Purva Prakash Kekan

Boston, MA | (857)-869-4863 | kekan.p@northeastern.edu | LinkedIn | Portfolio

SUMMARY

Data Analyst with hands-on experience in banking and pharmaceuticals sector, specializing in predictive modelling, dashboard development, financial reporting, inventory analysis and risk assessment. Recently completed Masters in Analytics from Northeastern University demonstrating skills in using tools like SQL, Python, Spark, Power BI, and Tableau to analyze complex data and develop insightful dashboards

EDUCATION

Northeastern University, **Boston, MA**
Masters in Analytics | GPA: 3.92/4.0 *March 2025*
Courses: Predictive Analytics, Big Data, Risk Management, Data Mining, Decision Support/Business Intelligence, Data Warehouse, Enterprise Analytics, Visual Data Analytics, Healthcare/Pharma Data and Applications, Introduction to Artificial Intelligence

Mumbai University **Mumbai, India**
Bachelor of Science in Information Technology | GPA: 8.92/10 *May 2023*
Courses: Data Structures, Python, Business Intelligence, NoSQL Database, Project Management, Probability and Statistics

TECHNICAL SKILLS

- **Data and Visualization:** Python (NumPy, Pandas, PySpark, Scikit-learn, Matplotlib, Seaborn), R, Tableau, PowerBI, Machine Learning Algorithms, GenAI Applications, A/B testing, Workflow Automation, MS Excel (Macros, VLOOKUP, VBA Scripting)
- **Data Management and Big Data**: ETL Pipelines (Alteryx, Spark), SQL (SQL Server, MySQL, PostgreSQL), NoSQL (MongoDB, DynamoDB), Oracle, AWS (Glue, S3, Redshift, Lambda, EMR), Hadoop, Cloudera, Databricks, Airflow, Data Lakes
- **Project Management and Workflow Tools**: Agile, Scrum, Jira, Git, Monday, Confluence

PROFESSIONAL EXPERIENCE

HDFC Bank Ltd (via TalentPro India HR Private Limited) **Mumbai, India**
Data Analyst *January 2023 – June 2023*

- Created **financial models** and dashboards to track **loan portfolio performance** maximizing capital market data solutions using **Looker** & **Power BI** to identify trends with attention to details improving marketing campaign returns by **15%**
- Handled large financial datasets by building **Spark applications** using **Python** & **PySpark** on **AWS** to handle over **10M records** daily which enhanced **budget analysis**, **business performance** and expanded **credit risk** assessment accuracy to **87%**
- Automated **data workflows** using **Alteryx** & **Python** to streamline **ETL workflows** and integrating **Snowflake** & **Databricks** for analysis which reduced financial data processing time by **30%** and refined **risk prediction** accuracy by **25%**
- Conducted **cross-functional collaboration** to analyze user engagement and financial processes by using **A/B testing** with **Google Analytics 4** and generating reports using **SQL** escalating user engagement by **20%,** for better financial procedures

Acute Bioscience **Gujarat, India**
Data Analyst *July 2022 – December 2022*

- Designed **Tableau dashboards** to analyze **product performance** by applying trend and **KPI analysis** which boosted data-driven **decision-making** by **20%** and helped management in resource allocation
- Developed **forecasting models** for **inventory management** by applying **statistical analysis** on past **sales data** which achieved the prediction accuracy of **92%** and optimized the **production planning** for better **business performance**
- Improved **data gathering** and data quality by utilizing **SQL Server stored procedures** which upgraded retrieval speed and minimized manual entry errors for more effective analysis
- Streamlined **documentation** and **data migration** by implementing **Excel macros**, **VBA scripting**, and **SQL queries** which raised order cycle times by **20%** and lessen **supply chain** mismatches by **25%,** driving operational **transformations**

ACADEMIC PROJECTS

- **Opioids Concern - Healthcare Analytics:** Developed a Power BI dashboard on the Boston 311 Syringe Requests data with 3M+ records, using geospatial analysis, predictive analytics through Python and seasonal trend analysis to uncover the patterns in demands of syringe requests, allocating resources optimally and SLA compliance [**GitHub**]

- **Text Mining Sentiment Analysis:** Performed NLP analysis on the "I Have a Dream" speech using R's tm and SnowballC packages with stemming, tokenization, and frequency analysis, visualizing key themes like "freedom" and "dream" through ggplot2 and Plotly, gaining 83% noise reduction and improved sentiment analysis accuracy [**GitHub**]

- **ML-Based Sports Analytics:** Predicted NFL play types using statistics and feature engineering by constructing machine learning models and using score differential analysis to optimize decision-making in gameplay, achieving 76.18% accuracy with 67.8% precision in pass play prediction through Python, Scikit-learn and data preprocessing [**GitHub**]