**Progress Report**

1. **Principal Investigator**

Damien Brandel – [brandeld1@newpaltz.edu](mailto:brandeld1@newpaltz.edu)

Evan Finke – [finkee1@newpaltz.edu](mailto:finkee1@newpaltz.edu)

1. **Title of Project**

Popularity and Ratings of Movie Genres

1. **Project Progress Summary**

* Combined/Cleaned movie datasets
* Dividing genre into individual columns.
* Created ratings column.
* Separated production year into separate column.
* Created a positive or negative sentiment via a sentiment analysis algorithm using review tags from dataset.
* Genre popularity sorting
* Year-based movie genre popularity sorting

1. **Activities**

* Combined/Cleaned movie datasets – 10/5
* Dividing genre & ratings – 10/12
* Separated production year & more data organizing – 10/19
* Sentiment analysis – 10/26
* Genre popularity sorting - 10/29
* Year-based popularity sorting & Progress report – 11/2

1. **Remaining Work**

* Genre rating sorting/prediction algorithm
* Hadoop data process
* MapReduce

1. **Difficulties**

Data cleaning & Run time

1. **Significant Changes to Proposal (if any)**

Added sentiment analysis feature