"Introduction to pandas "

What is Pandas and what is it used for?
How to install Pandas?

What is Pandas?

pandas

What is Pandas in Python?

Pandas is a Python library that specializes in data management, analysis and processing. To do this, it is based on the data structures or arrays of the NumPy library.

pandas

How to install Pandas?

Use the command py **-m pip --version** to see if you have the PIP package manager installed. Use the command py -m **pip install** numpy to install NumPy and finally use **py -m pip install** pandas to install Pandas.

pandas

How to install Pandas?

Step	Command	Description
0	py --version	(Optional) This is just to ensure that Python is installed and to know which version is installed. If you are sure you have Python, this step is not necessary.
1	py -m pip --version	PIP is the most popular package manager for Python. In this step you make sure you have it installed and know the installed version.
2	py -m pip install --upgrade pip	(Optional) Update the PIP version just in case.
3	py -m pip install numpy	Use PIP to install NumPy, the library Pandas relies on for data structures.
4	py -m pip install pandas	Install Pandas with PIP. You could just run this command from the start, but you better make sure you have the other packages installed so you don't run into any errors.
5	py -m pip install matplotlib	(Optional) Install Matplotlib, a library that uses Pandas in the plot() function to create graphs from the data being processed.

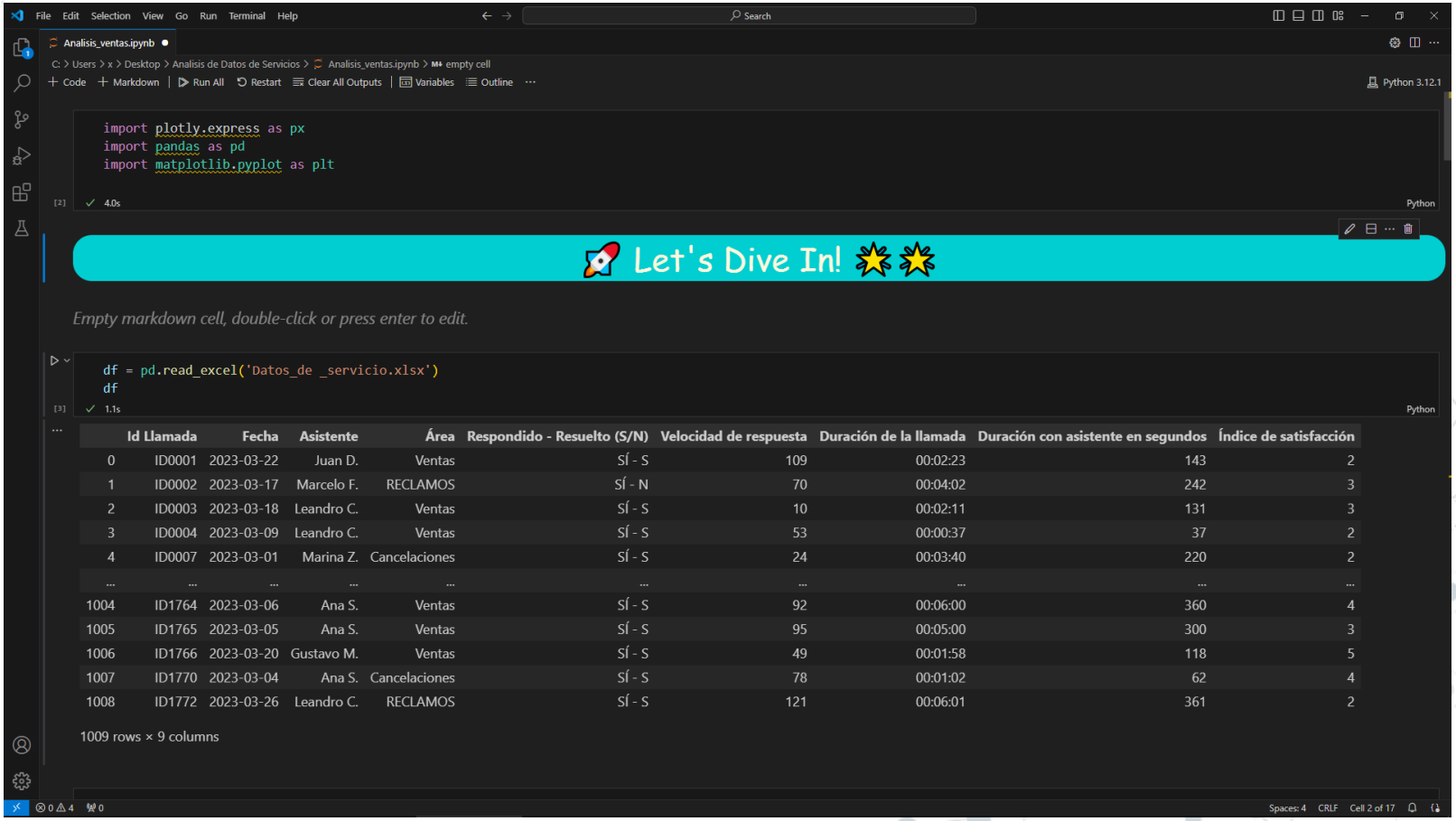
pandas

First we are going to import the libraries we need, such as pandas, matplotlib.
We are going to read our file, in this case it is called “Service_Data.xlsx”, as we do with:

df *- Es el dataframe*

df=pd.read_Excel('Datos_de_servicios.xlsx')

pd.read *- Lee el archivo de Excel*



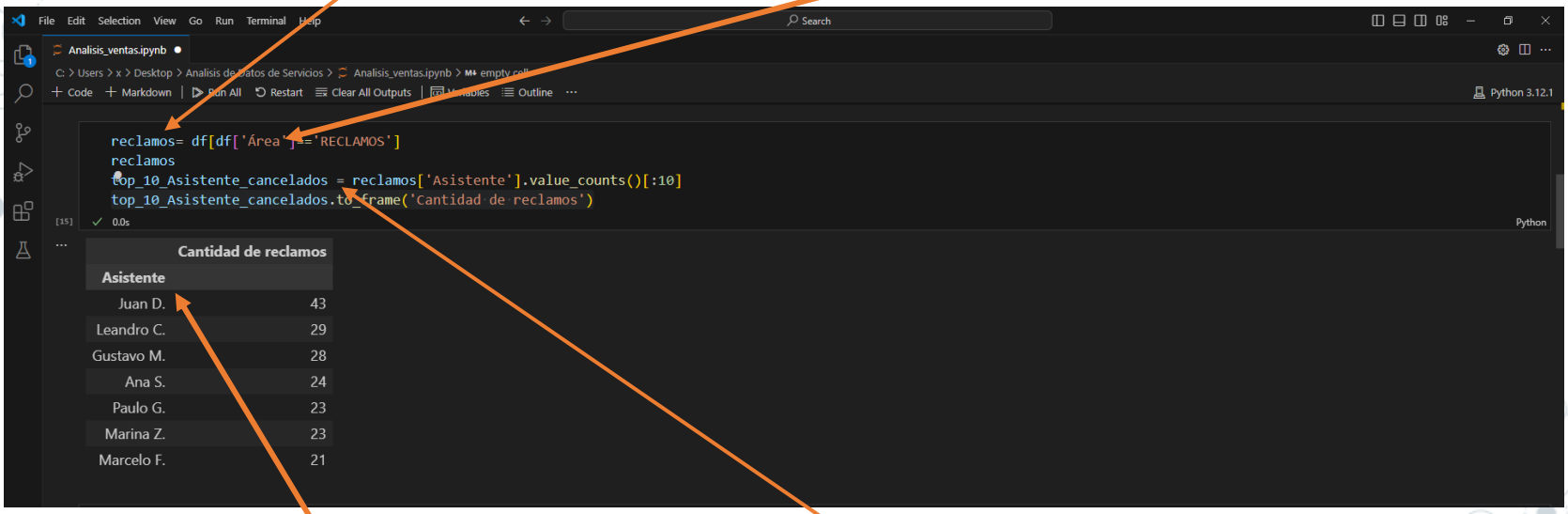
	Id Llamada	Fecha	Asistente	Área	Respondido - Resuelto (S/N)	Velocidad de respuesta	Duración de la llamada	Duración con asistente en segundos	Índice de satisfacción
0	ID0001	2023-03-22	Juan D.	Ventas	SÍ - S	109	00:02:23	143	2
1	ID0002	2023-03-17	Marcelo F.	RECLAMOS	SÍ - N	70	00:04:02	242	3
2	ID0003	2023-03-18	Leandro C.	Ventas	SÍ - S	10	00:02:11	131	3
3	ID0004	2023-03-09	Leandro C.	Ventas	SÍ - S	53	00:00:37	37	2
4	ID0007	2023-03-01	Marina Z.	Cancelaciones	SÍ - S	24	00:03:40	220	2
...	...	...	...	...	...	...	...	...	...
1004	ID1764	2023-03-06	Ana S.	Ventas	SÍ - S	92	00:06:00	360	4
1005	ID1765	2023-03-05	Ana S.	Ventas	SÍ - S	95	00:05:00	300	3
1006	ID1766	2023-03-20	Gustavo M.	Ventas	SÍ - S	49	00:01:58	118	5
1007	ID1770	2023-03-04	Ana S.	Cancelaciones	SÍ - S	78	00:01:02	62	4
1008	ID1772	2023-03-26	Leandro C.	RECLAMOS	SÍ - S	121	00:06:01	361	2

1009 rows × 9 columns

We analyze the data we have in the file and we want to obtain the data.
In this case we are going to analyze the top 10 complaints. How do we do it:

Name of the column and variable where the data is stored

Name of the column to be analyzed

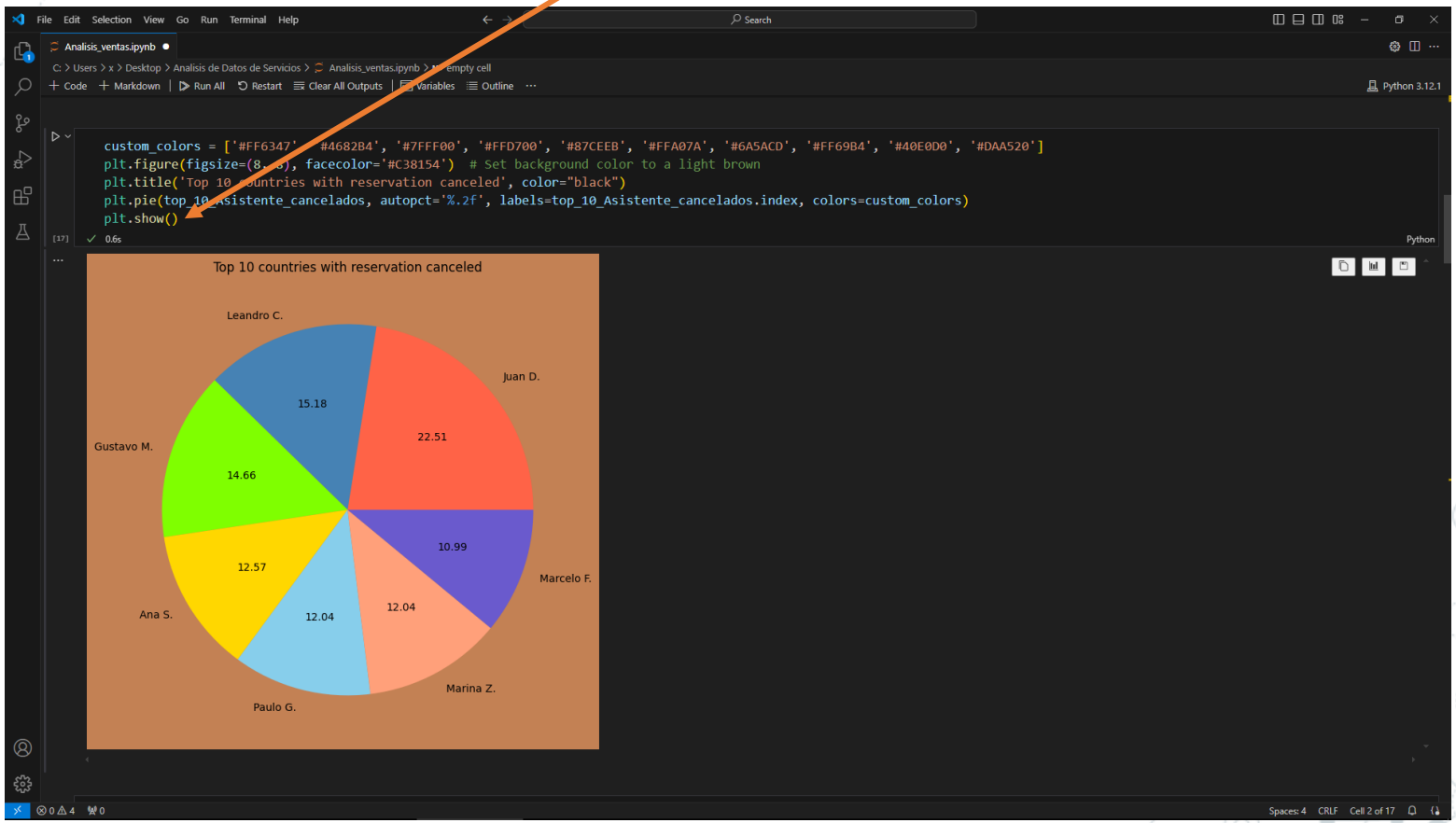


The 10 Names of the Attendees

A variable is created where we store theTop 10.

After analyzing the data, we visualize it.We use matplotlib.
To visualize the graph we write:

plt.show()



THANKS

https://github.com/Glaramos