



## AKASH SINGH

Course : Msc Chem. (Hons.)      Graduation year : 2018  
Institute: BITS PILANI      Email : f2014548p@alumni.bits-pilani.ac.in      Mobile : 9398681252



### ACADEMIC DETAILS

COURSE	SPECIALIZATION	INSTITUTE/COLLEGE	BOARD/UNIVERSITY	% CGPA	YEAR
XII	Science	Sri Chaitanya junior college	APBSE	92.9	2013
X	General	NAVY CHILDREN SCHOOL	CBSE	100	2011

### ELECTIVES/TECHNICAL PROFICIENCY

**Electives:**    Graphs And Networks, Discrete Structures for Computer Science, Copywriting, Computer Organ And Arch

**Technical Proficiency:** • **Certification** in PYTHON programming for data science    • **Certified** in JAVA web development by HSBC training program  
• **Certification** in computer organisation and architecture    • Google cloud platform (Associate Google Cloud Engineer and Professional google cloud data engineer **certification**)  
• PYTHON/JAVA / SQL /SHELL SCRIPT/C++    • Yellow belt certification in secure code warrior coding  
• Python full stack developer (pursuing **certification**)    • BigQuery    • DataProc    • Pyspark    • Dataflow    • Cloud Composer/Airflow    • Cloud Functions    • DataStage    • Python data analysis and visualization    • Google compute engine    • Google Cloud Storage    • Google Kubernetes    • Data Studio    • Pub/Sub    • Hive

### SUMMER INTERNSHIP/WORK EXPERIENCE

**HONGKONG AND SHANGHAI BANKING CORPORATION – SOFTWARE DEVELOPER (3 years 7 months )**

**4 July, 2018 – 11 Jan, 2022**

- Worked as a software engineer for HSBC Technology India (HTI Hyderabad) for over 3 years in the data science department.

**ACCENTURE – Application Development Senior Analyst (2 years)**

**12 Jan, 2022 – 7 May, 2024**

- Working as a Specialist GCP developer at Accenture (Hyderabad).

**TREDENCE – Associate Manager**

**8 May, 2024 – till date**

- Working as an Associate Manager in Data Engineering Department

**VIZAG STEEL PLANT (summer internship – 2 months)**

**May 2016 - Jul 2017**

- VIZAG STEEL PLANT – manufacturing steel products

•

I aided in the data collection for the analysis of cooling system at MBC plant (mechanical, biological and chemical treatment of water) It helped in the identification of the deficiency in the flash unit and the double pipe heat exchangers of the cooling system

•

I also did the calculation and ultimately made recommendations that will bring in more efficiency of heat transfer in condensers for effective working of flash unit.

**PROJECTS**

**Migration from Teradata to Google Cloud (HSBC)** : Worked on the migration of HSBC data from Teradata to Google Cloud and in the process got to work on DataStage the ETL tool, GCP (BigQuery, Compute Engine etc.), SQL conversion, Data Studio, Jenkins pipeline etc. Also worked on writing parsers in python such as the SQL parser for the automated conversion of the SQL scripts from TD to BQ compatible, DDL conversion Parser, BigQuery view creation python script etc.

Built a pipeline for copying BigQuery tables and Datasets by making a datastage job and a python script.

Ingestion of data to BigQuery using Juniper's TD to BigQuery ingestion pattern that was built by the team.

**DATA ON CLOUD (HSBC)**

Worked on BigQuery cost optimization.

Worked on Google Cloud Storage cost optimization. Configured DLP API for the identification of PII fields in BigQuery and publishing the result to GCP's native Data Catalogue. Moving data across GCP project in BigQuery by developing SQL scripts for data transfer. **Coordinated Data**

**Sourcing (HSBC)** Worked on Collibra.

**Migration from Hadoop and spark to GCP  
(Accenture – CVS HealthCare Client)**

Working on the Application and code migration from on-prem (HDFS) to GCP(BigQuery)

Building end to end data pipelines using dataproc/dataflow (ETL) and Cloud Composer/Airflow (Workflow Orchestration Service)

Automation of hql (Hive sql) scripts to BigQuery compatible sql scripts.

Using google cloud pub/sub to distribute change events from BigQuery.

### **Migrating (from Oracle) and building new GCP pipelines using GCP native services and BigQuery as database (Accenture – Verizon Communications Client)**

Designed and built several data processing/ETL pipelines in Cloud Composer (Airflow) using BigQuery and GCS as primary storage services.

Created several python and BigQuery SQL jobs and orchestrated them in cloud composer by writing DAGs.

Created cloud functions for gcs event-based file ingestions that would trigger the appropriate dag for the ingestion file processing as soon as it lands in the staging GCS bucket.

**Migrating from 1010 to GCP and building new GCP pipelines using Cloud Composer(Airflow) and BigQuery and Cloud SQL(postgreSQL) as database (Tredence – Dollar General Client)**

Designed, Developed and implemented the Audit, Alert& Notifications framework along with Data Quality and Data Validation framework

**Developing ETL Pipelines in pyspark on DataProc and orchestrating the pipeline via Airflow on Astronomer using Hive, BigQuery and GCS as databases(Tredence – Walmart Client)**

Optimized the Pyspark pipelines for 2 major applications by restructuring the schema, reducing the number of tables and reducing the number of Pyspark scripts.

Altered the spark config for the Pyspark jobs to make it run faster. As a result of the above changes the runtime of the app was brought down by 6.5 hours.

I also trained 2 teammates on Airflow DAGs and Pyspark on DataProc.

**Migration from Azure to GCP Databricks  
(Tredence – SAM’s Client)**

Migrating the Databricks pipelines (Workflows, Notebooks, tables/views) on Azure to Databricks on GCP.

Managing a team of 4 and resolving blockers.

Created the pyspark script for data validation post migration and data load.

Optimized the spark config for the pyspark workflows to handle shuffle heavy jobs on smaller clusters post migration.

**ACHIEVEMENTS AND AWARDS**

Certification in PYTHON PROGRAMMING FOR DATA SCIENCE	<b>2020</b>
Certified in JAVA WEB DEVELOPMENT by HSBC training program.	<b>2018</b>
Got best team award (ETL team which works on datastage.) for GCP migration project at the HSBC annual festival.	<b>2019</b>
Participated in and won many public speaking competitions (Public debate, Elocution etc.) in my school.	<b>2009</b>
Successfully completed NTPEL online course of computer organisation and architecture with an elite score of 84%.	<b>2017</b>
Professional google cloud data engineer	<b>2020</b>

#### EXTRA CURRICULAR ACTIVITIES

Teach math on weekends, Singing and playing guitar, played national level table tennis

INTERNAL