

ANDE SRIKANTH

Telangana, India

☎ +91-7660803965

✉ srikanthande76@gmail@gmail.com

[Linkedin](#)

🐙 [GitHub](#)

OBJECTIVE

Aspiring Data Engineer with hands-on training in Python, SQL, Hadoop, Spark, and Google Cloud Platform. Passionate about designing scalable data pipelines and transforming raw data into actionable insights to drive data-informed decisions in real-world applications.

COURSEWORK / TECHNICAL SKILLS

- **Languages:** Python, Java, C, JavaScript, SQL
- **Big Data Tools:** Hadoop, Hive, PySpark
- **Cloud:** Google Cloud Platform (BigQuery, Dataproc, Dataflow)
- **Web Tech:** HTML5, CSS3, React.js, Bootstrap
- **Tools:** Git, VS Code, Jupyter
- **Concepts:** OOP, DSA, ETL, Data Warehousing, Agile, SDLC

EXPERIENCE

Data Engineering Training Program — Revature

April 2025 – Present

- Completed intensive training in core data engineering technologies and tools.
- Gained hands-on experience in Python, SQL, Hadoop, Hive, and PySpark.
- Built end-to-end ETL pipelines and real-time batch workflows using Google Cloud Platform (BigQuery, Dataflow, Dataproc).
- Practiced Agile methodology with peer code reviews and sprint-based development.
- Applied data cleaning, transformation, and warehousing techniques in real-world scenarios.
- Developed projects focused on big data analytics and scalable cloud-native solutions.

PROJECTS

- **New Zealand Traffic Crash Analysis** [\[GitHub\]](#) *Data Analyst*
 - Performed end-to-end data ingestion, processing, and analysis on New Zealand traffic crash dataset using Python, SQL, Hadoop, Hive, PySpark, and Google BigQuery.
 - Developed ETL pipelines to clean, transform, and load large datasets efficiently into BigQuery for real-time analytics.
 - Utilized PySpark to perform distributed data processing and aggregation on crash-related metrics like fatality counts, injury counts, and road conditions.
 - Created data visualizations and reports to identify critical accident trends, contributing to data-driven decision making.
 - Gained hands-on experience with cloud data warehousing, big data processing frameworks, and scalable data engineering best practices.
- **ETL - Advanced Pipeline Development** [\[GitHub\]](#) *Data engineer*
 - Designed and implemented a modern serverless ETL data pipeline on Google Cloud Platform (GCP) leveraging Cloud Storage, Cloud Composer, Pub/Sub, Dataflow (Apache Beam), and BigQuery.
 - Developed an Apache Airflow DAG in Cloud Composer to orchestrate workflow; scheduled periodic scans of Cloud Storage buckets for new CSV files and published file metadata to Pub/Sub topics.
 - Configured Pub/Sub topics and subscriptions to trigger Dataflow pipelines for real-time ingestion, validation, transformation, and loading of CSV data into BigQuery.
 - Created scalable Dataflow pipelines in Python using Apache Beam to process CSV data: including parsing, cleaning, type casting, and formatting before appending to BigQuery tables.

- Set up partitioned BigQuery tables optimized for analytical queries and downstream reporting.
 - Automated pipeline orchestration with conditional job execution based on file detection and previous job success; implemented Pub/Sub-driven email alerts for job start and completion status.
 - Collaborated on documenting architecture, workflows, and troubleshooting procedures to facilitate team knowledge sharing.
-
- **TechM-DA: Amazon Best Seller Analysis with Python and MySQL** [\[GitHub\]](#) *Data Analyst Role*
 - Collected, cleaned, and transformed Amazon Best Seller Software dataset using Python (Pandas, NumPy) for robust exploratory data analysis (EDA).
 - Designed and implemented an optimized MySQL relational database schema to efficiently store product, category, and ratings data.
 - Performed analytical computations to identify top-rated products, pricing trends, category review distributions, and correlations between ratings and reviews.
 - Automated data cleaning workflows and handled file I/O errors for reliable and repeatable data ingestion.
 - Created visualizations using Matplotlib and Seaborn (bar charts, heatmaps, distribution plots) to present actionable insights.
 - Optimized SQL queries for fast retrieval of summary statistics and category-specific trends.
 - Documented and presented findings via Jupyter Notebook dashboard to mentors, ensuring clarity and reproducibility.

EDUCATION

Vidya Jyothi Institute of technology, Hyderabad

B.Tech - Artificial Intelligence -CGPA: 7.1

2020-2024

Hyderabad, Telangana

Certifications

- **Data Structures and Algorithms** – GeeksForGeeks
- **CLA: Programming Essentials in C** – Cisco
- **Python for Data Science** – Udemy
- **Basics of Blockchain** – Coursera

Achievements

- **Web Developer,GDSC-Vidya Jyothi Institute of Technology** – Contributed to web development projects, conducted peer sessions, and collaborated on building responsive websites using HTML, CSS, JavaScript, and React.
- **Top 10 in HackerRank Coding Contest** – Ranked among the top 10 at Anurag University (2023)
- **Shark Tank Winner – Revature Training Program (2025)** – Presented a data-driven business solution and secured top position in the internal innovation challenge