Nathaniel Jones

Data Scientist

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

Focus:

Detail-focused Data Scientist with knowledge in numerical methods, parametric and nonparametric methods, binary classification and logistical 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 Big Data.

EDUCATION:

Kennesaw State University — Graduated: **2021—Dec**
Bachelor of Science in Applied Mathematics
- Minor in Applied Statistics and Data Analysis

Kennesaw State University — Expected Graduation: **2023—Dec**
Master of Science in Applied Statistics

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).

Research Presentations & Projects

"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 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.

WORK EXPERIENCE:

Tyme Global—*Remote Data Entry Agent* — **2020—Jul** *to* **2020—Oct**
- Entered client information into databases with speed and accuracy and managed documents by organizing forms, filing records, and creating agent reports.
- Verified accuracy and validity of data entered in databases and corrected any data entry error to prevent later issues such as duplication or data degradation.

Dynata—*Remote Data Entry Agent* — **2020—Apr** *to* **2020—Aug**
- Collected responses to surveys and polls on behalf of various Research and Data collection firms and engaged in the data collection process where I quickly assessed new surveys and asked respondents for their input to the survey question.

SBK International—*Warehouse Data Manager* — **2020—Jan** *to* **2020—Jun**
- Evaluated inventory and supplies to check for quality and quantity issues while handling the day-to-day shipping of outgoing wholesale orders and receiving of incoming inventory. In addition, I redesigned the layout of the warehouse to accommodate new equipment, maximize space, and improve process efficiency.
- Oversaw and motivated team of three employees in warehouse to increase the value of the bi-weekly purchase order by 15% while reducing the time needed to collect and process an order by 30%.

CODING LANGUAGES:

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

SKILLS:

- **Data Science and Analytics**
- **Exploratory Data Analysis**
- **Machine Learning and Statistics**
- **Feature Engineering**
- **Numerical Analysis**
- **Nonparametric and Parametric Methods**
- **Binary Classification and Logistical Regression**
- **Creative Problem Solver**
- **Critical Thinker**
- **Organized**
- **Creative**
- **Goal-focused**
- **Perceptive**
- **Innovative**
- **Adaptable**
- **Active Learner**
- **Microsoft Office: Excel, Word, Powerpoint and Teams**
- **Number Wizard**
- **Git/GitHub**

PACKAGES:

- **Python**
  - **pandas, NumPy**
  - **scikit-learn, scipy, shap, dalex**
  - **sklearn**: model_selection, metrics, feature_selection, ensemble, svm, neural_network, linear_model
  - **matplotlib, seaborn, plotnine**
- **R**
  - **tidyverse, tidymodels, magrittr**
  - **ggplot2, sf, geofacet, ggpubr, ggh4x, tigris, tidycensus**
  - **dplyr, stats, perm, jmuOutlier, feather, DataExplorer**