Michael Willy, Data Engineer

michaelwilly.com | linkedin.com/in/michaelwilly/ | github.com/chriswilly

- Holds over five years of experience developing data projects across diverse teams and compute platforms.
- Develops automated models with Python & SQL to deliver research analytics. Familiar joining complex data sets from data warehouses, web APIs, and raw file formats. Deployment using version control, automated testing & formatting, and code review.
- Automates models on high-dimensional data consumed by notebook and Power BI or Dash-scripted dashboards. Works with customer to tailor the analytics story using visualizations for the the target audience.

Tools

Python (Polars SQL, Pandas, NumPy, Scikit-Learn, SQLAlchemy, Plotly), Git, Power BI, LaTeX
SQL (Oracle, MS SQL, SQLite), Amazon AWS S3, Bash & PowerShell, Cron, Airflow, Docker, Markdown

Experience

(1) **Data Engineer Applications**, Research Chemical Company 2023
- Operation and data automation of corporate R&D Electronic Lab Notebook (ELN)
- SQL accuracy validation for migrations, partitioning updates, and data sanitization for developers
- Build automated Python Polars SQL-context csv report generation on Linux & Microsoft platforms
- SQL & Python LDAP and regex reports for user patterns presented in Power BI visual dashboards

(2) **Data Science & Engineer**, Research Chemical Company 2020
- Developed Python & SQL pipelines to extract lab documents transformed into database model
- Built Fast Fourier Transform (FFT) analysis & plotting pipeline command line and notebook interface
- Built visualization with Dash Plotly, Jupyter Notebooks, and Microsoft Power BI to create a process visualization valued by another vendor as a $300k quote

(3) **Pilot Project Engineer**, Research Chemical Company 2018
- Integrated and automated time series and part wise database reporting using SQL & Python on virtual plant environment
- Designed machine user interfaces, instrumentation, and the data model for a new lab factory
- Managed contract equipment installation and start up for nine pilot machines valued at $800k capital

(4) **Research Engineer**, Research Chemical Company 2014
- Scaled-up organic light emitting diode (OLED) purification and programmed robotic device fab

Work Projects

- Fast Fourier Transform Numpy & SciPy signal processing automation with Plotly dashboard, JSON API, and runtime SQLite database November 2022 - June 2023
- Developed a graph Euclidian distance semi-supervised machine learning model to classify then estimate profiles using sets of process inputs trained on semiconductor wafer processing March - May 2022
- Data pipeline from distributed labs with regular expression file tree search. User view into database was PHP site producing a parameterized SQL report via a command line Python API. January - September 2022

Education

M.S. Applied & Computational Mathematics, University of Washington, Seattle, WA December 2023
B.S. Chemical Engineering, Minor in Mathematics, Drexel University, Philadelphia, PA June 2011

Date: December 21, 2023.