

# Guía para la entrega taller 2 - Análisis de datos y visualización con Pandas

Por:

Victor Hugo Contreras Ordoñez

[vhcontreraso@unal.edu.co](mailto:vhcontreraso@unal.edu.co)

Para realizar la entrega del taller correctamente por favor siga las siguientes instrucciones:

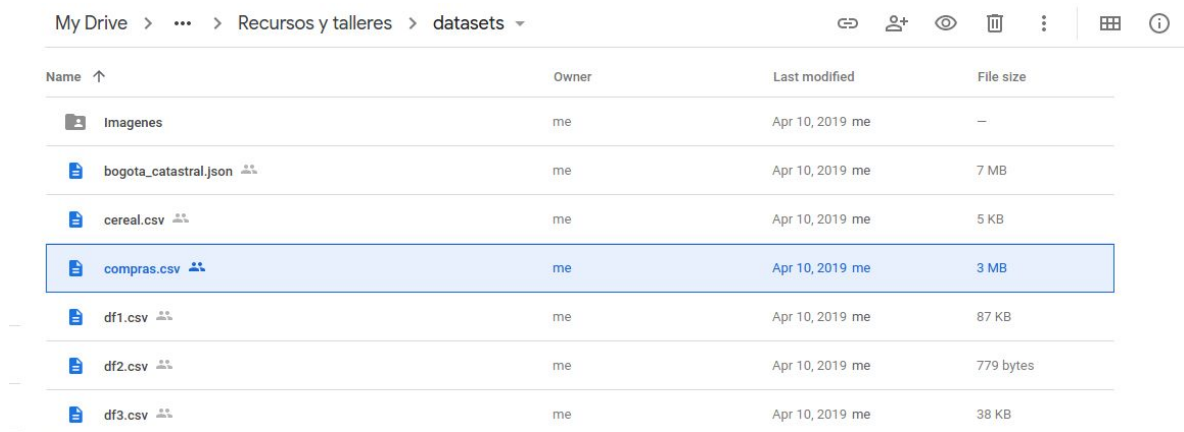
1. Abra el notebook titulado “ejercicio\_compras.ipynb” que se encuentra en la carpeta “M1-S5-Pandas\_2nda\_parte” dentro de la carpeta “Recursos y talleres” del curso.



My Drive > ... > Recursos y talleres > M1-S5-Pandas\_2nda Parte

Name	Owner	Last modified	File size
Ejercicios_compras.ipynb	me	May 2, 2019 me	80 KB
Ejercicios_salarios.ipynb	me	Apr 30, 2019 me	39 KB

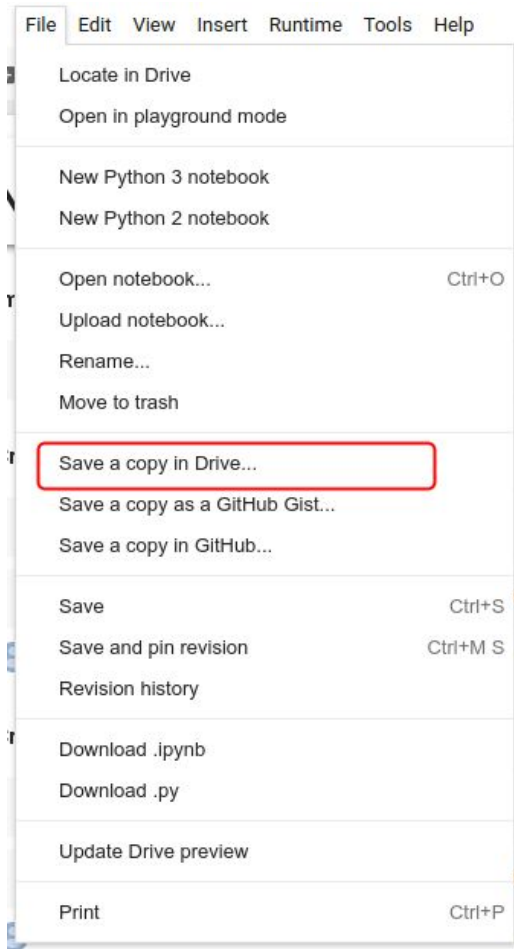
2. Escriba las soluciones a cada uno de los puntos solicitados en el taller (los datasets y archivos necesarios para la solución se encuentran en la carpeta “datasets”), los puntos marcados como “Extra” no son obligatorios pero se sugiere tratar de solucionarlos, para afianzar los conceptos.



My Drive > ... > Recursos y talleres > datasets

Name	Owner	Last modified	File size
Imagenes	me	Apr 10, 2019 me	—
bogota_catastral.json	me	Apr 10, 2019 me	7 MB
cereal.csv	me	Apr 10, 2019 me	5 KB
compras.csv	me	Apr 10, 2019 me	3 MB
df1.csv	me	Apr 10, 2019 me	87 KB
df2.csv	me	Apr 10, 2019 me	779 bytes
df3.csv	me	Apr 10, 2019 me	38 KB

3. Una vez haya solucionado el taller por favor guarde el archivo en su carpeta personal de drive con el comando “file->save” y por favor cambie el nombre del archivo en base a la siguiente estructura: ejercicio\_compras\_XXXXX\_YYYYY.ipynb, donde XXXXX representa su nombre y YYYYYY representa su apellido.



4. Abra el archivo titulado “s6.Visualizacion de datos con Pandas - Ejercicios.ipynb” que se encuentra en la carpeta “M1-S6-Pandas\_3ra Parte-vis” dentro de la carpeta “Recursos y talleres” del curso.

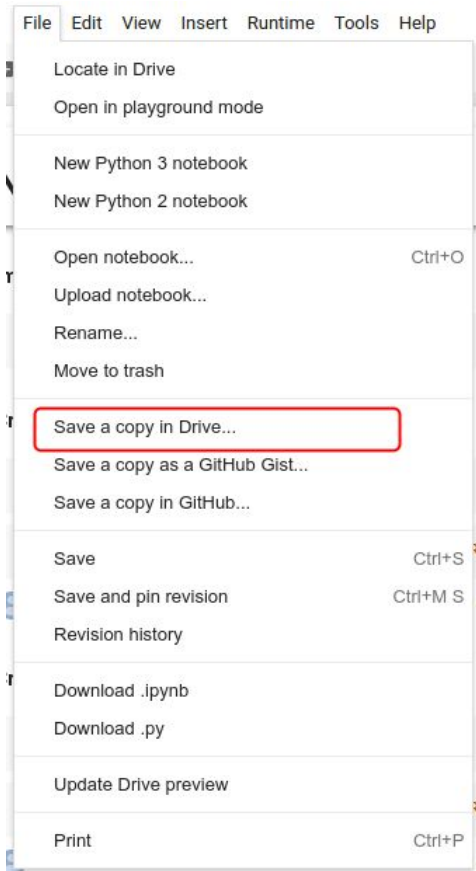
My Drive > ... > Recursos y talleres > M1-S6-Pandas_3ra Parte-vis				🔗	👤	👁	🗑	⋮	📄	ℹ
Name ↑	Owner	Last modified	File size							
s6.Visualizacion de datos con Pandas - Ejercicios.ipynb	me	May 2, 2019 me	391 KB							
s6.Visualizacion de datos con Pandas.ipynb	me	May 2, 2019 me	537 KB							

5. Solucione cada uno de los puntos solicitados en el notebook (recuerde que los datasets necesarios para solucionar el taller se encuentran en la carpeta “datasets”). Los puntos marcados como “Extra” no son obligatorios pero se sugiere desarrollarlos para afianzar y profundizar en los conocimientos del tema visto.

My Drive > ... > Recursos y talleres > datasets

Name ↑	Owner	Last modified	File size
Imagenes	me	Apr 10, 2019 me	—
bogota_catastral.json	me	Apr 10, 2019 me	7 MB
cereal.csv	me	Apr 10, 2019 me	5 KB
compras.csv	me	Apr 10, 2019 me	3 MB
df1.csv	me	Apr 10, 2019 me	87 KB
df2.csv	me	Apr 10, 2019 me	779 bytes
df3.csv	me	Apr 10, 2019 me	38 KB
flights.csv	me	Apr 10, 2019 me	2 KB
game_of_thrones.json	me	Apr 10, 2019 me	46 KB

6. Una vez finalizado el taller por favor guarde todos los cambios en su carpeta personal de drive con el comando “file->save” y por favor cambie el nombre del archivo en base a la siguiente estructura: “s6.Visualizacion de datos con Pandas - Ejercicios\_xxxx\_yyyy.ipyn” donde xxxx corresponde a su nombre y yyyy corresponde a su apellido.



7. Una vez finalizados y verificados los ejercicios previos por favor abra el siguiente enlace y realice la carga de los dos archivos con las soluciones descritos en los numerales 3 y 6: <https://www.dropbox.com/request/dCL8UrcAVZ0bv1gPhXIR>



8. Felicitaciones has finalizado tu segundo taller.

## Recursos

- Introducción a pandas [https://programacion.net/articulo/introduccion\\_a\\_pandas\\_1632](https://programacion.net/articulo/introduccion_a_pandas_1632)
- Pandas documentation <https://pandas.pydata.org/>
- Python basic tutorial <https://www.learnpython.org/es/Pandas%20Basics>
- Introduction to pandas library  
<https://towardsdatascience.com/a-quick-introduction-to-the-pandas-python-library-f1b678f34673>
- Change point size in scatter plot  
<https://stackoverflow.com/questions/14827650/pyplot-scatter-plot-marker-size>
- Matplotlib documentation [https://matplotlib.org/api/\\_as\\_gen/matplotlib.pyplot.scatter.html](https://matplotlib.org/api/_as_gen/matplotlib.pyplot.scatter.html)
- Scatter plot  
<https://jakevdp.github.io/PythonDataScienceHandbook/04.02-simple-scatter-plots.html>
- Visualización con pandas  
[https://pandas.pydata.org/pandas-docs/stable/user\\_guide/visualization.html](https://pandas.pydata.org/pandas-docs/stable/user_guide/visualization.html)
- Data visualization and exploration with pandas  
<https://towardsdatascience.com/data-visualization-exploration-using-pandas-only-beginner-a0a52eb723d5>

- Pandas and matplotlib data exploration  
<https://towardsdatascience.com/a-guide-to-pandas-and-matplotlib-for-data-exploration-56fad95f951c>