

Yokesh S

yokesh.srinivasan04@gmail.com

Mob: +91-9176277675

Objectives

To work in a globally competitive environment on challenging assignments that shall yield the twin benefits of the job satisfaction and a steady-paced professional growth.

Professional Summary

- Around 5.5 Years of IT experience in Analysis, Design, Development, Implementation, Testing and Support of Data Warehousing and Data Integration Solutions using **GCP, PySpark and Microsoft Azure**.
- **Having good work experience in Cloud Platforms such as GCP & Microsoft Azure**. Also, Expertise in managing the Cloud Services (IaaS, PaaS and SaaS).
- Have extensively worked developing ETL Business Transformation Logic in **ADF and ADB**. Expertise with database as follows: **Azure DW, SQL SERVER, BigQuery, Teradata and Oracle**.
- Also worked with Business Users for gathering the ETL & Reporting requirements from their end for Creating ETL Pipelines & Reporting Summary Datasets for **PowerBI** Reports.
- Have extensively worked on **Python, Unix, Windows Batch Scripts, Powershell script** tools for scheduling the ETL Jobs & various Automation Process.
- Ability to meet deadlines and handle multiple tasks, decisive with strong leadership qualities, flexible in work schedules and possess good communication skills.
- Good knowledge on PySpark, Hadoop Platform, Databricks and DeltaLake.
- Team player, Motivated, able to grasp things quickly with analytical and problem-solving skills.

Experience Summary

- Currently working in EY India, Chennai as Consultant from May 2021 to till date.
- Previously worked in Deloitte, Chennai as Consultant from May 2020 to May 2021
- Previously worked in Capgemini Technology Services, Chennai as an Associate Consultant from August 2017 to May 2020.

Educational Profile

- Graduation: BE in Computer Science and Engineering from Veltech Multitech Dr RR Dr SR Engineering College, Chennai (2013-2017) with 72%.
- HSC: HSC in S R D K Vivekananda Vidyalaya – CBSE (2011-2012) with 83%.

ASSIGNMENTS

Client Name	Levi Strauss & Co.
Company	EY India
Project Name	GDO Solar – Migration and Remediation
Project Duration	June 2022 – Till Now
Role/Title	Lead Data Engineer
Work Location	Chennai
Project Description	Migration of Current AWS Data to GCP and create GCP Data Lake as the driver data for Data Science models. Remediation of existing POS sales system.
Responsibilities	<ul style="list-style-type: none">➤ Understanding the existing logics and pipelines available in AWS and create reusable framework in GCP Environment.➤ Remediation of Existing POS sales system with the New one, so understanding the logics and creating alternate pipelines for Data Science track.➤ Understanding the business requirements and creating end to end pipelines for the new tables.➤ Working on the Databricks Notebooks and creating packages for the same.➤ Working on DAGs in Airflow for scheduling.
Software / Special Tools	GCP Big query, Databricks, Airflow, GCS, Postgres
Scripting and Databases	PySpark, SQL, Python, Shell.

ASSIGNMENTS:

Client Name	AB-InBev
Company	EY India
Project Name	Procurement Analytics - Payments.
Project Duration	May 2021 – June 2022
Role/Title	Lead Data Engineer
Work Location	Chennai
Project Description	Procurement Analytics – Payments where vendor risk scores are calculated based on ML models and Salesforce Dashboard is loaded.
Responsibilities	<ul style="list-style-type: none">➤ Integration of sources from Syspro and flowcentric to existing pipelines in Azure.➤ Constantly improving performance of pipelines.➤ Data cleansing and data availability for ML and closely working with Data scientists to improve the data quality.➤ Design and Implementation of new approaches like Alert Mechanism, Pyspark implementation and global parametrization.➤ Reduced the run time of the pipeline from 100 hours to 50 hours ~50% improvements.
Software / Special Tools	Azure ADF, Databricks, Azure Synapse, Azure blob storage, Python
Scripting and Databases	PySpark, SQL, Python, Azure DWH.

Client Name	Loblaws Limited
Company	Deloitte India
Project Name	Data Modernization.
Project Duration	June 2020 – May-2021
Role/Title	Consultant
Work Location	Chennai
Project Description	Data Modernization is an initiative from Loblaw's to Migrate their EDW from Teradata to GCP Bigquery. It also includes the development of integrations in GCP Cloud native environment.
Responsibilities	<ul style="list-style-type: none">➤ Creating scripts like Bteqs, TPTs and Fast Exports in Existing Teradata System in order to decommission the existing Integration tool Informatica.➤ Requirement gathering from Customers to set up the existing integrations to help in migration.➤ Analysis and Research on Different Cloud possible solutions in GCP to get the best fit for the requirement.➤ Working on POC Projects to fix the architecture for the original development.➤ Working on enhancements in current Teradata System to remove the dependency of Integration tool like Informatica.
Software / Special Tools	Airflow, Gcloud shell, Dataflow, Python.
Database	Google Bigquery, Teradata.

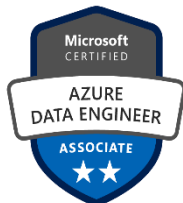
Client Name	Univision Media
Company	Capgemini
Project Name	Local Media Consumer Hub Application Development and Support.
Project Duration	June 2018 – May 2020
Role/Title	Lead Data Engineer
Work Location	Chennai
Project Description	Consumer Hub Application is an Univision Initiative that is used to gather the PII information and the personal interests of consumers. The Collected Data Insights can be used by various Business users to advertise their products to consumers based on consumers Subscription.

Responsibilities	<ul style="list-style-type: none"> ➤ Creating Event Hubs and Stream Analytic jobs in Microsoft Azure to pull the streaming data into Data Warehouse. ➤ Contacting Customer Directly to get the Business requirements and provide the best fit solution for the requirement and manage the team to provide the solution in promised timeline. ➤ Creation and Implementation of Various resources like Azure DB, Azure Synapse Analytics, Data Factory, Azure VM from Azure Cloud Environment required for Project. ➤ Developing complex ETL Business Transformation Logic in ETL Pipelines & creating Audit log capture logic for the business. ➤ Research on various Types of Databases to find the best fist for implementation. ➤ Creating Python Scripts to take care of CCPA from archive standpoint. ➤ Automation implementation to trigger the pipelines as per the flow and to acknowledge the process via Email. ➤ Audit logic and capture implementation for Request to Know, Delete and Optout.
Software / Special Tools	Azure Data Factory, Automation, Event Hubs, Stream Analytics, Python, SSIS, PowerShell.
Database	SQL Server, SFTP.

Client Name	CBS Media (Columbia Broadcasting System)
Company	Capgemini
Project Name	GCP Migration & Event Management Application
Project Duration	Nov 2017 - May 2018
Role/Title	GCP and ETL Developer
Work Location	Chennai
Project Description	Migration of On-Premise Enterprise data warehouse & BI Reports to GCP Cloud Platform. It's about migration of Datawarehouse from Teradata to Google Bigquery using Matillion & Apache Airflow. So, we have replaced the Informatica ETL, Teradata, Autosys, Cognos & On- premise Platform to Google Cloud platform. Web Application & Dashboards are built in GCP Platform using ASP.NET on top of the Bigquery Datawarehouse.
Responsibilities	<ul style="list-style-type: none"> ➤ Designing complex ELT Pipelines on Matillion for Data Ingestion & Data Validation. Conducting design and code walkthrough with the team. ➤ Preparing technical documentation such as design approach documents, Low Level Design documents, Unit Test Preparation, Integration Testing with Code Review Checklist. ➤ Developing ETL Business Transformation Logic in GCP Bigquery and also Developed Semantic Business View Queries on Bigquery. ➤ Automated the Balance Audit Process for every Data Load. Automated the batch control process and created DAG Scripts for scheduling the Data Load in Apache Airflow. ➤ Deployment of various changes made in each sprint. Review Tracker was maintained from end to end.
Software / Special Tools	GCP, Informatica, Matillion, Apache Airflow, Python & Shell Scripting
Database	Teradata, Google Bigquery.

Certifications:

- Exam DP-203: Azure Data Engineer Associate



- Oracle Certified Diploma in Software Engineering (Java Specialization).
- Diploma in Software Engineering (Java Specialization) from NIIT.

Professional Trainings & Achievements:

- Got 'Extra Mile', 'Kudos' Recognition Award for the excellence proved in the account level.
- Got 'Rising Star' Recognition Award for the Innovations & Process Improvements in Projects.
- Loblaw 'Wall of Fame' for handling critical timeline.