Nathaniel Jones

linkedin.com/in/Nathaniel-Jones711/
github.com/njones738
nathaniel.jones711@gmail.com

Data Scientist

FOCUS:

Detail-focused Data Scientist with knowledge in numerical methods, parametric and nonparametric methods, binary classification and logistical modeling, multilevel/hierarchical statistical modeling, machine learning, data cleaning and variable reduction. Proven ability in Python, R, and SAS programming to visualize and describe data into actionable project plans. Dedicated and hard-working with passion for Data Science.

EDUCATION:

Kennesaw State University
Bachelor of Science in Applied Mathematics
- Minor in Applied Statistics and Data Analysis

Graduated: **2021—Dec**

Kennesaw State University
Master of Science in Applied Statistics

Expected Graduation: **2023—Dec**

RESEARCH PROJECTS & PRESENTATIONS:

"Why Withdraw?"—*R, Python, SAS* **2022—Dec**
- Continued research into the CollegeScorecard with a longitudinal analysis of the two-year withdraw rate at 4-year post-secondary schools in Georgia using a multilevel statistical model.
- Comments and Ratings left by reviewers were scraped from RateMyProfessor. A sentiment analysis of the comments was conducted for use in the model.

"Modelling User Chatbot Experience"—*R, Python* **2022—Dec**
- User interactions with an internal Southern Company chatbot were analyzed with modern text analysis methods to gain insight in improving the chatbots functionality and overall user experience. Insights and recommendations were packaged and presented to Southern Company.
- Python Libraries such as SpaCy and Transformers were used to process and create a topic model of user inputs and program responses.

"Does the Pell Grant come with a Price?"—*R, Python* **2021—Dec**
- Awarded 3rd place at KSU's 2021 Analytics Day and was selected to be presented at:
  - Posters on the Hill 2022 (1 of 88 posters)
  - Harvard National Collegiate Research Conference 2022
  - Posters at the GA capitol 2022
- Continued my research into the CollegeScorecard with a spatial look at the institutions.

"Classification of Pell Institutions"—*R, Python* **2021—Dec**
- Classified Post-Secondary institutions using the binary indicator created in "Access to Higher Education". Many different models were created including XGBoost, Principal Component Analysis, Random Forest, and Logistical Regression.

"Access to Higher Education"—*R* **2021—Apr**
- Conducted parametric and nonparametric analyses on the features of post-secondary institutions with a majority undergraduate population receiving a Pell Grant.

Two-Layer Neural Network—*Python* **2021—May**
- Created a an up to 2-layer multilayer perceptron in Python as a class object.

Using Logistic Regression to Build Credit Scores—*Python, SAS* **2021—May**
- Using 1.2 million observations and over 300 features, a logistic regression model predicted whether a customer was considered a good or bad credit risk. An analysis on the profitability of the model was conducted to find the best model variant.

CERTIFICATIONS:

Research Data Services @ Georgia State University Data Certification **2021—May**
- Completed workshops on data analysis tools (SAS, Python, and R), data analysis methods (Mixed Methods), and finding data (Marketing Data).

CODING LANGUAGES:

- **Python**
- **SQL**
- **R**
- **JAVA**
- **SAS**
- **HTML**
- **MatLab**

SKILLS:

- **Data Science and Analytics**
- **Machine Learning**
- **Text Analytics**
- **Topic Modeling**
- **Nonparametric and Parametric Methods**
- **Multilevel/Hierarchical Statistical Modeling**
- **Feature Engineering/Selection**
- **Numerical Analysis**
- **Binary Classification and Logistical Regression**
- **Longitudinal Data Analysis**
- **Statistical Learning**
- **Microsoft Office**
- **Number Wizard**
- **Git/GitHub**

FAVORITE PACKAGES:

- **R**
  - **magrittr**
  - **DataExplorer**
  - **skimr**
  - **patchwork**
- **Python**
  - **SpaCy**
  - **pyttsx3**
  - **BeautifulSoup**
  - **Transformers**