

Sameer Enjapuri

Tampa, FL-33613 | (813)389-8792 | enjapurisameer@gmail.com | [LinkedIn](#) | [Tableau](#)

SUMMARY

Detail-oriented and results-driven professional with 5 years of expertise in Data Engineering and Analytics. Proficient in designing, building, and optimizing data pipelines for efficient extraction, transformation, and loading (ETL) processes. Excels in interpersonal skills and comfort working with stakeholders across disciplines, propelled by an unwavering dedication to seamless data integration.

EDUCATION

University of South Florida

Master of Science in Business Analytics and Information Systems

Expected May 2025 | **CGPA – 3.87/4**

Shah & Anchor Kutchhi Engineering College

Bachelor of Engineering, Electronics Engineering

Jul 2014 - Jun 2018 | **CGPA – 8.37/10**

TECHNICAL SKILLS

Languages: Python (Pandas, NumPy, Scikit Learn etc.), R, C, SQL, HTML, CSS

Databases: Oracle, MySQL, PostgreSQL, MongoDB, MS SQL

Tools: GitHub, Jenkins (CI/CD), Jupyter Notebook, RStudio, Pivot tables, Apache Airflow, Docker, Kubernetes, AWS (Amazon S3, AWS Glue, AWS Lambda, AWS Athena)

Data Warehousing: ETL (Extract, Transform, Load), MS SQL server, Apache Kafka, Confluent Kafka, Informatica PowerCenter

Data Visualization: Tableau, MS Power BI, MS Excel, Qlik Sense, Looker Studio, Amazon QuickSight

Data Analytics: Exploratory Data Analysis, Data cleaning, Data Visualization, Data Modeling, Statistical Analysis

Project Management: Agile, Power Point, Scrum, JIRA, Kanban, Confluence, SDLC

WORK EXPERIENCE

Graduate Research Assistant (University of South Florida)

Jan 2024 - Present

- Spearheaded the development of comprehensive reports using Tableau to **analyze Tampa Bay's economic performance** in areas such as economic outcomes, affordability, and talent pipeline relative to 19 Metropolitan Statistical Areas (MSAs).
- Collaborated with faculty and peers demonstrating strong communication skills to **gather, analyze, and visualize data**, providing insights that informed policy recommendations for enhancing the region's economic standing.
- Demonstrated proficiency in Tableau for **data presentation**, creating visually compelling dashboards and reports, effectively conveying complex economic data.

Capgemini (Client – Länsförsäkringar Bank)

Feb 2022 - Jul 2023

Consultant – Data Engineer

- Analyzed and processed data streams using Python-based Producer/Consumer APIs and **Confluent Kafka** connectors, resulting in a **40% reduction** in data processing time. This enhanced efficiency in data acquisition and preparation for downstream analysis.
- Engineered and implemented a **real-time data integration** solution using Confluent Kafka, resulting in a **25% increase** in data streaming efficiency and enabling continuous data streaming and analytics for a banking client. This Change Data Capture (CDC) system facilitated timely insights into market trends and customer behaviors.
- Utilized **Tableau** to design and deploy **interactive dashboards and reports**, integrating live data streams from Kafka. These visualizations provided stakeholders with actionable insights and predictive analytics, resulting in a **30% improvement** in decision-making processes.
- Maintained and optimized dashboards and reports to monitor **key performance indicators (KPIs)** and support data-driven decision-making across the organization.
- Applied **agile methodologies** to data analysis projects, emphasizing iterative development and continuous improvement. This approach resulted in a **30% reduction** in project delivery time, ensuring timely delivery of analytical solutions to meet business objectives.

Tata Consultancy Services (Client – Avis Budget Group)

Sep 2018 - Jan 2022

System Engineer – Data Engineer

- Orchestrated the entire Software Development Life Cycle (SDLC) for **ETL data pipelines** in Informatica PowerCenter, from requirements gathering to production deployment. Managed a cross-functional team of four members, achieving a **95% success rate** in **transitioning ETL workflows to Kafka**, ensuring seamless data integration and processing.
- Utilized advanced data modeling techniques and SQL Trace to **optimize database performance by 15%**, enhancing the efficiency of data retrieval and analysis processes. Applied **business intelligence** principles to ensure effective data modeling and mapping, facilitating accurate and insightful reporting.
- Analyzed rental car market trends using statistical methods and data visualization tools, resulting in strategic recommendations that drove a **15% increase** in revenue. This involved extracting, transforming, and **analyzing large datasets** to identify key trends and opportunities for business growth.
- Engineered and fine-tuned ETL processes and Kafka topics to ingest real-time data from IoT sensors on rental vehicles, **reducing data inaccuracies by 30%** and improving data quality for subsequent analytics and reporting tasks. This initiative enhanced the reliability and relevance of **analytical insights** derived from IoT-generated data streams.
- Cultivated and nurtured client relationships, leveraging strong networking skills to facilitate project success and satisfaction. Engaged with **stakeholders** to understand their data needs and requirements, ensuring alignment with project objectives, and delivering actionable insights that drive business value.

ACADEMIC PROJECTS

Data Engineering YouTube Analysis (AWS, ETL)

- Developed scalable data pipeline on AWS integrating services like S3, Glue, Lambda, and Athena for ingestion, ETL, and querying. Established a centralized data lake for storage and analysis of YouTube video data. Created a dashboard using Amazon QuickSight for reporting and extracting insights from video metrics.

Analyzing Spotify Songs Dataset (Databricks)

- Performed analysis, visualization and implemented ML models on Spotify dataset stored in a distributed file system. Employed PySpark for data processing, querying, and building ML pipeline on Databricks platform. Used K-Means clustering to cluster songs into similar playlists. To classify songs into genres, used a Decision Tree Classifier.

Hospital Management Advanced Database Management Project (SQL)

- Designed ERD, EERD, RS and created hospital management database with 14 tables in SQL SSMS. Created Stored procedures, Views, Triggers, UDFs to implement business logic. Deployed the database to Azure remote server.