**Azure Project Doc –**

**Fetch the list of videos from YouTube channel (Attached dataset) with the highest views and likes to promote advertisements in the channel which has maximum traffic and present same in POWER BI.**

1. **Download the dataset from web**
2. **Create ADLS Gen2 storage account upload the data in container.**
3. **Create data bricks account and open the data bricks workspace and create cluster.**
4. **To access ADLS storage in data bricks generate SAS token in share access key section.**
5. **Once SAS token generated, create notebook in data bricks.**
6. **Configure the SAS token using python with container name and storage name.**
7. **Once token is generated you are now able to access your storage account from data bricks.**
8. **Open and read the file using pyspark command.**
9. **Transform the data for both files (**minimum views must be 89,00,000 and minimum likes must be 300).
10. **Save the transform data by new name in same storage with two different filename.**
11. **Once transformation data saved, open power bi to visual the dashboard**
12. **Schedule the transformation for the two datasets and automate the generation of the list on a 5-minute basis.**