

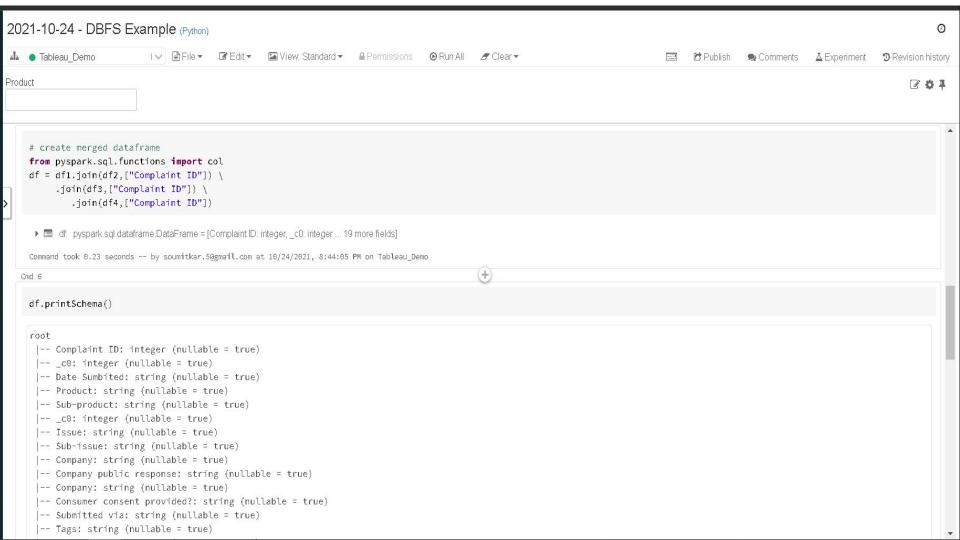
```
2021-10-24 - DBFS Example (Python)
       ♣ Tableau Demo
                                   I ✓ File ▼

☑ Edit ▼

                                                          Mary View: Standard ▼
                                                                            ♠ Permissions

    Clear ▼

                                                                                          Run All
D •
      Product
0
Cmd 2
(1)
           # Mount from Azure Blob
           dbutils.fs.mount(
Q
             source = "wasbs://<container-name>@<storage-account-name>.blob.core.windows.net",
             mount_point = "/mnt/<mount-name>",
             extra_configs = {"<conf-key>":dbutils.secrets.get(scope = "<scope-name>", key = "<key-name>")})
숆
£
           ∃java.lang.IllegalStateException: Secrets API is not enabled for this workspace.
           Command took 3.54 seconds -- by soumitkar.5@gmail.com at 10/24/2021, 8:42:57 PM on Tableau Demo
         Cmd 3
           # File location and type
           file_location1 = "/mnt/<mount-file/Financial_Complaints_Product.csv"</pre>
           file location2 = "/mnt/<mount-file/Financial Complaints Issue.csv"
           file_location3 = "/mnt/<mount-file/Financial_Complaints_Response.csv"</pre>
           file location4 = "/mnt/<mount-file/Financial Complaints Region.csv"</pre>
           file_type = "csv"
@
           # The applied options are for CSV files. For other file types, these will be ignored.
63
           df1 = spark.read.csv(file_location1, header=True, inferSchema=True)
           df2 = spark.read.csv(file_location2, header=True, inferSchema=True)
           df3 = spark.read.csv(file_location3, header=True, inferSchema=True)
           df4 = spark.read.csv(file_location4, header=True, inferSchema=True)
▶ (8) Spark Johe
```



```
2021-10-24 - DBFS Example (Python)
       ₼ Tableau Demo

▲ Experime

                                                                                                                                          Publish ...
                                                                                                                                                     Comments
D •
      Product
0
#conver string to date format
          from pyspark.sql.functions import unix_timestamp, from_unixtime
(1)
          df = df.select(
               'Date Sumbited',
Q
               from_unixtime(unix_timestamp('Date Sumbited', 'MM/dd/yyy')).alias('Date_Sumbited'), 'Date Received',
               from_unixtime(unix_timestamp('Date Received', 'MM/dd/yyy')).alias('Date_Received'))
솖
           ▶ ■ df. pyspark.sql.dataframe.DataFrame = [Date Sumbited: string, Date Sumbited: string ... 2 more fields]
3
          Command took 0.14 seconds -- by soumitkar.5@gmail.com at 10/24/2021, 10:19:37 PM on Tableau_Demo
         Cmd 8
          #selected columns
          column_list = ['Complaint ID', 'Date_Sumbited', 'Product', 'Sub-product', 'Issue',
                  'Sub-issue', 'Company public response', 'State', 'ZIP code',
                  'Tags', 'Consumer consent provided?', 'Submitted via', 'Date_Received',
                  'Company response to consumer', 'Timely response?',
                  'Consumer disputed?']
          df = df.select([column for column in df.columns if column in column_list])
?
            ▶ ■ df: pyspark.sql.dataframe.DataFrame = [Date Sumbited: string, Date Received: string]
(6)
          Command took 0.10 seconds -- by soumitkar.5@gmail.com at 10/24/2021, 10:20:34 PM on Tableau_Demo
         Cmd 9
8
           #Load to ADLS
df.write.csv("/mnt/<mount-name>"/Trans_Financial_Complaints, header = True, mode='overwrite')
```