

# YASH RAYTHATHA

Email: [yashbraythatha@gmail.com](mailto:yashbraythatha@gmail.com) || LinkedIn: [Linkedin](#) || Mobile: (973)380-6309 || Website: [Portfolio](#)

## SUMMARY

Experienced data professional with a strong background in data analysis, business intelligence, and reporting. Proficient in using SQL and BI tools to deliver strategic insights and support data-driven decision-making. Adept at collaborating cross-functionally to gather requirements and develop automated reports, driving business growth and operational efficiency. Seeking to leverage analytical expertise and technical skills in a dynamic, innovative environment.

## SKILLS

**Programming Languages:** Python, C/C++, HTML, CSS

**Frameworks & Tools:** Jupyter Notebook, Google Collab, ETL, PowerShell, AWS.

**Data Analysis & Visualization:** Pandas, NumPy, Matplotlib, Seaborn, SQL, Excel, Power BI, Data Cleaning, Tableau, quicksight, EDA.

**Big Data & Cloud Technologies:** Hadoop, Spark, AWS, Cloud Computing, Snowflake

**ML Techniques:** Model development, Deep Learning, NLP, Unsupervised and supervised learning, Parameter Efficient Fine tuning, Transfer learning

**Database:** MySQL, SQL, Spark SQL, Oracle SQL, PL/SQL.

## EXPERIENCE

### Data Engineer & Analyst

HCL Technologies August 2019 – March 2023

- Developed scalable ETL pipelines using Apache Spark and PySpark for both batch and streaming workflows, improving data processing efficiency by 20% and reducing latency in downstream data lake and warehouse analytics.
- Optimized Spark SQL transformations for large-scale datasets (10TB+), reducing query runtime from 2 hours to 10 minutes, saving 100+ compute hours/month and lowering AWS infrastructure costs by 30%.
- Automated data ingestion and transformation workflows using AWS Lambda, S3, and Step Functions, cutting manual intervention by 80% and improving pipeline reliability and performance.
- Designed and deployed cloud-native data pipelines to populate interactive dashboards (Power BI, Tableau) with real-time KPIs, supporting operational and executive reporting across departments.
- Built monitoring, validation, and alerting mechanisms to enforce data quality, ensure SLA compliance, and align with enterprise data governance standards in production pipelines.
- Maintained end-to-end documentation of data pipelines, schemas, and integrations to support onboarding and stakeholder collaboration.

### Junior Data Analyst

Savants Solutions Pvt Ltd November 2018 – August 2019

- Increased reporting efficiency by 40% by automating data extraction and visualization using Power BI and Tableau, enabling faster decision-making for stakeholders through KPIs.
- Improved data accuracy by 35% by optimizing SQL queries and implementing data validation checks, ensuring reliable business insights.
- Simplified complex data insights for non-technical stakeholders, improving communication and facilitating strategic decision making.

## PROJECTS

### Energy Consumption Analysis Dashboard [Link](#)

- Built a comprehensive Power BI dashboard to analyze electricity, gas, and water consumption and cost across cities and buildings.
- Visualized year-wise trends and usage breakdowns, enabling stakeholders to identify inefficiencies and optimize energy consumption.
- Performed data modeling and transformation to support interactive filtering and comparative insights across 11 buildings and 3 energy types.

### Superstore Data Analysis Dashboard [Link](#)

- Designed an end-to-end Power BI solution using DAX and Power Query to analyze Superstore performance
- Performed data cleansing and transformations to prepare the dataset for effective visualization and analysis.
- Developed insights into sales performance, return rates, and category trends, enabling identification of growth opportunities.

### Laptop Price Predictor [Link](#)

- Leveraged a Kaggle dataset, utilized Pandas for data cleaning, and employed Matplotlib and Seaborn for insightful visualizations to drive feature selection for predictive modeling.
- Developed and refined multiple machine learning algorithms to accurately predict laptop prices, focusing on optimizing model performance for accuracy.
- Deployed the predictive model using Streamlit, creating an interactive web application that enables real-time, user-driven laptop price estimations.

## CERTIFICATIONS [Link](#)

Azure Administrator (AZ-104), Oracle SQL Database Associate (1ZO-071), Microsoft Power Bi Data Analyst Associate (PL-300).

## EDUCATION

### Bachelor of Engineering: Computer Technology

May 2017

Rajiv Gandhi College of Engineering Research & Technology - Chandrapur, Maharashtra, India

Relevant Coursework: C/C++, Java, Data structure & Algorithms, Management Information system.

### Master of Science: Data Science

May 2025

University of New Haven – West Haven, Connecticut, USA

Relevant Coursework: Python, Machine learning, AI, Deep Learning, Natural language processing (NLP), Power Bi.