Om Patel

omrpatel@outlook.com | (647) 231-8391 | /omrpatel | /ompat5 | omrpatel.ca

Programming Languages: Python, SQL, R
Libraries: Pandas, Numpy, PySpark, PyTorch, Scikit-Learn, Matplotlib
Technologies: Spark, Snowflake, Looker, MySQL, Oracle DB, Hadoop, Power BI, Fivetran, DBT, Git

EDUCATION

University of Waterloo Sept 2021 - Apr 2026

Bachelor of Mathematics, Honours Data Science (Co-op)

- **Courses**: Data Analysis, Stochastic Processes, Generalized Linear Models, Sampling and Experimental Design, Distributed Systems

WORK EXPERIENCE

Data Scientist Intern @ Synopsys May 2025 - Dec 2025

- Designed and implemented a full-stack **data transformation engine** using **Qt**, featuring a custom **Lark** mathematical parser to evaluate user-defined functions on tabular data.
- Enabled **in-platform data manipulation** capabilities for statistical modeling and reduced analysis time by **20%** by eliminating data import/export workflows.
- Integrated **predictive modeling** features using **scikit-learn** to implement **ML algorithms** including Logistic Regression, GLMs and K-Means clustering.
- Streamlined the **experiment-analysis-reporting** workflow by providing in-platform statistical modeling and data preprocessing, removing reliance on external tools.

Data Scientist Intern @ Ontario Ministry of Education Sept 2024 - Dec 2024

- Supported future policy analysis by implementing schema changes on **normalized relational** and **star schema dimensional** databases, using **dbt** for impact analysis, backward compatibility checks, and data integrity validation.
- Decreased data preprocessing time from **1 hour** to **5 seconds** (**99.86%** faster) by building an automated **Pandas** pipeline for **300K+** urban planning entries, eliminating the client's manual workflow and enabling future housing market modeling.
- Led technical discussions with external clients to define preprocessing requirements for urban planning data files.

Data Scientist Intern @ Darwin CX Sept 2023 - Apr 2024

- Improved forecasting accuracy by **30%** by implementing a **K-Means clustering** model to segment sales data and account for seasonal fluctuations in purchasing behaviour.
- Reduced Looker report loading time by **98%** by establishing **persistent derived tables** that precompute large joins and aggregations.
- Decreased Analytics Division operations costs by **10%** by optimizing **Snowflake** compute resources and refactoring SQL queries.
- Built **Medallion architecture** data pipelines using **dbt** for silver/gold layer transformations, data validation, and **job orchestration**.

Data Analyst Intern @ Darwin CX Apr 2023 - Sept 2023

- Developed **20+ Looker** reports to transform raw data into actionable insights for client-specific requirements.
- Reduced audit resolution time by **50%** by creating SQL queries for faster data integrity verification.

Data Analyst Intern @ Ontario Ministry of Transportation Sept 2022 - Dec 2022

- Created **10 dynamic Power BI** reports on IT application usage to improve organizational business decisions.

PROJECTS

Mobile Game A/B Test Analysis | Product Analytics • Kohavi's Experiment Framework

- Designed and ran an A/B test on **90K+** users to evaluate gate placement effects on **retention** and **monetization proxy metrics**.
- Validated experiment integrity by including **SRM tests** and **guardrail metrics** to ensure engagement volume was not cannibalized.
- Recommended a variant that increased retention with no drop in engagement by using **proportions z-tests** and **permutation tests**.

Billboard Campaign Impact Analysis | Marketing Analytics • Causal Inference

- Measured the causal impact of a regional billboard campaign using **Difference-in-Differences** when A/B testing was not feasible.
- Fit an **Ordinary Least Squares** regression with time and treatment interactions to estimate a **6.52-unit** lift in customer deposits.