

# Nishant Notani

---

## Data Engineer

Ahmedabad, Gujarat

Phone: 9664870514 | Email: [nishantnotani.nn@gmail.com](mailto:nishantnotani.nn@gmail.com) | LinkedIn:-[Nishant Notani](#)

## Professional Summary

As a Data Engineer, I have 4 years of solid expertise in designing, developing, testing, and maintaining data management and processing systems. I have a thorough understanding of technologies like Python, SQL, NoSQL databases (MongoDB, Redis), cloud platforms (AWS, Azure), Spark, Spark-SQL, Airflow, Azure Synapse, Django, and Docker. I am eager to explore many more technologies.

## Work Experience

### Data Engineer | Dxfactor

April 2023 - Current

#### 1) Data Warehouse Project - Azure Synapse Analytics

April 2023 – Current

- Implemented a Data Warehouse solution by extracting data from AWS S3 / Azure Blob through Azure Synapse Ingestion and inserting it into Azure SQL Server.
- Fetched data from Mailchimp/SharePoint using Azure Synapse Notebooks.
- Inserted and updated records in Mailchimp for bulk emailing to gym members using Azure Synapse Notebooks.
- Utilized Python and basic Spark SQL in Azure Synapse notebooks for ETL processing, including data cleaning, transformation, sorting, and filtering.

#### 2) Tableau Bridge Project

April 2023 – Current

- Set up the Tableau Bridge client to connect the client's incompatible database with Tableau Online, ensuring daily updates.
- Simplified complex data generation using Tableau Prep, creating multiple Tableau flows.
- Designed and developed over 50 reports related to gym data using the Tableau flows.

#### 3) The Warehouse Project on AWS

January 2024 – Current

- Created a raw data layer using AWS DMS, fetching data from Azure SQL Server and inserting it into an S3 raw bucket.
- Developed curated layer pipelines using AWS Lambda and AWS Glue.
- Processed data in Glue pipelines, fetching it from the raw S3 bucket and inserting it into a curated S3 bucket.
- Created Glue crawlers in the curated layer for Athena access.
- Built Glue pipelines to process data from the curated layer and insert it into the analytical layer in PostgreSQL.

#### 4) Data Lake Project - AWS QuickSight

April 2023 – April 2024

- Connected processed data stored in S3 with AWS QuickSight to build insightful dashboards.
- Developed a user-friendly page within the client's application to set and update goals for the dashboard.
- Integrated dashboards seamlessly into the client's application.

### Data Engineer | Infosense Global

June 2021 - March 2023

#### 1) Mlsense

June 2021 – November 2022

- Integrated multiple data sources (PostgreSQL, Azure SQL Server, MongoDB) into the Mlsense application.
- Applied strong analytical skills to work with SQL/NoSQL databases, primarily Azure SQL Server.
- Implemented various data imputations, encoding, outlier detection, and transformation techniques.
- Optimized data detail, feature engineering, data exploration, and feature selection, improving performance by more than 40%.
- Converted ML feature engineering logic to Spark SQL.
- Integrated antivirus functionality and automation intelligence features in the Mlsense application.
- Extensively used SQL, NumPy, Pandas, and Spark for data analysis.

#### 2) Data Migration Project

October 2021 – June 2022

- Utilized AWS DMS to migrate Azure SQL Server data to AWS S3.
- Created ETL pipelines using PySpark and AWS Glue for application development.
- Implemented AWS services such as Glue, DMS, and Athena.
- Gained practical knowledge of AWS services, including Amazon S3, Amazon RDS, IAM, Glue, Lambda, Athena, and Azure services like ADF, Synapse, Azure SQL Server, and Databricks Notebooks.

### 3) Data Visualization Project

October 2021 – June 2022

- Created multiple dashboards that simplified complex data, delivering insights to clients using area charts, bar charts, heat maps, scatter plots, and pie charts.
- Pulled all data from Azure Synapse SQL Server, where analytical pipelines were created.
- Worked extensively with BI tools including Tableau, Power BI, AWS QuickSight, and Sisense.

### 4) Data Fusion Extraction

December 2022 – Current

- Added functionality to the application that allows for dynamic data source changes without rewriting code.
- Implemented CSV and PDF download functionality.

## Data Science Trainee | Fingertips

October 2020 - March 2021

- Learned data ingestion, processing, and machine learning model predictions.
- Gained experience with tools like Power BI and Tableau, and libraries like Pandas, NumPy, and DB connectors.

### 1) Telecom Churn Analysis (March 2021)

- GitHub Link: <https://github.com/Nishant-Notani/Final-project-telecom-churn>
- Language used: Python.
- Machine Learning-based project to predict customer churn.

### 2) Loan Status Classification (February 2021)

- GitHub Link: <https://github.com/Nishant-Notani/loan-status-classification>
- Language used: Python.
- Machine Learning-based project for classifying loans based on default risk.

### 3) Home Price Prediction (January 2021)

- GitHub Link: <https://github.com/Nishant-Notani/home-price-prediction>
- Language used: Python.
- Machine Learning-based project to predict home prices using various features.

## **Skills**

- Programming Languages: Python
- Data Processing & ETL: Spark, Spark-SQL, PySpark, ETL Development
- Web Frameworks: Django (Rest Framework: Backend)
- Cloud Platforms: AWS (S3, Glue, Lambda, Athena, DMS, SNS, CloudWatch), Azure (Azure Data Factory, Azure Synapse Analytics, Azure Databricks Notebooks)
- Databases: SQL (PostgreSQL, MySQL, SQL Server, Oracle), NoSQL (MongoDB, Redis)
- Orchestration & Automation: Airflow
- Data Visualization: Tableau, Power BI, AWS QuickSight, Sisense

## **Education**

Bachelor of Engineering in Information Technology, 2020

GTU AIT - Ahmedabad, Gujarat