1. Write a draft of your Capstone project milestone report with the following included:
2. Define the problem
3. Identify your client
4. Describe your data set, and how you cleaned/wrangled it
5. List other potential data sets you could use
6. Explain your initial findings
7. Share the Capstone Project 1 code and milestone report related to Github repository

Title: Utilizing meteorological measurements to predict mountain snowfall

Ojbective: Utilize meteorological measuresments in conjunction with supervised learning to predict snowfall.

Problem:

Dataset descriptions

Copper Mountain SNOTEL site (hourly)

Leadville Mountain ASOS site (hourly)

Grand Junction upper air data (12 hour)

Data cleaning/wrangling

Filled hourly data up to three hours using interpolation

Due to sporadic nature of snowdepth measurements, snowfall was calculated at 12 hr intervals

Surface data was plotted and obvious outliers were removed

data capture figure

Initial Findings

Table of general snowfall stats (e.g. max, count, mean, etc) by year

Table of R squared, p values, and % data capture as compared to snowfall for each variable

Linear Regression approach

Utilized statmodel OLS alogirthm

Table of training set results (adjusted R squared, F stat p value)

Present surface data only results

Present surface + upper air results

Figure of training set residuals

Figure of predicted vs actual snowfall utilizing test dataset (2006-2007)