## 

```
(https://databricks.com)
                                                           1
     storage_account_name = "ncplstorageact"
     container_name = "silver"
     storage_account_name = "ncplstorageact"
     container_name = "gold"
                                                           3
     # Define your storage account and container details
     storage_account_name = "ncplstorageact"
     container_name = "silver"
     mount_point = f"/mnt/{container_name}"
     # Check if the container is already mounted
     if not any(mount.mountPoint == mount_point for mount in dbutils.fs.mounts()):
         dbutils.fs.mount(
             source=f``wasbs://\{container\_name\}@\{storage\_account\_name\}.blob.core.windows.net",
             mount_point=mount_point,
             extra_configs={
                 \verb|f"fs.azure.account.key.{storage\_account\_name}.blob.core.windows.net": dbutils.secrets.get|\\
                 ('testScope', 'secret-ncplstorageact')
             }
         )
     # Mount the Silver container
     mount_point = "/mnt/silver"
                                                           5
     # Mount the Silver container
     mount_point = "/mnt/gold"
                                                           6
     # Read the data from the silver container
     accounts_df = spark.read.format("delta").load("/mnt/silver/accounts")
     customers_df = spark.read.format("delta").load("/mnt/silver/customers")
     loan_payments_df = spark.read.format("delta").load("/mnt/silver/loan_payments")
     loans_df = spark.read.format("delta").load("/mnt/silver/loans")
     transactions_df = spark.read.format("delta").load("/mnt/silver/transactions")
```

""" Write a query to calculate the total balance across all accounts for each customer, ensuring that all columns from the accounts and customers tables are selected and included """

```
8
from pyspark.sql import functions as F
                                                    9
# Join customers_df and accounts_df on customer_id
joined_df = customers_df.join(accounts_df, on="customer_id", how="inner")
                                                    10
# Group by customer_id, first_name, last_name, etc., and calculate total balance
result_df = joined_df.groupBy("customer_id", "first_name", "last_name", "address", "city", "state", "zip") \
    .agg(F.sum("balance").alias("total_balance"))
                                                    11
# Add total_balance column to the original joined data
final_df = joined_df.join(result_df.select("customer_id", "total_balance"), on="customer_id", how="inner")
                                                    12
final_df.show()
```

```
only showing top 20 rows
                                                   13
   # Write the DataFrame to the gold container as a Delta table
   final_df.write.format("delta").mode("overwrite").save("/mnt/gold/customer_total_balance")
                                                    14
   # Read back the data to validate
   validated_df = spark.read.format("delta").load("/mnt/gold/customer_total_balance")
   validated_df.show()
5|
       1300.25
         41
                Matthew|
                                    4040 Ash Blvd
                                                     Georgetown|
                                                                  ON|L7G0A1|
                                                                                   34
                                                                                         Checking | 3500.
                           Cooper|
        3500.5
5
         52
                Abigail| Henderson|5151 Cypress Ave| Mount Albert|
                                                                  ON|L0G0A1|
                                                                                   61
                                                                                          Savings | 500.2
        500.25
5|
         42| Charlotte|Richardson| 4141 Beech Dr|
                                                      Newmarket
                                                                  ON|L3Y0A1|
                                                                                         Checking | 5500.
                                                                                   54
5|
        5500.5
          4|
                 Emily|
                            Davis|
                                     101 Pine Rd
                                                        Calgary|
                                                                  AB | T2A0A1 |
                                                                                   78
                                                                                         Checking | 7900.
5|
        7900.5
         69
                 Joseph|
                            Diaz | 6868 Ash Blvd | Port McNicoll |
                                                                  ON|L0K0A1|
                                                                                   65
                                                                                          Savings | 550.2
5
        550.25
         60
                Olivia|Washington|
                                      5959 Oak Dr
                                                      Tottenham
                                                                  ON|L0G0A1|
                                                                                   95|
                                                                                          Savings | 925.7
5
        925.75
                 David| Sanchez| 3030 Maple Ave|
                                                                                         Checking | 5100.
                                                                  ON P1B0A1
                                                                                   50 l
         31
                                                      North Bay
        5100.5
5
                 Andrew | Mitchell | 2020 Spruce Ln
                                                                  ON|L8P0A1|
                                                                                   20
                                                                                         Checking | 2000.
         21
                                                      Hamilton|
0
       10700.5
                         only showing top 20 rows
                                                    15
   # List files in the gold container to confirm the mount and the data path
   dbutils.fs.ls("/mnt/gold")
[FileInfo(path='dbfs:/mnt/gold/customer_total_balance/', name='customer_total_balance/', size=0, modificationTime=
1731300279000)]
                                                   16
```