




Step 4: Clean Your Data

Alright, now you're getting to the exciting parts! You've hopefully learned a lot up until this point and you've invested a chunk of your time to learning Python and how to work with these vital tools. Now you're going to start working with **your data**. This is where your hard work starts really paying off. At this point, you should have completed the following sections:

- Step 1: Download Tools 
- Step 2: Learn Python Basics 
- Step 3: Python in Jupyter Notebooks 

Now you are going to learn techniques to do the seven most important things that you'll need to do while working with data:

1. **Import packages** into your notebook.
2. **Import your raw data** into your notebook.
3. Understand **what the data are** and **how the data are stored**.
4. **Change data** that are not stored in the way that best fits your needs.
5. **Delete data** that you do not want.
6. **Create data** that we need.
7. **Export the cleaned data** to a file.

As in step 3, you'll work through workbooks that are Jupyter Notebooks full of instruction and guidance. Please experiment! If you ever wonder "what would happen if I did this?", try it and see what happens! Experimentation, making mistakes, and figuring out what went wrong is a huge part of the learning process.

To get started...

You want to open the workbook in JupyterLab. To open the workbook in JupyterLab:

1. Open Anaconda.
2. Open JupyterLab.
3. Use the file navigator on the left to find the `step4_workbook.ipynb` in the `step4_clean_your_data` folder.
4. When you open the workbook file, you may see a pop-up that says "Select Kernel". A **kernel** is your coding environment. You should select `python_for_uxr_env`. This is the coding environment that you set up in Step 1.

When you're done...

You've beat me to `step5_describe_your_data`. I still need to create the workbook for this. Please send me an email @ alexdsbreslav@gmail.com and tell me to hurry up!