Soham Prasad Ambekar

Dallas, Texas, 75252; 4692659450; sohamambekar21@gmail.com, LinkedIn; GitHub

SUMMARY

Skilled analyst with 3+ years of expertise in data analytics, business analysis, and cloud technologies.

EDUCATION

The University of Texas at Dallas, *MS Information Technology and Management*   August 2022 – May 2024
University of Mumbai, *Bachelor of Science, Computer Engineering*   June 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

- Propelled a 3% uplift in sales through data-driven demand forecasting for perishable items using Python pipelines with linear regression and clustering, optimizing space & floor planning in the merchandising space.
- Orchestrated Power BI dashboards integrated with SQL and spreadsheet data, enabling real-time KPI monitoring of sales per store, traffic, and summarized findings for stakeholders, realizing 4% cost savings.
- Delivered reports and insights using SQL for ad-hoc queries, resulting in a 10% increase in client satisfaction.

Ingram Micro, Mumbai, IN   June 2019 – April 2022
Business Systems Analyst/Data Analyst

- Spearheaded communication of $1B annual sales, optimizing pricing & promotions to attain a 2% profit margin.
- Discovered dead carts & implemented a 30-day auto-delete mechanism to prevent inventory holding on large orders.
- Articulated automated reporting processes using Python and created scripts to extract data from databases and generate executive reports (ETL), saving 10 hours per week of manual work.
- Pioneered 10+ complicated SQL views utilizing BigQuery that powered Tableau dashboards used by 4 internal teams.
- Leveraged SQL expertise & automated HTTP/SFTP protocols to process data to customers with 99% accuracy.
- Led IBM Sterling B2B Integrator application support for EMEA and APAC, resolving L2/L3 production issues by interpreting data failures and implementing unique solutions to ensure complete uptime of business processes.

PROJECTS

Credit Risk Analysis   April 2024

- Programmed credit risk prediction model using XGBoost, achieving a 0.72 recall score in identifying bad credit risks.
- 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.

Pricing Analysis of Airlines   September 2023

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

Natural Disasters Impacting Refineries   March 2023

- 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 Tableau map spotlighting vulnerable refineries by visualizing key performance indicators (KPI) metrics, enhancing disaster risk identification speed and accuracy by 20%.

Instacart Market Basket Analysis   February 2023

- Recognized the top 10 products based on the Instacart data set by visualizing the grocery data using Tableau.
- Focused on identifying peak traffic hours, best-selling products, customer visit cycles, and most frequently repurchased items through in-depth data analysis to optimize operations and delivery.
- Launched 7+ dashboards displaying the Service, Product, and Category evaluations & developed the data story.

SKILLS

Programming: Python, R, Java, PySpark, CI/CD, GitHub Actions, UNIX, Linux, Shell, Git, SAS
Databases: MySQL, SQL Server, MongoDB, Apache Cassandra, Redshift, DynamoDB, Data Modelling, DAX
Libraries: NumPy, Pandas, Scikit-learn, Matplotlib, PyTorch, Data Science, Keras, ggplot, seaborn, Plotly, Bokeh
Cloud: Microsoft Azure, AWS, Google Cloud Platform, Databricks, EC2, S3, Sagemaker, QuickSight
Tools: Tableau, Power BI, JIRA, Confluence, Git, Jupyter, ArcGIS, Excel, Pivot tables, SharePoint, Visio, SAP