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.

- Acquired and integrated sociodemographic, environmental, and health data from diverse sources using advanced techniques like web scraping and spatial data enrichment.
- Performed complex analytics including descriptive statistics, trend analysis, and data visualization utilizing Python, SQL, Excel, Power BI, and STATA.
- 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.
- Utilized 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 Companies | US **May 2023 - August 2023**

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

- Built scalable ETL pipeline for Change Data Capture (CDC) by processing 10 million user records using PySpark & AWS Glue from raw to transformed state.
- Streamlined data validation process in ETL pipeline, thereby reducing the data load failures by 40%.

Data Engineer | KPMG | IN **May 2021 - July 2022**

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

- Led the migration of Oracle SQL scripts to BigQuery, efficiently 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.
- Streamlined data validation on GCP cloud storage by developing a Python script for row count analysis, reducing processing time from 1 hour to 5 minutes thereby improving the efficiency of data validation tasks by 70%.

Data Engineer | Vodafone | IN **August 2018 - May 2021**

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

- Generated revenue of more than 1M £ by delivering Cross-sell, Up-sell, & Retention campaigns with Marketing team.
- Automated weekly SQL updates with Python after a critical database error, enhancing accuracy and 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, DBT, Databricks
Programming languages & ETL Frameworks: Python, Apache Spark, GitHub, CI/CD, Apache Airflow.
Cloud Infrastructures: 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, Apache Airflow fundamentals and DAG authoring, DBT fundamentals, Databricks

EDUCATION

Indiana University Bloomington | US **August 2022 - May 2024**

Master of Science in Data Science; GPA 3.62/4.0

University of Pune | IN **June 2013 - May 2017**

Bachelor of Engineering in Mechanical Engineering; GPA 3.3/4.0

PERSONAL PROJECTS

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

- Designed an end to end ETL Data pipeline in AWS for YouTube Data analysis and visualizing 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.***

- Improved the accuracy from 85% to 92% using ensemble methods 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.