

UNIVERSIDAD GALILEO
INSTITUTO DE INVESTIGACIÓN DE OPERACIONES
MAESTRÍA EN DATA SCIENCE



DESCRIPCIÓN DE
ANACONDA, GIT Y GITHUB

EDGAR RUBÉN ROJAS OVANDO

CARNÉ: 21000567

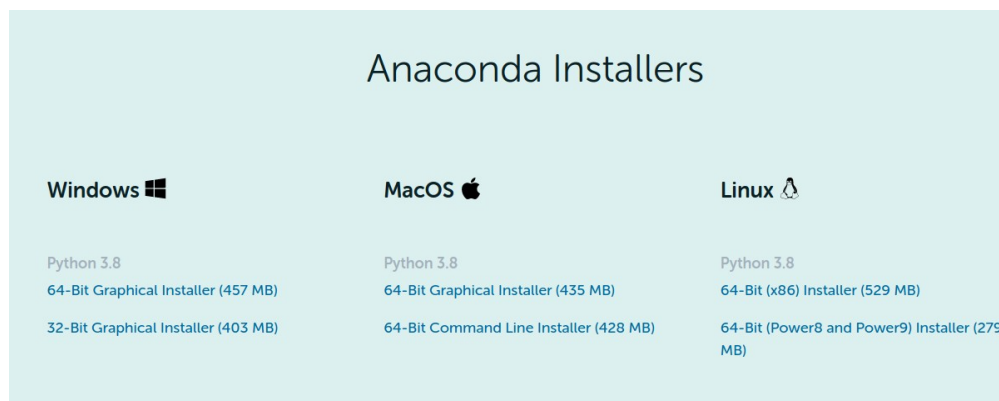
CIUDAD DE GUATEMALA, GUATEMALA, ENERO DE 2021

ANACONDA

Es una herramienta muy útil para el análisis de datos, siendo una distribución de Python. Esta herramienta contiene paquetes predefinidos que pueden ser utilizados conforme se necesiten. También cuenta con un administrador de ambientes que permite separar distintos proyectos con su propio entorno y configuraciones.

La instalación la realicé en un sistema operativo Linux. Realizando los siguientes pasos:

- Descargar la versión correspondiente (Es multiplataforma):



- Comprobar integridad del instalador:

```
$ sha256sum Anaconda3-2020.11-Linux-x86_64.sh
```

- Ejecutar instalador

```
$ bash Anaconda3-2020.11-Linux-x86_64.sh
```

- Aceptar el acuerdo de licencia presionando "ENTER"

- Especificar una ruta de instalación o precionar “ENTER” para instalar en la ruta por defecto.

Luego de realizar la instalación se procede a probar creando un ambiente, para ello situarse en la carpeta donde estarán los archivos de apoyo y ejecutar:

```
$ conda create -n pyt_galileo_2021 python=3 numpy  
jupyterlab
```

Siendo “pyt_galileo_2021” el nombre del ambiente, python en la versión 3, utilizando las herramientas numpy y jupyterlab.

Para listar los ambientes de conda ejecutar:

```
$ conda env list
```

En conclusión Anaconda es una de las mejores herramientas para Analisis de datos utilizado Python, con ella podemos utilizar los mejores paquetes sin necesidad de instalar uno a uno ahorrando tiempo de configuración y ejecución:

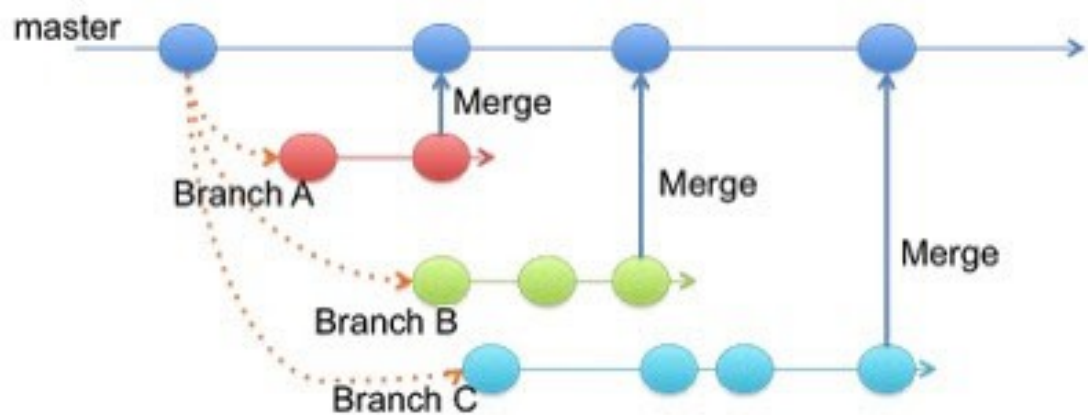


GIT

Es un sistema que sirve para el versionamiento, integración y organización de nuestro código independientemente del lenguaje de programación. Con git podemos ejecutar pruebas de integración y bifurcar ramas para desarrollar paralelo a los entornos de producción sin afectarlos, luego de esto nos facilita la integración de las versiones. Una característica muy importante es que nos permite volver atrás por si una versión no funciona como debiese.

Permite también el trabajo colaborativo en un mismo proyecto

La imagen siguiente detalla gráficamente el desarrollo de ramas e integración de un proyecto.



Para el manejo de repositorios en la nube las plataformas más populares son GitLab, GitHub y BitBucked.



En el transcurso del curso se utilizará GitHub como repositorio.

A continuación detallaré el proceso de instalación en Linux:

- Instalar Git (Sirve para todos los mencionado anteriormente)

```
$ sudo apt install git
```

Crear cuenta en GitHub.

Luego “Hacer clic en crear repositorio”, nombrarlo como se desee.

Para poder hacer commit desde un equipo se necesita añadir la llave ssh del equipo en la cuenta de GitHub .

Para crear la llave ejecutar:

```
$ ssh-keygen -t rsa -b 4096 -C "tuCorreo@mail.com"
```

Se generará el archivo id_rsa en la ruta ~/.ssh/

- Luego en la cuenta de GitHub hacer clic en Settings, -> SSH and GPG keys y pegar el contenido del archivo creado anteriormente con la key ssh.

Con esto hemos asociado el equipo a nuestra cuenta y solo basta con clonar el repositorio.

En conclusión Git es el mejor software para repositorios en la actualidad siendo el más utilizado a nivel mundial por encima de subversion.

Es esencial para el desarrollo de aplicaciones con un equipo colaborativo.