Colab Tutorial

Session TS00: Colab Basics

Objective

- Begin developing Python / Google Colab skills for statistical analysis.

Engineering Scenario

- A sample of 25 batteries have been tested to failure on a deep discharge cycle test.

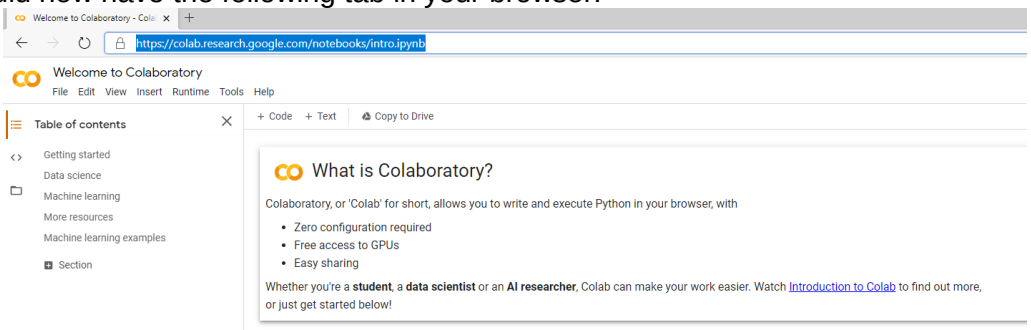
Tutorial Task

- Load the data for the batteries into Colab, plot it, edit it, and download the edited data.

Tutorial Guideline

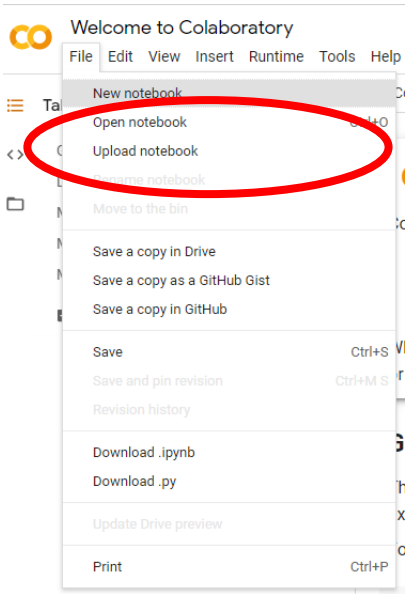
Colab is a cloud-based data analysis system and this link will take you to the introductory Colab notebook:
https://colab.research.google.com/notebooks/intro.ipynb

You should now have the following tab in your browser:

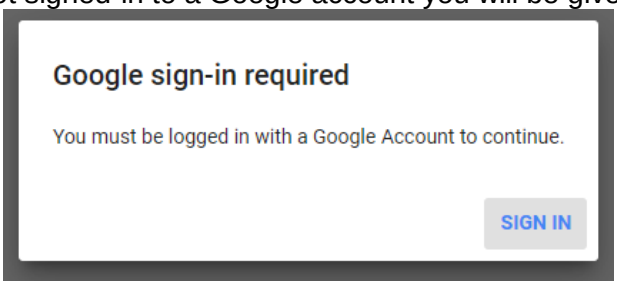


Scrolling through this notebook, you will find some simple code cells that you can execute for yourself and some links to Colab documentation.

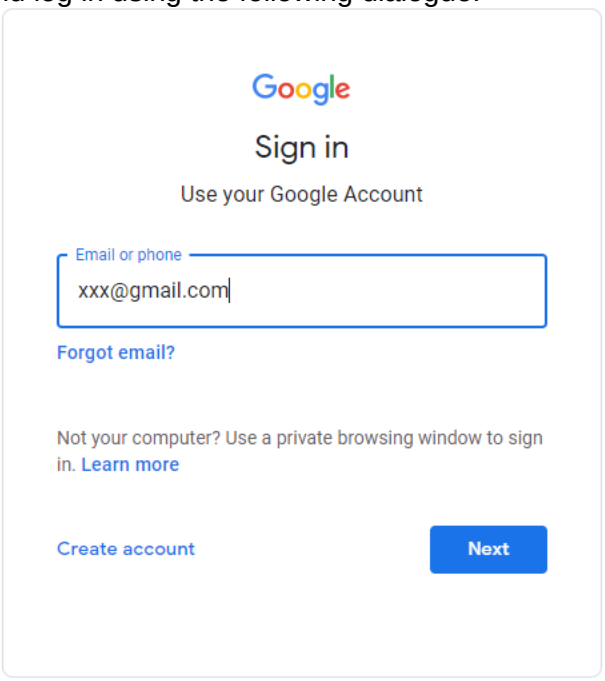
The first task is to create a new Colab notebook by selecting it from the menu, "**File > New notebook**":



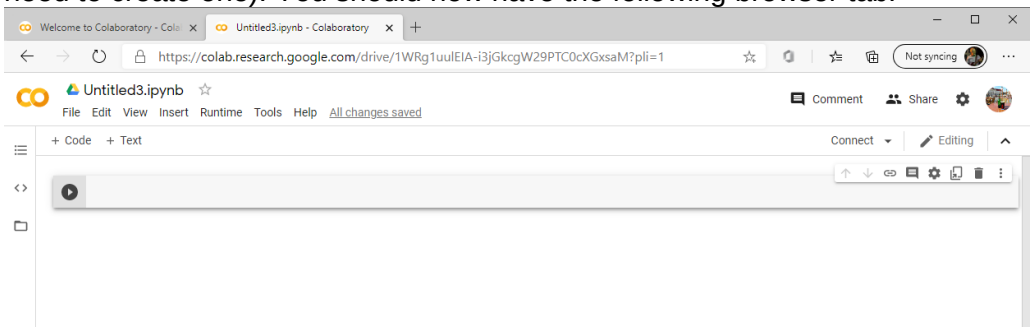
At this point, if you are not signed-in to a Google account you will be given the following prompt:



Click the sign in button and log in using the following dialogue:



Click the "**Next**" button and enter your password (if you don't have a Google account you will need to create one). You should now have the following browser tab:



Don't worry if the notebook has a different number in its title because we are going to rename it anyway by choosing "**File > Rename notebook**" and replacing "Untitled3.ipynb" with "Tutorial00.ipynb".

Basics of Colab Python Scripts

(1) Print hello world

```
print ('Hello World')
```

```
Hello World
```

(2) Print your Name and age

```
Name = "Your name"
age = 36
print (Name,age)
```

```
Your name 36
```

(3) Using input function to insert your name

```
input("What is Your Name? ")
```

```
What is Your Name?
```

(4) Insert your date of birth, then calculate your age.

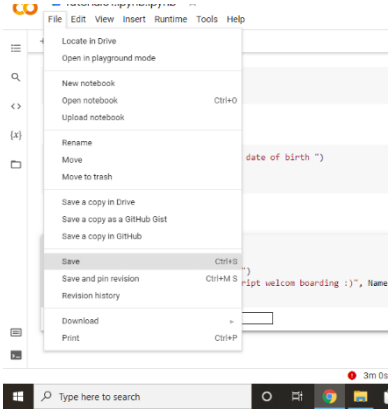
```
year_of_birth = input("What is your year of birth ")
Age = 2022 - int(year_of_birth)
print (Age)
```

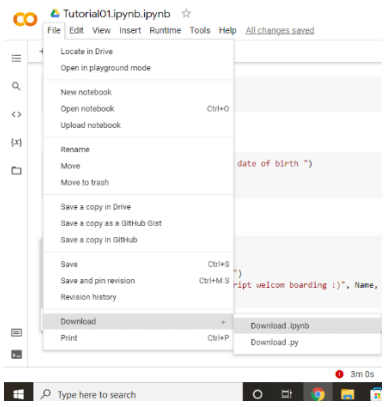
```
What is your year of birth
```

(5) Simple example - Welcome on aboard

```
[5]  Name = input("What is your name? ")
     Age = input("What is your age? ")
     Day = input("What is the date of today?")
     print("Welcome on board \n", "Name \n", Name, "Age \n", Age, "Date of today \n", Day)
```

(6) Now let's save and download the script to share it

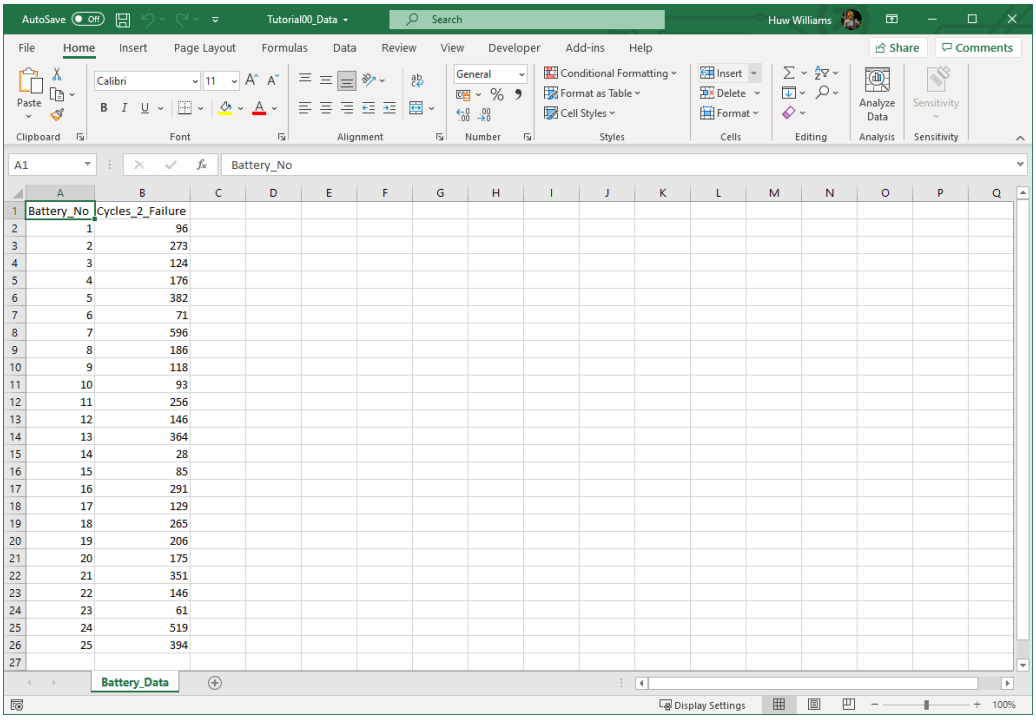




Download the script as ipynb

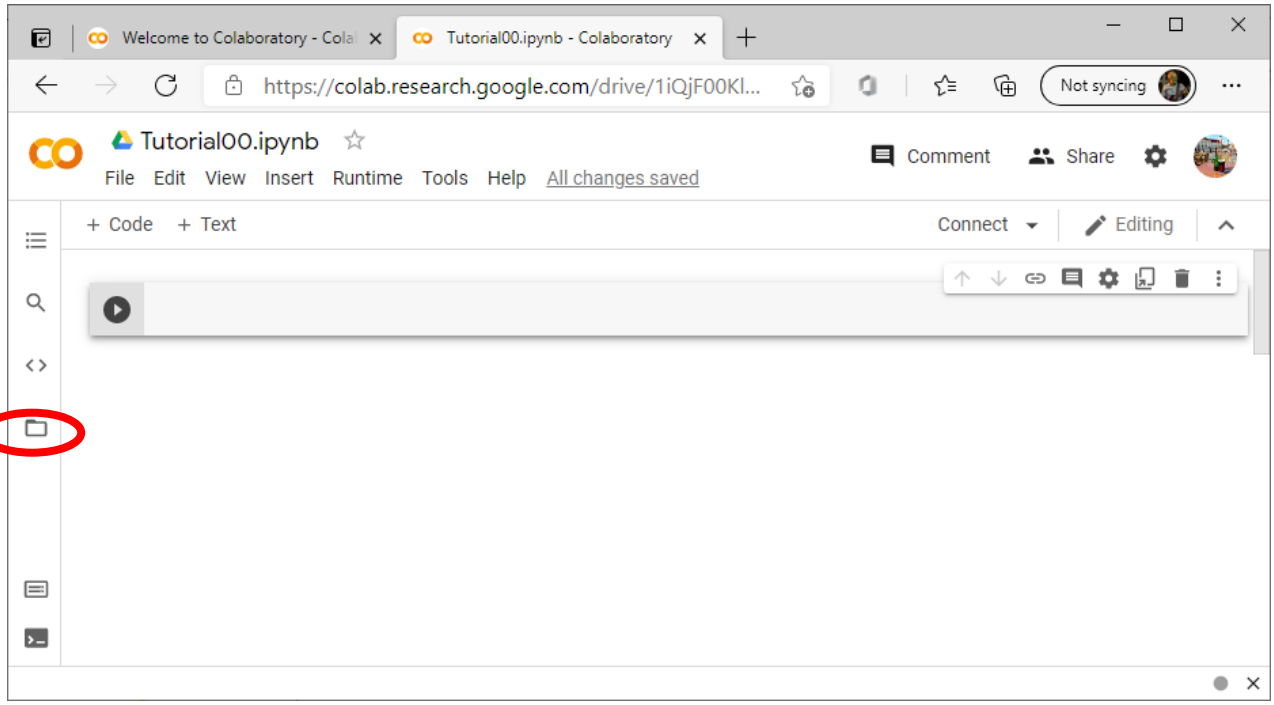
Engineering Scenario – Battery Test Data

We are now going to upload some data to Colab so that we can analyse it. The data is in the Excel workbook “Tutorial00_Data”. You need to download the file from the Teams / SharePoint link provided to you and save it into a folder (either local drive or network location) on your system!
If you open the workbook in Excel you will see the data as below:

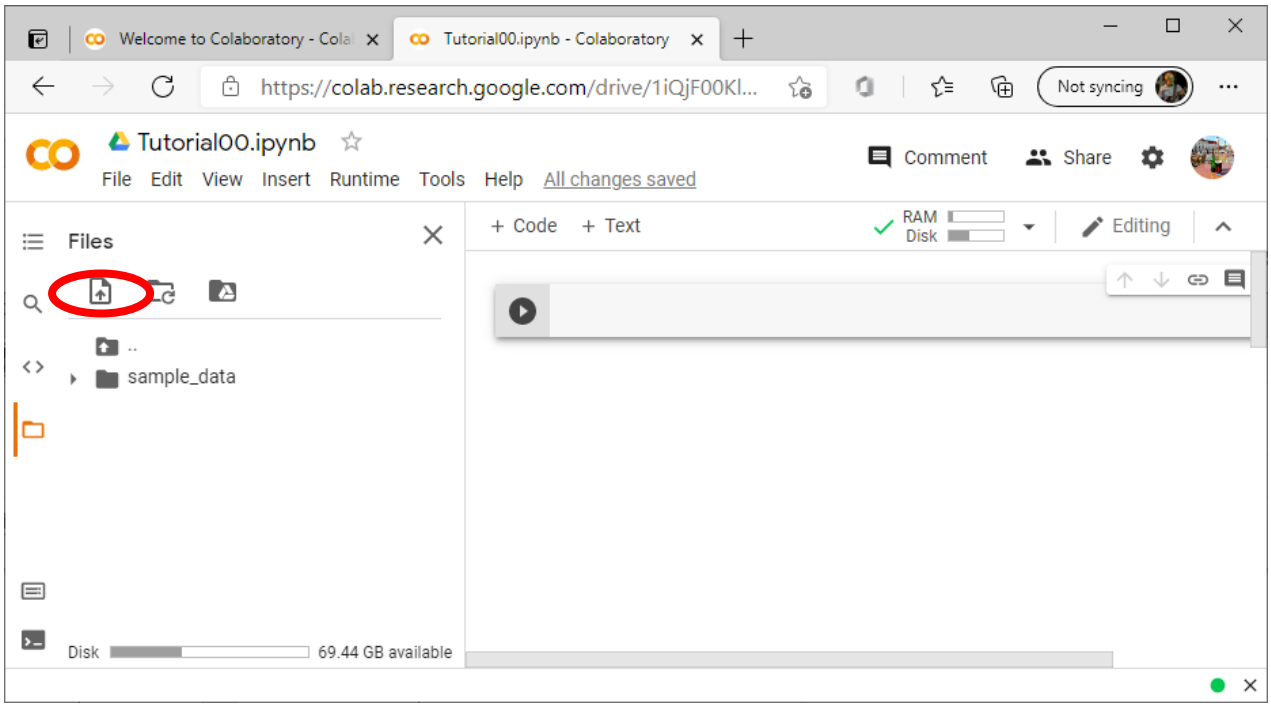


Battery_No	Cycles_2_Failure
1	96
2	273
3	124
4	176
5	382
6	71
7	596
8	186
9	118
10	93
11	256
12	146
13	364
14	28
15	85
16	291
17	129
18	265
19	206
20	175
21	351
22	146
23	61
24	519
25	394

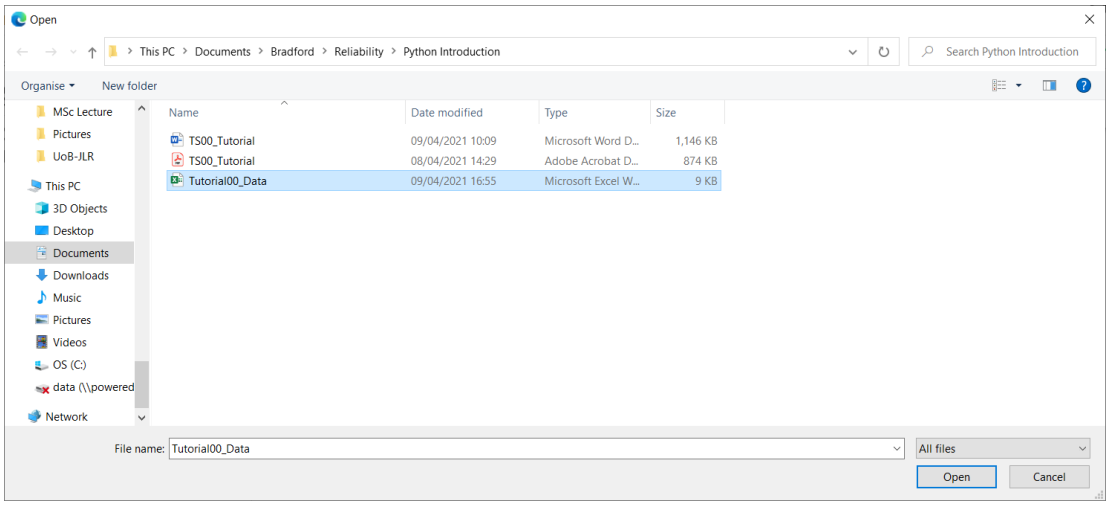
The next step is to upload the Excel file to Colab and we do this by first clicking on the Colab file icon:



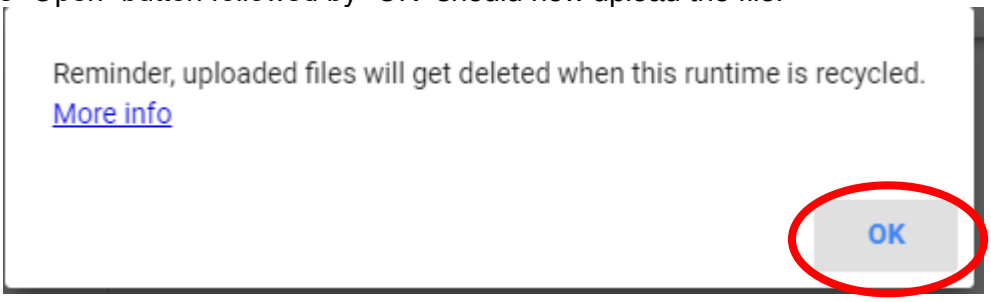
You will get this screen:

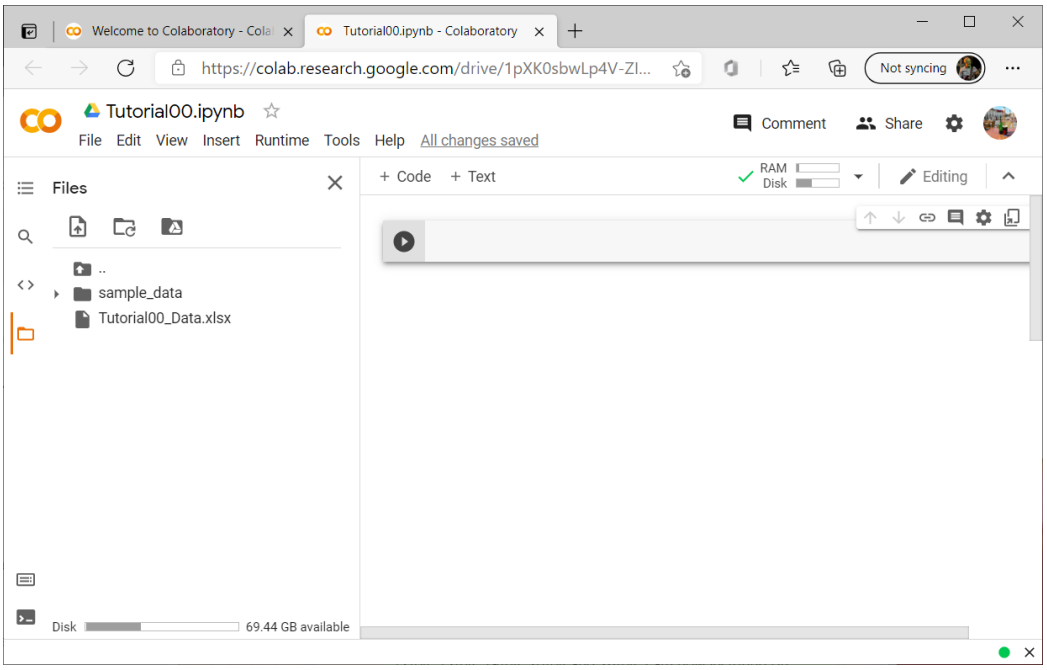


You should now click on the “upload” icon and navigate to the Excel file that you have just viewed:



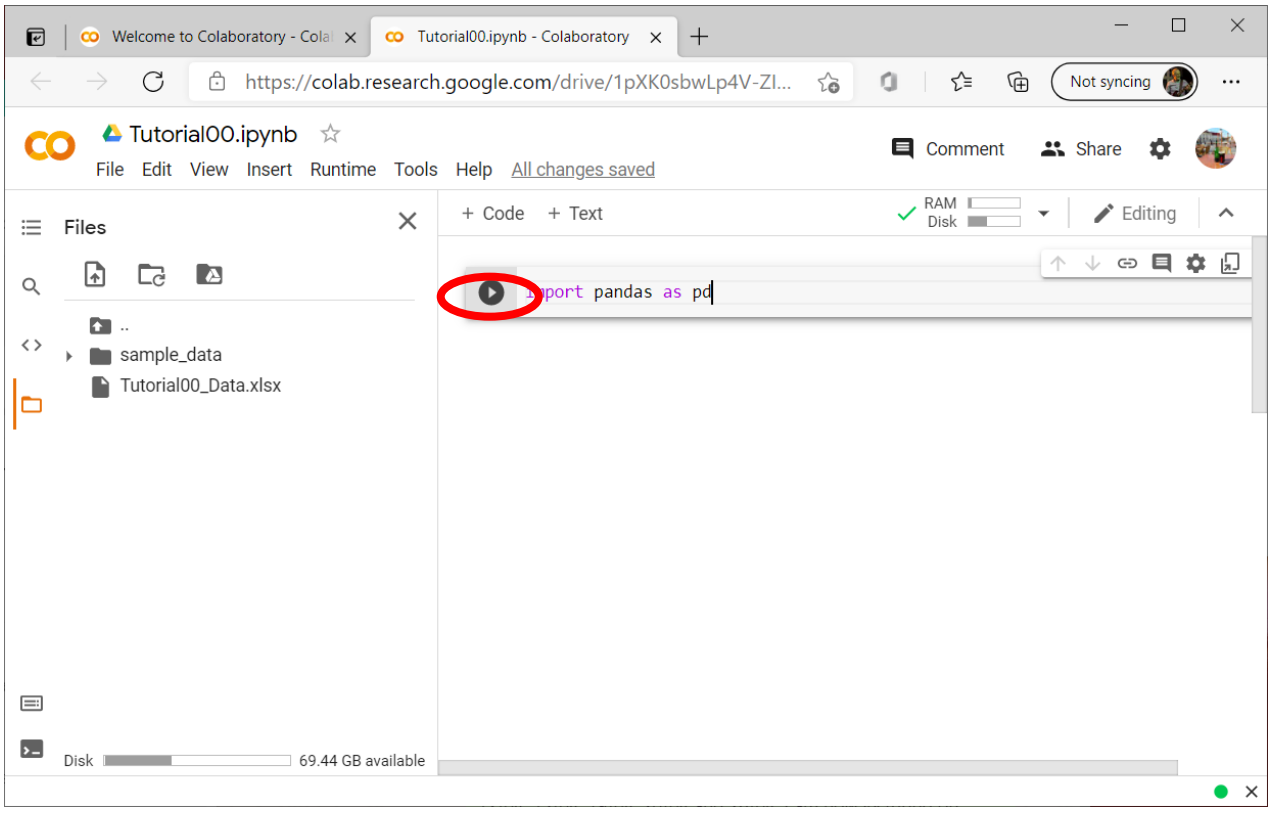
Clicking the “Open” button followed by “OK” should now upload the file:





The Excel file is now in the Colab cloud storage (if you want to keep it there more permanently you will need to mount a Google drive, which is free with your Google account).

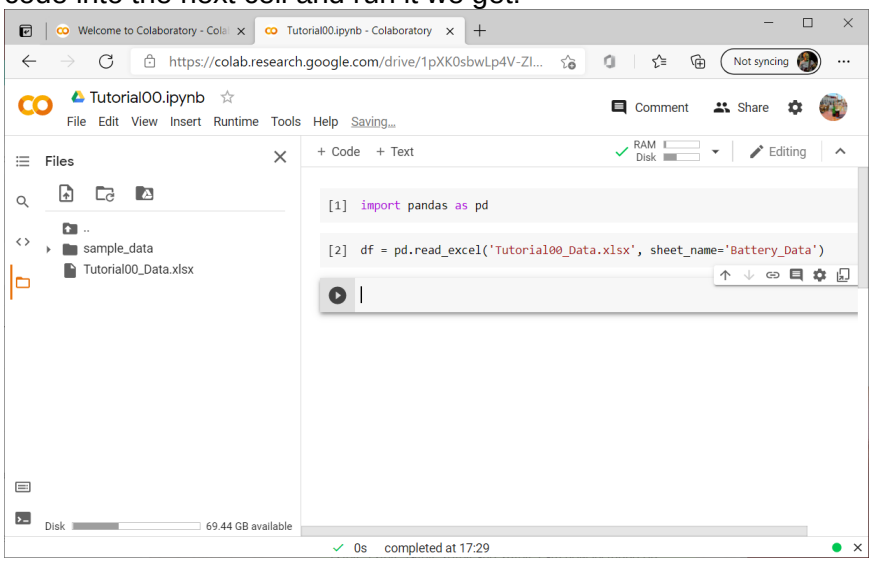
We are now going to use a Python package called “Pandas” (the name comes from panel data and is nothing to do with monochrome bears) by typing this code into the code cell and clicking on the “run” button:



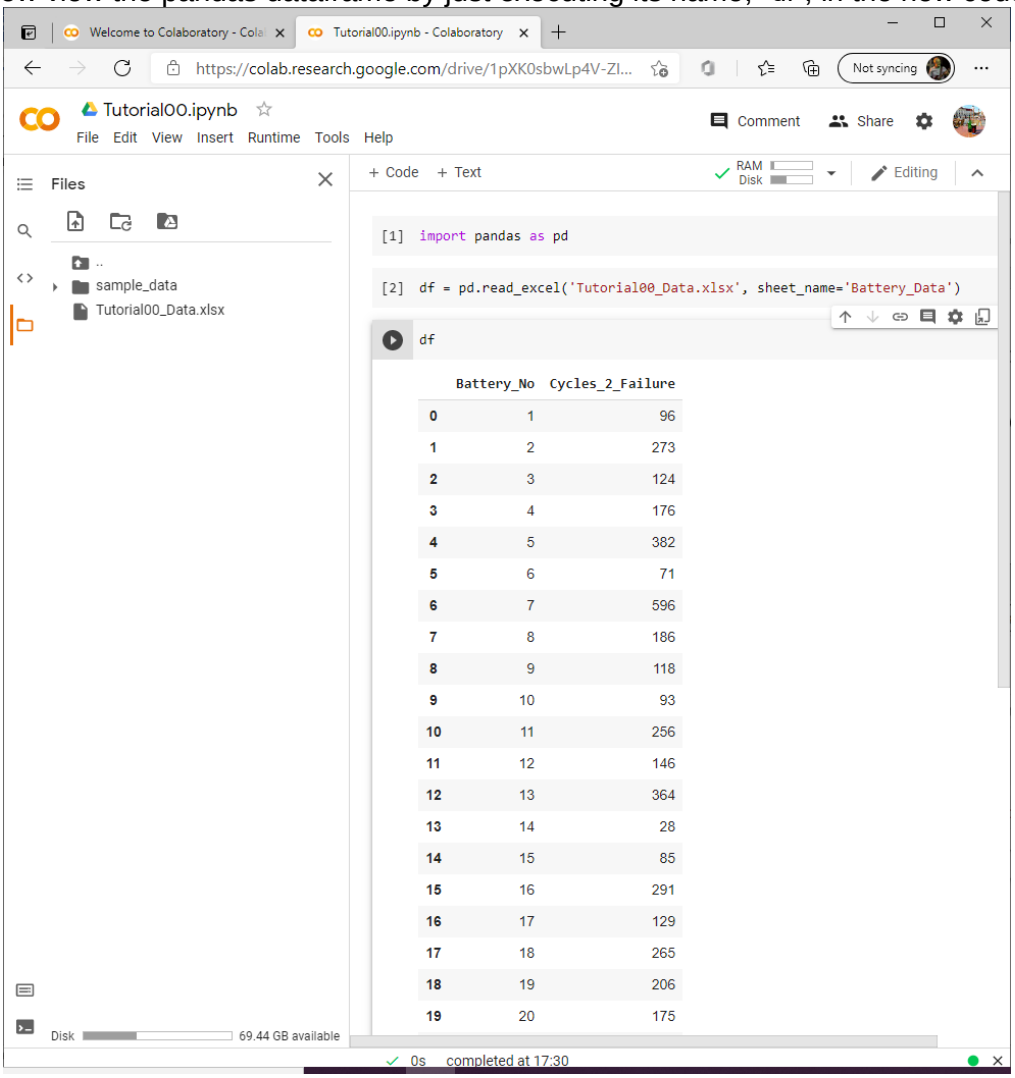
Pandas provide a broad range of mathematical functions (similar to Excel) based on its dataframe structure. We will now use Pandas to read the Excel worksheet into a pandas dataframe object (after clicking on “+ Code” to add a new code cell to the notebook). The code to do this is:

```
df = pd.read_excel('Tutorial00_Data.xlsx', sheet_name='Battery_Data')
```

If we enter this code into the next cell and run it we get:

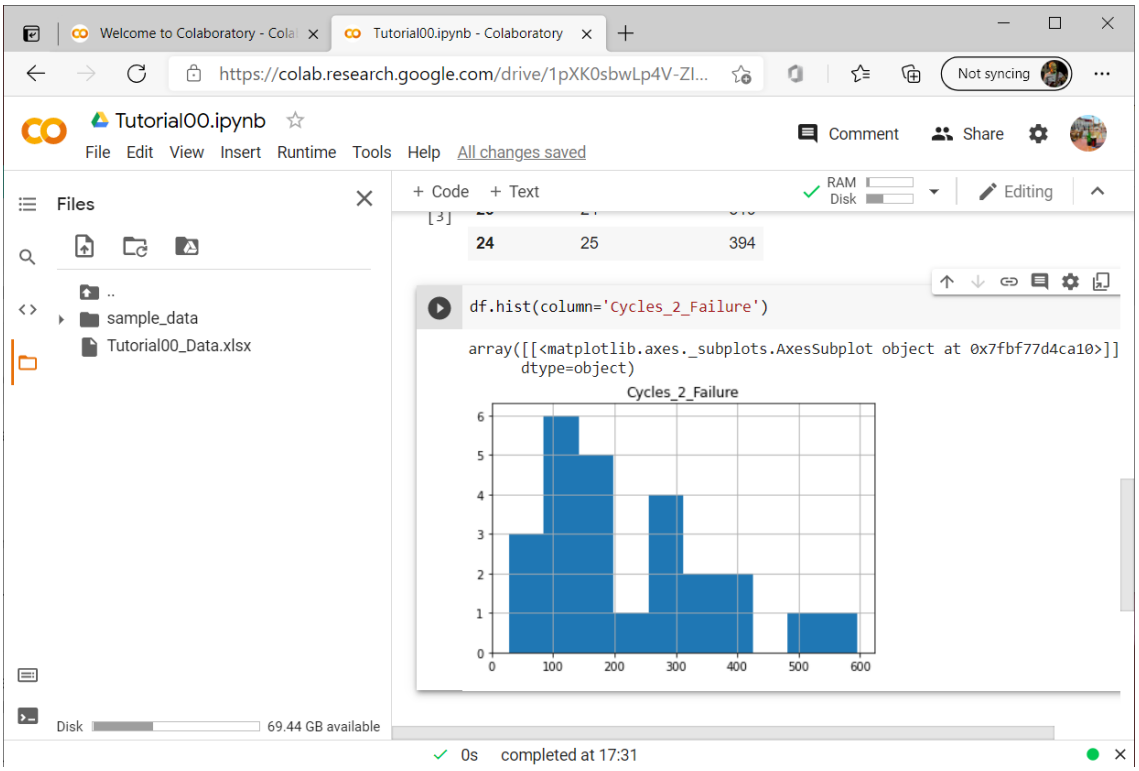


The screenshot above is what you see after running the cell followed by opening a new code cell. We can now view the pandas dataframe by just executing its name, “df”, in the new code cell.

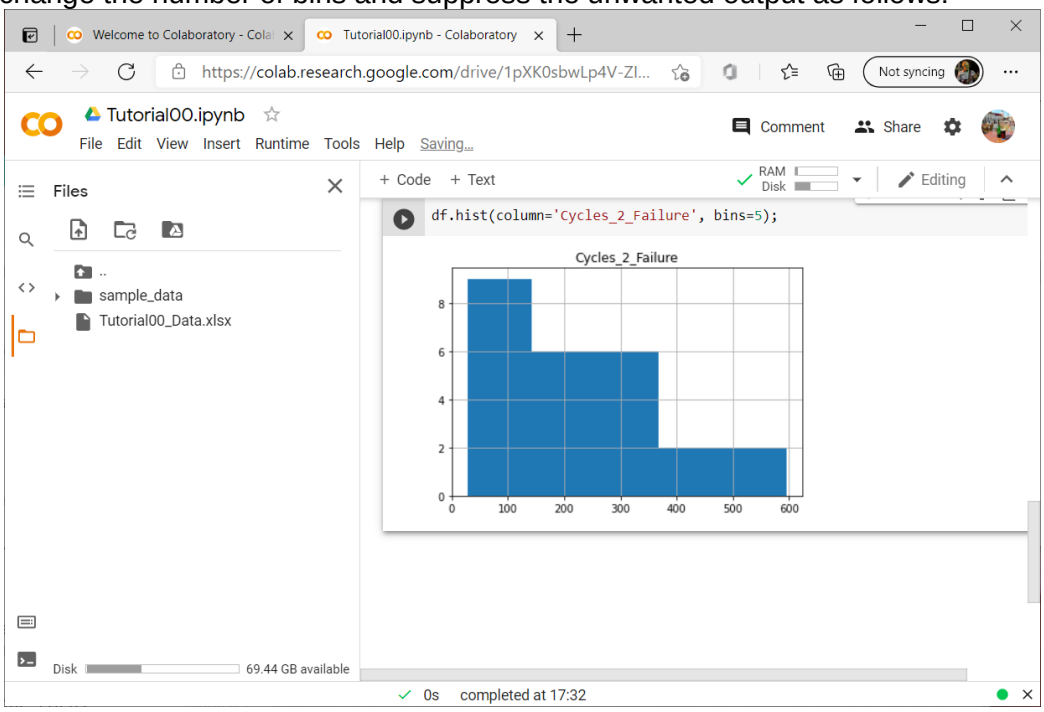


	Battery_No	Cycles_2_Failure
0	1	96
1	2	273
2	3	124
3	4	176
4	5	382
5	6	71
6	7	596
7	8	186
8	9	118
9	10	93
10	11	256
11	12	146
12	13	364
13	14	28
14	15	85
15	16	291
16	17	129
17	18	265
18	19	206
19	20	175

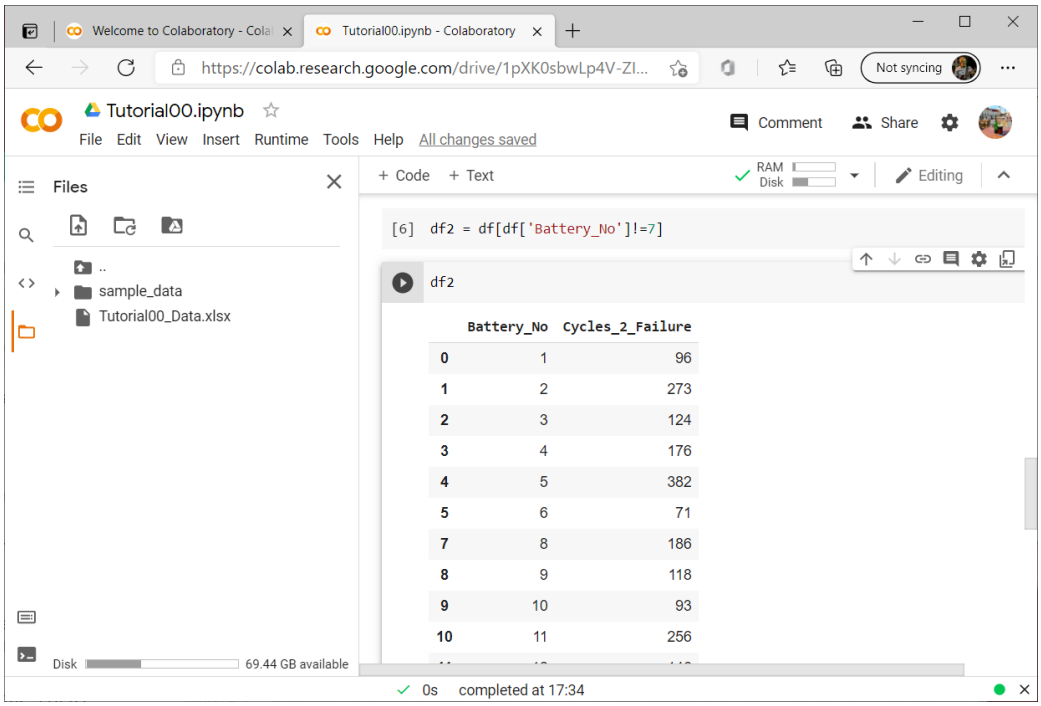
Note that the rows are numbered from 0 to 24 (not 1 to 25): this is inherited from Python’s numbering convention, which is very sensible once you have digested it.
Pandas has a simple command for plotting histograms:



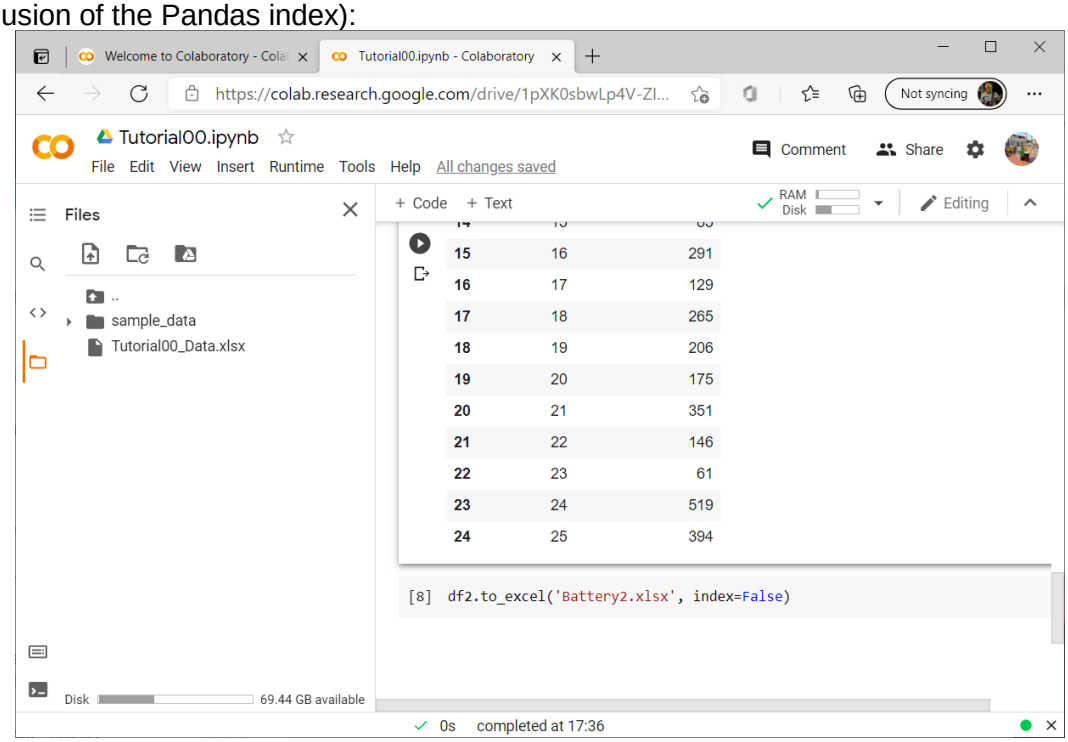
We can change the number of bins and suppress the unwanted output as follows:



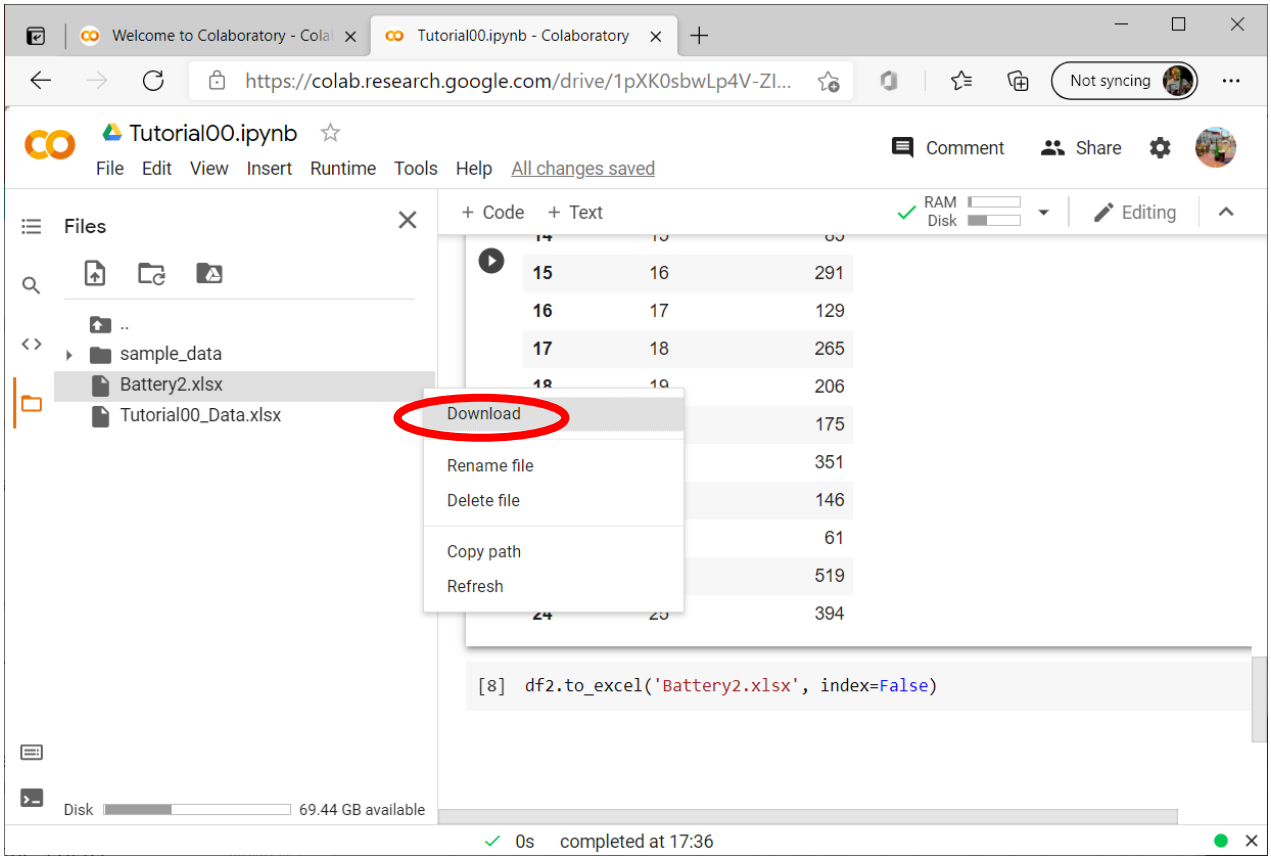
We are going to edit the dataframe by making a copy that excludes battery no. 7:



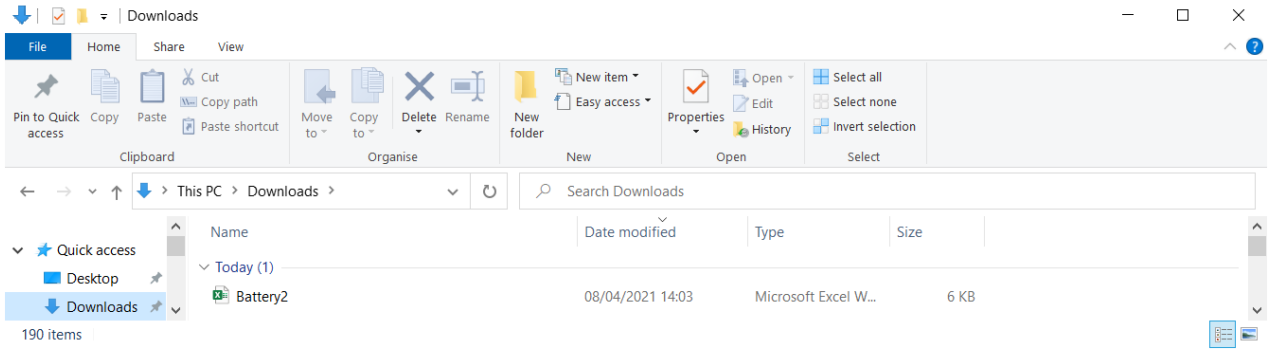
We will now save the amended dataframe as an Excel file (using the index=False option to avoid the inclusion of the Pandas index):



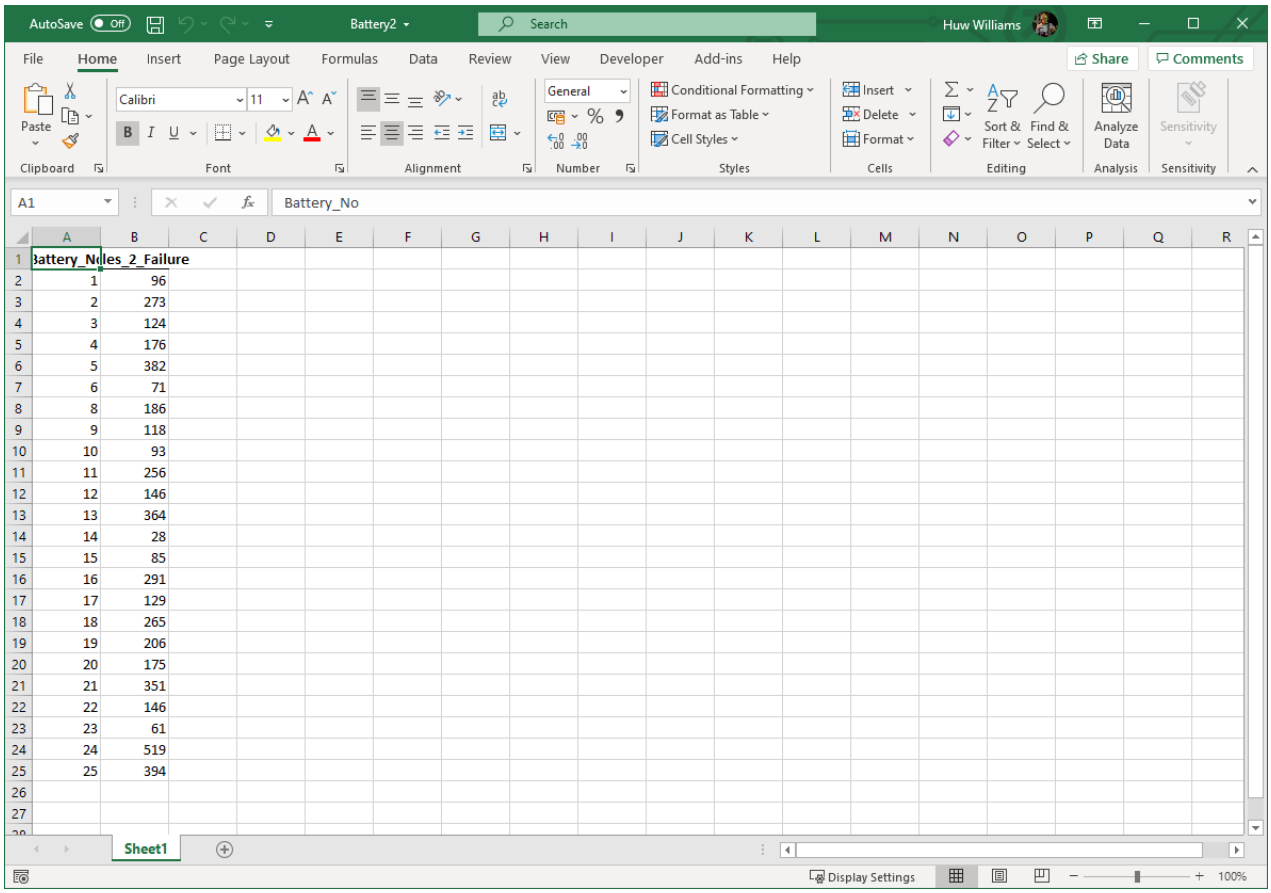
The new file will appear on the left of the screen. The last part of this tutorial is to download the new file to your PC. This is accomplished by clicking on the '…' that appears when the mouse pointer hovers over the file name and choosing 'Download':



The file should then appear in your Downloads folder:



Opening it in Excel shows that the line for battery 7 has been removed:



It is much more usual to import data into Colab rather than export it to Excel but we have included exporting in case you want to export some analysis results into the Excel environment.

You should now have the basic skills to get data in and out of Colab. If you are interested in learning more about Pandas try this link:

https://pandas.pydata.org/pandas-docs/stable/getting_started/tutorials.html