

# Md Abdul Rahim

Bigdata/Azure ETL Developer

**Address** KAKINADA AP

**Phone** 8125264046

**E-mail** Abdul121512@gmail.com

Highly skilled data engineer with over 6+ years of experience designing, building, and maintaining large-scale data systems. Proficient in various programming languages including Python, SQL, and Pyspark, with extensive experience in data warehousing, Azure ETL, and data modeling. Adept at troubleshooting complex data issues and implementing innovative solutions to optimize system performance.



## Skills

Databricks



Databases SQL, Oracle, Azure SQL



Data Manipulation



Hadoop



Data Modeling



Python SQL



Hive



GCP I



DBT



Spark Pyspark



ASTRO



Bigquery dataproc



HDFS



## Work History

2023-01 -  
Current

### Senior Data Engineer

Birla soft limited (payroll Altimetrik)

- Performed data cleaning on unstructured information using various Hadoop tools.
- Experience in NoSQL HBase, Hive data warehouse, CI/CD tools like Jenkins, GitHub, Hadoop distribution cloudera, Jira for hours and tasks tracking.

- Experience in manipulating/analysing large datasets and insights within structured and unstructured text data.
- Strong analytical, communication and organizational skills with the ability to analyze business practices and define and document optimal procedures or practices.
- Expertise in Apache Spark & Python.  
Proficient in developing optimized solutions in Hadoop environments.
- Created entire applications and patchable connectors for use with cloud-based platforms such as GCP

2022-06 -  
2022-12

## **Azure Data Engineer**

Advance auto parts(PayrollTEKsystems)

- Real time data is censored from Can Bus and will be batched into a group of data and sent into the hub
- Write Azure service bus topic and Azure functions when abnormal data was found in streaming analytics service
- Created SQL database for storing vehicle trip information's
- Created blob storage to save raw data sent from streaming analytics
- Constructed Azure Document DB to save the latest status of the target car
- Deployed data factory for creating data pipeline to orchestrate the data into SQL database
- Collaborated on ETL (Extract, Transform, Load) tasks, maintaining data integrity and verifying pipeline stability
- Designed and implemented effective database solutions (Azure blob storage) to store and retrieve data

2020-07 -  
2021-07

## **Azure Data Engineer**

Tech Mahindra

- Developed pipelines in azure data factory to fetch data from different sources and loaded it into Azure synapse analytics Extensively involved in writing transforming data in PySpark and pushing into ADLS
- Responsible for creating copy activity, lookup activity, Metadata activity, Monitor the pipelines are report fix the issues Creating data flows to transforming the data to Azure using Azure data factory
- Roles and Responsibilities Azure Data Engineers are responsible for designing data solutions that meet the organization's requirements
- They manage the process of collecting and ingesting data from various sources into Azure data storage solutions, such as Azure Data Lake Storage, Azure Blob Storage, or Azure SQLData Warehouse
- This includes data cleaning, transformation, and enrichment using tools like Azure Databricks, Azure Data Factory Data Flow, or Azure Functions
- Ensuring data security and compliance with regulations (such as GDPR or HIPAA) is a critical responsibility
- Optimizing the performance of Azure data solutions is crucial
- Logic Apps, Azure Functions, or scripting languages to streamline processes and improve efficiency.

2018-01 -  
2020-05

## **Azure Data Engineer**

IBM

- Strong desire in learning new technologies and methodologies. Strong Azure BI Development Experience (Azure Blob Storage, Azure Data Lake, Azure SQL Server, Azure Data factory)
- Working on Tasks such as Dataflow task, Execute SQL, Execute Package, File system and Containers like for loop, for each loop in Integration services development.
- Working with various transformations such as Lookup, Conditional Split, Derived column, merge join etc.
- Hands on experience in creating, deploying, trouble shooting, scheduling, and monitoring SSIS Packages.
- Good understanding on SQL concepts like DML, DDL operations, functions, stored procedures, and joins.
- Extensive experience in writing and debugging stored procedures, views, Functions and triggers.



## Education

### **Bachelor of Arts: Accounting And Computer Science**

Arni Unvercity - Himchal Pradesh



## Accomplishments

- Working with various transformations such as Lookup, Conditional Split, Derived column, merge join etc
- Extensive experience in writing and debugging stored procedures, views, Functions and triggers
- A good knowledge on indexes both clustered and non-clustered
- Connecting data sources, importing data and transforming data for Business intelligence
- Experience (Azure Blob Storage, Azure Data Lake, Azure SQL Server, Azure Data factory)
- Good understating to store and read data from cloud storages like azure blob storage and data lake
- Scheduled the ETL packages at desired frequencies using SQL server agent jobs
- Have experience migrating SSIS packages from on premises to Azure Data factory.



## Languages

English, Telugu,Urdu,Hindi



## Projects

## **PROJECT :1**

Title : GSK - Cleo

Client : GSK

Role : ETL Data engineer: GSK Consumer Healthcare

Company: Tech mahindra

Description:

GSK Consumer Healthcare India is a leading fast-moving consumer healthcare company with over 500 employees in India

and category-leading brands such as Sensodyne, Eno, Crocin and Otrivin. Its Mission Health program touches the lives of approximately 90,000 people by focusing on health and well being with projects dedicated to neglected tropical diseases, dengue management, cleft and palate surgeries, and plastic waste management.

Responsibilities:

- Developed ETL pipelines
- Developed complex PySpark scripts
- Checked code to Git and used Jenkins as CI/CD pipeline
- Worked on snowflake, Sagemaker for analytics workloads.
- Worked on TPT data ingestion frame work to import data from Teradata to hdfs.
- Environment: Agile, Databricks, , ETL,Bigquery, PySpark, Spark, Python. Airflow,DBT

## **PROJECT :2**

Project Title :Market trend analysis

Client : Birla Soft Limited

Environment : Hadoop, HDFS, PySpark, Hive, ETL, Python, HBase, Linux.

Role : BigData engineer

Company :Altimetrik

Description:

Analyse market trends by calculating the monthly sales growth for each product category. As your business becomes more established, you will be able to compare data and identify trends in financial performance, competitor movement and growth, manufacturing efficiency, new or emerging technologies, customer complaints, staff performance reviews

Roles and Responsibilities:

- Implementing Spark And Python jobs to transform and store data.
- Extracting the Data from the Hive and HBase based on the user requirements.
- Migrating the Hive tables from one file system to other based the project requirement.
- Perform big data processing using Hadoop, Sqoop, Hive and Spark Core and Spark SQL.
- Coordinating with team members for system design, integration, application maintenance.
- Working on different file formats like ORC, JSON and CSV.

Environment : ETL, Azure ,Data Lake ,Spark,Hadoop ,  
Airflow,pyspark

