

# Data Olympics: Go for Gold!

Gold is the epitome of sporting achievement. Sports people all over the world aspire to win a gold medal in the Olympics. When they first start their training, they will create a plan to set out short, medium and long-term goals. Short term goals might be to compete in a local event and finish in the top three. While a long-term goal might be to achieve gold in the Olympics. Regardless of where they are in their career, a significant amount of time is spent planning, training and racing in order to have the best possible chance of achieving their goals.

How does this relate to data? In a delta lakehouse environment, the terminology might be slightly different, but the process is very similar. Planning as we know, can be created by business analysts. This can set out what we aim to do. As an example, it might be to move all data used in a single report into a delta lakehouse. Races can be thought of as projects, while the goals can be related to each project just as the sports person will relate them to each race.

In moving data to a delta lakehouse, there will be three main goals and multiple minor goals. The main goals will be to achieve Bronze, Silver, and Gold in their entirety, as per the project plan. The minor goals might be to finish the migration of a single data source into the Bronze tier of the lakehouse. In an agile environment, this may take place in a single sprint. Just like most Gold medal winning Olympians, this isn't where the journey ends. New projects will come up, or new data will be discovered, and you will want to continue to "Go for Gold!".

A brief overview of the Bronze, Silver and Gold layer is below.

## Bronze Layer

Bronze can be classified as raw data with no modifications. There will be some additional meta data added, but the data should match up 100% to the source system, regardless of whether it is a CSV file, database table or any other source. The Bronze layer provides an historical archive which allows for reprocessing without needing to contact the source system again. End users should not have access to this layer. There are no controls or transformations that are happening at this layer.

## Silver Layer

The Silver layer is where data cleansing, deduplication and conformation happens. Data is brought together from multiple sources to form an overall enterprise view. One point to note, prior to the delta lakehouse, data was typically done in an ETL fashion. Extract, transform and load. Following the

delta lakehouse pattern, data is now performed as ELT. Extract, load and transform. This means that Silver should have minimal transformations applied to the data. Business logic and complex transformations live inside Gold.

## Gold Layer

The Gold layer has the final transformation and data quality rules applied. As Gold is used for end user reporting purposes, it uses denormalised and read efficient models. Kimball and Inmon are both excellent choices for this layer. Aggregate tables are also a great fit here and help to provide lightning quick reporting capabilities.

Many companies are currently "Going for Gold!". Data is now at the forefront of decision making. Would you like to get access to your data faster and with more confidence? Fujitsu Data & AI has many specialists that can help design, build and deploy your delta lakehouse.

To find out more, please contact our Fujitsu Data & AI specialist now.

## Contact

Fujitsu Data & AI  
+61 3 9924 3000

© Fujitsu 2022. All rights reserved. Fujitsu and Fujitsu logo are trademarks of Fujitsu Limited registered in many jurisdictions worldwide. Other product, service and company names mentioned herein may be trademarks of Fujitsu or other companies. This document is current as of the initial date of publication and subject to be changed by Fujitsu without notice. This material is provided for information purposes only and Fujitsu assumes no liability related to its use.