**Analytical Pipeline**

**Pre-requisite:**

|  |  |  |  |
| --- | --- | --- | --- |
| No | Pre-requisite | Sample name of resource | Reference Link/compute details |
| 1 | Provision ADLS Gen 2 account and add container named “default” | <azsynapsesa> | <https://docs.microsoft.com/en-us/azure/synapse-analytics/get-started-create-workspace> |
| 2 | Provision Azure Synapse workspace | <azsynapsews1> |
| 3 | Provision SQL Pool | SQLPool1 | DW100c |
| 4 | Provision Spark Pool | SparkPool1 | Small Cluster/Medium Cluster |

**Note**: In the scripts, notebooks, wherever storage account, container or folder is referred, replace it with appropriate your storage account, container name and folder name.

**Setting up demo:**

|  |  |  |
| --- | --- | --- |
| No | Steps | Reference Link / Comment |
| 1 | Create containers named “rawdata” and “curateddata” in ADLS Gen 2 account |  |
| 2 | Create Folder “WWI” in rawdata container |  |
| 3 | Download compressed files from github link | <https://github.com/prlangad/AnalyticsPipeline/blob/master/Data/WWI.zip> |
| 4 | Extract WWI csv files from compressed file on github repo link and load it into “WWI” folder using Azure Storage Explorer |  |
| 5 | Download SQL scripts and Notebook from github repo link and import it in Azure Synapse Workspace | <https://github.com/prlangad/AnalyticsPipeline/tree/master/Develop> |
| 6 | Run SQL Script to load data into SQL Pool in the order as prefix.  After the execution, show case select top 100 \* from any of the table. | Refer SQL script:  00 CREATE WWI TABLES.sql |
| 7 | In script 01 COPY INTO WWI TABLES.sql, replace <ADLSGen2 storage account> with your storage account name.  On execution of SQL script, data will be loaded into SQL pool tables. | Refer SQL script:  01 COPY INTO WWI TABLES.sql |
| 8 | Showcase Synapse SQL pool Security features | Refer Script:  02 Set up RLS DDM  03 Example run for RLS DDM |
| 9 | Synapse Spark data analysis | Refer Notebook:  04 Analyze data with Scala |
| 10 | SQL on-demand data exploration | Refer Script:  05 Data Exploration with SQLOD |
| 11 | Download PowerBI Report from github repo link and publish it in your PowerBI Workspace | <https://github.com/prlangad/AnalyticsPipeline/tree/master/Develop>  Update connection string to your Synapse SQL end point. |
| 12 | Create a linked service to your PowerBI workspace in Synapse | Go to Manage Hub -> Linked Services -> Add -> Power BI -> <Select your workspace> |
| 13 | Showcase edit PowerBI report in Synapse workspace |  |