

K. HAREESH KUMAR

DATA ENGINEER

CAREER OBJECTIVE

Looking for employment as a Data Engineer to work for a great company in order to act on my skills and past experience.

CONTACTS

📍 MoghalRajPuram,
Vijayawada, Andhra
Pradesh.

📞 8499909711

✉ harish.cde1@gmail.com

EDUCATIONAL BACKGROUND

Master's Degree in Computer
Science and Engineering

JNTU Kakinada, Andhra
Pradesh, India

TECHNICAL SKILLS

- **BIGDATA Technologies:**
 - Hadoop (HDFS, Hive), Kafka
- **Cloud Platforms:**
 - GCS, Big Query, Dataflow, Dataproc, Pub/Sub, Airflow
- **Build Tools:**
 - Jupyter Notebook
- **Languages:**
 - Java, Python, SQL, Spark
- **Data Formats:**
 - CSV, JSON, XML, ORC, Avro, Parquet
- **Project Execution Tools:**
 - **Agile** – Sprint, Scrum, JIRA, Confluence, GitHub

PROFESSIONAL SUMMERY

- I have 4+ years of experience in IT industry as a Data Engineer on Google Cloud components and Test Engineer.
- A qualified master of technology with good amount of exposure in cloud data engineer utilizing technologies like GCP, Big Query, SQL and Python (Scripting).
- Great analytical skills and good command over GCP components like Big Query, Pub/Sub, Kafka, GCS and Cloud Composer/Airflow.
- I have working experience with Big Query to load data to Stage, History and Audit tables.
- In addition to that, I can optimize the performance of BigQuery using Partitioning and Clustering techniques.
- Established Authorized views from various tables.
- Driven on creating DAGs and tasks to deploy and run PySpark jobs in Airflow scheduling tool and automated using python scripts.
- Awareness in performance tuning, deployment procedures, troubleshooting and solving database issues.
- I have working experience on python language.
- Used GIT to manage and deploy the code version.
- Before my data engineering role, I have a widespread experience in Manual Testing.
- Having experience in the Software Development Life Cycle(SDLC) and Software Testing Life Cycle(STLC).
- Thoroughly involved in different types of testing like System Integration Testing, Re-Testing, Browser compatibility Testing.
- Working experience on Linux commands and Windows environments.
- Good experience in all phases of Agile and participated in all daily scrum meetings with cross teams.

WORK EXPERIENCE

Currently working for GLITZ Technologies Private Limited in Bangalore from 06.05.2018 to till date.

Project #1:

Project : Retail & Consumer Goods Platform
Client : **TBInt.**
Duration : October 2021 to till date
Role : GCP Data Engineer
Technologies : GCS, Big Query, Composer, Dataflow, Dataproc
Pod Size : 7

Description:


TB International was established in 2005, in Ober-Ramstadt, Germany. Today, it is one of Germany's leading fashion wholesalers. In little more than a decade, TB International has risen from a startup to becoming one of Germany's leading clothing wholesalers. With a vertically integrated supply chain that handles everything from manufacturing to logistics, the company has won the trust of more than 15,000 retailers worldwide.

By migrating to Google Cloud infrastructure, one of Germany's leading fashion wholesalers supports rapid growth in the highly competitive world of international clothing retail.

Roles and Responsibilities:

- Agile methodology was followed during this project.
- Responsible to manage data coming from different sources.
- Developed producer scripts to read CSV files from the source system and transfer data to Kafka.
- Involved in the development of creating data pipeline to read data from Kafka and load data into BigQuery.
- Composed DAG's and tasks to deploy and run Pyspark jobs in airflow scheduling tool called composer.
- Using Source Code Management (SCM) as GitHub.
- Working Experience on Linux commands and Windows environments. skills and good command over GCP components.
- Responsible for Unit Testing.
- Created authorized views in BigQuery to provide required data to down streams (ML/Analytics/Dashboards)
- Worked on different file formats like text, Sequence, XML, JSON, AVRO, ORC, and Parquet.
- Involved in loading data from LINUX file system to HDFS.
- Actively participated in demo sessions to the client.

Project #2:

Project : KYC
Client :  ZERODHA
Duration : July 2020 to September 2021
Role : GCP Data Engineer
Technologies : GCS, Big Query, Composer, Python
Pod Size : 9

Description:

Zerodha is one of India's largest stock brokers, serving millions of users. The company offers investors attractive brokerage rates for all types of investments, including stocks, mutual funds, and commodities. Each day, Zerodha receives around 25,000 account opening applications, all of which undergo a Know-Your-Customer (KYC) process as required by SEBI in order to minimize fraudulent activities such as identity theft and money laundering. Customers must provide KYC documents such as their Permanent Account Numbers (PAN) cards, to verify their identities.

Roles and Responsibilities:

- Implemented partitioning, dynamic partitions and bucketing in HIVE and Map Side Join's.
- Involved in creating Hive tables loading data, and writing queries that will run internally in MapReduce
- Involved in loading data from LINUX file system to HDFS.
- Involved in creating Hive External tables for HDFS data.
- Worked on Creating DAG's and tasks to deploy and run Pyspark jobs in airflow scheduling tool.
- Using Source Code Management (SCM) as GitHub.
- Working experience on Linux commands and Windows environments. skills and good command over GCP components
- Agile methodology was followed during project.
- Responsible to manage data coming from different sources
- Developed producer scripts to read CSV files from the source system and transfer data to Kafka.
- Involved in the development of creating data pipeline to read data from Kafka and load data to Big Query.
- Responsible for Unit Testing.
- Created authorized views in BigQuery to provide required data to down streams (ML/Analytics/Dashboards)
- Worked on different file formats like text, Sequence, XML, JSON, AVRO, ORC, and Parquet.

Project #3:

Project Name: Online Retailer

Client : Blue Tree

Role : Quality Analyst in Manual Testing

Tools : REST, POSTMAN, SQL-server, JIRA, JSON

Description:

Blue Tree is an American online retailer of indie and vintage-inspired women's clothing. Blue Tree sells jewelry, clothing, antiques, perfume, candles, art, photography, books, vintage LPs, and stuffed animals. Blue Tree uses these pop-up stores to promote existing online and social media services such as Fit for Me and the Style Gallery.

Roles and Responsibilities:

- Good Understanding the business Requirements and Analyzing the requirement
- Able to understand story cards in Product Backlog and Sprint Backlog.
- Able to understand Product owner prioritization.
- Participated in Sprint planning meeting to identify sprint Backlog.
- Involved in Story analysis.
- Experience in identifying test scenarios and implementing Test Cases in JIRA.
- Executing the test cases to identifying the defects and logging in to JIRA.
- Experience on sprint planning tasks to execute during the particular sprint.
- Involved in Identifying Regression suite and execution of Regression Test suite for each sprint.
- Involved in preparation of defect reports, Test reports and Test Metrics.
- Participated in Sprint Retrospective meetings.
- Analyzing the services documents to understand API .
- Experience to test REST services using Postman tool.
- Good understanding on URI, JSON, HTTP Methods and HTTP Status codes.
- Good experience to understand and use the JSON payload.
- Experience to use Headers on postman tool.
- Hands on Experience in POSTMAN –Variables, Pre-request step and SNIPPETS
- Experience to perform Data Driven Testing using POSTMAN Tool
- Validate the JSON Response using POSTMAN Snippet and JSONPath Online Evaluator
- Good Experience to setup the different Environments in POSTMAN tool to perform API Testing
- Log the Bugs in JIRA and reporting the same
- Participate in Daily stand-up calls, Bug Triage calls and status meetings