In this project we aim to estimate and predict quantities related to the perception of the audience on movie actors and an (qualitative?quantitative? (we can say at which quartile of budget the actor belongs to)) estimation of actor’s hiring-cost. This information is critical at cast-selection time as it plays an important role on the success of the movie. However, few datasets provide data related to actors, most of them have information only about movies.

Therefore, we propose first to generate an estimation of this information by doing an aggregation on movie data. Then, we propose a linear regression model to predict these quantities. We enhance the linear regression model by leveraging information from the social network graph of actors. We hypothesize that the social relationships have a major influence on these quantities. Hence, we will measure the importance of the social relationships by how much our predictions improve when using the graph compared with baseline models.

As the number of actors is around (54000) we restrict our analysis only to the actors that have been protagonists in at least one movie. We also remove movies that have relevant quantities (e.g. budget, revenue, popularity) equal to 0. Thus, we have around 3000 valid movies.