*1a) Major activities/milestones planned for this week (in the timeline that you had going into the start of the week).*

*1b) Major activities/milestones accomplished this week (note: this may or may not actually be what was stated in 1a -- that's okay, as long as you're making progress and learning!).*

*2a) Open challenges and questions (including what -- if anything -- are the challenges that Ethan or I can help provide feedback or pointers on?)*

*2b) Major changes to research plan (if any, based on what you've learned or accomplished thus far, and the unexpected challenges you've faced this week)*

*3) Revised week-by-week timeline (this can be copy and pasted exactly from what you previously had if there are no changes, but assuming you are doing some course corrections along the way, it's good to have a revised timeline to refer to!)*

October 25 – Check-In #1

Lowell Deschenes & Matthew Thompson

**Major activities/milestones planned for this week:**

* Complete scraping data from MWAC archives
* Acquire all data

**Major activities/milestones planned for this week:**

* Matthew got the data weather, snow, and avalanche data from representatives at the Mount Washington Observatory (MWOBS) and the Mount Washington Avalanche Center (MWAC). This meant we did not need to scape avalanche forecast data from MWAC archives.
* He also explored the snow and avalanche data from MWAC, discovering its avalanche occurrence skewness and some indecipherable variable and value names.
* Lowell completed data cleaning on the weather data from MWOBS. He made sure there were no null values in the data frame and made sure there were no obvious bad entries in the nominal weather features by looking at their 5 number summaries.
* He made 3 new variables from the date feature. The features are day, month, and year.
* We also looked back over the project proposal to make sure the goals outlined in it are reasonable for two people to complete. We decided the goals are doable without being too simple.

**Open challenges and questions:**

* The data Matthew acquired was more comprehensive than expected and made it so we did not have to scape avalanche forecast data. With Matthew taking this course for graduate credit, we need to find a way to incorporate scraped data or maybe an advanced data cleaning technique.
* The merging of the two datasets will take much more cleaning to make sure they are compatible.
* The avalanche occurrence data we have is very skewed. It appears that we have many more instances when an avalanche occurred than when an avalanche did not occur. This will increase the difficulty in making an accurate model.
* We need to obtain meta data for both datasets as some variable and value names do not make sense to use

**Major changes to research plan:**

* In order to fix the skewedness of the snow data we will either have to use models which work well with skewed data or will have to make a new boundary to change what we are predicting such as high volume of avalanches vs low volume of avalanches.
* We will have to find a new feature to scrape from the internet or come up with another plan for Matthew to complete graduate level data obtaining/cleaning work.
* Matthew will be in touch with both MWAC and MWOBS again to obtain meta data.

**Revised week by week timeline:**

October 25 – Check-In #1:

* Acquired all data
* Revise our project proposal
* Start cleaning the snow and weather data

November 8 – Check-In #2:

* Completed data cleaning process

November 15 – Check-In #3:

* Completed data exploration
* Completed model implementation

November 22 – Check-In #4:

* Completed data and model analysis
* Completed visualizing results