6) Data Warehouse Implementation: After the E
successfully run, the data is ingested into Azure
as the Data Warehouse for the Soccer data. Azu
massively parallel processing (MPP) cloud-base
solution for large-scale data. It provides the fle
accommodate changing workloads, and its serv
the need for infrastructure management. Azure
built-in machine learning features and query op
processing. The data is ingested into Azure SQL
developed data warehouse model, including an
a group table as a dimension.
Below is a picture to show how the data looks i
insertion.
3 7) ETL Process (Extraction, Transformation, a
analyzing the soccer dataset in the OLTP and d
transformations to prepare the data for loading
(Data Warehouse) using Azure Data Factory, da
aggregation were needed. We combined the lea
table into a dimension table called DimLeague.
transformation was decided, ETL was performe
requirements to populate the dimensions and fa

urce:

ading strategy for dedi...

us/azure/synapse-analytics/s...

strategy for dedicated SQL p...

Performance

Text score: 95 out of 100. This score represents the quality of writing in this document. You can increase it by addressing Grammarly's suggestions.

95

Word count

Characters	25,604	Reading time	16 min 24 sec
Words	4,102	Speaking time	31 min 33 sec
Sentences	242		

Readability

Metrics compared to other Grammarly users

Word length	5	Above average
Sentence length	17	Above average
Readability score	47	

Your text compares in readability to The New York Times. It is likely to be understood by a reader who has at least a 10th-grade education (age 16).

DOWNLOAD PDF REPORT

Close

6) Data Warehouse Implementation: After the ETL pipelines have successfully run, the data is ingested into Azure SQL Pool which is used as the Data Warehouse for the Soccer data. Azure SQL Pool is a massively parallel processing (MPP) cloud-based data warehousing solution for large-scale data. It provides the flexibility of scale to accommodate changing workloads, and its serverless nature eliminates the need for infrastructure management. Azure SQL Pool also includes built-in machine learning features and query optimization for faster data processing. The data is ingested into Azure SQL Pool according to the developed data warehouse model, including an event table as a fact and a group table as a dimension.

Below is a picture to show how the data looks in Cosmos DB after insertion.

3 7) ETL Process (Extraction, Transformation, and Loading): After analyzing the soccer dataset in the OLTP and defining the necessary transformations to prepare the data for loading into the Azure SQL Pool (Data Warehouse) using Azure Data Factory, data enrichment, and data aggregation were needed. We combined the league and Country OLTP table into a dimension table called DimLeague. Once the data transformation was decided, ETL was performed based on the business requirements to populate the dimensions and facts.