

**YASH VARDHAN YADAV**

# Business Analyst

+[917565888485](tel:917565888485) |

[MAILTO:YASHZOB@GMAIL.COM](mailto:YASHZOB@GMAIL.COM) |

[GITHUB](#) |

[LINKEDIN](#)

## PROFESSIONAL SUMMARY

Experienced Data Scientist specializing in data cleaning, preprocessing, and exploratory data analysis (EDA), with a strong background in machine learning, AI-driven applications, and big data processing. Proven expertise in handling large datasets, optimizing data pipelines, and ensuring data integrity for high-quality insights. Passionate about data-driven decision-making and automation

## EXPERIENCE

### CONSISTENT INFOSYSTEMS - DELHI (APRIL 2024 – PRESENT)

- Developed and implemented data cleaning pipelines to preprocess large datasets, improving model accuracy by 25%.
- Applied outlier detection techniques (IQR, Z-score, DBSCAN) to refine datasets and remove noise, leading to better predictive accuracy.
- Built automated data quality checks using Python and Pandas, reducing missing values and inconsistencies by 40%.
- Utilized feature engineering techniques to enhance predictive model performance, including imputation methods (KNN, MICE, mean/median).
- Worked closely with business teams to understand key objectives and translated them into actionable data insights, resulting in optimized decision-making processes.
- Performed time-series data cleaning and transformation, ensuring accurate forecasting using ARIMA, Prophet, and LSTM models.
- Designed a data validation framework using Python and PySpark to ensure consistency across multiple data sources.
- Created an automated missing value imputation system using regression-based and iterative techniques for structured datasets.

### PROJECTS & TECHNICAL WORK

#### Customer Data Cleaning & Segmentation

- Processed 100,000+ customer records, handling missing and inconsistent data with advanced imputation methods.
- Applied data normalization and standardization techniques to ensure dataset compatibility for clustering models.
- Implemented unsupervised learning techniques such as KMeans and DBSCAN for effective customer segmentation.

### E-COMMERCE SALES DATA CLEANING & FORECASTING

- Cleaned and analyzed 5 years of sales data, identifying missing values and inconsistencies using data profiling tools.
- Used data interpolation and smoothing techniques to correct irregularities in time-series sales data.
- Applied seasonality decomposition and anomaly detection to improve forecasting model performance.

### FINANCIAL DATA PREPROCESSING & ANOMALY DETECTION

- Built a data pipeline for fraud detection, preprocessing financial transactions and identifying anomalies with isolation forests.
- Applied log transformation and feature scaling to improve model efficiency.
- Developed a real-time data cleaning tool for transaction processing using Pandas and PySpark.

## SKILLS

- Programming Languages: Python
- Data Cleaning & Processing: Pandas, NumPy, PySpark, OpenRefine
- **Feature Engineering** & Data Transformation: Scikit-learn, Statsmodels, Data Profiling
- Data Visualization: Matplotlib, Seaborn, Plotly
- Machine Learning & Statistical Techniques: Supervised & Unsupervised Learning, Outlier Detection, Data Imputation
- Business Analyst: Power Bi and MS Excel

## PREVIOUS EXPERIENCE

### PYTHON TEACHER & MENTOR

- Sanskaar Int. School (July 2021 – Oct 2023)
- Guided students in data preprocessing techniques, handling missing data, and feature engineering in Python.
- Computer Instructor
- Kendriya Vidyalaya Lakhimpur Kheri (Sep 2019 – March 2020)
- Developed educational programs focused on data analysis and cleaning methodologies.
- Java Trainer
- Net Tech Computer Institute (Jan 2014 – March 2018, Evening)
- Trained students in data structures and preprocessing techniques.
- IT Manager & Data Analyst.