THANH VU

Data Analyst | SQL · Python · Power BI · Excel

Lancaster, PA | vulethanh07112002@gmail.com | (717) 869-9502 | linkedin.com/in/thanh25 | github.com/thanhvu911 | Portfolio

SUMMARY

Data Analyst experienced in economic development, deal sourcing, and retail sales analytics. Builds automated KPI dashboards and reporting with SQL, Python, Excel, and Power BI that save teams 15+ hours a month.

EDUCATION

Franklin & Marshall College — Lancaster, PA

Bachelor of Arts in Computer Science | *GPA: 3.72/4.0* — May 2026

- Magna Cum Laude; Honors List (5 semesters); Dean's List (Spring 2022)
- Coursework: Data Science, Statistical Modeling, Probability & Statistics, Applied Regression, Machine Learning, Algorithms

EXPERIENCE

Data Analyst Fellow | EDC Lancaster County (Center for Regional Analysis), *Lancaster, PA* — Aug 2026 – Present

- Automated 25+ BLS and BEA labor, GDP, and wage series into a star-schema Power BI dashboard with 20 years of history and daily refresh, replacing manual data pulls and saving the team **15+ hours per month**.
- Designed a Google Sheets–backed web dashboard with time-series filters, automated imports, and a Gemini AI assistant that summarizes trends and flags anomalies, letting **6+ non-technical staff** update data without code.
- Delivered **4+ dashboards and analyses**, including a downtown economic analysis, by gathering requirements with the Center's Director and documenting each dataset in a data dictionary.

Data Science Intern | KCC Capital Partners, *Sheridan, WY* — Jan 2026 – May 2026

- Cleaned and consolidated data on **100+ precious-metal and mining companies**, markets, and transactions, and enriched **500+ contact records** with Seamless.AI, giving the deal team a reliable sourcing database.
- Screened companies for active mandates with the firm's built-in AI tools, shortlisting **100+ likely buyers** for outreach.
- Launched automated weekly valuation and market-forecasting reports and dashboards, saving the deal team **5+ hours per week**.

Data Engineer Intern | KPIM Joint Stock Company, *Hanoi, Vietnam* — Jan 2021 – Apr 2021

- Streamlined legacy client data workflows with Python (Pandas, NumPy) and SQL Server, cutting processing time by **30%**, eliminating manual entry errors, and feeding Power BI reports validated with the analysis team.

Technology Support Assistant | Quantitative Science Center, Franklin & Marshall College — Sep 2021 – Sep 2022

- Resolved **50+ technical incidents per week**, maintaining **99% classroom uptime**, and cut semester setup time by **40%** by standardizing computer images.

PROJECTS

Can Mail Data Predict Housing Vacancy? | *Microsoft Fabric, Dataflow Gen2, Power BI, DAX* — Sep 2026

- Tested whether quarterly USPS mail data can signal the annual Census vacancy rate across 6 counties and 84 county-years, building a Fabric Lakehouse pipeline into an interactive Power BI report.
- Measured agreement with a DAX-built Pearson correlation and directional accuracy, finding a pooled **r = 0.77** but only **57.7%** directional agreement, so USPS data is a county-dependent signal, not a substitute.

Grocery Transaction Data Warehouse | *SQL Server, T-SQL* — Jul 2025 – Aug 2025

- Built a Bronze–Silver–Gold warehouse over **131M+ transaction lines** with stored-procedure ETL and 21 automated data-quality checks, then showed that the top 20% of customers generate **44.7%** of $3.4B in revenue.

Sales Planning Analysis | *Power BI, DAX, SQL Server* — Aug 2025

- Developed a KPI dashboard comparing actual and planned revenue across **58K+ transactions** and 36 months, surfacing district plan completion of **89%–98%** and year-over-year swings of +37% and −15%.

RFM Customer Segmentation | *Power BI, DAX* — 2025

- Segmented **18.4K customers** into 9 RFM groups with DAX calculated columns that score recency, frequency, and monetary value from 1 to 4, pairing each segment with a recommended marketing action.

SKILLS & CERTIFICATIONS

Analysis & BI: Power BI (DAX, Power Query), Excel (pivot tables, lookups), data visualization, KPI reporting, statistics, data cleaning
Languages: SQL (T-SQL, Spark SQL, MySQL), Python (Pandas, NumPy, Matplotlib), R (ggplot2), KQL, Google Apps Script
Platforms & Data: Microsoft Fabric, SQL Server, Power Automate, Git/GitHub; BLS, BEA, Census ACS, and USPS data
Certifications: Microsoft Certified: Fabric Data Engineer Associate (DP-700), Oct 2026; HackerRank SQL (Advanced), Sep 2025

LEADERSHIP

Technology Committee | Diplomatic Congress, Franklin & Marshall College — Aug 2021 – Aug 2022

- Cut classroom technology costs by **18%** through IT policy work and centralized license management across **15+ departments**.