Samyuktha Kapoor

Boston, MA | +1 (857) 867-1929 | rajeshkapoor.s@northeastern.edu | [LinkedIn](#) | [Github](#) | [Portfolio](#)

SUMMARY

Data analyst with expertise in SQL-driven ETL processes and BI tools, experienced in standardizing disparate data for actionable insights. Proven track record in automating data workflows and optimizing reporting tools to enhance scalability and decision-making. Adept at adapting implementations post go-live and collaborating with cross-functional teams to ensure seamless data delivery.

EDUCATION

Northeastern University   **Sep 2023 - Dec 2025**
Master of Science, Data Analytics Engineering (GPA: 3.8)   *Boston, MA*
- **Achievements:** Teaching Assistant, IE 7275 - Data Mining in Engineering
- **Coursework:** Data Management, Data Warehousing & Integration , MLOps, Project Management, Computation & Visualization

Anna University   **Aug 2019 - Apr 2023**
Bachelor of Engineering, Electronics & Communication (GPA: 3.7)   *Chennai, India*
- **Coursework:** Data Structures, Machine Learning, Microprocesors, Digital Signal Systems, Probability & Statistics, Mathematics

EXPERIENCE

Northeastern Univeristy | *Data Science Research Assistant*   **Aug 2025 - Present**
- Developed automated healthcare data analysis tools using MNE-Python and Claude Code to streamline Visual Evoked Potential (VEP) signal analysis, reducing neurophysiological diagnostic processing time by 60% and accelerating clinical decision-making
- Engineered Python-based signal processing pipelines with m-sequence deconvolution algorithms to improve P100 wave detection accuracy, enhancing early-stage diagnostic precision and showcasing skills in data integration and advanced analytics

Massachusetts Clean Energy Centre | *Data & Business Intelligence Analyst*   **Jan 2025 - Sep 2025**
- Engineered interactive BI dashboards using Power BI and Excel for organizational metrics and grantee tracking, reducing reporting time by 40% and increasing data accessibility for stakeholders, emphasizing effective data visualization and analysis
- Developed comprehensive Power BI solutions that resolved performance bottlenecks in Power Query and optimized data models with DAX calculations, enabling real-time decision-making across 15+ cross-functional teams.
- Presented Organizational and Climatetech dashboards to board members, translating complex investment KPIs into strategic recommendations that influenced $5M+ funding allocations, highlighting strong business analysis skills
- Implemented SQL-based ETL integration processes to eliminate data silos between CRM and analytics platforms, and delivered technical training for 12 data liaisons, thereby improving data accuracy by 30% and ensuring compliance with grant reporting standards
- Collaborated on geospatial investment tracking by developing an ArcGIS dashboard with MassMappers, mapping $3.8M across 22 clean energy project sites to aid policymakers in optimizing future funding strategies

Rudra Enterprises | *Business Data Analyst*   **Aug 2022 - Aug 2023**
- Designed sales analytics dashboard integrated with ERP systems using Excel and time series forecasting to resolve inventory challenges for 50+ retail clients, reducing stockouts by 38% and optimizing procurement processes, demonstrating process reengineering
- Executed SQL queries to analyze supply chain data and developed customer segmentation models that enhanced customer retention by 24% and increased average order value by 15%, emphasizing strong expertise in SQL programming and data integration

Indian Institute of Technology | *Data Research Analyst*   **Apr 2021 - Jul 2021**
- Standardized and organized unstructured breast cancer diagnostic data using Pandas and Excel, achieving 97% data validation accuracy and facilitating reliable image classification for clinical applications, underscoring skills in ETL processes and data standardization.
- Conducted deep learning research on diagnostic accuracy for early breast cancer detection and published findings in IRO Journal, providing internal stakeholders with advanced insights and reinforcing an analytical approach to healthcare data.

ACADEMIC PROJECTS

AudioSEEK ML Audio Pipeline | [Link](#)
- Built an ML pipeline using Apache Airflow to enable timestamp-grounded question answering for audiobooks, implementing transcription, semantic chunking, and RAG-based retrieval that returns clickable timestamps, reducing audio navigation time by 75%

Payroll Management System | [Link](#)
- Developed end-to-end data warehouse by integrating Talend ETL pipelines into PostgreSQL with Tableau dashboards analyzing compensation trends, attrition rates, and overtime patterns, enabling HR leadership to identify $50K in payroll optimization opportunities

YouTube Performance Analysis Pipeline | [Link](#)
- Engineered scalable ETL pipeline on AWS (S3, Lambda, Glue, Athena) with automated data transformations and Tableau visualizations to analyze engagement metrics across 500+ videos, identifying content strategies that increased average view duration by 18%

Boston Airbnb Data Exploration Dashboard | [Link](#)
- Developed Tableau dashboard with advanced filters analyzing 3,000+ listings to uncover pricing trends and neighborhood dynamics, revealing seasonal patterns and optimal pricing strategies that informed host recommendations for 15% revenue increase potential

SKILLS

- **Languages & Frameworks:** Python (Pandas, NumPy, Scikit-Learn), R, SQL, HTML, CSS, PowerShell
- **Data & Cloud Platforms:** MS SQL, PostgreSQL, MongoDB, Talend, AWS (S3, Lambda, Glue, Athena), GCP, Salesforce
- **MLOps & DevOps:** Apache Airflow, Docker, Kubernetes, Git, Talend, ETL Processes
- **Analytics & Visualization:** Tableau, Power BI, Looker, Excel, ArcGIS, Jupyter Notebook, VS Code