**Agenda**

1. Questions to be discussed with partners:
   1. A lot of different data sets - **What’s QB’s preference?** - more machine learning aligned (CV, reservoir computing, deep neural nets)? Or more statistical analysis aligned (time series modeling, forecasting)? Can we combine datasets intelligently?

QB: Decision on harvard team.

Team: do statistical forecasting as baseline. Paper about comparison of models, such as adding CV

QB: very limited time.

* + 1. Currently, we are more inclined to do a paper
    2. We have information for forecasting using satellite images and using data from soil moisture, agriculture, etc.
  1. We intend to focus on **a few regions within the continental US**- do you have any thoughts on which regions would be best? We thought potentially...

(train on a sub-area and test if can extend to larger area/worldwide) Michael, can you show the plot you found yesterday :)

QB: sampling distribution should be representative of the population we care about. Or introducing bias through random sampling (FS). Before sampling, take a few cities, run it, see the sampling is working or not (VN).

* 1. What features in satellite imagery might be best for us to control for across regions? Masking? Cloud coverage?

QB: Eliminate images once we have idea of what we want to investigate

* A lot of off the shelf networks (image analysis in a more geneal sense) to speed up process of image processing, if images can be correlated over time point to point pixel
* Time signal directly from pixels of images

1. Narrow down with QB the problem statement with specifics (including time steps, area, inputs, etc)
   1. We will have to do this based on our data
2. Consideration of “actionable variables” - things like agricultural usage, urbanization, etc. that are human actions causally linked to drought **(what’s QB’s vision on this project? What kind of program does QB want it to become?**
   1. Read literature on this - dramatic events, forest fires, natural disasters