

RISHABH VAIBHAV

+91 7061864929 [✉ Rishabhvaibhav7@gmail.com](mailto:Rishabhvaibhav7@gmail.com) [in LinkedIn](#) [G Github](#)

Education

Haldia Institute of Technology, Haldia

B.Tech in Computer Science & Engineering

2018 - 2022

7.9/10 CGPA

Work Experience

ALTEN india

July 2022 – present

Data Engineer

WFO

- Worked on Development of backend services using **Python, Django and SQL**
- Worked with python Libraries like : **Beautiful Soup, Numpy, Pandas dataframe, Matplotlib, urllib2.**
- Experienced in developing web-based applications using **Python, Pyspark, Django, CSS, HTML, JavaScript**
- Leveraged **Pyspark** and **AWS** for efficient data cleaning tasks at scale

Projects

SQL-Integrated Data Dashboard

[Source Code](#)

- Established and managed connections to **SQL databases**, for efficient data retrieval in dashboard.
- Implemented a secure login system using **Django's authentication framework**
- Designed and coded the front-end using **HTML, CSS, and JavaScript** for enhanced user experience.
- Integrated **data visualization tools** to transform complex data sets into easily understandable **graphical representations**, enhancing user experience and decision-making processes.

Advanced Python Data Cleaning

[Source Code](#)

- Researched and sourced extensive data, cleaned, and transformed it using Python libraries like **Pandas and NumPy**.
- Developed, created, and maintained **ETL** processes in complex **transformations and extractions of data**.
- Design and optimization of graphical visualizations of data using **Matplotlib** toward improving the **decision-making process**.
- Managed and manipulated big datasets to provide high-quality results while maintaining data quality and integrity.

Efficient Data Transformation with PySpark

[Source Code](#)

- Developed and implemented **exploratory data analysis** using **PySpark** for building and **maintaining ETL processes** that convert raw data into exploitable forms
- **Cleaning and transforming data** from various sources using PySpark and **SQL queries**
- Implemented caching and fine-tuned Spark configurations for optimal performance and scalability in the **data processing pipeline**.
- Applied a variety of data transformation techniques to the data, such as **filtering, aggregation, and summarizing**, in preparation for analysis and visualization

Technical Skills

Core Programming: Python, SQL,

Libraries & Framework: Pyspark, Pandas, NumPy, Matplotlib, Django

Frontend: HTML, CSS, JavaScript

Databases & Cloud Services: MySQL, SQL server, PostgreSQL, Oracle 11g, AWS (including Redshift, Cloud9, S3)

Development Tools: Git, GitHub, Anaconda, Jupyter Notebook, DataBricks, Google Colab, Cloud 9

Honours and Awards

Member of Computer Society of INDIA

march 2019 – march 2022

Batch of 2022

HIT

- Facilitated and organized a technical workshop.
- Participated as a member of the college fest Prayukti.
- Successfully managed all aspects of the workshop, including logistics, scheduling, and resource allocation.

Algorithmic Certification

- Workshop on **AWS, GIT and Networking**
- Workshop on **Open Source Software (Linux)**
- Certification of Python **Udemy**