E-Sports Tournament Pipeline

Abstract

Esports tournaments generate large volumes of structured data, including tournament details, participating teams, game titles, dates, and prize pools. Efficiently managing this data requires a structured data engineering approach. The Esports Tournament Pipeline project addresses this challenge by applying key data engineering principles to create a robust and scalable system that manages the end-to-end flow of tournament information.

The pipeline begins with the ingestion of raw tournament data from sources such as tournament databases and historical records. This data is processed through ETL (Extract, Transform, Load) techniques, where it is cleaned, validated, and transformed into structured formats suitable for analysis. By incorporating batch processing pipelines, the system ensures that historical tournament records are processed seamlessly, enabling timely insights and efficient data management throughout the tournament lifecycle.

The cleaned and processed data is stored in optimized storage solutions, where it becomes readily available for downstream applications such as leaderboard generation, performance analytics, and high-level tournament statistics. This allows organizers to accurately track team performance, compare tournament outcomes, and generate meaningful insights from historical trends.

By demonstrating practical applications of concepts such as data modeling, pipeline orchestration, and data quality assurance, the Esports Tournament Pipeline highlights the importance of systematic data handling in managing structured esports data. Ultimately, this project showcases how transforming raw tournament data into actionable insights can support data-driven decision-making and foster a professional and organized esports ecosystem.

- Shreya Avinash Phadke (2320030194)
- Dikshitha B (2320030212)
- Manikanta (2320030020)