

Michael Walker

michaelsvwalker@gmail.com | 503-329-6143 | msvwalker-portoflio.com | github.com/MSVWalker | linkedin.com/in/msvwalker

Professional Summary: Data & AI professional with 9+ years turning data into shipped models, dashboards, and decisions. Have built reliable pipelines (dbt, Airflow), automated reporting, and shipped ML/LLM apps (XGBoost, RAG).

Skills

Data & ML: Python (pandas, numpy), SQL, scikit-learn, XGBoost, Monte Carlo, time series, A/B testing, DOE
Data Eng & Platforms: dbt, Airflow, ETL, Snowflake, BigQuery, PostgreSQL/Neon, Docker, AWS (EC2, S3), Git/GitHub
Apps, BI & LLM: FastAPI, Next.js/React, Tailwind, Recharts, Tableau, Power BI, Looker, LangChain, RAG, OpenAI API

Experience

Data Scientist, Teal Harbor Properties - Portland, ME (Remote) Jan 2022 - Present

- Automated SQL ETL pipelines to Snowflake reporting for ~600 properties, **freeing 100+ hrs/mo** for finance & ops.
- Standardized 18 KPIs** across properties; retired 15 legacy spreadsheets and cut reconciliations ~80%.
- Deployed Monte Carlo and XGBoost models to score acquisitions; adopted in 98% of investment memos.
- Shipped self-serve KPI dashboards (Tableau/Looker) standardizing finance metrics; cut underwriting from 4h to 30m and accelerated investment-committee decisions from monthly to weekly.

Tech: Python, SQL, dbt, Airflow, Snowflake, Looker, Tableau, Git

Staff Data Analyst, Hitachi High Tech - Hillsboro, OR June 2018 - Jan 2022

- Led 8 engineers/analysts across 35 projects; contributed to \$20M in new product revenue.
- Built unified SQL data models for 52k plasma-etch experiments (multi-source), cutting development time 40%.
- Trained ML models in scikit-learn to predict process outcomes and reduce run-to-run variance by 90%.
- Delivered Power BI dashboards and Python reports used by R&D, Product, and Execs for roadmap decisions.

Tech: Python, SQL, scikit-learn, Power BI, DOE, Git, Hypothesis Testing

Data Analyst, Hitachi High Tech - Dallas, TX Jan 2016 - June 2018

- Invented and patented two plasma-etch processes, helping secure multimillion-dollar client contracts.
- Synthesized insights from 300+ technical papers to guide product direction and experiment design.
- Designed A/B and DOE across 6k+ trials, boosting throughput & quality by 6.7× (+670%).

Tech: Python, SQL, DOE, JMP

Independent Projects - Remote

- Bellingham WA Weather Analytics App (2025)** - Full-stack
Ingests daily weather to Postgres/Neon; models with dbt; scheduled via Airflow; FastAPI backend; Next.js/Tailwind frontend (Recharts). Focus on historical data, anomalies & rolling trends; mobile-friendly dashboard.
Stack: Airflow, dbt, Postgres/Neon, FastAPI, Next.js, Tailwind
- Tenant Response Assistant (2025)** - RAG/LLM
Piloted a WhatsApp assistant + VA dashboard (RAG over leases/work orders/SOPs), improving first-response time ~45% (VA surveys) and standardizing replies.
Stack: LangChain, OpenAI, vector store, FastAPI
- Glucopredict (2024)** - Predictive Health with Machine Learning
Feature-engineered CGM sequences and lifestyle signals; 86% accuracy with SHAP-based interpretability and input validation; shipped a Streamlit app for scenario testing (“what-if” changes to sleep, meals, activity).
Stack: Python, XGBoost, pandas, scikit-learn, Streamlit

Education & Certifications

B.S. Physics, Math Minor | University of Florida – GPA 3.59/4.00 Aug 2010 - June 2014
Data Science Career Track, Springboard | 600+ hours, weekly 1:1 professional mentorship Apr 2024 - Nov 2024