

AKASH GANGADHARAN

US • akash.gangadharan94@gmail.com • (812) 778-4269 • [Linkedin](#) • [Portfolio](#)

PROFESSIONAL EXPERIENCE

Data Research Assistant | Indiana University School of Medicine | US

January 2024 - May 2024

Skills: SQL, Python, Power BI, Data Extraction, STATA, ETL.

- Integrated over 1 million records of CDC/govt health data from disparate sources into a unified dataset using scraping and spatial techniques to enable advanced analytics and visualization.
- Performed data analysis with Python, SQL, STATA, and Power BI to create visualizations and dashboards, increasing people participation by 20% for coalition groups.
- Built dashboards, and presentations to summarize key findings for stakeholders and leadership.

Microsoft Ambassador | University Information Technology Services | US

October 2023 - December 2023

Skills: Power BI, Microsoft forms, MSSQL Server.

- Led 50+ workshops on Microsoft Office 365 among students and promoting the adoption of Microsoft Office 365.
- Revamped workshops by utilizing Microsoft forms and Power BI to analyze event participation, driving data-informed decisions resulting in a 2X increase in student adoption of Microsoft technologies.

Data Engineer Intern | Rocket Central | US

May 2023 - August 2023

Skills: AWS Glue, PySpark, Python, CI/CD, GitHub, AWS S3, AWS Athena, ETL pipeline.

- Engineered PySpark ETL pipeline leveraging AWS to process 10 million records for Change Data Capture (CDC), reducing processing time 15% and improving reliability.
- Implemented streamlined data validation protocols that reduced data load failures by 20% in the ETL pipeline, ensuring higher data quality and improving overall operational efficiency.

Data Engineer | KPMG | IN

May 2021 - July 2022

Skills: GCP BigQuery, Apache Airflow, SQL, Python, CI/CD, GitHub, SSIS, SSMS, ETL pipeline.

- Led the migration of Oracle SQL scripts to BigQuery, orchestrating data loading and achieving project completion 15 days ahead of schedule.
- Transformed 15+ financial reports by automating end-to-end data pipelines in SSIS, cutting delivery time from 2 hours to 10 minutes, enabling Excel exports, and automating stakeholder notifications.
- Developed Python script for GCP data validation enabling row count analysis, cutting processing from 1 hour to 5 minutes and improving efficiency by 70%.

Data Engineer | Vodafone | IN

August 2018 - May 2021

Skills: SQL, Teradata, Python, Excel, VBA, Qlik sense, Data extraction, ETL pipeline.

- Generated revenue of more than 1M £ by delivering Cross-sell, Up-sell, & Retention campaigns with Marketing team.
- Automated weekly SQL updates with Python scripts to optimize database performance, improving efficiency by 50%.
- Streamlined data processing by automating SQL and Salesforce integration, cutting task time from 5 days to 30 minutes and boosting campaign efficiency.

SKILLS

Business Intelligence and Visualization Tools: SQL, MS Excel, Tableau, Power BI, Data Modeling, Teradata.

Programming & ETL Tools: Python, Spark, Hive, GitHub, CI/CD, Airflow, Snowflake, Hadoop, Databricks, DBT, Kafka.

Cloud Platform: AWS (S3, Athena, Glue, EMR, Redshift), GCP (BigQuery, Cloud Storage, Composer, Compute Engine).

Certifications: Google Cloud Professional Data Engineer, Google Cloud Professional Machine Learning Engineer, Microsoft Azure Data Scientist Associate, Airflow fundamentals and DAG authoring, DBT fundamentals, Databricks fundamentals.

EDUCATION

Indiana University Bloomington | US

August 2022 - May 2024

Master of Science in Data Science

PERSONAL PROJECTS

YouTube Data Engineering and Analysis using AWS | Skills: AWS Glue, S3, AWS Athena, PySpark, QuickSight.

- Architected and deployed ETL Data pipeline in AWS for YouTube Data analysis and visualization in AWS QuickSight.

Uber Data analysis pipeline using GCP | Skills: Python, Data Extraction, Data transformation, GCP, Mage AI, Looker.

- Built an end to end ETL Data pipeline using GCP and Mage AI for Uber Dataset and visualizing in Looker Studio.

Home Credit Default Risk Analysis | Skills: Python, Machine learning pipeline, ETL, Data Visualization, SVM.

- Improved the accuracy from 85% to 92% using Support Vector Machines model in an end-to-end machine learning pipeline to detect whether the customer will default on the home loan.

ACHIEVEMENTS

- Secured **Third place** in Grant Thornton Idea case competition and **awarded \$2000**.
- Secured All India Rank of **549** out of 200,000 students in MH-CET Exam (**Top 0.5%**)