Bronze & Silver Guiding Documentation

Greetings adventurer,

So you seek to build a bronze layer with Azure Synapse Mapping Dataflows! Excellent choice! With the power of the Synapse Analytics workspace, you can create robust data solutions and achieve great success in your quest.

To begin, you will need to initialize the Delta Table in a spark notebook. This will allow you to create a table with built-in version control and other advanced features that will make managing your data much easier.

Once your Delta Table is set up, you can move on to creating your mapping data flows (<https://learn.microsoft.com/en-us/azure/data-factory/concepts-data-flow-overview>). You will need to use two sources (<https://learn.microsoft.com/en-us/azure/data-factory/data-flow-source>) for your data, which you can configure in the data flow settings. Be sure to select the appropriate connection type and credentials for your sources.

Next, you will want to add derived columns(<https://learn.microsoft.com/en-us/azure/data-factory/data-flow-derived-column>) to your data flow. This will allow you to transform your data and create new columns based on existing ones. You can use a variety of functions and expressions to accomplish this, depending on your specific needs. Make sure to have a look at the hash function which is a very valuable tool in your set.

To join your two sources, you can use the Join(<https://learn.microsoft.com/en-us/azure/data-factory/data-flow-join>) transformation. This will allow you to combine data from multiple sources based on a common key or set of keys. You can also choose the type of join you want to perform, such as inner join, left outer join, etc.

If you need to alter specific rows in your data, you can use the Alter Rows (<https://learn.microsoft.com/en-us/azure/data-factory/data-flow-alter-row>)transformation. This will allow you to apply specific changes to certain rows based on conditions you set. For example, you could use this to replace null values with a default value, or to apply a formula to specific rows.

To select specific columns from your data, you can use the Select (<https://learn.microsoft.com/en-us/azure/data-factory/data-flow-select>) transformation. This will allow you to choose which columns you want to keep in your data flow, and which ones you want to exclude. This can be useful for creating a more streamlined data set that is easier to work with.

Finally, you will need to specify a sink (<https://learn.microsoft.com/en-us/azure/data-factory/data-flow-sink>) for your data flow. This is where your data will be stored after it has been processed by your mapping data flow. You can choose from a variety of sink types, such as Azure Blob Storage, Azure Data Lake Storage, or Azure Synapse Analytics.

And there you have it, adventurer! With these tools at your disposal, you can create a powerful bronze layer that will help you achieve your data goals. Good luck on your journey!