What I did:

• Downloaded data from the ENCODE project

• Obtained “tilings” data [Gilbert 2014]

• Started data preprocessing:

◦ Converting file formats

◦ Merging data files

◦ Filtering data for variables we are interested in using within our feature set

• Started building skeleton script for deep learning algorithm

What I learned:

• Learned how to work with different file formats that I was not previously familiar with (e.g. bed files, bam files)

Plan for the following week:

• Finish downloading all the data available from the ENCODE project

• Finish data preprocessing

• Read the “DeepCRISPR” paper

• Continue working on model