Yash Khatavkar

yashkhatavkar27@gmail.com | (469)951-3468 | linkedin.com/in/yashkhatavkar24 | github.com/yashkhatavkar

PROFESSIONAL EXPERIENCE

Data Analyst Intern, Sedin Technologies Inc, Sunnyvale, USA   **May 2024– Jul 2024**

- Automated sales forecast dealer business report in Tableau, improving financial planning and budgeting accuracy by 30%.
- Designed ETL scripts to streamline and process sales data from CSV to Azure Data Factory for 130+ tables every 15 minutes.
- Created and optimized complex stored procedures for daily partner relationship-management reports and business operations.
- Ingested data from varied tables, executed SQL queries to validate vendor list and successfully reduced data duplicity by 80%.
- Leveraged Python for seamless integration of sales data to existing databases for robust data analysis of third-party vendors.

Data Analyst and Engineering Intern, BlueStar India Limited, India   **Nov 2022– May 2023**

- Standardized text data using Pandas, Regex to improve data integrity & processing throughput achieving 90% data accuracy.
- Conducted geospatial analysis using ArcGIS API to geocode nearest customers and branches and reduced service time by 10 hours.
- Leveraged AWS Textract to extract invoice specifications, enhancing product documentation management for 100K+ records.
- Applied advanced data cleaning and preprocessing techniques using Pandas and NumPy to refine application and vendor data.
- Collaborated with cross-functional teams to create a Tableau dashboard for customer segmentation clearing 6-month backlog.

Data Analytics Engineer, Mitibase Technologies Private Limited, India   **Feb 2021 – Oct 2022**

- Engineered a Neo4j graph database for 500K+ records, reducing interconnection data visualization reporting time by 40%.
- Collaborated with business & cross-functional teams to deliver ad-hoc reports using SQL queries, reducing response time by 25%.
- Implemented the integration of CI/CD pipelines using Jenkins and Git for existing standalone and web applications efficiently.
- Directed detailed root cause analysis and implemented SQL scripts to resolve critical tickets, improving resolution time by 60%.
- Cleaned and preprocessed complex JSON data using Python, ensuring data quality & consistency in analyzing Indian startups.

Product Data Analyst, Bajaj Allianz General Insurance Company Limited, India   **Aug 2019 – Jan 2021**

- Devised Tableau Product Scorecard of 10+ KPIs for real-time operational analysis; saving 17 hours of manual work each week.
- Executed competitive data analysis using advanced Excel functions: VLOOKUP, XLOOKUP & Pivot tables to get valuable insights.
- Built 4 Qlik Sense dashboards covering performance metrics for 200 insurance agents, enhancing performance tracking by 30%.
- Benchmarked 31 competitors' financial statements, identifying business insights to optimize premiums & underwriting strategies.
- Conducted market and competitor analysis for bite-sized insurance, identifying key customer segments & distribution channels.
- Collaborated with marketing and underwriting teams to design product growth plans by refining rate & coverage strategies.
- Managed product life cycle testing and conducted UAT to ensure successful digital enhancements and feature improvements.

ACADEMIC PROJECTS

Bank Loan Analysis *Tableau, Microsoft SQL Server, Microsoft Excel*

- Cleaned and preprocessed data in MS SQL Server to ensure quality and consistency, leading to accurate actionable insights.
- Developed an interactive Tableau dashboard with calculated fields and slicers to monitor KPIs and facilitate data exploration.

Credit Card Fraud Detection *Python, Jupyter Notebook, Pandas, NumPy, Matplotlib, scikit-learn*

- Performed data cleaning, data preprocessing and exploratory data analysis on a 600k-record dataset to ensure data quality.
- Utilized Logistic Regression, Decision Tree Classifier and Gaussian Naive Bayes. Achieved 99.98% accuracy in Random Forest.

EDUCATION

University of Texas at Dallas, USA, *Master of Science in Business Analytics*   **Aug 2023 – May 2025**
Coursework: Database Management, Spreadsheet Modelling, Machine Learning, Big Data, Data Modeling, Data Warehousing

IIIT Bangalore, Bengaluru, India, *Technical Diploma in Data Science*   **May 2022 – Aug 2023**
Coursework: Data Visualization, Exploratory Data Analysis, Inferential Statistics, Time Series Forecasting, Data Strategy

TECHNICAL SKILLS

Programming Languages– Python, R, SQL, SAS
Databases and Cloud– SQL Server, PostgreSQL, MySQL, SSIS, SSMS, Apache Spark, AWS S3, Redshift, Oracle, NoSQL
BI and Marketing– Tableau, Power BI, Qlik Sense, Excel, Google Analytics, MS-Access, Adobe Analytics, ArcGIS, Neo4j
Soft Skills– Collaboration, Problem Solving, Adaptability, Strategic Planning, Decision Making, Communication, Time-Management