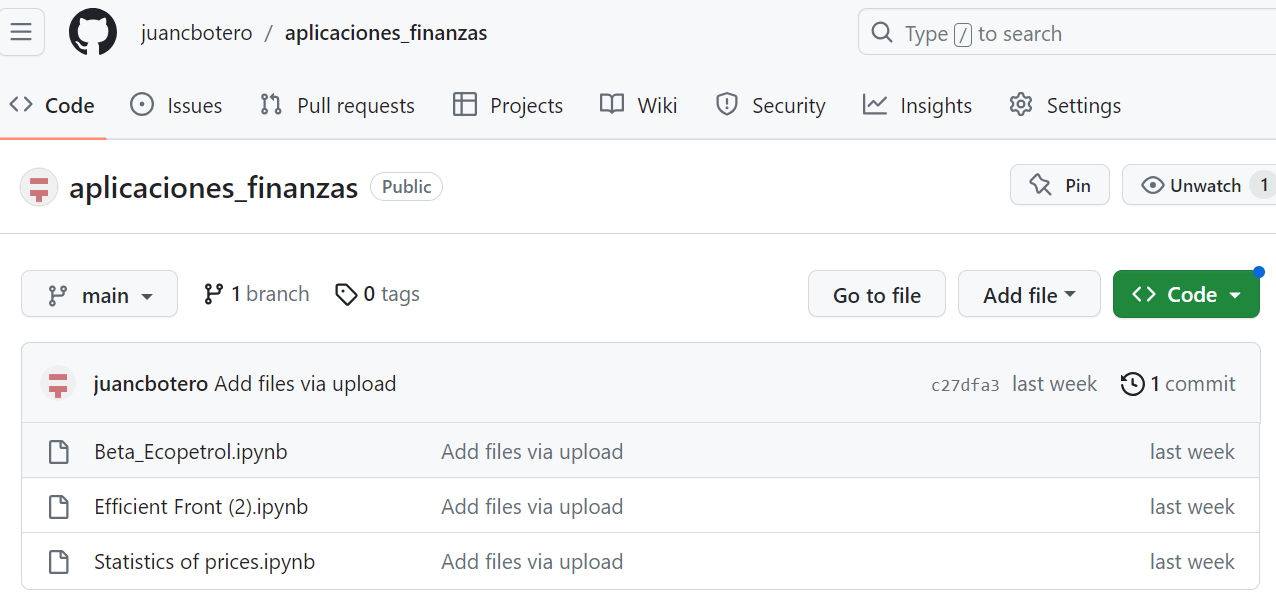
# Instrucciones Para Descargar Archivos de Python

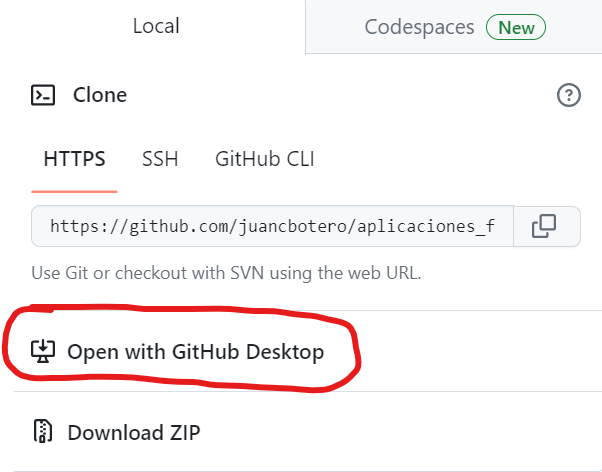
Ingresen a mi GitHub cuya dirección es: [juancbotero/aplicaciones\_finanzas (github.com)](https://github.com/juancbotero/aplicaciones_finanzas)

Para lo anterior les recomiendo registrarse en GitHub (<https://github.com>) y crear su propio espacio. Ahí pueden guardar estos archivos, más cualquier otro código que desarrollen más adelante.

Los archivos que utilizaremos son: 1) [Beta\_Ecopetrol.ipynb](https://github.com/juancbotero/aplicaciones_finanzas/blob/main/Beta_Ecopetrol.ipynb); 2) [Efficient Front (2).ipynb](https://github.com/juancbotero/aplicaciones_finanzas/blob/main/Efficient%20Front%20(2).ipynb); y 3) [Statistics of prices.ipynb](https://github.com/juancbotero/aplicaciones_finanzas/blob/main/Statistics%20of%20prices.ipynb)

Para importar estos archivos al archivo Python que estemos trabajando, les sugiero descargar la aplicación GitHub Desktop. Cuando estén en mi repositorio de GitHub, le dan *click* a el botón verde “Code” y ahí escogen la opción “Open with GitHub Desktop”. Cada vez que se actualice un archivo en mi GitHub, a ustedes se les actualizará también.





Estos archivos se descargarán en el disco duro de cada uno y desde ahí los pueden invocar cuando estén trabajando en Jupyter.

# Utilización del Colab de Google

Si aún no tienen instalada la suite de Anaconda, pueden trabajar en el Colab de Google. Para esto vayan a la URL de Colab: <https://colab.research.google.com/>

Allí se deben log in, lo cual pueden hacer con su cuenta de Google, aunque también lo pueden hacer creando una cuenta.

Una vez allí pueden ir a la dirección de GitHub que está arriba y cargar los archivos de Python que allí se encuentran.

Cuando cargan esos archivos de Python desde Colab, éstos quedan automáticamente guardados en Google Drive. Es recomendable que también suban a Google Drive los archivos de Excel o csv con que vayan a trabajar. Como los archivos de Python que están en GitHub no fueron creados desde Colab, es importante modificar ligeramente el código para que no intente leer los archivos de Excel donde están los datos desde el disco duro, sino desde el sitio de Google Drive donde están. Igualmente, yo les mostraré en clase cómo queda ese código modificado, pero es necesario que primero monten los archivos de Excel con que vamos a trabajar en una carpeta dentro de Google Drive.