

A Tarun Teja | Data Engineer | 2+ years (Experience)  
+91 8309603732  
[tarun.aerva.teja@gmail.com](mailto:tarun.aerva.teja@gmail.com)  
Optum Global Solutions (subsidiary of UnitedHealth Group)

### Professional Summary

- An enthusiast Data Engineer and Analyst with a work experience of 2+ years in the HealthCare domain
- Good experience working with the On-Shore clients in understanding and gathering the Business requirements and preparing process flows
- Strong experience working on Big Data solutions such as Hadoop, HDFS, Map Reduce, Sqoop, Hive, Spark, PySpark, Scala Spark on MapR Distribution
- Successfully implemented various automation processes using Shell Scripting and Python to improve and optimise performance and decrease manual process
- Good exposure to Data Science solutions such as Statistics, Machine Learning techniques like Regression, Classification

### Technical Skills

Languages	: Python, Scala, Shell Scripting
Databases	: MYSQL, SQL Server
Big Data Technologies	: Hadoop, MapR Distribution, HDFS, YARN, Map Reduce, Hive, Sqoop, Spark, Spark SQL, PySpark, Scala Spark
Data Science	: Linear and Logistic Regression, Decision Trees, K-Means Clustering and Hierarchical Clustering, Time Series Modelling, evaluation metrics using Classification matrix, Accuracy, Precision, ROC
Tools	: Jupyter Notebook, Tableau, Visio, Excel
Project Management	: Swift Kanban, Azure Management Tool

### Projects Profile

**Designation: Assoc Data Engineering Analyst**

**Project: Benefit Assist Program**

**Client: UnitedHealthcare**

**[July 2019 - Dec 2019]**

- Identify claims of potential members from 40 different employers segregated into platforms having UHC Medical Insurance who are eligible to avail special benefits and to tag these claims and generate the report using which Business can make a outreach to the member to notify the member of the potential benefit
- Managed the process and provided support during the quarterly change requests and code updates with end to end testing without effecting the timely deliverables
- Optimized the application code by incorporating parallel process execution for different platforms thereby reducing the runtime significantly from seven hours to two hours
- Re designed the process flow to Sqoop in the medical feed from DB2 Application thereby reducing the effort of business to manually generate the same and providing the feed to us via Shared Drives
- Enhanced the report generation by incorporating the automation process that writes the final Hive table to an excel spreadsheet using Scala Spark and that shares the report with business via Shared Drive using Shell Scripting thereby eliminating the manual intervention

**Skills Set: Spark - Scala & SQL, Hive, Sqoop, SQL Server, Shell Scripting**

**Designation: Assoc Software Engineer**

**Project: HCA HealthCare**

**Client: UnitedHealthcare**

**[Sep 2020 - Jan 2021]**

- Identify the claims of the employees of HCA Healthcare group having both UHC and Aetna Medical Health Coverages to tag the Supplemental Benefits to the member claims along with provider benefits and make an outreach to the member notifying them of their potential benefit
- Initiated discussions with Business to design an end to end process flow to depict the Business logics and understanding for sign off

**A Tarun Teja | Data Engineer | 2+ years (Experience)**

**+91 8309603732**

[tarun.aerva.teja@gmail.com](mailto:tarun.aerva.teja@gmail.com)

**Optum Global Solutions (subsidiary of UnitedHealth Group)**

- Developed an ETL process to ingest the UHC and Aetna Medical data of the HCA employees shared by Business via Shared Drive into BDPaaS and loading the same into Hive Tables using Scala Spark
- Set up a full load Sqoop process to ingest the Financial Eligibility of the HCA Members present in MY SQL Server into BDPaaS on a regular basis
- Developed ETL scripts for both UHC and Aetna individually using Spark SQL to prepare the final data as per the business logics to tag the benefit flags to the claims
- Integrated the automation process that writes the final Hive table to an excel spreadsheet using Scala Spark and that shares the report with business via Shared Drive using Shell Scripting thereby eliminating the manual intervention
- Developed different automation processes for UHC and Aetna using Shell scripting that wraps up the mentioned processes to run parallelly
- Scheduled the process in TWS that would trigger during the business defined dates to generate the reports which will be utilised by business to make a member outreach

**Skills Set: Spark - Scala & SQL, Shell Scripting - Automation, Sqoop - Data Importing, Excel - Reporting**

**Designation: Software Engineer**

**Project: Claims Data Source Ingestion**

**[April 2021 - June 2021]**

- Developed a process to ingest the raw claims files (seven every day) into Hive tables on daily basis
- Outlined the end to end workflow
- Developed Spark script to create partitioned hive tables leveraging these files
- Incorporated the process of creating Denorm and historical tables from the base hive tables using PySpark which would read the data from the Hive tables and create denorm and historical views
- Developed a shell script to create a process to archive the raw claims files that would archive all the files received one month back and would compress these files
- Moved the process to production where the scripts are scheduled to run seven times every day

**Skills Set: Python - Dataframes, Spark - SQL, PySpark, Shell Scripting – Automation**

#### ***Professional Achievements***

- Promoted by one level (**24-L1 to 24-L2**) within six months of induction
- Promoted from **Grade 24 to 25** (Assoc Data Engineering Analyst to Software Engineer)
- Accoladed with eight Bravo's by the Operations leads, Business Analysts and Managers for Collaboration, Performance and Contribution
- Recognised for conducting a Technical Session on Python (Functions, NumPy, Pandas) to the entire team

#### ***Educational Summary***

- Graduated in Mechanical Engineering from Amrita Vishwa Vidyapeetham with a GPA of 7.77, 2014 - 2018
- Pursued Post-Graduation Program in Data Science Engineering from Great Lakes, 2018 - 2019