Aaron Allen

P.O. Box 684, Cheney, WA, 99004
(509) 336-9542
atastats@gmail.com
linkedin.com/in/atastats/

I am actively searching for a challenging and interesting position in or related to the fields of data science and software development. With my background in Math, Statistics, and Computer Science, as well as my ability to quickly learn and understand new concepts, I believe I have what it takes to be a major contributor.

EDUCATION

MS in Statistics (Data Science and Machine Learning) GPA: 3.5 – Washington State University Dec. 2019

BS in Applied Mathematics (Computer Science) GPA: 3.8 – University of Idaho May 2017

Relevant Course Work: Advanced Linear Algebra, Data Science, Statistical Computing, Applied Linear Modeling, Statistical Theory, Data Structures, Combinatorics/Graph Theory, Probability Theory, Numerical Methods, Computability, Design of Experiments, Epidemiology and Biostatistics, and Multivariate Analysis

Programming Tools:

- **Proficient**: Python (numpy, scipy, pandas, matpotlib), R (ggplot, dplyr, tidyr)
- **Intermediate** Linux, SQL, Tensorflow, scikit-learn, Azure, SAS, MATLAB
- **Exposure**: Java, C

RELEVANT EXPERIENCE

Control Charts Based on Machine and Statistical Learning Techniques (Masters Project)

- Used Support Vector Data Description, Discriminant Analysis, and Logistic regression for classification of time series disease counts from a Bivariate Negative Binomial Distribution with extensions to Multivariate cases.
    - Other methods include Maximum Likelihood Estimation, distribution simulation, Markov Chains, and Bayes' Classification

Analysis of K-means and Hierarchical Clustering (Course Project)

- Used K-means and Hierarchical Clustering in an analysis of simulated Gaussian data to investigate performances and classified vehicle data through cluster analysis.

A Waiting Period Approach for Control Charting Using Markov Chains (Course Project)

- Applied Markov chains (specifically transition matrices) to develop a waiting period approach to classical control charting in hopes of achieving better classification between in and out of control data.

NFL Data Analysis and Prediction (Course Project)

- Cleaned large NFL API data set using dplyr in R, used various linear modeling methods for game winning predictions.

Twitter Scraping (Independent Project)

- Developed a Python script to pull tweets and other desired account data and clean for a format allowing future analysis.

Graduate Teaching Assistant – Washington State University, Pullman, WA Aug. 2017 – Dec. 2019

- Lab Instruction, material creation, grading, tutoring, and proctoring. Course topics included Sampling, Correlation, Regression, Central Tendency, Hypothesis Testing, Independence, Goodness of Fit, and ANOVA.

HONORS, CERTIFICATES, AND PROFESSIONAL AFFILIATIONS

Student Memberships

- Pi Mu Epsilon Honorary Mathematical Society (Idaho Alpha Chapter) Inducted May 2016

Dataquest.io Certificates

- Pandas and Numpy Fundamentals June 2020
- Elements of the Command Line June 2020
- SQL Fundamentals June 2020

References Provided Upon Request