



Curso de Python para la Ciencia de Datos

Dr. Ing. Rodrigo Salas (rodrigo.salas@uv.cl)
Prof. Titular
Universidad de Valparaíso

Las actividades del curso de Python para la Ciencia de Datos se realizarán utilizando el Jupyter implementado en **Colab de Google**

<https://colab.research.google.com>

Por lo cual se sugiere crear una cuenta en Gmail con el fin de tener respaldo de los códigos, no obstante, se puede utilizar sin tener la necesidad de crear la cuenta.

En caso de requerir instalar los SW en el computador local, a continuación se encuentra la información relacionada con la instalación de los software requeridos para que los estudiantes puedan realizar las actividades prácticas.

Instalación de Python

Instalar python 3.9 mediante el framework de ANACONDA

<https://www.anaconda.com/download/>

Proceder a actualizar anaconda

```
conda update conda
conda update --all
```

Además utilizar el comando conda para instalar los siguientes paquetes (se debe usar un terminal de comandos):

```
conda install numpy
```



```
conda install scipy
conda install matplotlib
conda install pandas
conda install scikit-learn
conda install scikit-image
conda install networkx
conda install statsmodels
conda install seaborn
conda install nltk
pip install wordcloud
conda install tensorflow
conda install keras
```

Para la programación utilizaremos el notebook de Jupyter que viene instalado en anaconda.

Si los paquetes ya se encuentran instalados, se utiliza el comando

```
conda update nombre_paquete
```

para actualizarlos.

La utilidad de algunos de los paquetes es la siguiente:

- **Numpy:** utilizado para manipulación matricial
- **Scipy:** posee herramientas estadísticas y de computación científica
- **Matplotlib:** utilizado para graficar



- **Pandas:** estructuras de datos para hacer estudios estadísticos
- **Scikit-learn:** Herramienta para machine learning y minería de datos
- **Scikit-image:** Herramienta para computer vision y procesamiento de imágenes.
- **Seaborn:** Herramienta para visualización de gráficos
- **Nltk:** Herramienta para lenguaje natural
- **Networkx:** Herramienta para trabajar con grafos.
- **Statmodels:** Herramienta estadística para ajustar modelos de regresión y de series de tiempo.
- **Tensorflow y Keras:** Herramientas para Deep Learning