

MANIKA AGARWAL

9057898024 , m.agarwal0499@gmail.com , [linkedin.com/in/manika-agarwal-5183231a0](https://www.linkedin.com/in/manika-agarwal-5183231a0)

Immediate Joiner

Results-driven Data Engineering Analyst with 3.8years of experience with strong foundation in data interpretation and ETL process. Proficient in leveraging tools such as Excel, SQL, and Python to extract actionable insights from complex datasets. Possess knowledge and hands-on experience in both Google Cloud Platform (GCP) and Amazon Web Services (AWS). Proven ability to translate data into strategic recommendations for business improvement. Seeking to contribute analytical expertise and drive informed decision-making within a dynamic team environment.

KEY COMPETENCIES

ETL	Cloud data fusion	Data Visualization	Apache airflow
GCP	Cloud storage	AWS	Data Modeling
Power Bi	Python	Numpy & Pandas	Data cleansing
Big query	GitHub	SQL	Excel
Data Analysis	Alerts & Monitoring	Scrum & Agile , Jira	Reporting and Dashbaords

PROFESSIONAL EXPERIENCE

Mckinsey & Company January2025-Present

- **Project Onboarding & Initial Contribution**
 - Participated in comprehensive onboarding training sessions to familiarize with company processes, tools, and technologies.
 - Engaged in project explanation sessions, gaining an understanding of project goals, timelines, and team responsibilities.
 - Contributed to the development of JavaScript code for automation tasks within ongoing projects, enhancing overall efficiency.

Group M Nexus

Data Engineer July 2024-January2025

- **Custom Apache Airflow Operators Development**
 - Designed and implemented custom Apache Airflow operators to automate and optimize data ingestion pipelines for client data feeds, ensuring seamless integration of new data sources.
- **ETL Workflow Orchestration**
 - Utilized Python and SQL to build, orchestrate, and optimize ETL workflows, ensuring efficient data transformation and loading processes. These workflows were designed to handle large volumes of data while maintaining scalability and performance.
- **Unit Testing & Workflow Validation**
 - Conducted unit testing for custom Airflow operators using mocking techniques to ensure reliability and correctness of workflows, decoupling them from external systems.
- **Data Ingestion Application Development**
 - Developed a data ingestion application that accepts a JSON file as a template. The application dynamically creates a DAG file in Python format based on the provided template.
 - Once the DAG file is pushed to the main branch, it triggers a Cloud Function, which automatically places the DAG in the GCS (Google Cloud Storage) directory. The DAG is then visible in the Airflow UI for execution, streamlining the process of managing and deploying new workflows.

Accenture

Data Engineering Analyst September 2022-July 2024

- **Namespace Comparison DAG Development:**
 - Developed a Namespace Comparison DAG in Apache Airflow, utilizing Cloud Data Fusion API.
 - Enhanced the DAG to include an Excel sheet displaying plugin frequency, alongside extracting and displaying the number of pipelines under specific namespaces.

- ## Accenture

July 2021-August2022

- Creation of Insert statements for BigQuery through the use of NumPy and Pandas. Proficient in automating Excel-to-View processes, optimizing ongoing operations for BigQuery efficiency. Extensive hands-on experience with Datafusion, showcasing the ability to create pipelines that resolve complex, multilevel nested JSON files and efficiently transfer data to BigQuery tables.
- In the realm of workflow optimization, successfully implemented a housekeeping DAG for BigQuery tables using Apache Airflow. Leveraged BigQuery UI to create stored procedures for extracting essential data.
- Facilitating seamless Dashboard creation using Data Studio.
- Demonstrated proficiency in orchestrating DAGs to transmit vital data from BigQuery to Cloud Monitoring, culminating in the creation of monitoring dashboards for comprehensive performance tracking.
- Conducted a Proof of Concept (POC) on Cloud Data Fusion labels. Established system profiles and labels, tracking their impact in a Dataproc job. Investigated the effects of labels on billing.

STRENGTH

Data Analysis

Analyzed and interpreted data to drive business decisions resulting in increase in revenue.

Problem Solving

Developed solutions to complex problems in a team setting resulting in increased efficiency.

Ensured accuracy in data collection and analysis resulting in error reduction.