

Data Acquisition and Description

We got our data from Oikolab. They get their data from various sources, such as NOAA, GEFS, CHIRPS, and ERA5. The data came in a csv file, with a variety of variables to choose from, all of which we explored in our autoregression file, in which we decided which variables to focus on in our correlation matrix.

Milestones

Linear regression models:

Previously done

1. Preprocessing data
2. Encoded categorical data
3. OLS Summary for each city
4. Mean snowfall for February by year by city graphs

Need to do

1. Add a regression line to plots
2. Change x-axis ticks
3. Add comments to this code

Auto regression models:

Previously done

1. Preprocessing data
2. Encoded categorical data
3. Fit an auto-regression model with the SLC all Param data
4. Get a predicted snowfall for Feb 10 2034 (which day is arbitrary)

Need to do

1. Get an accuracy for regression model/other statistical analysis of the autoregression
2. Find another model to make a comparison
3. Add bullet for prediction for feb 2034 to other plots
4. Possibly run other city datasets through the auto regression model.
5. Clean code

Data:

Previously done

1. Reformat data/other preprocessing.
2. Create a correlation matrix to identify which variable correlates most with Snowfall in SLC
3. Find historic mean, upper limit, lower limit for Feb 1994-2024 in SLC

Need to do

1. Show statistical significance or insignificance of data and plots.
2. Make sure all datasets cover the same years