

# TONATIUH DE LEON

**Location:** Rexburg, Idaho

**LinkedIn Profile** [Link](#) • **Phone:** (208) 948-3742

**GitHub Projects** [Link](#) • **Email:** kayashy.de.leon@gmail.com

---

## RELATED EXPERIENCE

### Data Engineer | The Church of Jesus Christ of Latter-Day Saints | Utah 2024

- Validated data integrity post-migration from SQL servers to Databricks by using Python HashMap testing, ensuring 100% table accuracy for critical ETL jobs.
- Optimized 2 SQL queries using Python functional programming, reducing code length by 80% and doubling runtime efficiency for ETL processes.
- Developed Python widgets and leveraged Azure Blobs to automate a continuous updating system of a live streaming table in Databricks, maintaining data integrity for the Analysts team.

### Data Analyst | BYU-Idaho 2023 - 2024

- Developed a decision-support app for the VP that enabled strategic planning by providing easy access to 12 institutional KPIs, leveraging advanced SQL, R, Power-BI, and DAX logic.
- Led 4 stakeholder meetings employing user-centered design principles to prioritize requirements. Such collaboration led to agreements on key app features, enhancing user functionality.
- Automated department Excel data sharing system, saving 40+ weekly hours using VBA & macros.
- Integrated Workday data into R; automated cleaning & reporting process of 250 employee records.
- Maintained 5+ reports in Power-BI by updating 50+ DAX measures for accuracy & decision-making.

### Analytics Consultant | Aspire365 Mental Health Provider | Connecticut 2024

- Designed report with 20+ KPIs for company data utilization audit, leading client to pass audit.
- Conducted statistical analysis on patient depression/anxiety data, identifying 4 focus groups for therapy enhancement, increasing treatment effectiveness.
- Collaborated with a 6-member team to develop a statistical analysis app in R Shiny, streamlining data insights for mental health professionals.
- Enhanced visualization tool in app, optimizing it to generate 100+ professional quality graphs.

### Spotify Regression | Personal Data Science Project ([link here](#)) 2024

- Extracted Spotify API data into SQL, cleaned data for 1M+ artists, enabling precise analysis.
- Modeled a multiple linear regression in R, predicting daily artist streams with  $\sim .95 R^2$  accuracy.
- Validated model with an  $R^2$  error variance of 0.01, ensuring high reliability.

---

## EDUCATION

**B.S. Statistics | BYU-Idaho** 2021 - Graduating April 2025

**Minor:** Data Science • Economics

**Certificates:** Machine Learning • Business Analytics • Mathematics

**GPA:** 3.5 Overall • 4.0 Major

---

## SKILLS

**Programming Languages:** 2 Years of experience on average

- R, SQL, Python, DAX, BVA.

**Data Tools and Platforms:**

- Databricks, Apache, Azure, Power-BI, Tableau, Git, Excel.

**Technical Skills:**

- Machine Learning, Regression/Time Series Analysis, Experimental Design, Database Design/Management, Cloud Data Warehouse Practices. Education/health/survey Data Experience.

**Languages:**

English, Spanish, Haitian Creole.