Project 1: Trending Topics

YouTube maintains a list of the top trending videos. YouTube uses a combination of factors, including measuring the number of views, shares, comments, and likes to make commercials and determine trending videos. This dataset is a daily record of the top trending YouTube videos and the factors that influences them. Analyzing this dataset, we can look at the main factors that are included in determining the videos that eventually are considered trending such as the likes, dislikes, and comments. One project is to determine if a video will eventually become a trending video based on specific characteristics. There is a YouTube data set on kaggle.com that will be useful when completing this project.

Project 2: Degrees That Pay

Picking a major is very important for a college student as the decision affects the student's livelihood and career choices after college. It is optimal for students to select a degree that is both useful and stimulates their interest. I am more interested in this project. The data set includes a year-long survey of 1.2 million college graduates that have earned a bachelor's degree. PayScale Inc conducted the study. It provides the salary for students in that fall within three categories, one year, five years, and ten years after graduation. They also include the factor of which school. We can observe which major earns the highest salary and the major that earns the lowest salary. The data set contains elements that can affect salary.

Conclusion

The predictive models that I think will be most appropriate for the projects are logistic regression for the first project and linear regression for the second. The deliverables for both projects are a slide deck and a paper outlining the code for the project.