Trey Capps

Raleigh, NC | 919.306.8508 | trey.capps@gmail.com | treycapps.com | Github | LinkedIn

EDUCATION

NORTH CAROLINA STATE UNIVERSITY
Raleigh, NC

Bachelor of Science
Aug 2020 - May 2022

Major in Statistics; Minor in Mathematics

Relevant Coursework: Regression Analysis; Statistical Learning; Experimental Design; Statistical Data Management

SANDHILLS COMMUNITY COLLEGE
Southern Pines, NC

Associate of Science
Aug 2018 - May 2020

Committed 30 hours per week to Sandhills Community College Golf Team

PROFESSIONAL EXPERIENCE

INSTITUTE FOR TRANSPORTATION RESEARCH AND EDUCATION
Raleigh, NC

Statistical Analyst Intern
Aug 2021 – May 2022

- Independently designed experiments that leveraged regression techniques and hypothesis testing to evaluate intervention effectiveness from over a decade of time series data spanning multiple tables
- Implemented various time series techniques to determine high-impact violations which provided insights into ways to increase future intervention effectiveness
- Collaborated with a group of 3 GIS analysts to introduce and present program effectiveness to 20+ clients in a conference setting

PROJECTS

SPOT-IT: SONG RECOMMENDER
May 2022 – Current

https://github.com/trey-capps/spot-it

- Developing a web application allowing Spotify users to find similar songs based on various Reddit communities, creating a seamless integration of the two platforms
- Created Python scripts to implement custom class modules to automate ETL pipelines allowing for extraction of data across multiple APIs and storage of data in NoSQL databases
- Implemented a RESTful API to provide multiple endpoints for data retrieval and model predictions
- Utilized clustering and linear algebra techniques to generate similarity metrics providing users with song recommendations through a web application

UNDERGRADUATE RESEARCH
Raleigh, NC

North Carolina State University - Department of Statistics
Jan 2022 – May 2022

- Utilized logistic regression to identify key factors contributing to student success providing stakeholders with insights to make department specific policy changes
- Collaborated with a group of 4 students to present results at the research symposium

PATIENT SAFETY TOPIC MODELING
Oct 2021 – Dec 2021

- Utilized unsupervised learning to extract topics from device reports reducing the manual review time by more than one-half
- Suggested future problem areas using time series analysis and forecasting methods

SKILLS

Programming: Python, R, Git, SAS
Database: SQL (Postgres, MSSQL), NoSQL (MongoDB)
Visualization: Streamlit, Tableau, R-Shiny
Machine Learning: Regression, Classification, Clustering
Statistics: Time series forecasting, Hypothesis testing, Experimental Design
Some Experience: AWS (S3, Redshift), PySpark, Airflow

AWARDS

- 2019 NJCAA Division III Men's Golf National Champion - June 2019