Lesson1

Setting up Reproducible Workflow

Start by navigating to a folder that will serve as the host of the main repository, e.g. `/rob/research/projects`
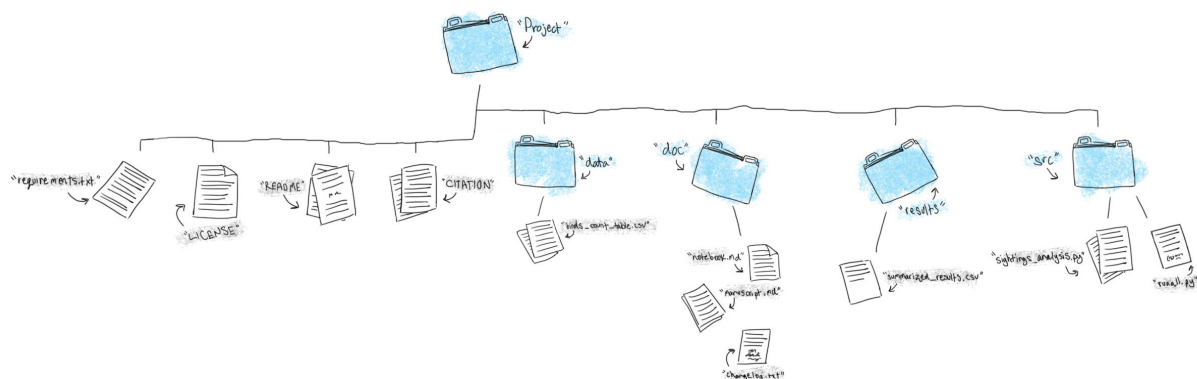Once there, replicate this folder structure.

Figure 1:

I recommend doing this on the command line first, just to get some familiarity, but if you want to do it with Windows Explorer or Finder, that's fine too. Here are the folders:

```
mkdir myrepo
cd myrepo
mkdir doc
mkdir src
mkdir results
mkdir data
```

Now make some files at the root level of the directory:

```
touch LICENSE
touch README
touch CITATION
touch requirements.txt
```

Ok - all done - for now. Let's talk about what goes into each, and then we'll start to populate the folders a bit.

Start With Raw Data

I've given you a raw dataset from the Bahamas Marine Mammal Research Organization. They are in csv format. Download them from box, and put them into your `data` folder. You can look at them with the head command. Note where you are on the command line first, otherwise you may be looking for data in all the wrong places.

```
head bbmroData.csv --lines=2
```

A Few Scripts

Ok, at this point we have a file structure, and raw data - nothing else. But let's keep going with a simple R script to read in the data. Make a script in the `src` folder and call it something clever like `readData.R` Try this first on the command line.

```
touch readData.R
```

While we're at it, let's make three more empty files:

```
touch summarizeData.R
touch plotData.R
touch runRegression.R
```

And then we'll make a controller script, since that will be useful.

```
touch runAll.R
```

That will make the files, but they're going to be empty just now.