

Instructions

August 14, 2021

1 Installing dependencies

To generate graphs, we will use Python and Jupyter Notebooks to make them. Install Python from the Python website, <https://www.python.org/>.

Install the following dependencies by running the following commands in your terminal:

```
pip install matplotlib
```

```
pip install pandas
```

```
pip install jupyterlab
```

```
pip install pyppeteer
```

2 Cleaning your data

To be able to have Python read your csvs, we need to clean the data first. In the Attendee Report, delete all line from the beginning to the line that says ‘Attendee Details’ (including the line with ‘Attendee Details’!). Now we can start.

3 Using the Notebook Templates

If you are using `attend.ipynb`, `nonmembers.ipynb`, and `members.ipynb`, skip this section and change the `filepath` variable on the first line of the notebook to the filepath of the attendee report csv.

Copy the `base.ipynb` notebook and name it to something like `mynotebook.ipynb` (we will use this filename for the rest of the commands). Open the `mynotebook.ipynb` notebook in Jupyter Lab by running the command

```
jupyter-lab mynotebook.ipynb
```

The notebook already has a bunch of helper commands to make generating graphs easier. Notice that code are squared off into separate blocks. These blocks are called “cells”. Change the `filepath` variable in the first cell to the path of the Attendee Details csv.

3.1 Filtering by Membership, Attendance, etc.

The third cell with:

```
df = pd.read_csv(filepath, index_col=False).drop_duplicates('Email')
```

Does three things:

1. It reads our csv into a DataFrame that holds our data.
2. The `index_col` value says "Dont make the first column of the csv our index"
3. And drops all rows with duplicate values in the 'Email' column.

We can manipulate this DataFrame by filtering out rows based on values in the column for example:

```
df = df[df['NRCP Membership'] != 'Non-Member']  
df = df[df['Attended'] == 'Yes']
```

Filters out rows where: 1. the NRCP Membership column has the value "Non-Member" 2. The person did not attend the meeting

3.2 Getting the count and percentage of a value/column in a table

You can do this by creating a new cell by pressing the "+" button and typing this:

```
column = get_count(column, df) # df is from the third cell/part of the template  
get_table(column)
```

Jupyter automatically shows the output of a function when it is not an assignment ("=" operator is involved like the first line)

3.3 Graphing a plot

Make a new cell and either write:

```
column.plot.pie(**defaultpie)  
plt.show()
```

or

```
column.plot.bar()  
plt.show()
```

4 Converting a notebook to pdf

In the terminal run:

```
jupyter nbconvert mynotebook.ipynb --to webpdf --no-input --allow-chromium-download
```

A pdf named `mynotebook.pdf` will be created