

Plantilla en Quarto

Emmanuel Alcalá

14/07/2022

Requerimientos para usar Quarto

1. [Instalar quarto](#)
2. Instalar un editor de texto. Recomiendo [VScode](#). Si es en VScode, instalar la [extensión de Quarto para VScode](#).
- Para usar con RStudio, ver [aquí](#).
3. Si se va a trabajar con Python debe configurarse apropiadamente. En VScode seleccionar el intérprete de Python que se quiere usar (e.g., de conda). Presionar **ctrl+shift+p** (command palette) y escribir **Python interpreter**, seleccionar el intérprete deseado.

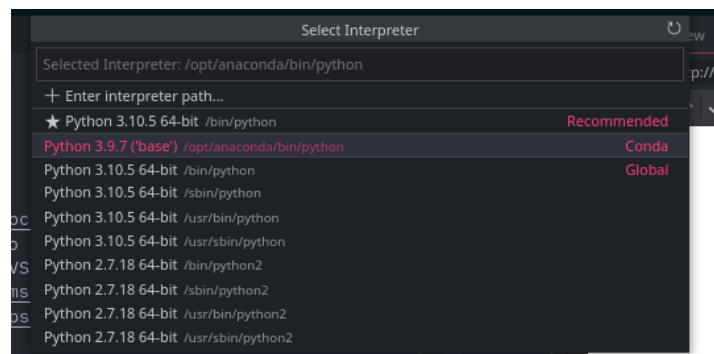


Figura 1: Intérprete de Python en VScode.

En RStudio, instalar primero la librería **reticulate** y enlistar los environments de conda

```
install.packages('reticulate')  
library(reticulate)  
conda_list()
```

Lo anterior da lo siguiente:

```
name python
1 base /opt/anaconda/bin/python
```

Escoger el environment

```
use_condaenv(condaenv = '/opt/anaconda/bin/python')
```

Luego de esto, en la pestaña de Tools de RStudio, seleccionamos Python, luego en Python interpreter Select..., en Conda Environments aparecerá

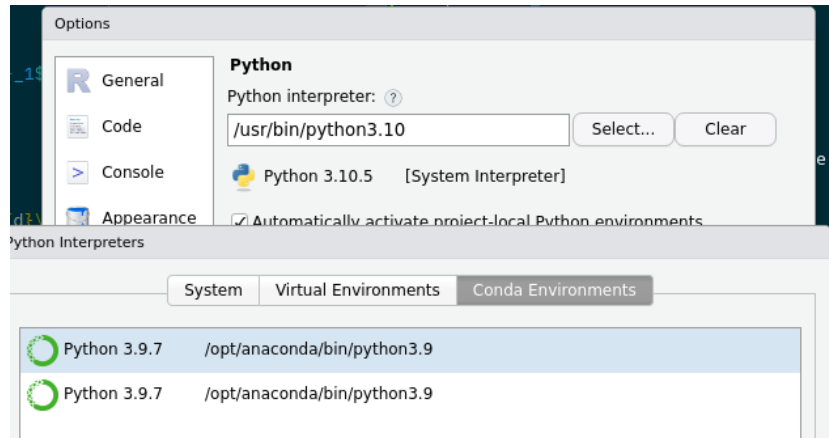


Figura 2: Intérprete de Python en RStudio.

4. Para renderizar en pdf debe instalarse `tinytex`. Abrir terminal y escribir

```
quarto tools install tinytex
```

5. El uso de código de python requiere configurar adecuadamente el PATH de `python` y el kernel de jupyter. En el encabezado, agregar

```
jupyter: python3
```

1 Título

El título de tu proyecto. Debe ser conciso y reflejar el argumento principal. En ocasiones es bueno plantearlo como pregunta, otras veces como una afirmación.

2 Introducción

Para estructurar un proyecto por primera vez es recomendable seguir alguna estructura particular, estandarizada.

La aplicación [Cookiecutter](#) provee una estructura estándar que parece adecuada para esto, sin embargo, puede ser demasiado. Yo prefiero una estructura como esta

```
parent_dir/
|_ project_name/
|   |_ data_processed/
|   |_ data_raw/
|   |_ docs/
|       |_ 00_analysis.qmd
|       |_ figures/
|       |_ scripts/
|           |_ python/
|           |_ r/
|_ Readme.md
```

En ‘project_name’ puede ser IDI-III, o lo que ustedes quieran. La jerarquía de directorios en linux/unix se podría crear fácilmente con el siguiente comando

```
mkdir -p project_name/{scripts/{r,python},data_raw,data_processed,figures,docs}
```

Probablemente en PowerShell de Windows se pueda hacer algo similar, pero actualmente es recomendable instalar el [Windows Subsystem for Linux \(WSL\)](#). Nunca es mala idea tener la potencia del shell de unix/linux a la mano.

Para algunos ejemplos de reportes técnicos de proyectos, revisa los proyectos finales del curso CS229: Machine Learning de Stanford.

- [2021](#).
- [2019](#).
- [2014](#).

2.1 Definición del problema o pregunta

De acuerdo con la primera sesión. Puedes encontrar las diapositivas [aquí](#).

3 Datos

Aquí describes los datos que usaste, sus fuentes, variables, etc. Deberá contener la siguiente información:

1. Justificación concisa de por qué el conjunto de datos elegido es relevante para el problema elegido.
2. Describir las fuentes de los datos.
3. Describir qué procesamiento se hizo para dejarlos en estado usable.
4. Describir las variables que contienen los datos. Por ejemplo, un codebook en caso de que se usen abreviaturas para las variables. Un codebook en formato csv se vería así

```
abreviatura,variable  
mif,mean income per family  
bmi,body mass index  
Q25,question 25  
etc,etc
```

Tabla 1: Ejemplo de codebook.

Abreviatura	Variable
mif	Mean income per family
bmi	Body mass index
Q25	Question 25

4 Métodos y Análisis

La figura [Figura 3](#) muestra el proceso que estaremos siguiendo en esta fase.

IDI-III tratará principalmente de estudiar el problema, limpiar y transformar los datos y seleccionar las variables, pero también comenzaremos a escribir y familiarizarnos con las herramientas de publicación (en este caso, Quarto).

1. Análisis exploratorio. Añadir tablas, gráficos exploratorios, etc. Esto no es propiamente un resultado, sino un análisis que se realiza para justificar otras decisiones.
2. Si se realizan transformaciones en una variable (e.g., se log-transformó, se exponenció, se escaló, se normalizó, etc) o cualquier ingeniería de características, extracción, etc., a partir de los datos exploratorios. Justificar la decisión.
3. Descripción de los métodos, como algoritmos, benchmarks, métricas de comparación (e.g., *RMSE*) etc. *No se colocan esos resultados aquí*, solo se menciona qu+e se utilizó.

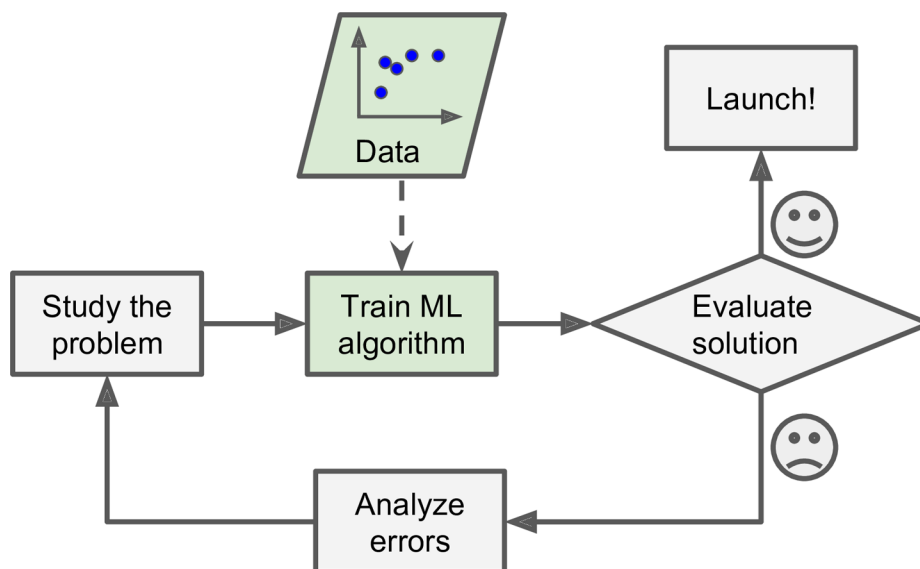


Figura 3: Proceso típico de proyecto de DS. Tomado de Géron (2019).

Quarto soporta renderización de ecuaciones usando la sintaxis de \LaTeX . Ver este [artículo](#) y [este](#) para ver cómo escribir matemáticas.

5 Resultados

En esta sección se colocan los resultados principales, como tablas de comparación, gráficos, etc. Por ejemplo, si se probó un algoritmo con respecto a otros tres, se coloca la ejecución en diversas métricas.

Las figuras y tablas deben ir acompañadas de una etiqueta y una breve descripción. Se pueden referenciar usando `@{label}` en donde `label` en este caso específico es `fig-polar`. La referencia renderizada se verá así Figura 4.

Las tablas pueden hacerse con Markdown

Tabla 2: Leyenda de tabla

Col1	Col2	Col3
A	B	C
E	F	G
A	G	G

Ver la tabla Tabla 2.

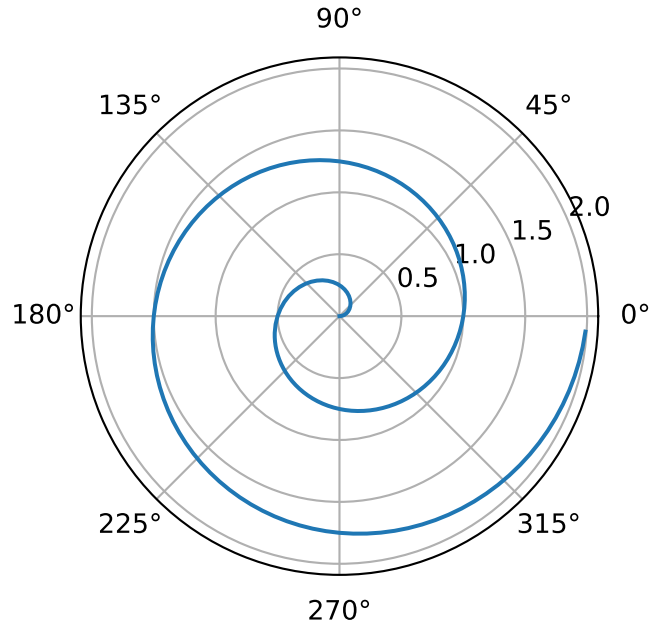


Figura 4: Esta es la leyenda de una figura.

6 Referencias

Para citar, usar (@alcala2021statistical) que se renderiza como (López-Cárdenas et al. (2021)). La entrada @alcala2021statistical debe estar tal cual en el archivo `referencias.bib`. Las referencias en formato de bibtex se pueden obtener desde Google Scholar. El formato tiene la siguiente estructura

```
@inproceedings{alcala2021statistical,
  author    = {López-Cárdenas, Patricia G. and Alcalá, Emmanuel and
              Sánchez-Torres, Juan Diego and Araujo, Elsie},
  booktitle = {2021 18th International Conference on Electrical
              Engineering, Computing Science and Automatic Control (CCE)},
  title     = {A Resampling Approach for the Data-Based Optimization
              of Nanosensors},
  year      = {2021},
  volume    = {},
  number    = {},
  pages     = {1-4},
  doi       = {10.1109/CCE53527.2021.9633114}
}
```

Para imprimir las referencias hay que colocar lo siguiente

```
::: {#refs}  
:::
```

Que se verá de la siguiente manera (notar que coloqué dos citas: Gerón y López-Cárdenas et al).

```
::: {#refs}thon :::
```

Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media.

López-Cárdenas, P. G., Alcalá, E., Sánchez-Torres, J. D., & Araujo, E. (2021). A Resampling Approach for the Data-Based Optimization of Nanosensors. *2021 18th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*, 1-4. <https://doi.org/10.1109/CCE53527.2021.9633114>