## Women of Data Science

# Netflix Entertainment 2007-2020

##### by Hope McCadden and Jazmin Barnett-Birdsong

Executive Summary

To utilize and explore the many tools and skills acquired in the Data Science program, students are assigned a six-week project. The project requirements allow students to choose the topic data, the programming language(s), and the appropriate analysis. The project requires students to work in teams to mimic the working relationships of professional data science teams.

The purpose of this document is to explain the project chosen by Hope and Jazmin. This document will introduce the topic, the statistical analysis, the scope of the project, the project requirements, and provide thoughtful insight into the assumptions, limitations, and risks associated with the undertaking of this project.

### Business Objective

Our objective is to elevate the difference in content that was streamed on Netflix between 2007 to 2020. We will be using applications for this project such as R, Python, and Tableau programs will be implemented to wrangle, analyze, and visualize the “Netflix Movies and TV Shows " dataset made available by Kaggle.

We aim to present the results of the following three questions. Which genre is the most lucrative? Which director released the most entertaining content? Was there significantly more content release for adults than for children?

Background

The topic of this project is the content released on a popular streaming platform, Netflix between 2007 and 2020. Hope and Jazmin chose the dataset “Netflix Movies and TV Shows” because of their shared interest in film and television. Their objective is to draw conclusions about the type of content released on the streaming platform. The information obtained is useful for viewers when choosing streaming platform subscriptions by their offered content genres, direction, and content rating.

Scope

For this project, Hope and Jazmin intend to focus on the tools and software that were learned through the Data Science program. Nearly all the tools that will be used are free-sourced. The choices of tools and software are based on project requirements and popular software throughout the larger data science community for the purposes of job-seeking. Hope and Jazmin may use other software for this project, but that is not a requirement of the project.

### Functional Requirements

Data Wrangling: Before analysis, the chosen dataset must be thoroughly cleaned, and missing data must either be removed or substituted. Some new column data may need to be added and some data may need to be recoded or re-structured for best results. Some data may also need to be converted to different data types to continue to the next step.

Data Analysis: Hope and Jazmin will need to pose questions and review the variable types of the chosen data. They will determine the best statistical analysis techniques as well as the best programming language(s) in which to perform the chosen analyses.

Data Visualization: After analysis, Hope and Jazmin will use data visualization software to highlight some of the more interesting findings of the data. These visuals will be used in the final PowerPoint presentation to help the audience better interpret Hope and Jazmin’s findings. They have chosen Tableau for the data visualization software for this project.

Presentation: Hope and Jazmin must present their project and its findings to their course instructors and their peers. The presentation must be given in about 20 minutes and the students must present the project clearly for all to understand. Hope and Jazmin will be professionally attired and will welcome feedback.

### Delivery Schedule

The suggestion time for project completion is six weeks. Portions of the project will be delivered weekly. Each week the tasks are listed below.

* Week 1:
  + Create a Kanban board.
  + Assign week 2 SCRUM master.
  + Choose evaluation questions from the selected dataset.
  + Assign programming languages to answer statical analyses questions.
* Week 2:
  + Review statistical analyses for Chi-Square.
  + Study the dataset and ask questions.
  + Discover what correlation exists.
  + Discuss what predictive models exist.
  + Begin Data Wrangling.
* Week 3:
  + Data Wrangling continues.
  + Implement the Chi-Square analyses.
* Week 4:
  + Work with Tableau
  + Review other final project examples
  + Review and validate findings from the previous week.
  + Draw insights and/or conclusions.
* Week 5:
  + Create a PowerPoint Presentation that displays project findings.
  + Go over the presentation with the instructor/friends/family members to ensure content simplicity.
* Week 6:
  + Make the final changes to the presentation.
  + Practice presentation presenting at least two to three times each and at least once with the instructor.

Other Requirements

All team members are skilled in the programs used.

### Assumptions

All hardware, software, and platforms are working, in-tact, free of bugs, current, and accessible.

### Limitations

With the time constraint and the unpredictability of their personal lives, Hope and Jazmin will periodically assess the established goals of this project to determine the likelihood of on-time completion. If an issue arises, one of the three evaluation questions may be eliminated to maintain the integrity of the rest of the work.

Risks

The risk of this project being incomplete is minimal. The expectation of completing the project successfully is great.