

Skills

[1em]0em

- Database: SQL Server (SSIS, SSAS, SQL Server Agent, sqlcmd, bcp), Oracle (APEX)
- Programming: R, T-SQL, Bash/Zsh
- Data Visualization: Tableau, R (ggplot)
- Version Control: Git, Apache Subversion (SVN)
- Documentation: LaTeX, Markdown, Jupyter Notebook, Lucidchart

Database: Vertica , SQL Server (SSIS, SSAS) , Oracle Cloud: AWS (Redshift, Athena, s3, Lake Formation, EC2, Lambda, Glue) Data Visualization: Tableau , Amplitude , R (ggplot) Programming: Bash/zsh , R, Python (working knowledge) Version Control: Git , Apache Subversion (SVN) Documentation: LaTeX, Markdown, Jupyter Notebook, Lucidchart Certifications: AWS Certified Cloud Practitioner

- Collaborated with Human Resources and the Office of Faculty Advancement on the analysis of faculty salary to ***optimally allocate over \$660K*** of salary increases to minimize salary compression and the number of inversion cases among faculty ranks
- Partnered with campus' IT to migrate the Institutional Research department's database from end-of-life hardware to a virtualized server
- Conducted multiple training sessions within the Institutional Research department on SQL and version control covering basic usage and best practices
- Presented at the 2019 California Association for Institutional Research Conference on the construction and design of a comprehensive longitudinal student data set and an accompanying data dictionary

Experience

Intuit April 2021 - Present Business Data Analyst Revamped the Turbo Tax Full Service (TTFS) Summary dashboard by reducing its data source refresh time from 45 min to 5 min Performed an exploratory analysis on customers who could not be serviced through the TTFS product by classifying its out-of-scope/detach reasons, raw text from Turbo Tax experts Developed a topic model from TTFS' Product Recommendation Score (PRS) survey data to identify any common themes among the customers' responses Automated the querying and moving of PRS data from AWS Athena to Vertica to join new and historical data Assisted with ad-hoc requests involving clickstream data and with any TTFS-related asks

San Diego State University San Diego August 2018 - April 2021

Data Scientist and Database Developer

Automated the ETL components of the campus' Data Lake pilot project which includes scripting AWS Lambda functions, configuring Glue crawlers, developing database objects in Redshift and establishing permissions for various data users in Lake Formation Collaborated with Human Resources and the Office of Faculty Advancement on the analysis of faculty salary to optimally allocate over \$660K of salary increases to minimize salary compression and the number of inversion cases among faculty ranks Partnered with campus' IT to migrate the Institutional Research department's database from end-of-life hardware to a virtualized server Devised a method for producing a data dictionary using command line tools, a database management system, and LaTeX/Markdown as an alternative to a \$39K per year third-party solution Presented at the 2019 California Association for Institutional Research Conference on the construction and design of a comprehensive longitudinal student data set and an accompanying data dictionary Developed a script to programmatically update the data source of over 100 Tableau dashboards

Education

San Diego State University Master of Science in Statistics December 2019 San Diego State University
Bachelor of Science in Statistics May 2015