

## **Data Science Project Proposal: Problem statement**

*Team: Jonas Pirkel, Nikhil Patil, Ashish Kumar*

1. What questions would you like to explore?
2. Why are these questions the right questions to consider?
3. What machine learning approach (approaches) you plan to use?

Netflix is dependent on its user's well-being during their stay on Netflix's website and in their applications. Rating systems help users to distinguish between "good" and "bad" movies. This directly affects how a user chooses a new movie or sitcom to watch.

1. We are considering multiple questions using:
  - What affects the rating of a user? (e.g. time of the year, genre, )
  - Can we predict a specific user's rating based on their last rating?
  - Which are the highest-rated shows on Netflix? Can we predict genre or type of future highest-rated shows based on metrics?
  - Can we predict user behavior during the different times of the year especially to check whether they give higher ratings to the movies during holiday times?
  - Can we predict which movie a user is watching next based on his history of movies?
  - Do ratings fluctuate seasonally? Can we predict the average rating of movies on the platform based on the season
2. These questions offer insights into rating behavior among all users. We hope to explain trends and forecast what kind of movie a user might like to see next. This will help Netflix to show relevant data to their customers, keep them engaged, and maintain the subscription to make profits.
3. We are planning to use supervised learning because the input data is polished and well known and in order to predict the future performance of shows, we need supervised learning. We want to make use of linear regression and classification and apply other methods that we are learning within the next few weeks.