

KUSH PATEL

Analytics Engineer | Cloud Data Warehousing | Dimensional Modelling | Power BI Semantic Layer | SQL, Python, PySpark
Toronto, ON (relocating from Vancouver, BC) | +1 (778) 637-1602 | dharma.patel552@gmail.com
[linkedin.com/in/kush-patel-48885719b](https://www.linkedin.com/in/kush-patel-48885719b) | github.com/KushPatel29 | kushpatel29.github.io

SUMMARY

- Analytics engineer with 4+ years across analytics engineering, data engineering, and business intelligence: converting raw omnichannel source data (ERP, POS, e-commerce, inventory, CRM, marketing) into clean, structured warehouse layers and the trusted data products built on them.
- Designs and builds data marts on cloud data warehouses (Microsoft Fabric, Azure Synapse, SQL Server, Snowflake) applying dimensional modelling best practices including star schemas, conformed dimensions, and slowly changing dimensions, with cloud-native ELT orchestration in Azure Data Factory, Fabric Data Pipelines, and Airflow.
- Delivers Power BI end to end, from semantic model design through polished, publish-ready reports, giving commercial, merchandising, finance, and supply chain teams self-service access to trusted data without analyst dependency.
- Treats data quality as engineering: automated testing frameworks, pipeline alerting, referential-integrity and control-total checks, and documentation that keeps the mart well understood. Every portfolio project ships with tests in CI and a published failure analysis: github.com/KushPatel29 | kushpatel29.github.io

CORE COMPETENCIES

Dimensional modelling (star schema, slowly changing dimensions Type 1/2), data mart design, semantic layer design and self-service enablement, cloud-native ELT/ETL orchestration, data quality testing frameworks, pipeline monitoring and alerting, omnichannel source integration (POS, e-commerce, ERP, CRM, REST APIs), ad hoc analysis and root-cause investigation, demand forecasting for planning and buying, cross-functional partnering with merchandising, planning, finance and commercial teams, metric definition and lineage, Git and CI, RLS/RBAC

TECHNICAL SKILLS

Cloud Data Warehousing: Microsoft Fabric (Lakehouse, Warehouse, OneLake, DirectLake), Azure Synapse Analytics, Snowflake, Azure Databricks, SQL Server, PostgreSQL

ELT / ETL & Orchestration: Azure Data Factory, Fabric Data Pipelines and Dataflows Gen2, Apache Airflow, SSIS, SQL Server Agent, incremental and CDC-style loads

BI & Semantic Layer: Power BI (DAX, Power Query, semantic modelling, relationship design, performance tuning), SSRS paginated reporting, Plotly Dash, Excel (Power Pivot, Power Query)

Languages: SQL (T-SQL), Python (pandas, NumPy, scikit-learn, SQLAlchemy, Flask), PySpark, DAX, REST APIs and JSON

Modelling & Governance: Kimball star schema, conformed dimensions, slowly changing dimensions, Bronze/Silver/Gold medallion layers, semantic layer design, metric dictionary, data lineage, cutover runbooks

Data Quality & Reliability: Automated test suites, schema-drift detection, null, duplicate and referential-integrity checks, control totals, anomaly thresholds, reconciliation dashboards, failure alerting and incident runbooks

Analytics & Forecasting: Time-series forecasting (Prophet, ARIMA), Random Forest, Isolation Forest anomaly detection, cohort and customer-behaviour analysis

Version Control & CI/CD: Git, GitHub, GitHub Actions, Jenkins, automated test pipelines

PROFESSIONAL EXPERIENCE

Two Rivers Specialty Meats - Vancouver, BC

Data Analyst, Operations, Logistics & Enterprise Reporting

Dec 2023 - Present

- Owned end-to-end data transformation converting raw omnichannel source data into clean, structured layers: integrated ERP, Square POS, Synerion workforce, and REST API feeds into curated Bronze/Silver/Gold datasets serving ad hoc analysis, BI dashboards, and operational reporting.
- Designed and built components of the enterprise data mart applying dimensional modelling best practices, including star schemas, conformed dimensions, and slowly changing dimensions, producing reliable, performant data products sliced by product/SKU, supplier, region, customer, and channel.
- Developed Power BI reporting end to end, from semantic model design and standardized DAX measure libraries through publish-ready reports and drill-through, giving Sales, Finance, and Supply Chain self-service access to trusted data without analyst dependency.
- Engineered cloud-native ELT pipelines in Microsoft Fabric and Azure Data Factory/Synapse using SQL and PySpark, with Airflow-orchestrated jobs, incremental loads, parameterized transformation logic, and structured error handling.
- Maintained data quality standards across the mart: automated schema, completeness, uniqueness, and referential-integrity tests, anomaly thresholds on key measures, exception logging, pipeline alerting, and documentation so the curated layer stayed trustworthy across the business.
- Built reconciliation controls comparing source against curated totals for GL/P&L and inventory valuation, and led root-cause investigations across ERP, POS, and API sources to resolve discrepancies at the feed level rather than patching the report.
- Performed ad hoc analysis in SQL and Python on margin, customer churn, and inventory movement, and delivered customer intelligence reporting (account inactivity signals, lost-account tracking, rep coverage) that informed planning, buying, and operational decisions.

- Built demand forecasting models (Prophet, ARIMA, Random Forest) deployed as Python web applications, giving planning and purchasing teams SKU-level forecasts to reduce stock-outs and perishable waste.
- Owned the legacy SQL Server estate through migration: maintained SSIS packages and SSRS paginated reports while refactoring priority pipelines to Fabric notebooks, validating cutover with row counts, checksums, and control totals so no reporting was disrupted.
- Partnered with Sales, Finance, and Supply Chain to define governed KPIs (revenue, COGS, gross margin, OTIF, inventory turns, days on hand), documenting metric definitions, acceptance criteria, and runbooks.
- Implemented secure-by-design reporting with RLS/RBAC, column-level filtering, and masking so sensitive cost and margin data stayed protected as self-serve access widened.

Shivam Investments - India / Remote

Financial Analyst (Data & Reporting)

Dec 2020 - Jul 2022

- Gathered reporting requirements from portfolio and risk stakeholders and delivered Power BI dashboards with drill-through and auditable measure definitions.
- Automated recurring analysis using Python (pandas) and SQL views feeding Power BI, replacing manual spreadsheet consolidation and improving reporting timeliness.
- Performed GL/P&L reconciliations across sources and investigated discrepancies to maintain finance-grade accuracy.

ANALYTICS ENGINEERING PROJECTS

Supply Chain Control Tower | *Microsoft Fabric Lakehouse, Azure Data Factory, PySpark, SQL, Power BI*

- Built a lakehouse-backed data mart (Bronze/Silver/Gold) with incremental loads, unit-of-measure conversion, and shelf-life logic over inventory and logistics feeds.
- Modelled curated star schemas with slowly changing dimensions and a governed semantic layer, applying RLS/RBAC so each function sees only authorized margin and inventory data.
- Delivered a KPI suite (inventory turns, days on hand, OTIF/fill rate, stock-out and expiry risk) with exception flags, drill-through to SKU and lot, and refresh monitoring.

Data Quality & Observability Framework | *Fabric, Azure Data Factory, SQL, Power BI*

- Operationalized data quality monitoring across warehouse layers: schema drift, null and duplicate checks, referential integrity, and anomaly thresholds on key measures.
- Built reconciliation dashboards comparing source against curated totals for GL/P&L and inventory valuation, and drove root-cause investigations to closure.
- Implemented alerting and incident runbooks that converted repeat pipeline failures into documented, one-touch fixes.

Transaction Risk & Alert Triage Engine | *Python, scikit-learn, PostgreSQL, Streamlit, GitHub Actions*

- Built a rules-plus-anomaly detection pipeline flagging suspicious transaction patterns including structuring, rapid layering, round-dollar anomalies, and dormant-account reactivation, with an Isolation Forest layer to surface behaviour the rules do not encode.
- Implemented cost-model threshold optimization weighing false-positive review cost against missed-detection risk, replacing an arbitrary cutoff with a defensible operating point.
- Shipped with a full test suite running in CI, documented assumptions, and a published failure-analysis section stating where the model underperforms and why.
- Delivered a Streamlit triage dashboard ranking alerts by risk score with drill-through to the contributing rule and feature values.

Legacy BI Modernization | *SQL Server, SSIS, SSRS, T-SQL to Microsoft Fabric*

- Maintained and enhanced a SQL Server BI estate: SSIS control and data flows with SCD patterns and SQL Agent schedules, plus SSRS paginated reports and subscriptions for finance-grade statements.
- Refactored priority pipelines to Fabric Data Pipelines, Dataflows Gen2, and PySpark notebooks, replacing stored-procedure-heavy transforms with scalable notebook patterns.
- Executed parallel-run validation (row counts, checksums, control totals) and documented cutover runbooks to retire technical debt without disrupting stakeholders.

EDUCATION

Northeastern University, Vancouver, BC - Master of Professional Studies in Analytics (Applied Machine Intelligence), GPA 3.76/4.0 | Sep 2022 - Apr 2024

Gujarat Technological University, India - Bachelor of Engineering in Computer Science | Aug 2017 - Aug 2021

CERTIFICATIONS

- Google Data Analytics; Power BI Data Modelling with DAX; Advanced SQL; IBM Python for Data Science, AI and Development
- SQL for Data Science: Data Pipelines, Database Management, Data Modeling; Microsoft Certified: Fabric Analytics Engineer Associate (DP-600), in progress