SHUBHAM RAVINDRA KOSHTI

315-952-7490 • srkoshti98@gmail.com • LinkedIn • GitHub • Portfolio

SUMMARY

A data-driven professional with over 2 years of experience in data analysis, specializing in developing and maintaining interactive dashboards and automated data models that drive business decisions. Proficient in SQL, Python, Databricks and BI tools, with a strong focus on ensuring data quality and delivering actionable insights through rigorous analysis and visualization. Adept at collaborating with cross-functional teams to align data initiatives with strategic goals, enhancing decision-making and operational efficiency

WORK EXPERIENCE

Analytics Engineer | iConsult Collaborative | Syracuse, NY   **August 2023 – Present**

- **Reduced manual search time** by identifying reliable sources and automating data fetching and preprocessing stages for 5 important metrics using Apache Airflow, Databricks and PySpark
- **Identified regional patterns** critical to conservation efforts by conducting exploratory data analysis (EDA) via collaborating on GitHub codespaces, leading to a **20% increase in reporting accuracy**
- **Collaborated with Ecology and Research departments** to develop a comprehensive semantics layer using Tableau, enhancing end-user adoption and facilitating deeper ecological insights
- **Managed ad-hoc ecological SQL report requests from stakeholders**, adapting to changing business decisions and ensuring alignment with strategic goals through cross-functional validation processes

Data Analyst | Infosys Limited | Pune, India   **June 2020 – April 2022**

- **Enhanced order efficiency** by identifying delays and redundancies at each stage of customer order journey through SQL queries on orders database and Python-driven analysis
- **Spearheaded customer segmentation analysis** using Python for gaining insights into customers' order purchasing behavior in the Singapore region, enabling more precise targeting in marketing campaigns
- **Reduced order processing time by 30%** by streamlining role and permission management within **Salesforce CRM**, leading to faster order fulfillment and improved customer satisfaction
- **Contributed to efforts of improving conversion rates by 12%** by analyzing impact of marketing strategies and recommending targeted adjustments based on customer segmentation insights

PROJECTS

FEMA Disaster Analysis: End-to-End Workflow Automation | GitHub   **September 2024 – December 2024**

- **Designed and implemented Apache Airflow DAGs** to automate weekly ingestion of disaster-related data from FEMA APIs into Azure Blob Storage, thus reducing manual intervention
- **Built a scalable pipeline architecture** integrating Airflow and Databricks for streamlining end-to-end data processing and analysis for 3 FEMA datasets having around a million records at present
- **Established ODBC connection between Tableau and Databricks**, enabling real-time, interactive dashboards for disaster trends, funding insights and public assistance analysis
- **Achieved efficient data processing and transformation** by utilizing Unity Catalog, PySpark and SparkSQL, resulting in consistent, high-quality datasets

Predictive Employee Attrition Analysis and Modelling | GitHub   **April 2024 – June 2024**

- **Achieved 96% accuracy** in predicting employee attrition using Logistic Regression and Random Forest models, helping management identify and address key attrition drivers
- **Outlined a potential 15% reduction in annual attrition** through recommended changes in work scheduling and job satisfaction programs, supporting long-term employee retention
- **Optimized analytics processes** by recommending data preprocessing enhancements and integrating qualitative feedback, improving model interpretability and decision-making efficiency

Adidas US Sales Performance and Profitability Analysis | GitHub   **October 2023 – December 2023**

- **Improved decision-making** for 5 key metrics developing dynamic Tableau dashboards, providing actionable insights
- **Projected 15% increase in digital sales** by proposing strategies to boost online sales in underperforming US regions
- **Enhanced profitability in the Northeast region** by recommending targeted digital marketing initiatives to stakeholders

TECHNICAL SKILLS

Analytics Tools: Tableau, Microsoft Excel (Statistical Modelling, Hypothesis Testing), Dash Plotly, Looker
Programming Languages: Python, R, SQL, Spark, Java, C++
Workflow Tools: Apache Airflow, ServiceNow, Kubernetes, Databricks Workflows
Database Management: Azure Data Studio, PostgreSQL, Docker, PowerApps
Cloud Platforms: AWS (Lambda, EC2, S3), Azure, Salesforce CRM, Databricks

EDUCATION

Syracuse University | Master of Science | Information Systems, Data Science | GPA: 3.8   **August 2022 – May 2024**
Graduate Teaching Assistant | Martin J. Whitman School of Management
Pune University, I2IT | **Bachelor's in Engineering | Electronics and Telecomm.** | GPA: 3.7   **August 2016 – May 2020**
Co-Secretary | IEEE Pune Chapter 2019