**Divya Vuduthala**

Ph: 3146829849| Email: [divya9.vu](mailto:divya9.vud@gmail.com)dhatala@gmail.com

LinkedIn: <https://www.linkedin.com/in/divya-v-81391a134>

# Summary

* An IT graduate with 10 plus years of industrial experience in all phases Software Development Life Cycle with expertise in ETL Tools, Alteryx, Qlik Replicate, Qlik Compose, Tableau, Microsoft Power BI, Python, SQL and Microsoft Azure.
* 7+ years of experience in Alteryx platform that involved data preparation, data blending, building custom macros and the creation of data models.
* 5+ years of experience in Data Manipulation, ExploratoryData Analysis, Reporting Using Alteryx, Microsoft Power BI, Python and SQL.
* 5+ years of experience in scheduled process automation of several Alteryx workflows on desktop and Server.
* 5+ years of experience in ETL using Alteryx to extract data from on-premise ODBCs like Oracle, Cloudera, Databricks and IBM Netezza.
* Build data pipelines in airflow in GCP for ETL related jobs using different airflow operators.
* Experience in GCP Dataproc, GCS, Cloud functions, BigQuery.
* Experience in moving data between GCP and Azure using Azure Data Factory.
* Experience in building power bi reports on Azure Analysis services for better performance.
* Used cloud shell SDK in GCP to configure the services Data Proc, Storage, BigQuer
* 2 years of experience in Data Auditing, Data Validation and Data Reconciliation for data from various sources and targets using Alteryx and developing dashboards for these reports using Microsoft Power BI.
* 2 years of experience in Attunity Qlik Replicate performing Data Replication, Ingestion and Streaming from Oracle to Azure Data Lake and Azure Data Warehouse.
* 2 years of experience in Attunity Qlik Compose associating Qlik Replicate tasks,mapping tables, performing data transformation and establishing change data capture in real time from Oracle to Azure Data Lake and Azure Data Warehouse.
* 2 years of experience in Azure Data Factory performing full load and scheduling pipelines for scheduled updates of data from Oracle and on-premise to Azure Data Lake Storage.
* 3+ years of experience in Tableau Desktop, Tableau Reader, Tableau Prep and experienced in analysis, modeling, design, and development of Tableau reports and dashboards for analytics.
* 3 years of experience in building dashboards, forecasting graphs, building DAX queries using Microsoft Power BI
* 4+ years of experience in Data Science using Alteryx, Python and Azure AutoML.
* 1+ years of experience in Web Development, performing statistical and sentimental analysis on the website users.
* 2 years of experience working with Agile Methodology, Sprint Planning, Scrums, managing features, user stories, tasks, bugs and backlog grooming in Microsoft Azure DevOps.
* Responsible for designing, building, testing and publishing workflows in Alteryx.
* Expertise in creating dashboards, stories and developing different charts styles including Bar Charts, Line Charts, Tree Maps, Pie Charts, Scatter Plots, Bullet Graphs, Histograms, Heat Maps, Geo Maps, Text Graphs and developed reports and using advanced analytics like Reference Lines, Reference Bands, Trend Lines and Box plots in Tableau and Microsoft Power BI.
* Experience in designing entity-relationship models using Microsoft Visio and LucidChart
* Experience in handling multiple ad-hoc requests and delivering them within the deadlines ensuring high quality of work.
* Strong written and oral communication skills, self-motivated, focused, creative, quick learner, ability to work with minimal supervision, adaptive to new technologies and platforms, passionate about solving business problems.
* Alteryx Designer Core certified.

**Technical Experience:**

* I have experience in software development, maintenance, testing, enhancement, production support and system solutions for banking, insurance and Telecommunications industries

# Technical Skills

|  |  |
| --- | --- |
| Operating Systems/ Virtualization | Windows9x/NT/2000/XP/2003/2008, Apple Macintosh, UNIX, LINUX, AIX, Windows 2012 Virtualization, VM Ware, MAC |
| Statistical Modeling | Descriptive statistics, Excel( VLOOKUP, Rand, Pivot tables, Data Analysis tool Pak), Hypothesis Testing, Regression (Linear, Random forest, Lasso, ridge), Classification methods (Logistic, Multinomial, Random forest, XG BOOST, Decision Trees, Naïve Bayes, KNN, SVM), Parameter tuning, Cross validation, Model evaluation (ROC, AUC, Sensitivity, Specificity), NLP (Text mining), Word embedding (CBOW,word2vec,Tf/IDF), Deep learning Neural Networks, AI Computer vision, A/B Testing. |
| Programming Languages | Python, Tableau, Alteryx, SQL, Splunk, JavaScript, HTML, CSS, Cobol, R, Hive, Pyspark |
| Databases | MYSQL, NOSQL, VSAM, Db2 ,   AWS-Dynamo DB,AWS-RDS-MySQL,NoSQL Mongo DB |
| Browsers | Internet Explorer, Chrome, Fire Fox, Netscape Navigator |
| Presentation Tools | Word, Excel, Power Point, Visio |
| Project Management/ Agile Tools | MS Project, RALLY, JIRA |

# Education / Certification / Achievements

**Education:**

1. Bachelor’s in engineering from Jawaharlal Nehru Technological University, India in 2008
2. Master’s in Data Science from Lewis University, Chicago IL, USA in 2017

**Certifications:**

1. ITIL® Foundation Certificate in IT Service Management
2. Design and Implement Big Data Analytics Solutions/ MS0616697386
3. IBM SPSS Statistics Level 1 v2/ IBM000182809
4. [Alteryx](https://www.linkedin.com/company/alteryx?trk=public_profile_result-card_subtitle-click) Designer Core Certified

**Experience**

**Roles: Data Engineer & Senior Data Analyst Dec2021-Till Date**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Valley National Bank**

**Wayne, New Jersey**

**Responsibilities:**

* Provide real-time support to data and product-related queries to support business and technical teams.
* Provide data solutions, tools, and capabilities to enable frameworks.
* Perform data analytics and reporting to generate insights and guide business decisions.
* Provide guidance to partner teams by establishing data dictionaries, data mapping and data patterns.
* Convert business requirements to design, identify tools and develop solutions.
* Support the teams in the selection and implementation of data analytics tools and frameworks that are suitable for their needs.
* Work with different teams to perform diagnostic analytics, prescriptive analytics, and descriptive analytics.
* Articulate with Data Conversion, Data Quality, complex Data Profiling, Performance Tuning, System Testing and implementing RDBMS features.
* Experienced with assisting teams in making important business improvement decisions by providing relevant analysis through audits for various site-specific and global programs.
* Excellent analytical skills, ability to work in a team and communicate effectively at all levels of the development process.
* Build data pipelines in airflow in GCP for ETL related jobs using different airflow operators.
* Experience in moving data between GCP and Azure using Azure Data Factory.
* Experience in building power bi reports on Azure Analysis services for better performance.
* Excellent knowledge of **Building Conceptual/Logical/Physical Architecture, Normalization, Data warehouse architectures, concepts and Technologies, Dimensional Star Schema and Snowflakes Schema methodologies and Data Quality,**
* Design and maintain well-structured relational database schemas. Create automated ETL pipelines for a variety of raw manufacturing data sources. Research and develop new data storage architectures for growing data volumes
* Designing a resilient framework using Amazon Web Services , which involves read/streaming data from the producer/Source through Apache Kafka process to ingest the data to Client specific cloud storage location (S3) . Inhouse data validation tools are integrated to take care of data quality checks.
* Setting up notification service through Amazon SNS (Simple Noti cation Service) which in-turn triggers AWS Lambda function to pick the data from one S3 to Centralized S3 location.
* Data load framework is being used to load the data from S3 location to Cloud Snowflake Data Warehouse . Ensure appropriate services and effective design in a cost-effective method.
* Intake Business manager/product owner’s request on their strategies and future vision and turn those business requirements into data visualization products using Tableau Business Intelligence Data Visualization tool .
* Integrating Snowflake SQL within Python code and setting up Apache AirFlow scheduler to automate the business reports to be generated on a regular basis (Daily/Weekly/Monthly) and push the output to Amazon Storage region (S3).
* Enforced **Data Governance, Data Quality** using in-house, script based and tool-based solution also storage and creation and storage of Metadata using tools.
* Designed and Co-ordinated with the Data Science team in implementing Advanced Analytical Models in Hadoop Cluster over large Datasets.
* Incorporated data warehouse solution for **Analytical, Aggregation and Historical needs.**
* Created operation **Data Stores for reporting and Analysis needs. Created Data Mart for domain specific Analytical requirements.**
* Real Time Data Streaming Analysis using AWS Technology Stack & Confluent KAKFA
* Use industry tools such as Alteryx and Power BI to generate metrics and develop dashboards for insights and data visualization.
* Develop algorithms to automate data reconciliation, data cleansing and data wrangling logics.
* Develop expertise on metadata, lineage, and business definitions to identify gaps and anticipate challenges.

**Sr. Data Analyst/Data Scientist Feb.2018-Dec-2021**

**Charter Communications, St. Louis MO.**

**Responsibilities:**

* Facilitated development life-cycle (SDLC) for Onboarding and Talent Acquisition software including gathering of agency requirements, documenting specifications, business process modeling, application design, quality assurance, implementation, and support.
* Involved in extensive adhoc reporting, routine operational reporting and data manipulation to produce routine metrics and dashboards for management
* Created parameters, action filters and calculated sets for preparing dashboards and worksheets in Tableau
* Participated in End to End data mining life cycle and used advanced data mining techniques to extract the data from different sources, conducted studies and generated rapid plots with different visualization tools.
* Interacting with other data scientists and architects, custom solutions for data visualization using tools like a **tableau**, **Splunk, pyspark** and **Packages** in **python**
* Developed **Python** modules for **machine learning** & **predictive analytics**
* Maintenance of large data sets, combining data from various sources by **Excel, Enterprise, and** **Access** and **SQL queries**
* Worked on **data cleaning, data preparation** and feature engineering with **Python**, including **Numpy, Scipy, Matplotlib, Seaborn, Pandas,** and **Scikit-learn**
* Created visualization dashboards in **Splunk** and **Tableau**
* Created Deep Learning model to predict on outages as well as other important information needed for DevOps teams
* Preprocessed data using PCA and created workflows other dimensionality reduction techniques as well as tools such as Alteryx
* Performed **SQL Testing** on **databases**
* Developed **SQL** scripts using **OLAP** functions to improve the query performance while pulling the data from large tables
* Collaborate with application and enterprise architects to define solution standards in data architecture and conceptual modeling for transaction, reporting, data warehouse databases with considerations to database availability, scalability manageability, and performance. Created various proof of concepts in Big Data Solution by sourcing data from Structured, Semi-structured and unstructured data i.e., Relational, Weblogs, External Data, Social media, Weblogs and various other data sources. Information converted to valuable data insights and data augmentation for decision, visualization and prediction process. Suggested latest upgrades and technologies for NoSQL databases.
* Performance tuning and optimizing the application.
* Publish Interactive dashboards and schedule auto-data refreshes
* Created Splunk queries that helped market analysts spot emerging trends by comparing incremental data with reference tables and historical metrics
* Machine learning process with Splunk and building models, testing, predicting /preventing the outages of the buyflow
* Created new scripts for Splunk scripted input for system, collecting **CPU** and **OS** data
* Built **Power BI** Data Visualization **Dashboard** for VoLTE and LTE trending
* Created automated ETL process utilizing Alteryx, SQL and Python to combine Sales/Orders data and Outage Data to provide insights for decision making and effective targeting of Outages reducing up to 90% of manual process work
* Created multiple Tableau dashboards for Released Offer Tracking, Outage Management Analysis & Customer Behavior
* Created ETL utilizing Alteryx and SQL to combine data from outage management system and AppDynamics and Splunk and provide reporting on Alerting setup to calculate response time in turn increasing the efficiency of DevOps team
* Created python scripts for tasks such as SFTP processes and managing Alteryx gallery using Python and R scripts
* Contributed in development of in-house R shiny application for improved management and reporting of system outages
* Creating Alteryx ETL to replace legacy script utilizing SQL which reduced runtime and increased overall efficiency

**Environment:** Alteryx, Python, R, Tableau, Splunk,9.2, SQL

**Data Analyst Feb.2017-June-2017**

**Medorizon, Romeoville IL.**

**Responsibilities:**

* Analyzed massive and highly complex data sets, performing ad-hoc analysis and data manipulation
* Wrote reports to reporting system to extract data for analysis using filters based on the analysis
* Worked on complex information model, logical relationships, and the data structures that support different jean brands
* Wrote several **SQL Queries** using **Teradata SQL** Assistant for Ad Hoc Data Pull request
* Performing **statistical data analysis** and **data visualization** using R and Python
* Worked on **creating filters** and **calculated sets** for preparing dashboards and worksheets in **Tableau**
* Identified risk level and eligibility of new insurance applicants with Machine Learning algorithms

**Software Engineer June 2012 – September 2016**

**IBM (CBSI, Hyderabad, India)**

**Responsibilities:**

* Conducted analysis in assessing customer consuming behaviors and discover value of customers with RMF analysis; applied customer segmentation with clustering algorithms such as **K-Means** Clustering and Hierarchical Clustering
* Collaborated with data engineers to implement **ETL process**, wrote and optimized **SQL queries** to perform data extraction and merging from Oracle
* Performed data integrity checks, data cleaning, exploratory analysis and feature engineer using **R** and **Python**
* Involved in converting **Hive/SQL** queries into Spark transformations using Spark RDDs, Spark SQL using Scala
* Merged data from different sources using Hive joins and performed Adhoc queries.
* Developed personalized product recommendation with **Machine Learning algorithms**, including **Gradient Boosting Tree** and **Collaborative filtering** to better meet the needs of existing customers and acquire new customers
* Used Python and Spark to implement different machine learning algorithms, including Generalized **Linear Model, Random Forest**, **SVM, Boosting** and **Neural Network**
* Worked on Spark and used PySpark, sparksql programming languages to process the large volumes of data.
* Worked on **data cleaning, data preparation** and feature engineering with **Python**, including **Numpy, Scipy, Matplotlib, Seaborn, Pandas,** and **Scikit-learn**
* Determined customer satisfaction and helped enhance customer using **NLP**
* Recommended and evaluated marketing approaches based on quality analytics on customer consuming behavior
* Performed data visualization and Designed dashboards with **Tableau** and provided complex reports, including charts, summaries, and graphs to interpret the findings to the team and stakeholders

**Environment:** R Studio, Python, Tableau, SQL Server 2012, 2014 and Oracle 10g, 11g

**Software Engineer Dec 2011 – May 2012**

**HCB Core Banking (HSBC), (DCS), Pune, India**

**Responsibilities:**

* Work with users to identify the most appropriate source of record required to define the asset data for financing
* Designed, developed, deployed and automated several data engineering projects using Alteryx, which eliminated human intervention and saved processing time.
* Designed, developed and deployed audits and reconciliation workflows in Alteryx for various data sources to validate with data on cloud and on-premise targets.
* Performed data extractions, transformations, exploratory data analysis, statistical analysis and machine learning in both Alteryx and Python.
* Developed various ETL and ELT jobs using Alteryx and Qlik Compose.
* Worked on Azure Auto ML to train & tune the models, democratize and empower machine learning model development process.
* Wrote various SQL scripts as a process of data extraction and data manipulation in Microsoft SQL, Impala SQL and Spark SQL environments.
* Replicated majority of Oracle data to Azure Data Warehouse and Azure Data Lake and established change data capture in real time using Attunity Qlik Replicate.
* Developed multiple Attunity Qlik Compose tasks associating Qlik Replicate tasks, mapping tables, performing data transformation and establishing change data capture in real time from Oracle to Azure Data Lake and Azure Data Warehouse.
* Migrated cloud and on-premise data from multiple sources to Azure Data Lake Storage using Azure Data Factory and validated the quality and precision of the migrated data in Azure Databricks.
* Experience in error handling and troubleshooting the scripts failure, to load the data into the database.
* Generated various Daily, Weekly, Monthly, Quarterly, Bi-Yearly, Yearly reports by different types of reports using Microsoft Power BI.
* Built and published several enterprise wide dashboards for leaders using Microsoft Power BI.
* Developed functionality to show hierarchies in Microsoft Power BI for all levels of employees.
* Responsible for deploying reports to the report manager and troubleshooting for any error occurring in execution.
* Designed the architecture and entity-relationship diagrams for all the Alteryx projects using Microsoft Visio.
* Handled multiple ad-hoc requests and delivered them within the deadlines ensuring high quality of work.
* Worked in an Agile environment by handling Sprint Planning, Scrums, managing features, user stories, tasks, bugs and backlog grooming in Microsoft Azure DevOps.

**Environment**: SQL/Server, R connector, Python, R, Tableau