Shekhar Suman

Data Engineer

🏠 shekhar.com   🐙 Github   in li-Shekhar   ✉ shekhar.dakshana17@gmail.com   📞 9113408368

EXPERIENCE

ATCS A NAGARRO COMPANY | DATA ENGINEER, MERCEDES-BENZ USA
Feb 2023 – July 2023 | Jaipur, Raj.

- ➔ Worked as a Data Engineer for Data migration from Legacy system(Mainframe DB2) to Azure Cloud Data Warehouse for Mercedes-Benz USA.
- ➔ Analyzing pre-existing data-stage jobs and replicating them in Azure Databricks using Pyspark.
- ➔ Created ADF pipeline to perform ETL jobs/processes, maintaining and scheduling them.
- ➔ Performed Data cleansing, data cataloging, and data transformation on Azure Databricks service using Pyspark.
- ➔ Performance optimization of the Azure Data Factory pipelines and provide fixes to reduce/minimize the ADF pipeline failures

ATCS A NAGARRO COMPANY. | ASSOCIATE DATA ENGINEER, GLOBAL INVESTMENT MANAGEMENT FIRM
July 2022 – Jan 2023 | Jaipur, Raj.

- ➔ Building an ETL (Extract, Transform, Load) pipeline for migrating on-premises CSV data from an SFTP server to AWS S3 and generating business data using AWS.
- ➔ Perform data transformation tasks as required by your business needs. This may include cleaning, filtering, joining, and aggregating the data.
- ➔ Organized the transformed data in the S3 bucket with appropriate prefixes and folders to facilitate easy access and management.
- ➔ Designed and implemented item tables in AWS Athena using Glue Data Catalog for efficient data querying and analysis.
- ➔ Performed on flask webserver which pass the request to flask server to trigger and tracked R-script inside EC2 and Scheduled EC2 state change to running or stopped using AWS event Bridge.

WOCKHARDT A PHARMA COMPANY | MANAGEMENT TRAINEE
July 2021 – June 2022 | Aurangabad, Maharashtra

- ➔ This project is based on the production of human insulin. Our role was that how to increase the efficiency of production's Yield.
- ➔ Undergone the whole downstream process of project human insulin Training along with plant visiting and MS Excel, and SQL as a part of Wockhardt's learning culture.
- ➔ Worked on SQL query and MS Excel to prepare a control chart of unit operation considering previous batch data to Analyse the Consistency of the process's Yield.
- ➔ Learned the Concept of Feature Engineering, Data Transformation, and Extract.

PROJECTS

IMPACT DATA ANALYSIS OF HOUSE PRICES DATASET | ML AND PYTHON LIBRARIES
Feb 2023

- ➔ Conducted thorough data preprocessing and cleaning to handle missing values, outliers, and inconsistencies in the House Prices dataset.
- ➔ Derived actionable insights from the model to understand how specific features positively or negatively influenced house prices.
- ➔ Created informative visualizations, including scatter plots, box plots, and heatmaps, to communicate the results of the impact data analysis effectively.

SKILLS

PROGRAMMING

Language:
Python • SQL • PostgreSQL

Cloud Services:
AWS Glue • AWS S3 • AWS Athena • AWS EMR • AWS EC2 • Azure Data Factory • Databricks • Azure Storage Account(Blob/Container)

Databases:
MySQL • PostgreSQL • Azure SQL Server • DynamoDB

LIBRARIES/FRAMEWORKS

Pandas • NumPy • Pyspark • Django and Rest Framework • flask

TOOLS/PLATFORMS

VS Code • Docker • ETL • Ms Excel • NoteBook • Spyder

EDUCATION

IIT BOMBAY
BTECH IN CHEMICAL ENGINEERING
June 2017 - June 2021 | Mumbai
Cum. CGPA: 6.21 / 10.0

J N V, BANGALORE URBAN
INTERMEDIATE IN PCM
March 2014 - March 2016 | Bangalore
School of Navodaya Vidyalaya Samiti
Percentage : 80

ACHIEVENTS

- Ranked in Top 0.41 percentile amongst 1.4 million participants in **IIT JEE Main**
- Participated in Inter IIT sports meet [Basketball]
- Represented with Patna Region team in **U19 National Basketball** organized by NVS