Projects: For all projects refer: [Your Repositories (github.com)](https://github.com/seanrodrigues1?tab=repositories)

**Vaccine Adverse Event Reporting: (Tableau)**

* Conducted data analysis of VAERS (Vaccine adverse event reporting system) data provided by CDC and designed an interactive Tableau dashboard to present the results to hospital management.
* Analyzed the adverse events (side effects) of covid-19 vaccine against the demographics of US population such as age & gender.

PROJECT LINK : [Vaccine Adverse Event Reporting | Tableau Public](https://public.tableau.com/app/profile/sean.rodrigues/viz/VaccineAdverseEventReporting_16937904732290/AdverseeventsreportingDashboard)

**Netflix Analysis Dashboard: (Tableau)**

* Designed an interactive dashboard to analyze the number of Netflix movies & shows released by year, country, genre and PG Rating.
* Created interactive filters to enable user to display the description & other details of a certain movie or show.

PROJECT LINK : [Netflix Analysis Dashboard | Tableau Public](https://public.tableau.com/app/profile/sean.rodrigues/viz/NetflixAnalysisDashboard_17038131979240/NetflixDashboard)

**Bank Customer Churn Prediction: (Predictive Analytics)**

* Developed a classification model to predict if a bank customer will discontinue his/her bank account or not based on factors such as credit score, tenure etc. with 87 % accuracy and a recall of 52%.
* Model and Evaluation: Evaluated several Machine Learning models such as logistic regression, SVM, random forest, XG Boost as well as an Artificial neural network Deep Learning model using Scikit-Learn &TensorFlowlibraries.

**Boston House Price Prediction Application (Python)**

* Developed a regression model to predict house prices in Boston based on factors such as crime rate, rooms etc. with a mean absolute error score of 0.58 % on the final XG Boost Regressor model.
* Created a frontend web application using Flask & HTML and utilized CI-CD pipelines to deploy the application on Heroku Cloud.

**Market Segmentation of Online Retail Customers (Unsupervised Machine Learning):**

* Developed a **clustering** model in python to segment retail customers based on the RFM model (Recency, Frequency, Monetary value) to help the company Identify and target its customers efficiently.
* Evaluated several clustering models such as K-Means Clustering, Hierarchical clustering and DBSCAN based on silhouette score.
* Result: 3 different clusters of customers were detected by the model and identified as – gold, potential and one-time customers

**Heart Disease Prediction: (R)**

* Conducted Exploratory data analysis using various R packages like carot, ggplot etc. and developed a classification model to predict if a patient has heart disease or not based on features such as blood pressure, Blood cholesterol level etc. with 95% accuracy.

**Clustering of Spotify Users: (Unsupervised Machine Learning)**

* Developed a Clustering model in Python to cluster Spotify users based on music/audio features such as acousticness and loudness.
* Evaluated K-Means, Hierarchical clustering and DBSCAN clustering models and selected K-Means as the final model with a silhouette score of 0.54.

**Impact of Multiplayer and Single Player gaming on Academics (Research Project)**

* Wrote a research paper to study the relationship between the type of game a student plays and the amount of time he spends playing and if there is an impact on academic performance.
* Developed a Google form questionnaire to survey students & conducted hypothesis testing (**t-tests)** in excel.

**Multi Cloud Data Migration:(DevOps)**

* Migrated the data of a hypothetical luxury hotel company to AWS S3, containerized & deployed the application to a Kubernetes cluster using **Kubernetes**, **Dockers** and Terraform on Google cloud platform.

**A/B Testing: (Statistical testing)**

* Developed an A/B test in python to test if a new webpage design for a hypothetical e-commerce company is more effective than the previous version of the webpage in terms of purchase conversion rate.
* Result: Obtained a Lift of -0.146% and a p-value of 0.22 after performing a chi-square test of independence.

**Sentiment Analysis using LLAMA2: (LLM/Gen AI)**

* Utilized the pretrained LLAMA2 chat **LLM** model performed prompt engineeringto conduct a sentiment analysis of a given input review/comment and returns the sentiment: positive, negative, or neutral. Implemented Streamlit to create the front end web app.