

## **Portfolio Milestone 1**

Harlan Wittlieff

Data Science, Bellevue University

DSC 680: Applied Data Science

Catherine Williams

July 3<sup>rd</sup>, 2022

## Project List

1. DSC 680 – Project 1 Placeholder
2. DSC 680 – Project 2 Placeholder
3. DSC 680 – Project 3 Placeholder
4. API Data Pulling
  - a. This project demonstrates how to pull data from an API, clean the data, and store it in a SQLite database.
5. Wine Quality Predicting – KNN & Random Forest with SMOTE
  - a. This project builds KNN and Random Forest models to predict wine quality. SMOTE is utilized to balance the predicted classes for training the model.
6. NLP with Logistic Regression & Random Forest
  - a. This project utilizes various techniques for natural language processing. After the raw text is processed, two models are built to predict positive or negative sentiment.
7. PCA & K Means Clustering
  - a. This project uses principal component analysis to reduce dimensionality and clusters ALS data.
8. R – Linear Regression
  - a. In this project a linear regression model is built in R to predict a Pokémon's power.
9. Recommender System
  - a. This project builds a recommender system for movies.
10. Web Scraping
  - a. This project leverages BeautifulSoup to scrap a table from a Wikipedia page.