

# Soham Prasad Ambekar

Dallas, Texas, 75252; 4692659450; [sohamambekar21@gmail.com](mailto:sohamambekar21@gmail.com), [LinkedIn](#); [GitHub](#)

## EDUCATION

**The University of Texas at Dallas**, MS, Information Technology and Management

August 2022 - May 2024

**University of Mumbai**, BE, Computer Engineering

July 2015 - May 2019

## CERTIFICATIONS

[Google Data Analytics](#); [Oracle Cloud Infrastructure](#); [Data Management](#); [Databricks Lakehouse Fundamentals](#)

## PROFESSIONAL EXPERIENCE

**Kroger**, Dallas, US

May 2023 – August 2023

*Technology Intern/Data Analyst*

- Operationalized business intelligence solutions and conducted advanced analytics to identify opportunities and strengths to meet business goals. Presented and communicated the findings in an understandable and actionable way.
- Orchestrated Power BI dashboards integrated with Snowflake and spreadsheet data, enabling updated KPI monitoring, reporting, and presenting findings to stakeholders, enhancing business-driven decision-making by 50%.
- Performed Python-based predictive modeling utilizing linear regression & clustering algorithms to forecast demand for perishable items for brick-and-mortar stores, optimizing statistics for inventory management by 9%.
- Delivered reports and insights using PowerPivot for ad-hoc analysis, resulting in a 10% increase in client satisfaction.

**Ingram Micro**, Mumbai, IN

June 2019 – April 2022

*Business Data Analyst*

- Created and maintained data visualizations using Power BI that educated, informed, and engaged business partners on key metrics and performance measures increasing stakeholder engagement by 25%.
- Discovered dead carts & proposed a 30-day auto-delete mechanism to prevent inventory holding on large orders.
- Implemented automated ETL pipelines using Python scripts to extract data from databases & generate executive reports, streamlining reporting processes, saving 10 hours per week of manual work, and enhancing report accuracy.
- Pioneered 10+ complicated views utilizing SQL that powered Power BI dashboards used by 4 internal teams.
- Leveraged functional expertise & automated HTTP/SFTP protocols to process data to customers with 99% accuracy.
- Led IBM Sterling B2B Integrator application support for EMEA and APAC, facilitating L2/L3 production issues by interpreting data failures and implementing unique solutions to ensure complete uptime of business processes.
- Leveraged MS Excel skills to enhance data management efficiency, including VLOOKUP and Pivot Tables.

## PROJECTS

### [Instacart Market Basket Analysis](#)

- Focused on identifying top 10 products, peak traffic hours, best-selling products, customer visit cycles, and most frequently repurchased items through time series analysis to uncover the bigger picture in customer buying trends.
- Launched 7+ dashboards displaying the Service, Product, and Category evaluations & designed a compelling data story.

### [Natural Disasters Impacting Refineries](#)

- Pinpointed tornadoes and flooding as top disproportionate risks, jointly accounting for 50% of disasters threatening Texas/Louisiana refineries through calculated data synthesis.
- Championed an interactive multi-map spotlighting vulnerable refineries, enhancing disaster risk identification by 20%.

### [Credit Risk Analysis](#)

- Demonstrated proficiency in handling imbalanced datasets, feature engineering, visualization libraries using cross validation, hyperparameter tuning, and interpreting machine learning models to derive valuable business insights.
- Programmed credit risk prediction model using XGBoost, achieving a 0.72 recall score in identifying bad credit risks.

### [Pricing Analysis of Airlines](#)

- Identified key pricing parameters across destinations by analyzing Excel flight datasets, reducing pricing errors by 11%.
- Formulated aggregate functions, pivot tables, and filter controls to uncover pricing trends, empowering informed business strategy and travel decisions through data-driven insights.

## SKILLS

**Programming:** Python, R, Java, PySpark, CI/CD, GitHub Actions, UNIX, Linux, Shell, SAS

**Databases:** MySQL, SQL Server, MongoDB, SSIS, SSRS, DynamoDB, Data Cleaning, Data Modelling, DAX, Snowflake

**Libraries:** NumPy, Pandas, sci-kit-learn, Matplotlib, PyTorch, TensorFlow, Keras, seaborn, Plotly, Bokeh

**Cloud:** Microsoft Azure, AWS, Google Cloud Platform, Data Science, EC2, S3, Sagemaker, QuickSight, SAP, Visio

**Tools:** Tableau, Power BI, JIRA, Confluence, Git, Jupyter, ArcGIS, Excel (PivotTables, V-Lookups, Macros)