

Curso de **Personalización Avanzada de Entorno para Ciencia de Datos**

Jesús Vélez Santiago
@jvelezmagic

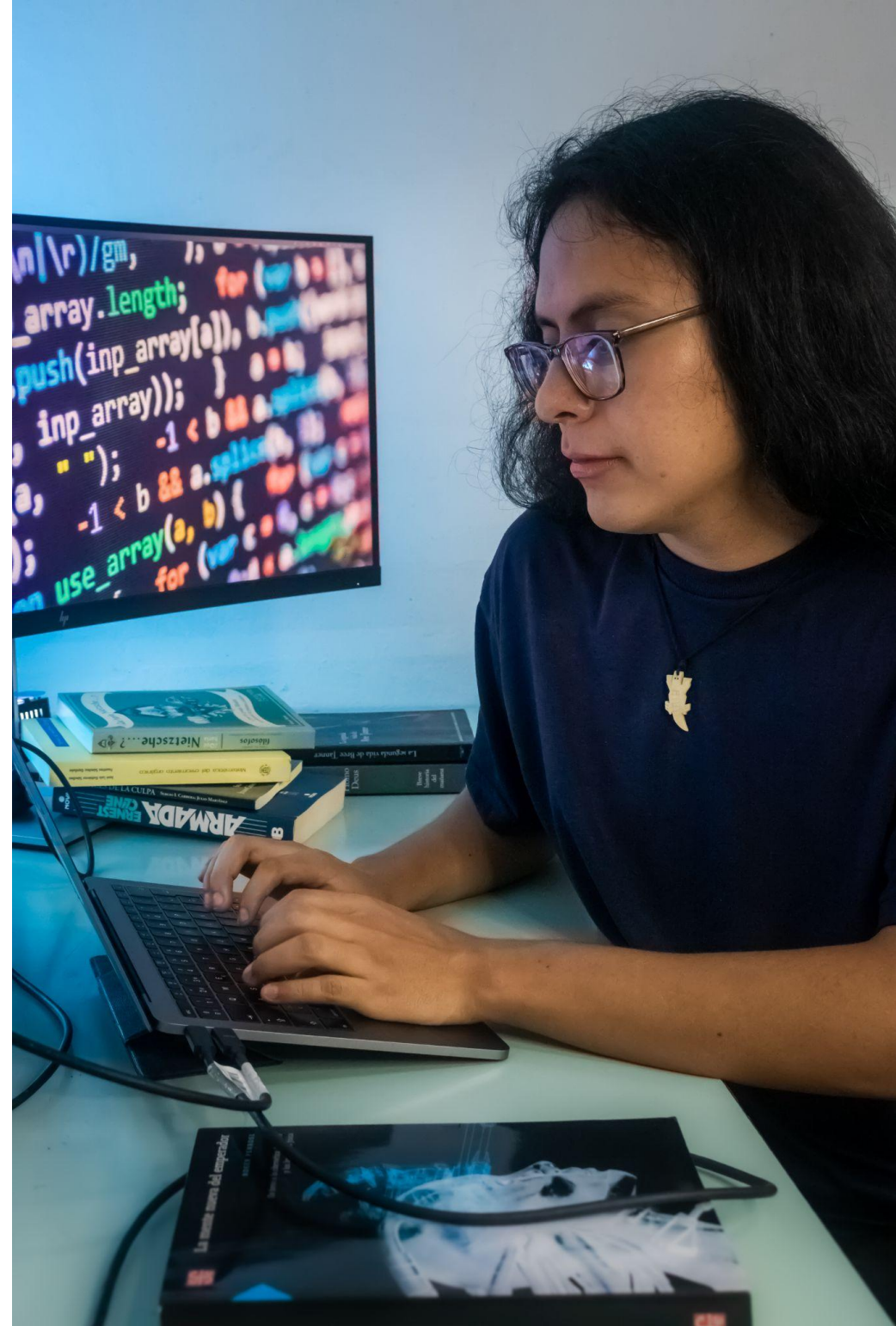


Científico genómico/datos

A través del:

- Desarrollo de software.
- Matemáticas.
- Estadística.

Busco **crear**
herramientas para
enseñar y ayudar al
desarrollo de la ciencia.



Requisitos

- Uso de Jupyter Notebooks.
- Manejo de ambientes y librerías con Conda.
- Python intermedio.
- Uso de la terminal y línea de comando.
- Manejo de Git y GitHub.



Personalización avanzada

- Crearás **plantillas de entornos** de trabajo personalizadas para tus análisis.
- Aprenderás a desarrollar **proyectos autocontenibles y multiplataforma**.

**¿Qué son y por
qué utilizar plantillas
de proyectos?**



“

**Medio que posibilita
portar o construir un
diseño predefinido.**

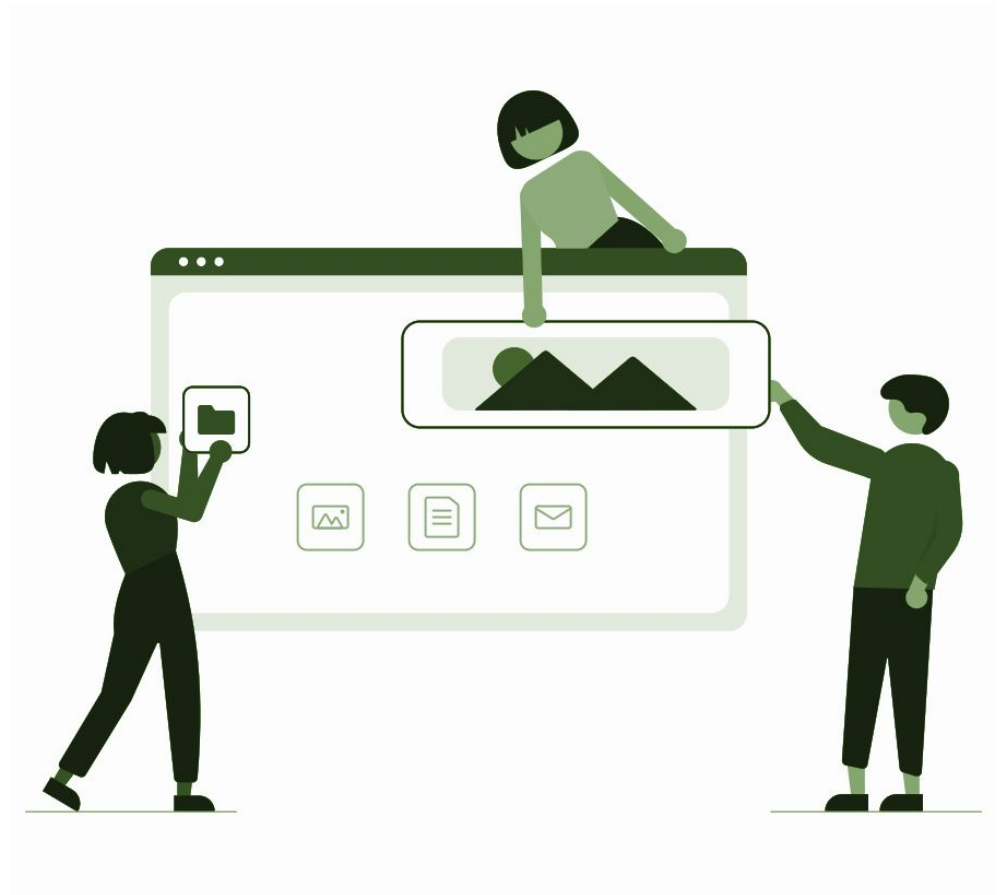
”

Julián Pérez Porto y María Merino - definicion.de



¿Por qué utilizarlas?

- Agilizará tu trabajo
- Te agradecerán
- Te agradecerás





¿Por qué utilizarlas?

- La rutina y automatización reducirán la fatiga por decisión.
- Personalizar es más fácil que construir de cero.
- La reproducibilidad se vuelve mucho más factible.
- Encontrar las cosas se vuelve sencillo.

Instalar Cookiecutter



Cookiecutter

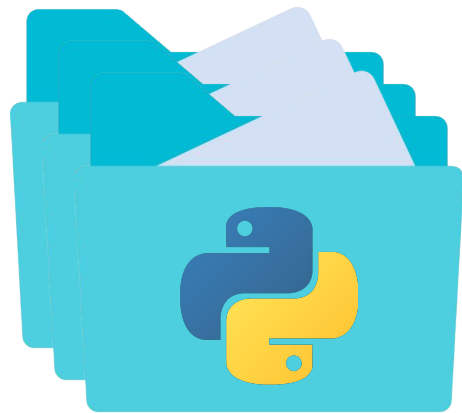




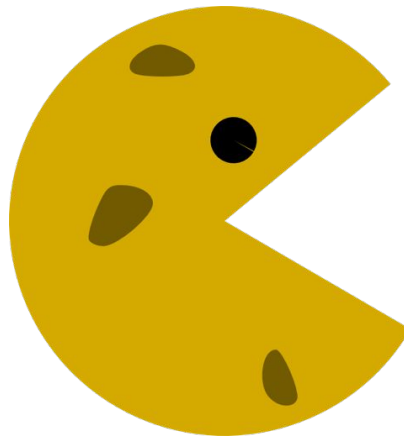
Cookiecutter

- Multiplataforma.
- Plantillas en cualquier lenguaje de programación o formato de marcado.
- Puede ser utilizado como una herramienta de línea de comandos o como una librería de Python.

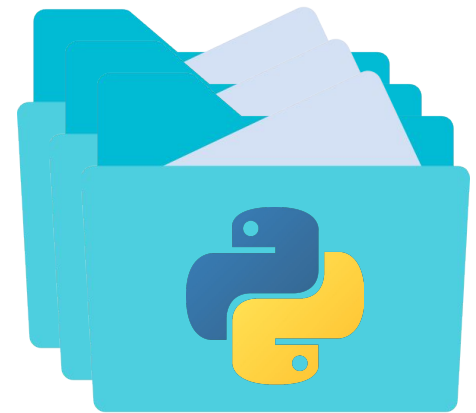
Cookiecutter



Plantilla en
línea o local



Procesamiento y
especificación de
parámetros



Un nuevo
proyecto



Crear plantilla personalizada





Crear plantilla personalizada

```
{{ cookiecutter.saluda }}
```

Hola, científico de datos

```
{% if cookiecutter.eres_asombroso %}
```

¡Genial, por su puesto que lo eres!

```
{% endif %}
```

```
{# ¡Has aprendido bastante! #}
```

Crear plantillas de proyecto personalizadas

Cookiecutter





Estructura de la plantilla

```
@jvelezmagic

cookiecutter-personal
├── README.md
├── cookiecutter.json
├── environment.yml
├── hooks
│   ├── post_gen_project.py
│   └── pre_gen_project.py
├── {{ cookiecutter.project_slug }}
│   ├── .gitignore
│   ├── README.md
│   ├── data
│   │   ├── processed
│   │   │   └── .gitkeep
│   │   └── raw
│   │       └── .gitkeep
│   ├── environment.yml
│   └── notebooks
│       └── 0.0-{{ cookiecutter.project_slug }}-introduction.ipynb
```



Crear plantilla personalizada

`{{ cookiecutter.enhorabuena }}`

¡Lo has hecho increíble!

**Ahora puedes crear plantillas
totalmente personalizadas.**

¡Disfruta creando!



```
@jvelezmagic

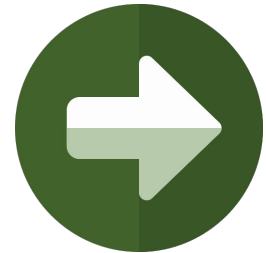
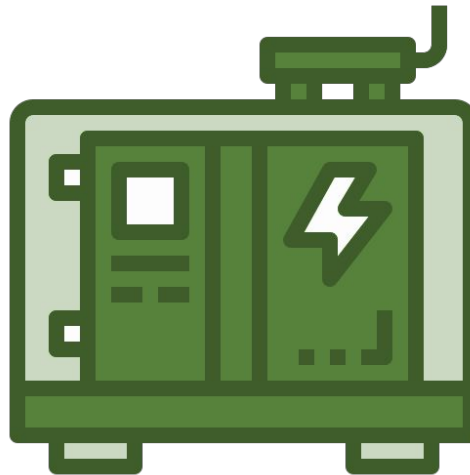
cookiecutter-personal
├── README.md
├── cookiecutter.json
├── environment.yml
├── hooks
│   ├── post_gen_project.py
│   └── pre_gen_project.py
├── {{ cookiecutter.project_slug }}
│   ├── .gitignore
│   ├── README.md
│   ├── data
│   │   ├── processed
│   │   │   └── .gitkeep
│   │   └── raw
│   │       └── .gitkeep
│   ├── environment.yml
│   └── notebooks
│       └── 0.0-{{ cookiecutter.project_slug }}-introduction.ipynb
```

Implementar hooks

Cookiecutter



Implementar hooks



Distribuir plantilla de proyecto



Manejo de rutas: problemática



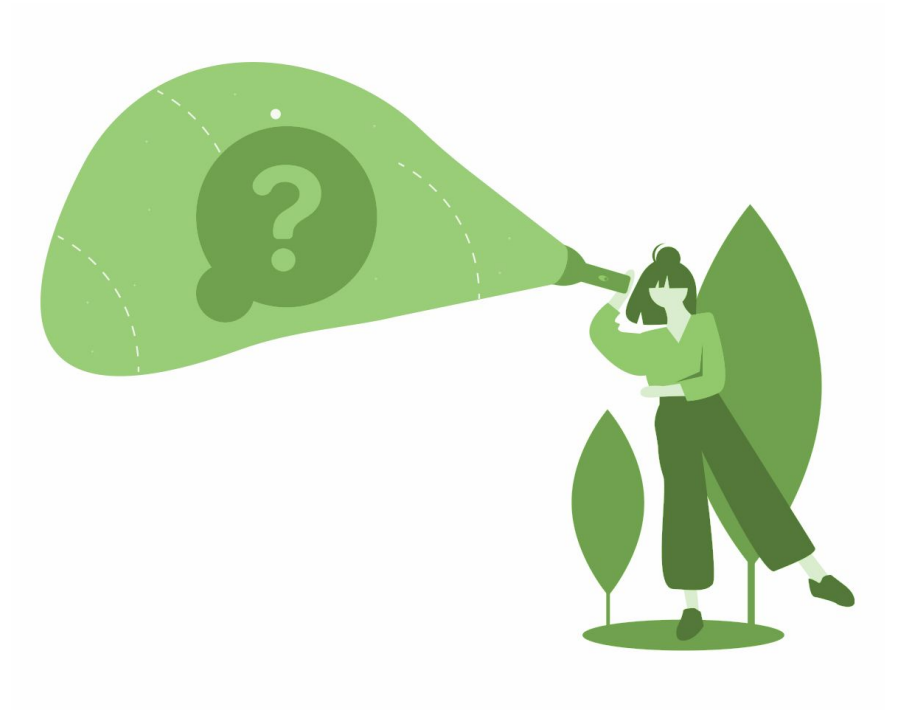
Problemática



Problemática



Problemática





Problemática



Rutas del sistema de archivos





Objetivo

Crear la siguiente ruta
independiente del sistema
operativo:

`./data/raw/`

Manejo de rutas del sistema: OS





○○○

```
os.path.join(".", "data", "raw")
```

Manejo de rutas del sistema: Pathlib





Pathlib

○○○

```
pathlib.Path(".", "data", "raw")
```

Manejo de rutas del sistema: PyFilesystem2





PyFilesystem2: fs

○○○

```
fs.open_fs("./data/raw/")
```

**¡Escoge el que
más se ajuste a
tus necesidades!**



Crear referencias relativas de archivos



Descarga de plantilla y configuración de ambiente virtual

Proyecto final



Utilizar tu proyecto como un módulo de Python

Proyecto final



Entender el flujo de trabajo de los notebooks

Proyecto final



Aplica estas herramientas





Conclusiones

- **Utiliza una plantilla de trabajo.** Agilizará tu productividad, te agradecerán y te agradecerás.
- **Emplea manejadores de rutas.** La manipulación de archivos será controlada y robusta.



Conclusiones

- Al utilizar **todos** los pasos en **armonía** lograrás unificar tu entorno de desarrollo.
- Ahora, podrás ampliar tus herramientas y continuar aprendiendo sin miedo alguno.
Nunca pares de aprender.

Cómo seguir

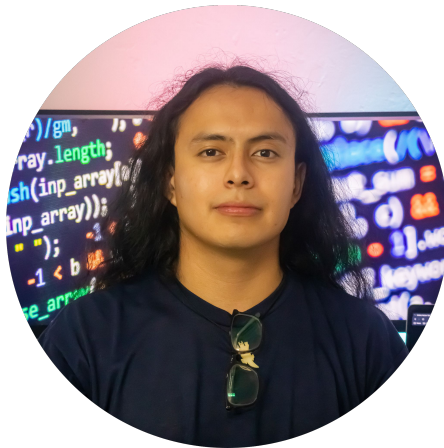
- Construye nuevos proyectos con estas herramientas.
- Avanza con tu aprendizaje de machine learning y deep learning.



🟩 ¡Felicidades!



Conclusiones



@jvelezmagic



jvelezmagic.com