AOS 575 Application Lab 1: Significance Testing and Bootstrapping
In class Thursday Sept. 16, 2020 and Tuesday Sept. 21, 2020

Notebook #1: Statistical Significance using Bootstrapping
aos575_applab1_bootstrapping.ipynb

Learning Goals
1) Working in an ipython notebook: read in csv file, make histogram plot
2) Assessing statistical significance using bootstrapping
3) Compare statistical significance via bootstrapping to significance via t-test

Underlying Data and Science
In the notebook we compare the tropical Pacifc Sea Surface Temperature (SST) anomalies to Madison total yearly snowfall. We will test the hypothesis that Pacific SST anomalies drive a change in Madison snowfall via a teleconnection initiated by the El Niño Southern Oscillation (ENSO). How could ENSO affect snowfall? During the warm phase of ENSO (El Niño), the midlatitude jet shifts southward and precipitation in the US is affected. During the cold phase of ENSO (La Niña) the opposite occurs. This notebook guides you through an analysis of how to statistically examine this relationship. We use the Niño 3.4 index and Madison snowfall from the following sources:

The Nino3.4 data are from:
https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Nino34/

The Madison snowfall is from:
http://www.aos.wisc.edu/~sco/clim-history/stations/msn/msn-snow.html

Questions and suggestions to guide your analysis of #1:
In addition to the occasional questions in the Jupyter Notebook, add the answers to these questions in markdown at the very end of the notebook, clearly labeling which questions go with which answers:

1. Complete the following table:

	Mean snowfall	**Std. Dev. Snowfall**	**N (# years)**
All years			
El Niño years			
La Niña years			

2. Use hypothesis testing to assess if the differences in snowpark are statistically significant. Write out the 5 steps and apply. Test your hypothesis using bootstrapping

How bootstrapping works: Say there are N years with El Niño conditions. Instead of averaging the Madison Snowfall in those N years, randomly grab N Madison Snowfall values and take their average. Then do this again, and again, and again 1000 times. In the end you will end up with a distribution of snowfall averages in the case of random sampling, i.e., the distribution you would expect if there was no physical relationship between Nino3.4 SST anomalies and Madison snowfall.

3. Test the sensitivity of the results. Change the number of bootstraps. Change the significance level. Change the temperature threshold for El Niño/La Niña to +/-0.5 degrees Celsius or +/-2 degrees Celsius. Comment on how things change and how your conclusions are altered.

Notebook #2: Statistical Significance using z/t tests:
aos575_applab1_ztest_ttest.ipynb

Learning Goals:
1) Working in an ipython notebook: read in netcdf file, make line plots and histograms
2) Calculate statistical significance of the changes in the normalized mean using a z-statistic and a t-statistic
3) Calculate confidence intervals on CESM-LE global warming using z-statistic and t-statistic

Underlying Data and Science
You will be plotting global mean surface temperature from the Community Earth System Model (CESM) Large Ensemble Project (Kay et al. 2015). The Large Ensemble project includes 42 ensemble members of fully coupled climate model simulations for 1920-2100. 1920-2005 use historical emissions forcing (greenhouse gas emissions, volcanoes, aerosols, etc.), while the future (2005-2100 here) uses the Representative Concentration Pathway 8.5 (RCP8.5) emissions forcing. This is a high emissions scenario. Each ensemble member begins in 1920 with roundoff level error differences in the initial conditions that cause the ensemble members to diverge from each other after a few weeks (i.e. the buttefly effect, same as for weather forecasts!). Each ensemble member represents one possible representation of past or future, giving us a sense of the natural variability in the climate system. The 1920 ensemble members start from a very long perpetual-1850 simulation (2000+ years!) that uses preindustrial levels of emissions. In this notebook we compare the global temperature during time period 2020-2030 to the global temperature in the 1850 simulation. When we include many representations of natural variability, do we see a statistically significant change in global temperature?

More information on the CESM Large Ensemble Project can be found at: http://www.cesm.ucar.edu/projects/community-projects/LENS/

Questions to guide your analysis of Notebook #2:
In addition to the occasional questions in the Jupyter Notebook, add the answers to these questions in markdown at the very end of the notebook, clearly labeling which questions go with which answers:

1. There are 2600 years in the 1850 control run. This is enough for a very stable population mean and standard deviation. What is the population mean and standard deviation for the 1850 control run? Standardize the 1850 surface temperature and then re-report the population mean and standard deviation. What are they now? Plot a histogram of the

1850 simulation surface temperature. Does it look like a normal distribution (Gaussian) to you?

2. Calculate the global warming in one ensemble member during 2020-2030 (as defined in the notebook). Compare this warming relative to the 1850 simulation. Is there statistically significant warming? Use hypothesis testing and state the 5 steps. What is the null hypothesis? What is the difference between using the z and t statistics? What is the probability that the warming in this one ensemble member occurred by chance? Play with the start and end years. See how that changes your answers. When does warming become statistically significant?

3. The power of a Large Ensemble approach is that it provides many potential realizations of the climate and allows a rigorous quantification of the effects of greenhouse gas emissions by "averaging out" the natural variability. Plot a histogram using many of the CESM-LE ensemble members. With such a large sample, does it look like global surface temperature is a normally distributed quantity? You can use these ensemble members to provide a confidence interval on the ensemble mean (i.e., the forced response) using the ensemble standard deviation (the natural variability). How do the confidence intervals change if you change the number of ensemble members (e.g., try 30, 10, 6, or 3 members). What's the difference between the 95% and 99% confidence intervals?

When you are done with the notebook, push it to your GitLab account.
Launch a terminal:

```
cd aos575
git add --all
git commit -m "Application Lab 1"
git push origin master
```

Be sure not to delete any files in your aos575 directory. Sometimes git gets a little finicky about this. Shoot me a message if git gets upset at you.