


# Charles Atchison

## DATA SCIENTIST

 [CharlesAtchison](#)

 [Charles.Atchison@outlook.com](mailto:Charles.Atchison@outlook.com)

 [charleslatchison](#)

 [\(863\)- 409 - 4987](tel:(863)-409-4987)

## TECHNICAL SKILLS

- Python
- Pandas
- SQL
- Airflow
- NLP Modeling
- MS Excel Expert
- Project Management
- Time-Series
- Anomaly Detection
- Data Pipeline

## ABOUT ME

Expertise in data analysis, report presentation, time-management, and teamwork, with extensive remote work experience. Known for achieving significant improvements to report automation and data analysis with visualization. Highly resourceful and effective critical thinker and team player with years of military and private business experience.

## ACADEMIC HISTORY

*Harvard - Master's in Data Science 2023 - 2025*

Graduate of Harvard's Master's in Data Science program, an elite and interdisciplinary curriculum renowned for its comprehensive focus on statistical methodologies, machine learning, and large-scale data analysis. This rigorous program, spearheaded by world-leading experts, emphasizes hands-on experience and the application of cutting-edge techniques in real-world settings. With a foundation rooted in both theory and practice, I've been meticulously trained to tackle complex data challenges and drive impactful decision-making.

*University of Arizona - BA in Psychology 2019*

## RELEVANT WORK HISTORY

*Intelliware Solutions - Data Engineer - Feb 2020 - Jan 2023*

Utilized Apache Airflow DAGs to Extract data from a PostgreSQL server hosted on an Amazon Web Server. The DAGs would then Transform the data and match the data to each of the over 1,000 accounts. Lastly, these matches would Load into self-constructing Excel spreadsheets, and then they would send to over 10,000 individuals.

Built a Python script within Django framework to consume FedEx data feeds that contained millions of shipment histories and live data. Using this, the script would monitor for shipment event anomalies and would cross-reference SQL servers and notify our customers of the anomalies of their packages, saving millions of dollars by preventing medications and other goods from spoiling and correcting errors.

## PROJECTS

### American Attitude Topic Sentiment Prediction

Acquiring survey data from the Pew Research Panel, our team will be exploring the drivers of pessimism in American Prospective Attitudes. Understanding what most likely drives pessimistic or optimistic thinking about the future will help business leaders clarify strategies for moving forward and guide expectations of future success in the customers they serve, products offered, investments made, in Marketing and Sales, and throughout their business organization.

### Github Programming Language Prediction

Constructed a script that used BeautifulSoup to scrape Github Repo README's and the associated programming language percentages. Our team cleaned and explored the data and utilized Natural Language Processing to construct models that would predict the programming language based on the README. This model achieved over a 94.8% accuracy at predicting the programming language based on the README.

### Bitcoin Price Prediction Time-Series Analysis

Analyzed Bitcoin's seasonality utilizing Holt-Winters' trend forecasting, via two-parameter modeling and associated linear exponential smoothing. Achieved a root means squared error of ~\$53.07 via time-series behavior prediction.

### Zillow Zestimate Error Analysis

Utilized 3 clustering metrics with data queried SQL database where the Zillow data was stored. Examined potential correlations between multiple features within the dataset and tested hypotheses of the associated features. Upon testing it was discovered that there was an association between some of the clusters and the error within the Zestimate.

### Telco Drivers of Churn

Acquired data via SQL and explored via bivariate and multivariate feature analysis. My hypothesis utilized a single-tail analysis on the churn rate via payment types and confirmed, that individuals who utilized electronic checks had a significantly higher rate of churn than other payment types. I achieved an ~80% accuracy of churn prediction via hyperparameter optimization and feature engineering.