

Harshit Gupta

+91 8340558673 | harshitgupta.gh53@gmail.com | LinkedIn: [@harshit053](#) | GitHub: [@harshit053](#) | Bangalore, India

WORK EXPERIENCE

Infosys Ltd, Bangalore, India

August 2022 – Present

Client: **American Express - Engineer II (Contractor)**

- Developed PySpark scripts to extract data from GCP BigQuery and convert it to .txt files for transfer to mainframe systems, aiding the Modernization of the Account Receivables System at American Express.
- Utilized GCP Cloud Composer (Airflow DAGs) to automate workflows and employed Dataproc Clusters for efficient job orchestration and Spark cluster management.
- Optimized job performance through in-depth analysis using Dataproc to achieve a 25% increase in overall job throughput and responsiveness, thus reducing execution time and enhancing efficiency.
- Ensured script reliability by implementing Pytest scripts and a Behaviour-Driven Development (BDD) Testing approach to ensure code quality and robustness.
- Developed Python-based SFTP scripts to securely transfer files from GCP to mainframe, adhering to Amex's specific SFTP protocols. Also configured the files using Amex's SFTP tools.
- Collaborated with cross-functional teams using Agile methodologies (Rally, Jira) to optimize workflows and enhance system performance.

Skills Acquired: Python3, PySpark, SQL, Google Cloud Platform (BigQuery, Cloud Composer, Dataproc), Airflow DAGs

PROJECTS

Data Processing Pipeline | **Python3 , PySpark , Apache Spark , SQLite**

[Website](#) | [GitHub](#)

- Demonstrated a simple data processing pipeline designed to clean, transform, and analyze data.
- Optimized Join Operation using broadcast join to efficiently join the resulting dataframes.
- Validated records by identifying suspect records and separated good records from bad ones.
- Created an appropriate data model in an SQLite database and ingested the cleaned data into it.

Skills Learnt: Data extraction, Data handling and manipulation and Data ingestion

YelpCamp | **MongoDB , ExpressJS , EJS , NodeJS , PassportJS , Bootstrap**

[Website](#) | [GitHub](#)

- Developed a full-stack campground review site with user accounts, campground listings, and review posting.
- Implemented authentication and authorization using Passport.js with session-based login.
- Built complete CRUD functionality for campgrounds and reviews, adhering to RESTful API principles.

Skills Learnt: Full-Stack Development, CRUD operations, Authentication and Authorization

DigiMenu | **HTML , CSS , JavaScript , MongoDB , NodeJS , ExpressJS , PassportJS**

[Website](#) | [GitHub](#)

- Built a digital menu application allowing users to browse and view available dishes and categories.
- Implemented core features for menu management, including item creation and display.
- Incorporated user authentication and login functionality using PassportJS.

Skills Learnt: CRUD operations, RESTful routing, Authentication and Authorization

SKILLS

- **Core Competencies:** Python3, Apache Spark and PySpark, SQL Queries, Google Cloud Platform, Pytest.
- **Web Development Skills:** HTML, CSS, JavaScript, MongoDB, NodeJS, ExpressJS.
- **Additional Skills:** Data Structures and Algorithms, Git, GitHub, MongoDB Atlas, Agile Methodology (Jira, Rally).

Certifications

- [Google Cloud Platform – Data Engineer, Big Data](#)

EDUCATION

Visvesvaraya Technological University, Bangalore, Karnataka, IN
B.E. in Electronics and Communication | Overall: **8.8 CGPA**

August 2018 – July 2022

Chinmaya Vidyalaya, Bokaro Steel City, Jharkhand
CBSE - Senior Secondary | Overall: **84%**

March 2016 – March 2017