Ritesh Ojha

Portfolio | LinkedIn | GitHub | | riteshojha2002@gmail.com | Pune, Maharashtra | +91 7354012014

Work Experience

Data Analyst Intern **Junior Robo** Jun 2023-Jan 2024

Remote

- Developed and implemented a data management system using Google Sheets
- Built 5+ interactive dashboard reports in Power BI to visualize data.
- Automated data storage processes using JavaScript, improving efficiency by 80 %.
- Developed an 8+ task automation using Google App Script, streamlining communication and data delivery.

Advanced Skills: Google Sheets, Power BI, App Script, JavaScript, HTML/CSS, Google Workspace API

Education

B Tech IT Data Science **Ajeenkya DY Patil University, Pune** Jun 2020-Jun 2024

- GPA: 8.74 /10.

Skills

- Languages and Runtimes: Python, JavaScript, Shell Scripting
- Big Data Technologies: Airflow, Apache Spark, Kafka, Talend,
- Developer Tools: Power BI, Excel, Tableau, Looker Studio,
- Cloud Platforms: Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft Azure
- Infrastructure Tools: Docker, Kubernetes
- Databases: MySQL, PostgreSQL, Cassandra
- Other Tools: Git/GitHub, GitHub Actions, Linux,

Projects

Smart City Data Pipeline *(Skills: Python, Kafka, Apache Spark, AWS S3, AWS Crawler, AWS Athena, Docker)*

- Simulated data streams (100 messages/minute) on vehicles, traffic, GPS (coordinates), & weather.
- Utilized a Kafka consumer to ingest the simulated data streams, enabling real-time processing.
- Processed data with Apache Spark for real-time analysis.
- Stored processed data in Parquet format on AWS S3 (scalable cloud storage).
- Enabled efficient querying with AWS Athena (serverless analytics).

Podcast Episodes Data Pipeline *(Skills: Python, Docker, Airflow, PostgreSQL)*

- Automated daily data retrieval for a Podcast Marketplace
- Developed Airflow DAGs to ingest data from 50+ daily podcasts.
- Extracted and processed podcast episode data using Python.
- Utilized Docker for containerization and streamlined service deployment.
- Loaded processed data into a PostgreSQL database for efficient querying.

CERTIFICATIONS

- AWS Academy Cloud Foundations
- AWS Academy Data Analytics
- Google Data Analytics - Coursera
- Power BI Guide - EnterpriseDNA
- Python Basic - HakerRank