**Group 5 Project Proposal**

Sahil Aggarwal, Aditya Chatterjee, Kevin McCraney, Altan Orhon, Umang Sehgal  
February 16, 2018

Our research project will investigate data related to scene length and frequency of cuts in a large dataset of several movies across the 20th century. We wish to answer the following (manifold) research question: is there a statistically significant and substantial change in film dynamics in the Soviet Union after 1934, when the policy of socialist realism—which discouraged abstract and experimental forms of expression (Taylor & Christie, 1991)—was introduced to regulate artistic production? Further, was there a change after Stalin died, in 1953?

This project sits squarely within the digital humanities. Statistical techniques have been used to identify anonymous authors (see Mosteller, 2009) or gauge demographic trends across similar works (see Moretti, 2015). Much work has gone into using text as data; unfortunately, there has been minimal scholarly work related to analyzing cinematic metrics in aggregate. Existing research has focused on bodies of data related to the oeuvre of one director, or similarities across a handful of genre films. It also serves to illuminate the deep ties between artistic production and ideology by quantifying impact on aesthetics.

Our data comes from Cinemetrics (Cinemetrics, n.d.), which comprises shot length measurements from over 18,000 films. We have 157 film records from the U.S.S.R. before 1934 (from 1908), 40 from 1934-1953, and 66 from 1954-1979. Because our intention is to judge whether differences in film dynamics before, during, and after socialist realism can be attributed to something other than chance or history effects, we will use inferential statistics. We operationalize change in film form as change in median shot length (Baxter, 2014). To assess the degree and statistical significance of changes, we will use analysis of variance with three groups (pre, during, and post). We may experiment with interrupted time series designs to capture evolving trends in filmmaking, or analyze the raw shot length and location data if median shot length is too imprecise.

**Bibliography**

Baxter, M. (2014). Notes on cinemetric data analysis [monograph]. Retrieved from <http://www.cinemetrics.lv/articles.php>

Cinemetrics. (n.d.). Retrieved from <http://www.cinemetrics.lv/database.php>

Moretti, F. (2015). *Distant reading*. London, U.K.: Verso.

Mosteller, F. (2009). Who Wrote the Disputed Federalist Papers, Hamilton or Madison? *The Pleasures of Statistics*, 47-67. doi:10.1007/978-0-387-77956-0\_4

Taylor, R., & Christie, I. (Eds.). (1991). *Inside the film factory*. London, U.K.: Routledge.