

Preparar el cuaderno de análisis de datos

1. Extrae el ZIP en una carpeta nueva y ábrela en VS Code. Utiliza Python 3.11 o posterior y las extensiones Python y Jupyter del entorno de aula.
2. Abre Terminal → Nueva terminal. Comprueba `python --version` (en Windows también puede ser `py`).
3. Crea un entorno: `python -m venv .env`. No hace falta activar scripts de PowerShell.
4. En Windows instala con `.env\Scripts\python.exe -m pip install -r requirements.txt`; en Linux/macOS, `.env/bin/python -m pip install -r requirements.txt`.
5. Abre `analisis.ipynb`. Arriba selecciona Kernel → Entornos de Python → el intérprete de `.env`. Si no aparece, abre la carpeta correcta y revisa que terminó la instalación.
6. Ejecuta la primera celda con su botón de reproducción. Descarga el Excel original desde UCI dentro de `datos/`. Requiere Internet la primera vez. Si el centro proporciona el archivo, guárdalo exactamente como `datos/Online Retail.xlsx` y la celda reutilizará esa copia.
7. Ejecuta la siguiente celda: deben aparecer 541909 filas y 8 columnas. Lee el diccionario antes de analizar. Si falla una celda, conserva el error y revisa ruta, kernel y paquetes; no conviertas el archivo a CSV para ocultarlo.
8. Las celdas Markdown son para tus decisiones. Edita una con doble clic y ejecútala para ver el texto formateado. Guarda con `Ctrl+S`.
9. Al terminar, reinicia el kernel y ejecuta todo. Esto comprueba que no dependes de variables creadas fuera de orden.

Comprueba que los paquetes estén instalados antes de ejecutar las celdas. También puedes utilizar el mismo notebook en el entorno Jupyter del aula; no necesitas haber cursado Servidor.

Fuente y licencia: <https://archive.ics.uci.edu/dataset/352/online-retail> · Daqing Chen (2015), DOI 10.24432/C5BW33, CC BY 4.0. No se distribuyen datos personales del alumnado. El archivo de origen se descarga de UCI, no se inventa una conversión llamada `online_retail.csv`.