

DARSHIL M KAPADIA

darshilkapadia97@gmail.com | [LinkedIn:darshilkapadia-us](https://www.linkedin.com/in/darshilkapadia-us/) | [github/darshilkapadia](https://github.com/darshilkapadia) | United States

SKILLS

- **Programming Languages:** Python, R, SQL, C++
- **Machine Learning/ Workflow Automation:** TensorFlow, Scikit-learn, PyTorch, Airflow, Mage AI, Git
- **Data Manipulation & ETL:** Pandas, NumPy, PySpark, DBT, Apache Spark
- **Data Visualization:** Power BI, Tableau, Looker, Matplotlib, Seaborn
- **Cloud Platforms:** Azure Data Lake, Azure Data Factory, Azure Synapse, AWS, Snowflake
- **Databases:** Google BigQuery, SQL Server, MySQL, MongoDB

WORK EXPERIENCE

Data Analyst, *Client-FIRST CHATHAM BANK*

06/2022 – Present

Tech Stack: Power BI, Azure Data Factory, Apache Spark, Snowflake, Azure Synapse, Python, Data Analysis

- Boosted lead conversion by 3.5% by developing Early Warning KPI Dashboard in Power BI, integrating real-time web traffic data and lead scoring enabling marketing to target customer segments and prevent a 10% ROI drop in campaign performance
- Saved 3 hours weekly by automating data pipelines with Azure Data Factory and with finance metrics into interactive Power BI dashboards, enabling real-time performance tracking, accelerating decision turnaround time by 20%
- Reinforced marketing campaign effectiveness by 12% by enabling scalable customer data insights through an Azure analytics platform (Azure Blob Storage, Data Factory, and Synapse Analytics) to provide a 360° view of customer interactions
- Lowered loan default rates by 18% by developing a credit risk model in Scikit-learn and deploying it on Azure ML, analyzing 50+ variables including income patterns, credit history, and repayment behavior for precise risk evaluation
- Generated actionable insights into campaign and social media metrics by building PowerBI dashboards on Snowflake models, exposing follower growth data correlations to optimize marketing initiatives for 25% engagement improvement

Data Integration Analyst, *KAVAYE INFOTECH SERVICES*

06/2019 - 11/2020

Tech Stack: Python, Excel, Scikit-learn, DBT, MySQL

- Reorganized data processing for retail customer inquiries with refactored code using libraries like Pandas and potentially Scikit-learn, enabling 20% faster processing of 29+ GB customer data and facilitating a 15% faster response time to inquiries
- Developed SQL transformations and data pipelines using DBT (data build tool) to cleanse, manipulate, and transform the data in Snowflake, ensuring high-quality and consistent data for downstream analytics and reporting
- Enabled efficient analytics data flow by integrating 10+ enterprise sources and developing scalable Python/Pandas ETL pipelines for streamlined extraction/transformation, ensuring on-time delivery through scheduling and distributed computing
- Maintained 30+ Excel workbooks by creating pivot tables for data exploration, macros to automate repetitive tasks, importing data from MySQL databases, and generating clear and informative reports

PROJECTS

AI BLOG KOINONOS Python, Streamlit, Google AI (Gemini Pro LLM), OpenAI (DALL-E 2)

- Developed a user-friendly Streamlit interface for an AI blog content generator powered by Google's Gemini Pro LLM
- Empowered users to customize parameters like title, keywords, word count, and image count within the Streamlit interface
- Integrated DALL-E 2 API for creating visually stunning images to accompany generated text
- Enabled users to customize blog length and integrate AI-generated visuals for enhanced engagement

NYC TAXI ANALYTICS ACCELERATOR Google BigQuery, Looker, Data Modeling, Mage AI

- Designed a scalable data model utilizing fact and dimension tables to efficiently store and analyze 3 GB of ride-hailing data
- Automated data pipelines with Mage (open-source data orchestration tool) to efficiently transform and prepare ride-hailing data in BigQuery for analysis of key metrics like peak times, hot locations, ride fares, and average tips
- Utilized Looker's drill-down functionality to enable users to explore ride-hailing data at various levels of granularity, facilitating deeper analysis of metrics like average trip fare and average tip amount by specific timeframes or locations

EDUCATION

Stevens Institute of Technology - Master of Computer Science

Amity University - Bachelor of Computer Science & Engg

CERTIFICATIONS

- Data Analysis using PySpark
- Database Design in PostgreSQL
- BigQuery for Data Analyst