

## Homework 0: Get Familiar with Google Colab

Through this homework, you will learn to import the local excel sheet to Colab, input data, and export the updated data frame.

Please run the code chunks one by one, and enter the code along with your comments in the **TODO** part of this notebook.

```
1 # We will rely heavily on the following packages for data processing
2 # After "import .. as ..", "pd" and "np" will represent the package "pandas" and "numpy" respectively
3
4 import pandas as pd
5 import numpy as np
```

To learn more about pandas and numpy, please refer to the following links:

Pandas: [https://pandas.pydata.org/docs/getting\\_started/index.html#intro-to-pandas](https://pandas.pydata.org/docs/getting_started/index.html#intro-to-pandas)

Numpy: [https://numpy.org/devdocs/user/absolute\\_beginners.html](https://numpy.org/devdocs/user/absolute_beginners.html)

Use page search to navigate to the functions you require.

## Data Import

```
1 # Import local excel sheet to Colab
2
3 from google.colab import files
4 file = files.upload() # the function "upload" from package "files" uploads local file into google colab session
5 df = pd.read_excel("Info Sheet.xlsx") # the function "read_excel" from package "pandas" transform the excel sheet to a data frame
6 df.head() # the imported data frame was given the name "df" in the previous line, "xx.head()" displays the first five entries of data fram
```

Choose Files Info Sheet.xlsx

- Info Sheet.xlsx(application/vnd.openxmlformats-officedocument.spreadsheetml.sheet) - 8683 bytes, last modified: 1/14/2023 - 100% done

Saving Info Sheet.xlsx to Info Sheet.xlsx

Name	Program	Email Address	Phone Number	Experience with Python	Expectation for this Course
------	---------	---------------	--------------	------------------------	-----------------------------

## Data Insert

```
1 # Add a new row to the dataframe
2
3 df.loc[0] = ["Jane Doe", "MSIE", "doe.j@gmail.com", "800-000-0000", "NA", "Learn basics about data mining."] #adding the first row to the
4 df.head() #check the updated dataframe
5
```

	Name	Program	Email Address	Phone Number	Experience with Python	Expectation for this Course
0	Jane Doe	MSIE	doe.j@gmail.com	800-000-0000	NA	Learn basics about data mining.

To learn more about setting values in a dataframe with "xx.loc()", please refer to the subsection "setting values" in the following link:

<https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.loc.html>

**TODO:** Eliminate Jane Doe's information and insert your own as the first row of the info sheet

```
1
2 (0)
3 ["Arijit Mukherjee", "MSDAE", "mukherjee.ari@northeastern.edu", "857-350-6469", "Moderate", "Learn basics about data mining and various app
4
5
```

	Name	Program	Email Address	Phone Number	Experience with Python	Expectation for this Course
0	Arijit	MSDAE	mukherjee.ari@northeastern.edu	857-350-	Moderate	Learn basics about data mining and

## ▼ Data Export

```
1 # Export dataframe "df" to an excel workbook and download to a local folder
2
3 df.to_excel("output.xlsx")
4 files.download("output.xlsx")
5
```

---

[Colab paid products - Cancel contracts here](#)

✓ 0s completed at 9:40 PM

