Mohit Ravindra Kamble

kamble.mo@northeastern.edu | (857) 891-8496 | linkedin.com/in/mohitravindrakamble | www.mohitkamble.com

Summary

Recent MS graduate in DA with expertise in developing ML pipelines, ETL workflows, and BI dashboards. Expert in Python, SQL, PySpark, cloud platforms, transforming healthcare data into AI-powered predictive insights for executive stakeholders.

Education

Northeastern University, Boston, MA **Apr 2025**
MS in Data Analytics & Statistical Modelling GPA 3.89
University of Mumbai, India **Jul 2021**
Bachelor of Engineering in Information Technology GPA 3.7

Skills

- **Programming Languages:** Python, SQL, R, MATLAB, VBA, T-SQL, HTML
- **Data Engineering Tools:** MySQL, PostgreSQL, SSMS, SSRS, Alteryx, Azure Databricks, PySpark, Snowflake, Oracle, SSIS, Azure Synapse Analytics, ERP, SPSS, SAS
- **Cloud Platforms:** AWS (S3, Glue, Sagemaker), GCP (BigQuery), Microsoft Azure
- **AI/ML, Analytics & Visualization:** Supervised & Unsupervised Learning, NLP, pandas, NumPy, scikit-learn, PyTorch, Matplotlib, seaborn, ggplot, tidyverse, Hypothesis testing, ANOVA, t-test, A/B testing, chi-square, LLM, Power BI (Power Query, DAX), Tableau, Google Data Studio, Looker, Domo, Qlik View, Amazon Quicksight
- **Certifications:** Alteryx Designer Core Certification, Google Ads Search Certification, Google Ads Measurement Certification, HIPAA Certification, HITECH Certification, CHPS Certification

Experience

Data Analyst, Cognizant, India **Aug 2021 – Aug 2023**

- Developed 20+ Power BI dashboards, Power Apps solutions, and Power Automate workflows while writing advanced T-SQL queries and stored procedures to analyze large healthcare datasets (>1M records), collaborating with stakeholders to translate business requirements into technical specifications that reduced administrative costs by 30%.
- Designed scalable data models in Azure Databricks/PySpark and implemented ETL processes via Azure Synapse Analytics and SSIS, integrating Oracle, SQL Server, and SAS data to build enterprise BI solutions using SSRS and Qlik View for Medicaid-similar healthcare analytics.
- Performed advanced data analysis to identify trends and patterns, presenting actionable insights to executive leadership that drove 16% improvement in business performance while mentoring 2 junior team members on T-SQL, Power Platform development, and stored procedure optimization.
- Implemented Agile methodologies with Azure DevOps for BI project delivery, conducting root cause analysis (RCA) to improve data quality and developing process documentation that ensured scalable, maintainable healthcare analytics solutions across the enterprise.

Projects

Boston Public Health Commission (Master's Capstone Project) | Python, Azure, SQL, Power BI **Jan 2025 – Mar 2025**

- Led 3-member team developing Power BI dashboards and T-SQL stored procedures in Azure Databricks for opioid overdose analysis, automating healthcare workflows with real-time visualizations for data-driven city decisions.
- Designed enterprise data models and ETL pipelines for overdose analytics while mentoring team on BI architecture.

FDA Drug Analytics Platform| Python, Dash, ML Models, Render Cloud **Jun 2025 – Jul 2025**

- Developed AI-powered FDA analytics platform leveraging ensemble ML algorithms achieving 98.8% drug approval prediction accuracy across 914K records and implemented automated anomaly detection AI system using Isolation Forest identifying 9,355 high-risk applications with prescriptive intelligence recommendations.

OTT Content Analytics Pipeline | GCP, Apache Airflow, PySpark, Power BI **May 2025 – Jul 2025**

- Architected GCP data pipeline processing 10M+ IMDb records using Dataflow, Dataproc, BigQuery, and Cloud Composer with automated change detection and metadata tracking.
- Built real-time streaming infrastructure with Pub/Sub, Apache Beam, and Dataflow Flex Templates, implementing zone failover and CI/CD via GitHub Actions.
- Developed Power BI dashboards for OTT analytics with KPI tracking and content metrics, leveraging BigQuery dimensional modeling and PySpark transformations.