





HEMANTH KUMAR REDDY BOGATHI

 hemanthreddy64207@gmail.com |  +44 7733094263 |  United Kingdom |  [LinkedIn Profile](#)

SUMMARY

Detail-oriented Data Science graduate with hands-on experience in data engineering, ETL/ELT pipelines, and cloud-based solutions (Azure). Skilled in SQL, Python, Power BI, and APIs, with a proven ability to transform raw data into actionable insights.

Education

Teesside University, UK Sep 2023 – May 2025
MSc Data Science With Advanced Practice

Relevant Coursework

Data Structures	Algorithms Analysis	Artificial intelligence	Machine Learning	Database Manage ment
Software Methodology	Project Management	Computer Architecture	Data Analytics	Data Visualization

Experience

Junior Software Intern, People Tech Group Aug 2022 – Aug 2023
Hyderabad, INDIA

- Specialized in designing and implementing gauges using C++, QT/QML, XML, and SVG rendering technologyI specialize in developing animations for instrument panel cluster Displays, Heads-Up Displays, and Climate screens for General Motors using the QT creator with C++, also create simulations for testing and conduct validation on Windows and hardware platforms like QNX.
- Enhanced UI functionality in automotive systems.
- Analyzed development for critical issues using debugging for Windows and QNX Momentics environments to inform strategic decision-making and drive business growth.
- Collaborating closely with functional teams like Design and Static Code analysis to ensure seamless integration.

Data Analyst Intern And Visualization Developer, Think Pacific (Remote) Sep 2024 – Jan 2024

- Designed and developed interactive data visualizations to represent a 10-year impact assessment for Think Pacific’s community-based projects using Power BI and Tableau.
- Analyzed and processed raw data from multiple project reports to create insightful charts, graphs, and dashboards highlighting key metrics like community engagement, volunteer impact, and project outcomes.
- Enhanced data presentation efficiency by transforming static reports into dynamic dashboards, enabling real-time exploration of project data.
- Collaborated with the Program Manager to translate complex data into visual storytelling, improving stakeholder understanding and decision-making processes.

Portfolio Projects: GitHub Profile

Supermarket Sales Analysis (SQL) | Data Analyst Project

- Performed end-to-end data cleaning and preprocessing in MySQL by addressing missing values, eliminating duplicates, and enforcing consistent naming conventions for improved data quality and reliability.
- Created complex SQL queries to analyze total sales, revenue fluctuations, and identify high-performing products and store branches over time.
- Applied customer segmentation strategies and performed statistical analysis to gain insights into customer behavior, helping shape data-driven business recommendations for marketing and sales teams.

GCP Real-Time E-Commerce Data Pipeline

- Real-Time Streaming Pipeline: Designed and implemented an event-driven architecture using Google Pub/Sub and Dataflow (Apache Beam) to ingest, process, and store e-commerce transactions instantly.
- Synthetic Event Simulation: Developed a Python-based framework to generate and simulate customer interactions, purchase events, and abandoned carts for comprehensive testing.
- Data Processing and Enrichment: Cleaned and transformed streaming data before loading into BigQuery, applying business logic enhancements for real-time insights.
- Dashboard Creation: Built interactive visualizations in Looker Studio to monitor user behavior, purchase patterns, and revenue trends dynamically.
- Cloud Computing Expertise: Gained hands-on experience with serverless GCP services, optimizing pipeline stability and ensuring scalability with auto-scaling architecture under high data loads.

Data Warehouse & ETL Pipeline Project | PostgreSQL, Python, Airflow

- Designed a scalable star schema data warehouse architecture in PostgreSQL to support analytical querying across different data domains.
- Developed automated ETL pipelines in Python to extract raw data, perform transformations, and load it efficiently into the warehouse.
- Leveraged Apache Airflow to orchestrate and schedule ETL jobs, with monitoring and logging features for workflow reliability and error handling.

Excel Dashboards for Business Insights

- Aggregated diverse data sources including orders, financials, customer addresses, and satisfaction surveys into a single, structured Excel dashboard for high-level analysis.
- Designed interactive dashboards using pivot tables, dynamic slicers, and advanced charts to create a seamless user experience and enable intuitive filtering of insights.
- Presented real-time performance metrics like revenue growth, customer trends, and order frequency, allowing management to make quicker, data-supported decisions.

Technical Skills

Data Analysis & Visualization: Python (Pandas, NumPy, Matplotlib), SQL, Power BI, Tableau , NLP
Database: MySQL, PostgreSQL, ETL
Cloud Platforms: GCP, AWS
Programming Languages: Python, C++, Java
Tools: Excel, Jupyter, GitHub, Kubernetes, QT/QML, XML, Apache