NTT DATA

P4DA

Python for Data Analysis

Pandas

FUTURE AT HEART

Python for Data Analysis

NTT DATA

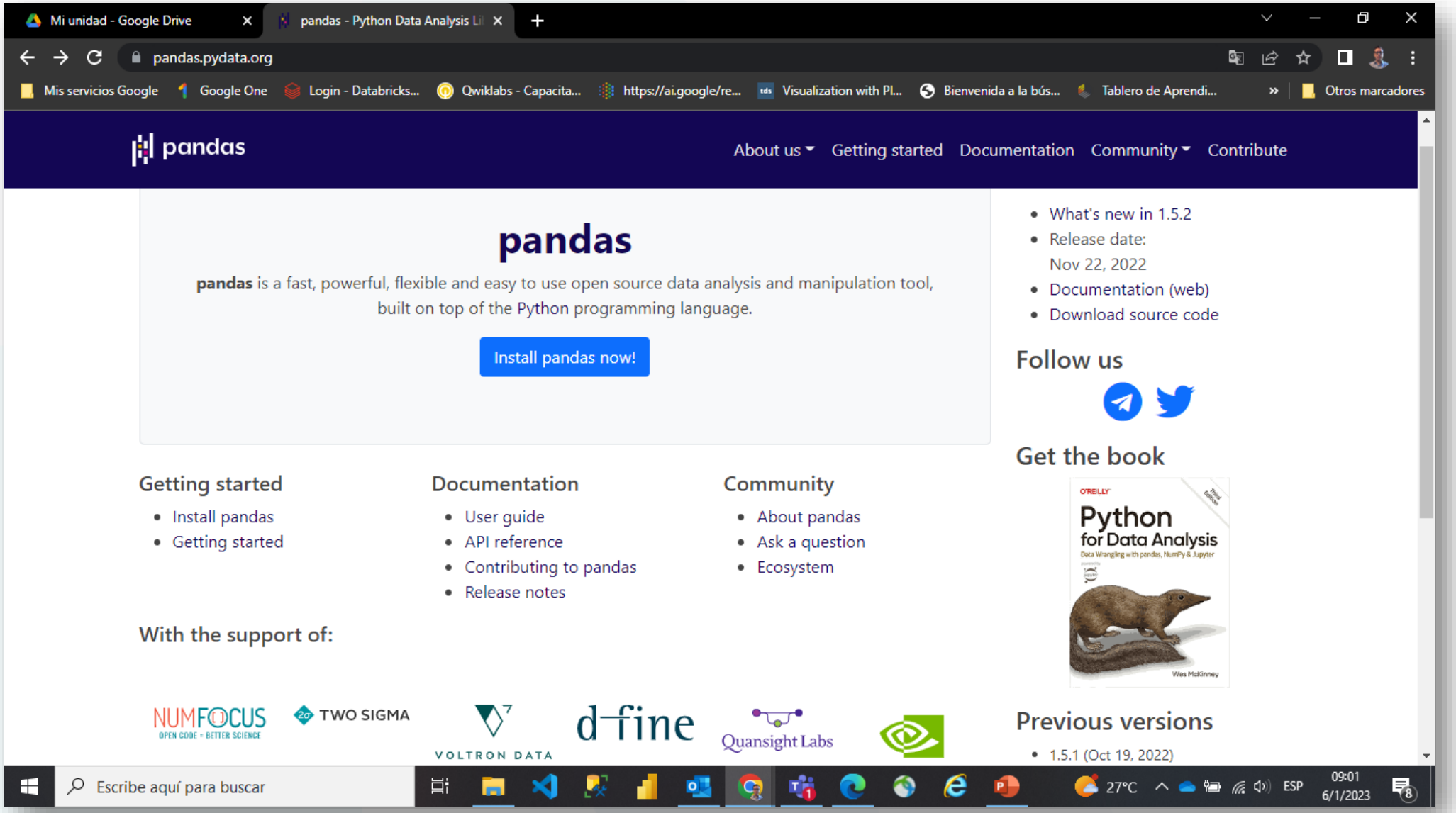
pandas

https://pandas.pydata.org/

Python for Data Analysis

pandas

NTT DATA



Python for Data Analysis

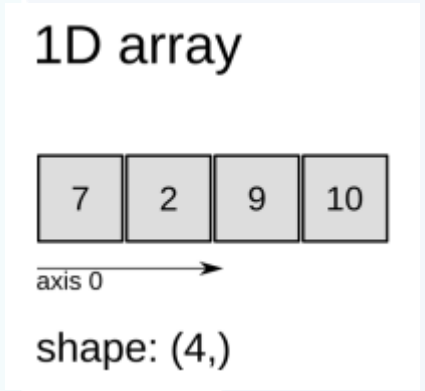
pandas

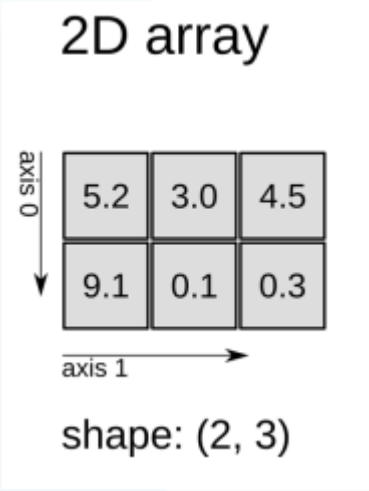
NTT DATA

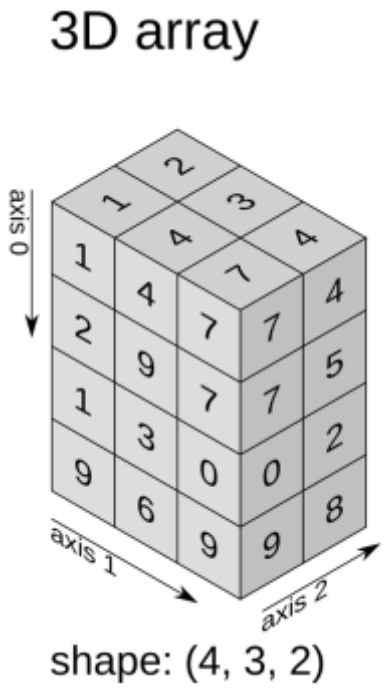
Wes Mckinney

wesmckinney.com/book

Pandas significa Panel de Datos y es una de las bibliotecas más usadas de Python para la manipulación y análisis de los datos, habiendo sido desarrollada inicialmente por Wes McKinney





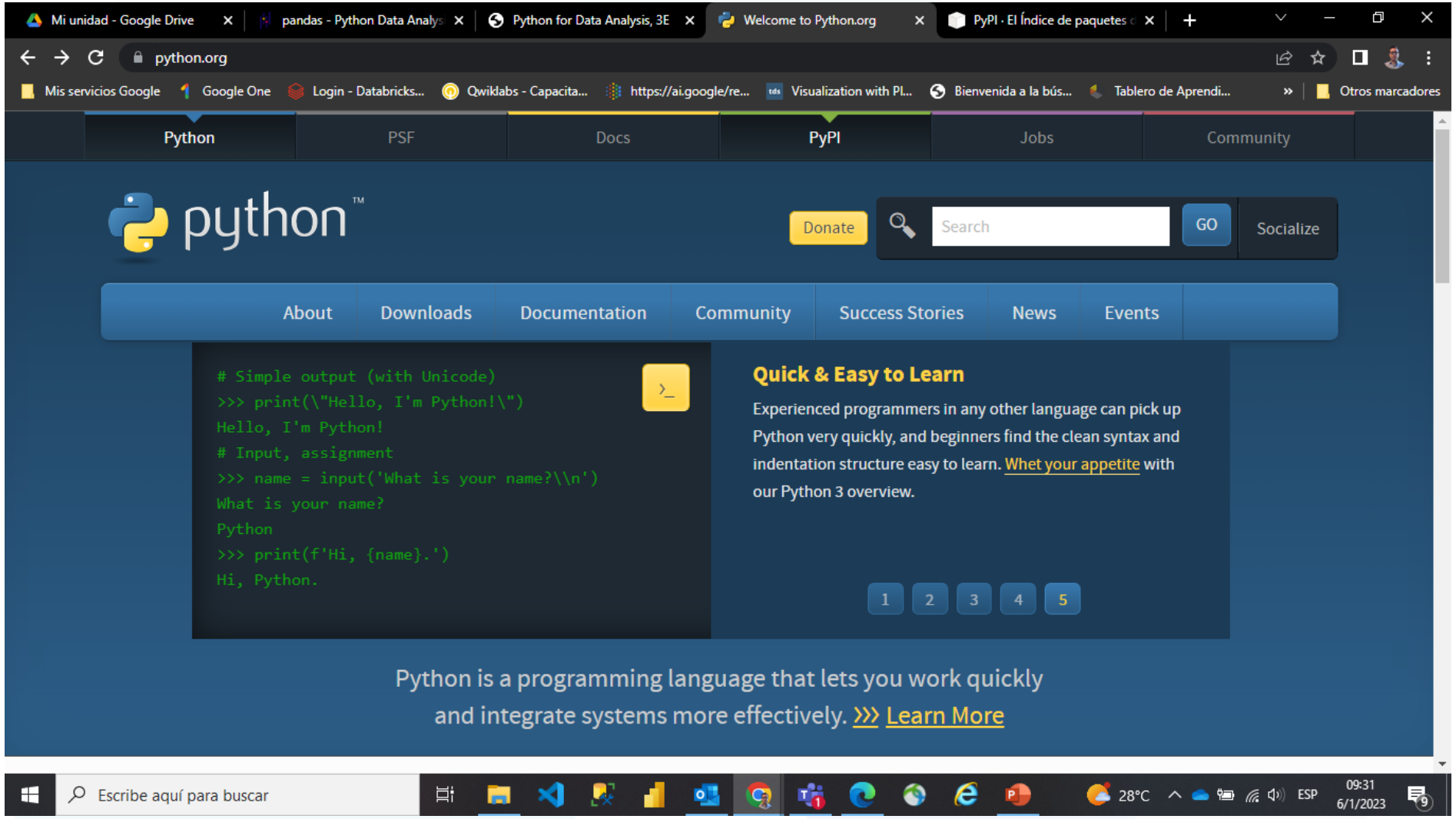


Estructura de Datos de pandas…

Python for Data Analysis

pandas

NTT DATA



¿Por qué Python para análisis de datos?

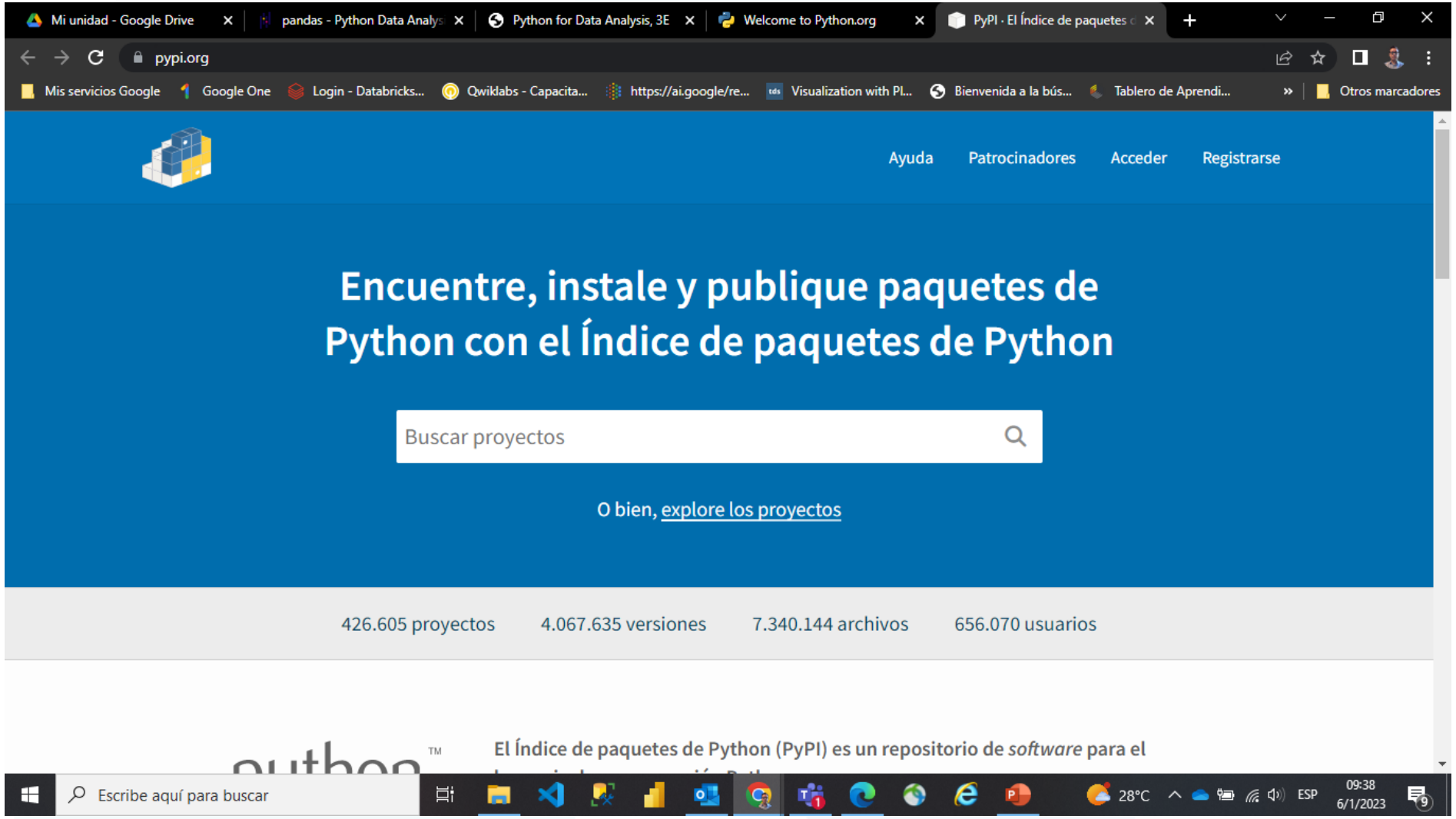
Porque ha desarrollado una gran comunidad científica activa de computación interactiva y análisis de datos (ciencia de datos, aprendizaje automático y desarrollo de soft)

python.org

Python for Data Analysis

pandas

NTT DATA



Paquetes o bibliotecas de Python

pypi.org

Python for Data Analysis

pandas

NTT DATA

Ecosistema de Pandas

Estadística y Machine Learning
- Sklearn-pandas
- Statsmodels
- Featuretools
- Compose

Bibliotecas Gráficas
- Altair
- Plotly
- Seaborn
- Bokeh
- QtPandas

IDE
- Ipython
- Jupyter Notebook
- Jupyter Lab
- Spyder

API
- Pandas Datareader
- pandaSDMX
- fredapi

Dominios específicos
- Geopandas
- Xarray

Otros
- Dask
- Dask-ML
- Koalas
- Odo
- Ray
- Vaex

Limpieza de datos y validación
- Pyjanitor
- Engarde

Extensión tipo de datos
- Ciberpandas
- Pandas-Genomics
- Pint-Pandas

Herramientas de Desarrollo
- Pandas-stubs
- Hamilton

pandas.pydata.org/community/ecosystem

Nos vemos en los Notebooks...

NTT DATA

FUTURE
AT HEART