



Intro to Google Colaboratory

PERE-PAU VÁZQUEZ & JOSE DÍAZ

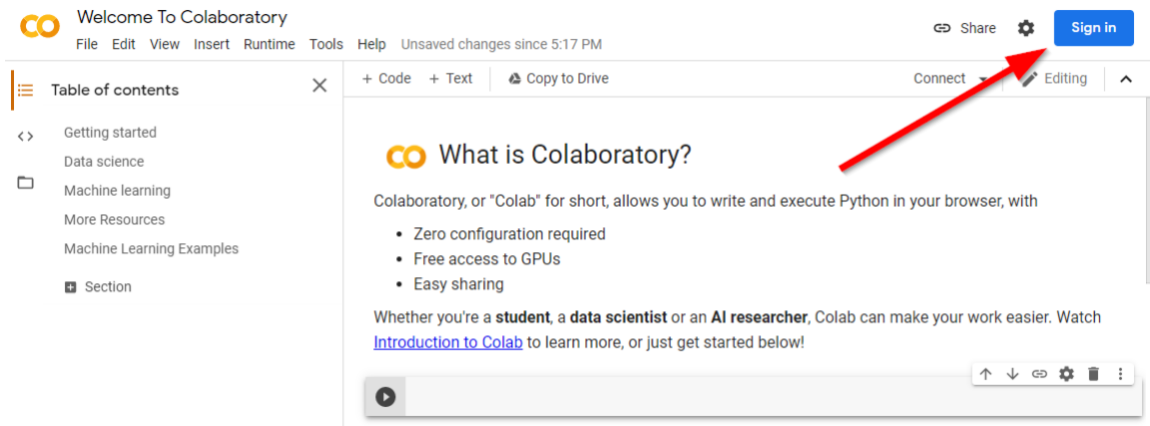
Information Visualization | 2021

Getting Started with Google Colaboratory

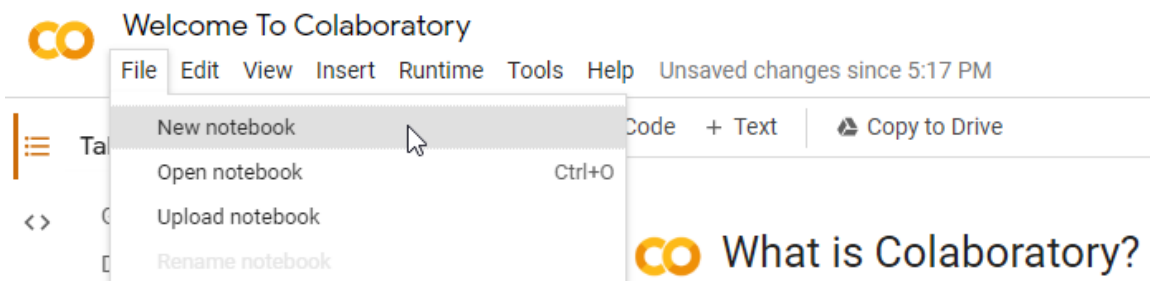
Google Colaboratory, or “Colab”, allows its users to write and execute Python code in the browser, with no need of tedious configurations, provides free access to GPUs and it’s easy to share. To work with Colab you just need a Google Account and a web browser to access the following website:

<https://colab.research.google.com/notebooks/intro.ipynb>

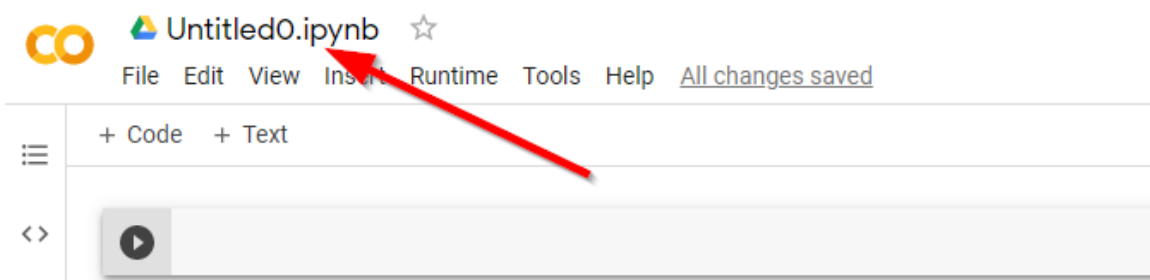
There, we find a quick tutorial and many examples that show the possibilities that Colab offers to the users. The first thing to do is signing in (top right corner).



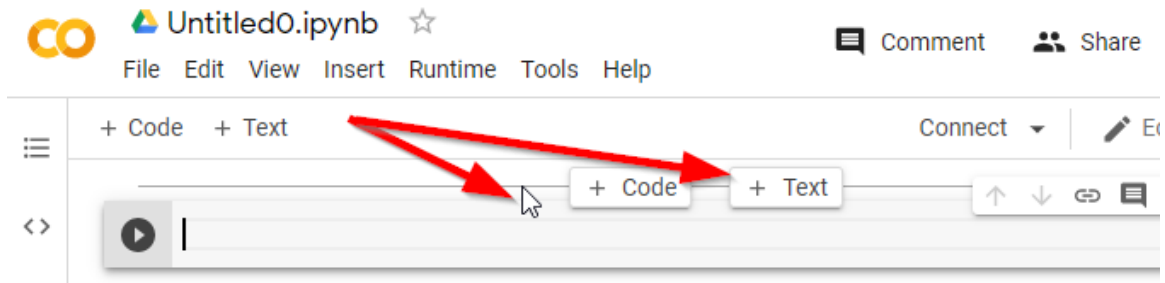
The work in Colab is organized in *notebooks*. These can be saved in Google Drive, Github, or be downloaded as python files to be executed by a Python interpreter. To create a notebook, we have to select *New notebook* in the *File* menu.



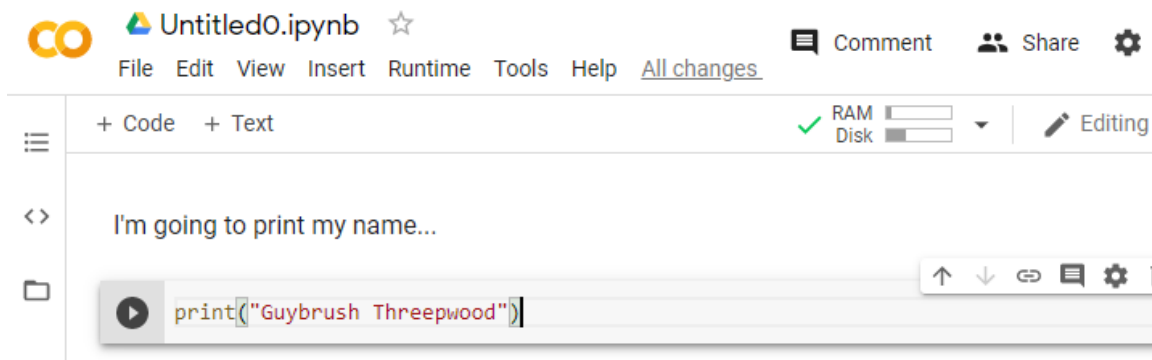
Each notebook consists of text labels and code snippets. We can insert them by using the buttons *+Text* and *+Code* respectively. As well, we can change the name of the notebook by clicking on the title.



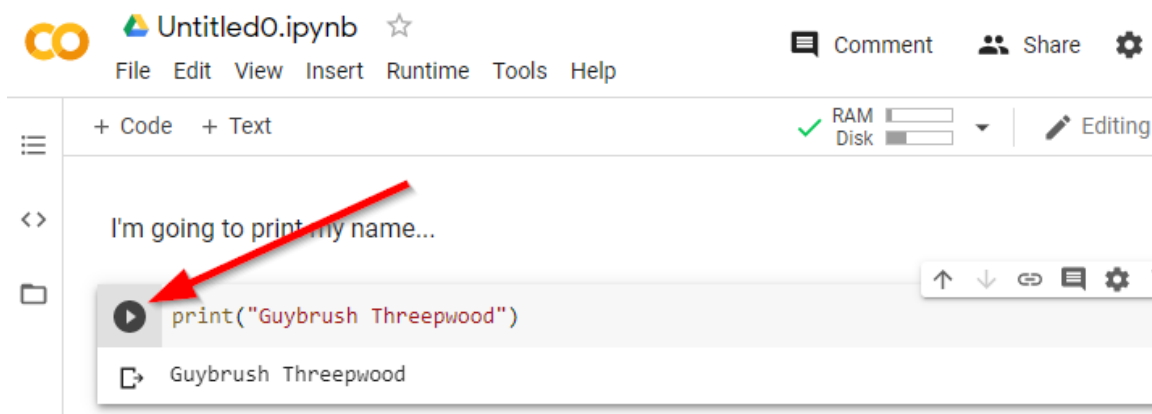
Let's try to add a text label to our notebook and insert some Python code to print your name. Note that text labels and code snippets may also be added placing the cursor at the desired position and clicking the buttons that appear.



Your notebook should be similar to the next image:

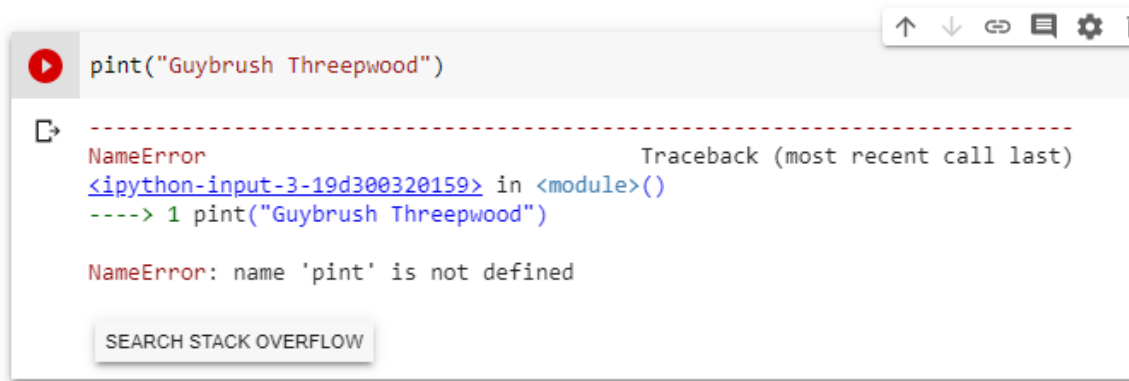


If we want to execute a code snippet, we have to press the play button, and the result appears below.



If the code is not correct, an error message appears when running the code.

I'm going to print my name...



The screenshot shows a Jupyter Notebook cell with a red play button icon. The code in the cell is `print("Guybrush Threepwood")`. Below the code, a red dashed line separates the code from the error message. The error message is a `NameError` with the text "name 'print' is not defined". Above the error message, it says "Traceback (most recent call last)". Below the error message, there is a button that says "SEARCH STACK OVERFLOW".

```
print("Guybrush Threepwood")
```

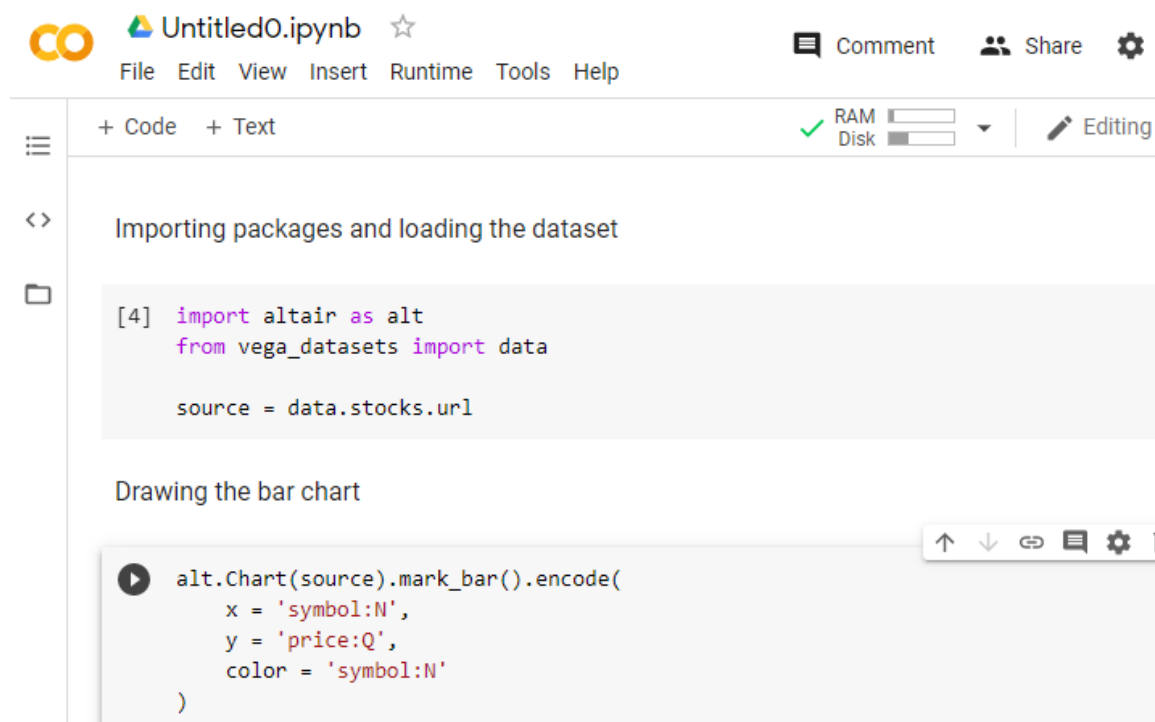
`NameError` Traceback (most recent call last)
`<ipython-input-3-19d300320159> in <module>()
----> 1 print("Guybrush Threepwood")`

`NameError: name 'print' is not defined`

SEARCH STACK OVERFLOW

Let's create a new code snippet and introduce the code to generate the chart shown in the previous section. If you do it correctly, you will see the chart of the image below the code snippet.

One last consideration: the code introduced in one snippet can be reused by the following ones. For instance, there is no need of importing Altair in every snippet. We can do it once at the beginning before running the rest of the snippets:



The screenshot shows the Jupyter Notebook interface. At the top, there is a toolbar with icons for File, Edit, View, Insert, Runtime, Tools, and Help. Below the toolbar, there is a sidebar on the left with icons for a file explorer, a code editor, and a text editor. The main area of the notebook contains two code snippets. The first snippet is titled "Importing packages and loading the dataset" and contains the following code:

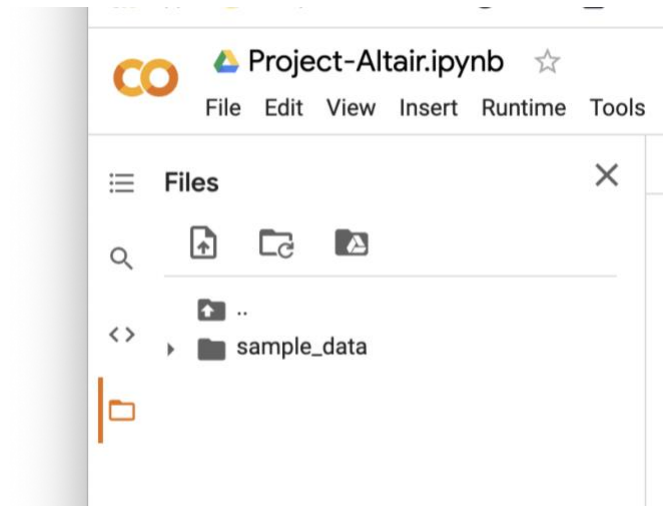
```
[4] import altair as alt  
    from vega_datasets import data  
  
    source = data.stocks.url
```

The second snippet is titled "Drawing the bar chart" and contains the following code:

```
alt.Chart(source).mark_bar().encode(  
    x = 'symbol:N',  
    y = 'price:Q',  
    color = 'symbol:N'  
)
```

IMPORTANT! It's mandatory to run the "import snippet" before executing the others that depend on it. Otherwise, you will get an error!!

There are different ways of uploading the files you want to use in your projects. The left panel, for example, will let you see a file system:



You can use the buttons to upload your data...

But this can also be achieved through code, by using the corresponding libraries:

```
[ ] import io
    from google.colab import files
    import pandas as pd
    import altair as alt
    from vega_datasets import data
```

And...

```
[ ] uploaded = files.upload()
```

Choose files No file chosen Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to show the upload widget.

Saving United_States_COVID-19_Cases_and_Deaths_by_State_over_Time.csv to United_States_COVID-19_Cases_and_Deaths_by_State_over_Time.csv

```
[ ] df = pd.read_csv(io.BytesIO(uploaded["United_States_COVID-19_Cases_and_Deaths_by_State_over_Time.csv"]), sep=";")
```