

Implementing this project presented several challenges, particularly in handling the integration of two datasets. One was retrieved via an API call and another from a local file. One of the primary difficulties was ensuring consistency between the two sources, as discrepancies in formatting, column names, and missing values made the data difficult to work with. Managing the API documentation and handling potential failures in data retrieval added an extra layer of complexity.

Some aspects of the project were easier than expected. Basic data transformations, such as matching columns names, filtering relevant statistics, and normalizing data formats, were straightforward thanks to the robust functionality provided by pandas. Merging the datasets, once properly preprocessed, was more manageable than anticipated due to well-structured keys linking the two sources.

Certain aspects proved to be more challenging than expected. Ensuring data quality and consistency across different sources required significant attention to detail, especially when dealing with potential missing or inconsistent player statistics. Designing effective and meaningful analyses from the combined dataset required thoughtful consideration of key metrics and relationships.

This utility could be incredibly useful for future data projects by serving as a framework for integrating multiple data sources. The ability to merge API data with local datasets allows for dynamic updates and deeper insights, making it applicable to real-world scenarios such as sports analytics, financial modeling, or business intelligence. By refining the process of data retrieval, cleaning, and integration, this approach enhances efficiency and reproducibility for future projects.