

Jeff Boerger

512-565-2373 – jeff.boerger@gmail.com - linkedin.com/in/jeffboerger - jeff.boerger.co - github.com/jeffboerger

Data Engineer - Python - SQL - Extract, Transform, Load (ETL) Pipelines

PROFESSIONAL SUMMARY

Data Engineer and Analyst with a BS in Computer Science and hands-on experience designing end-to-end ETL data pipelines, building cloud data infrastructure on Snowflake, orchestrating workflows with Apache Airflow, and deploying interactive analytics dashboards on real-world business data. Proven ability to translate complex datasets into actionable business insights. Background as a business co-founder and franchise operator brings a practical, results-first approach to data problems that most engineers lack.

TECHNICAL SKILLS

Languages: Python, SQL, R Programming, Bash

Libraries: Pandas, NumPy, Matplotlib, Seaborn, Plotly, Scikit-learn, SQLAlchemy, Requests, BeautifulSoup4

Data Engineering: Apache Airflow, Snowflake, dbt, ETL/ELT Pipelines, PostgreSQL, SQLite, Data Modeling, Data Warehousing

Tools and Platforms: Git, GitHub, Streamlit, Power BI, Jupyter, VS Code, Microsoft Office (Excel, PowerPoint, Word), Tableau

EXPERIENCE

General Manager, Freedom Fun USA, Houston, TX

Aug 2024 - Present

- Reduced monthly commission reconciliations from 3–5 hours per location to under 5 minutes by building a Python automation tool that scrapes live transaction data from the ERS booking platform via BeautifulSoup4, applies split-commission rules and non-commissionable exclusions per company policy, and exports formatted Excel, PDF, and CSV reports — designed for multi-location scalability across franchise operators
- Grew regional franchise revenue 300% from \$30K to \$120K by analyzing customer demand patterns, identifying seasonal trends, and executing targeted sales strategies
- Present operational data, revenue performance, and strategic recommendations directly to C-suite leadership including CEO and President in regular executive meetings
- Reduced emergency escalations to management by designing training protocols and operational checklists, leading a team across scheduling, training, and daily warehouse operations
- Eliminated manual payroll errors and reduced scheduling overhead by building Excel-based operational reporting systems for timesheets, scheduling, and commission tracking across the franchise

Undergraduate Data Analytics Researcher, Brookhaven National Laboratory, Upton, NY

Jun 2023 - Aug 2023

- Delivered anomaly detection analysis as the sole undergraduate on a supercomputing performance research team by parsing deeply nested JSON provenance data using recursive traversal algorithms — extracting call stack sequences, execution runtimes, and outlier severity scores across thousands of function entries
- Enabled multi-dimensional anomaly pattern recognition by developing Plotly visualizations — sunburst charts mapping call stack hierarchies by severity, violin plots of execution time distributions across compute ranks, and scatter plots correlating hashed call paths with normalized anomaly scores
- Accelerated pattern-based anomaly detection across high-volume supercomputer run data by applying SHA-256 hashing to call stack sequences to identify and track unique execution paths
- Contributed findings and visualizations to a CRA-CERP research poster

Clinic Manager and Co-founder, Body Traxx Chiropractic, Pflugerville, TX

Nov 2016 - Feb 2020

- Grew Body Traxx Chiropractic from a two-person startup to a three-person practice by co-founding the clinic and managing patient scheduling, billing workflows, and staff coordination — systematized operations to the point of making own role obsolete
- Protected patient data integrity across all billing and medical records by administering the ChiroTouch EHR system including user authentication, access control, and HIPAA-compliant data security
- Improved physician visibility into patient retention by building an Excel-based visit tracking system adopted by clinic doctors to monitor scheduling patterns and flag at-risk patients

Jeff Boerger

512-565-2373 – jeff.boerger@gmail.com - linkedin.com/in/jeffboerger - jeff.boerger.co - github.com/jeffboerger

EDUCATION AND CERTIFICATIONS

Bachelor of Science in Computer Science, Texas State University, San Marcos TX

Dec 2023

Relevant coursework: Data Structures and Algorithms, Introduction to Machine Learning, Software Engineering, Introduction to Probability and Statistics

Certifications: Google Business Intelligence (May 2024), Google Advanced Data Analytics (Mar 2024), Google Data Analytics (Feb 2024)

PROJECTS

Data Engineer - Independent Project, Clinic Revenue Pipeline

Aug 2024 - Present

- Engineered a Python ETL data pipeline extracting, cleaning, and loading 9 years of multi-source Excel data (3,287 rows) into PostgreSQL and Supabase - modular architecture with single-responsibility functions and a dbt transformation layer producing staging and mart models
- Iterated on data pipeline architecture over multiple agile development cycles - progressively migrating from SQLite to PostgreSQL to Supabase cloud, refactoring transformation logic, and expanding the dbt model layer as data complexity grew
- Quantified COVID impact on subscription revenue through statistical modeling and Exploratory Data Analysis (EDA) - discovered only a 6% revenue decline in 2020, validating the stickiness of the subscription model
- Identified and traced a \$9,645 single-day revenue outlier to a non-recurring bulk payment, preventing distortion of operational revenue benchmarks
- Deployed interactive Streamlit dashboard connected to Supabase cloud database - visualizing \$1.3M in revenue across 9 years with KPI metrics, YoY growth, billing cycle pattern analysis, and dynamic year filtering

Data Engineer - Independent Project, Theme Park Analytics Pipeline

Apr 2026 - Present

- Engineered a real-time data pipeline ingesting live ride wait times across all four Walt Disney World parks via the ThemeParks.wiki REST API - capturing standby waits, Lightning Lane status, and hourly forecasted wait times per attraction
- Orchestrated hourly data ingestion using Apache Airflow on a cron schedule, loading raw snapshots into a Snowflake cloud data warehouse with a timestamped time series architecture enabling trend analysis across hours, days, and attractions
- Built a dbt transformation layer with staging models cleaning raw API data and mart models aggregating average, max, and min wait times over rolling 24-hour windows by park and by ride
- Connected Snowflake data warehouse to Power BI Service for enterprise BI reporting layer integration - explored dashboard capabilities and documented migration path beyond free tier
- Deployed interactive Streamlit dashboard querying Snowflake directly - visualizing real-time park comparisons, ride-level wait trends, and Lightning Lane availability across Walt Disney World

Data Engineer - Independent Project, Resume Keyword Analyzer

Jun 2024 - Present

- Built a Python Natural Language Processing (NLP) tool scoring resume-to-job-description alignment using fuzzy string matching, set-based keyword comparison, and percentage gap scoring - supports .txt, .docx, and .pdf resume uploads
- Implemented multi-source job description ingestion supporting local file upload and live URL scraping via Requests and BeautifulSoup4 - enabling real-time analysis against any publicly posted job listing
- Curated role-specific keyword datasets from real 2026 job postings for Data Engineer, Data Analyst, and SWE roles; deployed as a live Streamlit web app; actively used to optimize own job applications