

Data Engineering

Welcome to my Data Engineering Projects repository! This repository showcases a collection of projects I have completed in the field of data engineering. Each project is organized into its respective folder, and each folder contains a detailed **README.md** file explaining the project and its components.

Repository Structure

The repository is structured as follows:

- **MapReduce:** Contains projects related to MapReduce, demonstrating distributed data processing techniques using the MapReduce programming model.
- **Apache Spark:** Includes projects leveraging Apache Spark for large-scale data processing.
- **Apache Pulsar:** Include a project demonstrating message production and consumption using the Apache Pulsar messaging system.
- **RayTune ML:** Contains a project that leverages **RayTune** for distributed hyperparameter tuning.