Johnny Zgombic

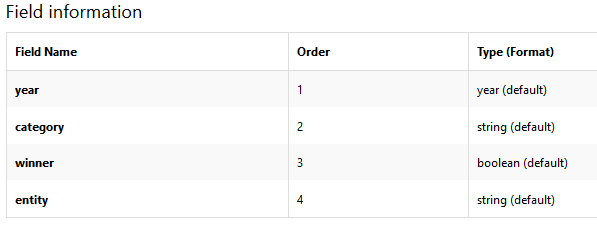
2019 Fall Term - IS 362 – Data Acquisition and Management

12 December 2019

**Final Project Proposal**

The project I will be undertaking, with your approval, is using the data from the Academy Awards Archives since 1927 to see whether we can predict Best Picture from Best Film Editing. I will be using data which can be found at <https://datahub.io/rufuspollock/oscars-nominees-and-winners>.

This data includes normalized JSON and CSV data for every year and category that is presented at the Academy Awards. The tables in question include the following information:



Along with determining the outcome above, there are several interesting trends I will look at and present including repeat winners and if there is any correlation between Best Actor and Best Picture as well.

By using the many tools available, I will take this data, clean it up, do several aggregations, grouping and statistical analysis to eventually come to a result which will tell me whether we can predict Best Picture successfully or if everything is indeed random.