

# Alan J Averett

832.856.4666 | ajaverett0@gmail.com | linkedin.com/in/ajaverett | ajaverett.github.io

## EDUCATION

### Brigham Young University- Idaho

B.Sc. Data Science, Statistics; *Major GPA: 3.8/4.0*

Rexburg, ID

Apr 2020 - Dec 2023

- **Certifications:** Web and Computer Programming (3.7/4.0); Machine Learning (4.0/4.0)
- **Leadership:** Chi Delta Society, Society of Hispanic Professional Engineers (SHPE), Data Science and AI Society

## EXPERIENCE

### WPA Intelligence

*Machine Learning Intern*

Washington D.C.

Jul 2022 - Nov 2022

- Decreased production time (1 hr to 1 min) by building an automated script *in SQL and R* to clean and transform voter data
- Prepared raw data for database integration using *tidyverse* for **data manipulation, cleansing, and transformation**
- Implemented ODBC connections in *R and SQL*, orchestrating the seamless integration of **over 100 million records** of transformed voter data from several spatial data sources into a database

### FamilySearch

*Statistician*

Salt Lake City, UT

Sep 2023 - Present

- Eliminated 50+ fraudulent contractors in the African History Project by **developing and automating** in-house tuned anomaly detection methods in *Python* over 200k+ genealogical submissions; regularly presented abstracts to management

### Mathematics Department BYU-Idaho

*Data Science Society President, Chief Lab Manager*

Rexburg, ID

Mar 2023 - Sep 2023

- Led and coordinated data science society activities and meetings; mentored and supported project teams including weekly bootcamps to teach **data analysis/statistics and ML hyper parameter tuning**, over 6 teams of ~ 6-12 people each
- Managed, and trained all tutors for school-wide DS Lab, teaching *Pandas, NumPy, PySpark, SQL*, in *DataBricks*

*Teacher's Assistant*

May 2022 - Present

- **DS Programming:** Tutored data analysis, statistics, and basic ML in *Python (Pandas, Sklearn)* and *SQL*
- **Data Wrangling and viz:** Tutored complex data analysis and visualization with *R tidyverse* (dplyr, ggplot2 etc.)
- **Big Data Programming:** Tutored big data computation and analysis with *PySpark and SQL* in *DataBricks*
- **Linear Algebra:** Tutored both non-computational and computational linear algebra with *Python's NumPy and SymPy*

## CONSULTING

### Baltimore Mission (Spanish Outreach)

*Technology Consultant and Regional Manager*

Baltimore, MD

Mar 2019 - Feb 2020

### Wilford Woodruff Papers

*Data Scientist*

Rexburg, ID

May 2023 - Dec 2023

- Provided actionable graphs and trends using **time-series analysis** and **NLP (sentiment analysis and textual clustering)**
- Delivered **interactive web app in Streamlit** that matched various corpus data together using text vectors (*NLTK, sklearn*)

### Ohio Voter Data Consulting

*Founder, Data Scientist*

Rexburg, ID

Sep 2023 - Present

- Democratized **big data** models for grassroots campaigns, **Google cloud computing** ML models for ~10 million rows in *Python's sklearn* for predictive analysis in voter behavior, visualizing data interactively with *R Shiny, Plotly, Quarto*
- Awarded First in nationally recognized MLH hackathon for innovation and creativity by optimizing campaign outreach

## SKILLS

- **Python:** Pandas, NumPy, Sci-kit Learn, PySpark, Polars, Keras / Tensorflow, BeautifulSoup4, DataBricks
- **R:** CAR, Mosaic, RMarkdown, Tidyverse (dplyr, tidyr, stringr, forcats, lubridate, purrr, etc.)
- **Visualization libraries:** Streamlit, ggplot2, Matplotlib / Seaborn, Plotly Express, Altair
- **Other Technical Skills:** Excel, Visual Basic for App. (VBA), SQL, MySQL, Git/GitHub, Quarto, LaTeX, Regex, Docker