**Tufts-ASA**

**Meeting Minutes**

**July 25, 2020**

1. **Minutes taken by:**

Kees Schipper

1. **Present:**

Kyle Monahan

Kees Schipper

…

1. **Absent:**

NA

1. **Agenda:**

- Presentation on how to find and download open source data using Tufts data lab and various other resources

1. **Business from Previous Meetings:**

NA

1. **New Business:**
	* First stop: data lab website at Tufts
		1. Virtual data lab assistants that can help with data needs
	* GIS catalogues at Tufts are available
	* Datalab-Support@elist.tufts.edu
		1. Email if you can’t find a certain dataset or if Tufts doesn’t have that data.
	* Kyle has created a list called Environmental Data Sources
		1. Many of these data sources are paid for by Tufts
		2. Links take you to Tufts data lab website so that you can gain access to the data through vpn
	* Kaggle
		1. Competitions and learning for finding example datasets
		2. Jupyter notebooks and example codes from other users can be viewed publicly. Good for learning new methods of analyses and learning about new datasets
	* Jupyter notebook
		1. Collection of code and results of that code in a file. Very helpful for seeing how code works in a step-by-step fashion.
	* Tidy Tuesdays
		1. Every Tuesday, the R community releases a dataset and conducts zoom calls where individuals analyze that dataset and publicly release their analyses.
		2. R users tend to be very open and happy to teach others about how to use the program
	* Google datasets
		1. Google “google datasets search”
		2. Is a google search engine for data.
		3. Good for screening data in different topics.
	* Github
		1. Data repository where people put their datasets for public release
		2. Place for people to “put code online”
		3. “Awesome Public Datasets”🡪search in Github: is a great Github entry that has a huge list of very cool datasets for analysis.
	* Data.gov
		1. Government data. Can have questionable quality. Not cleaned, etc…
	* Web scraping
		1. Can write program scripts that actively pull data from online. Curate your own datasets.
2. **Adjournment (When and by who?):**

Kees Schipper