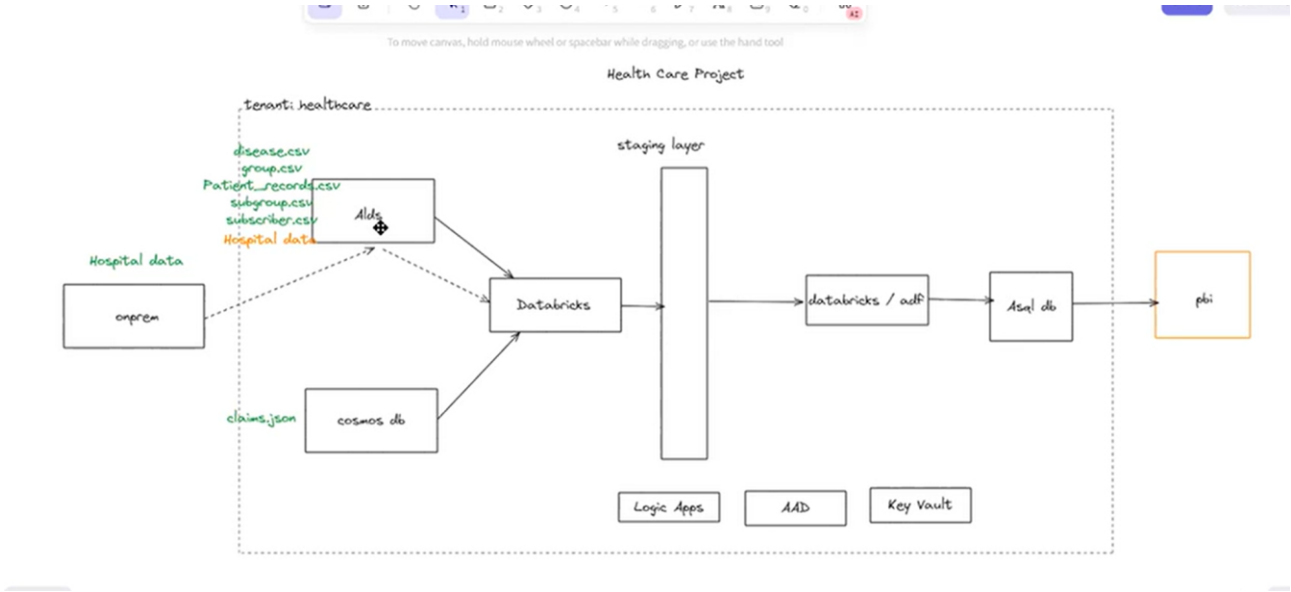
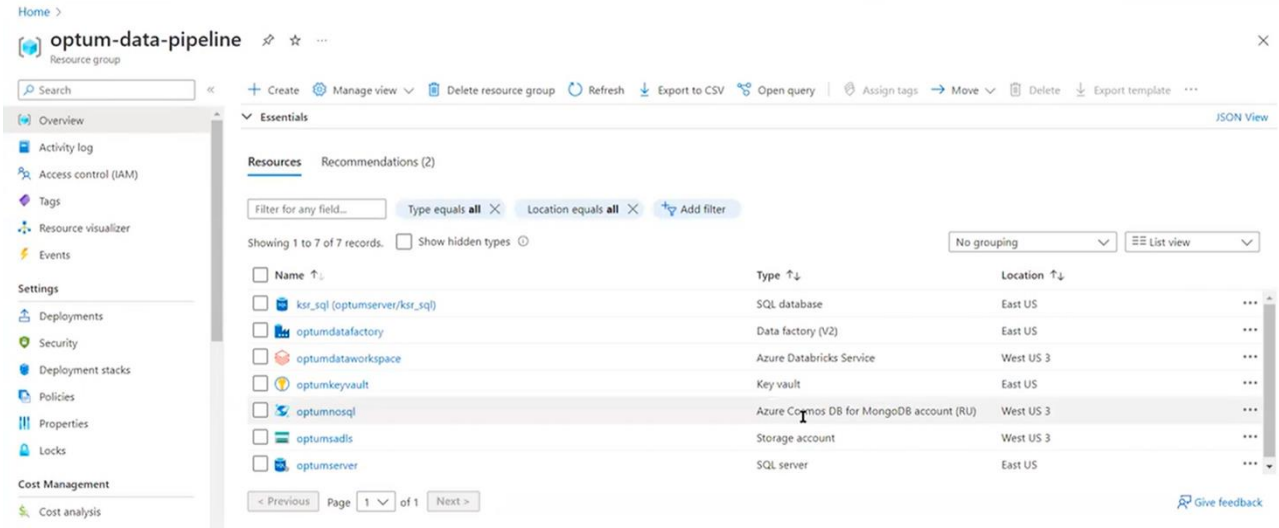
HEALTH CARE PROECT USIJNG AZURE DATA BRICKS

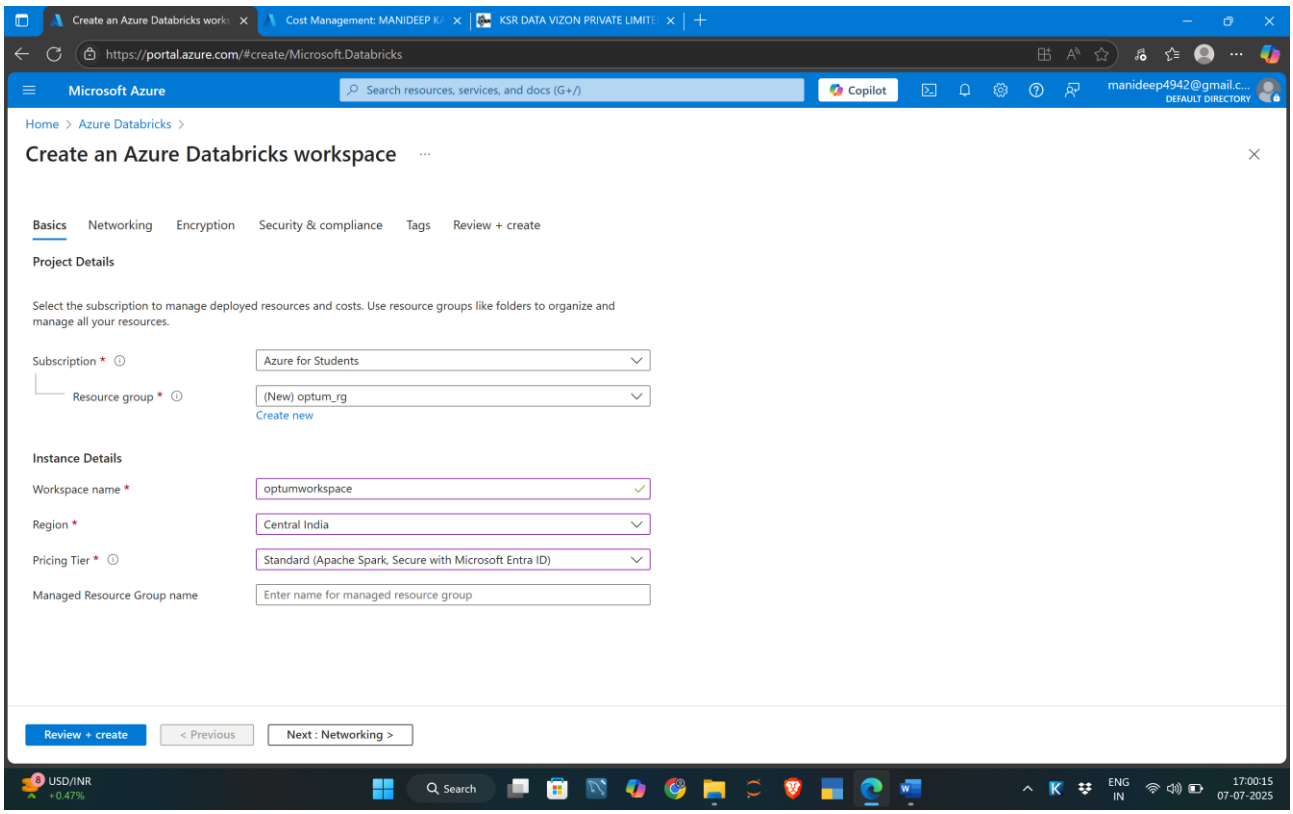
NEED SERVICES

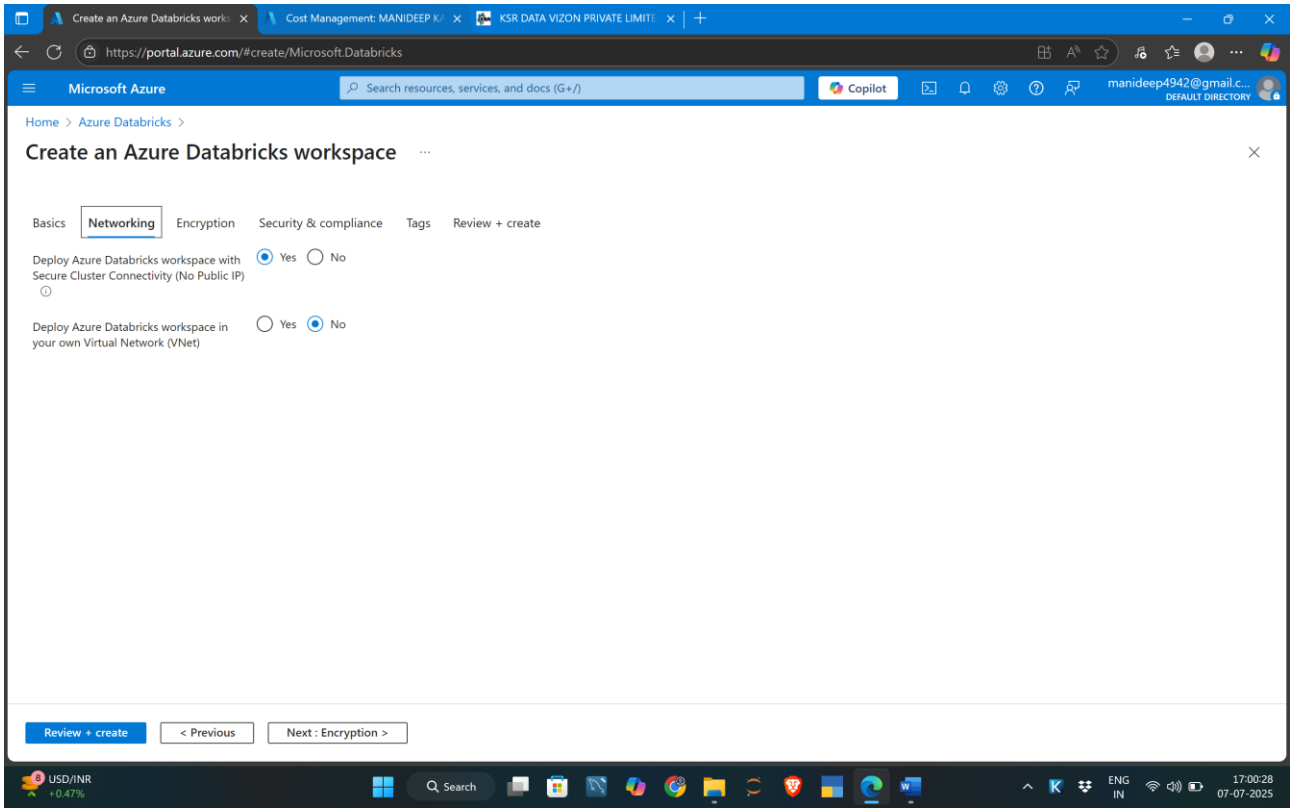
1. ADLS STORAGE
2. ADF
3. DATABRICKS
4. AZURE SQL SERVER
5. COSMOS DB
6. ON MYSQL WITH INTREGRATION RUN TIME
7. KEY VAULT AND IF NEEDED LOGIC APPS ALSO

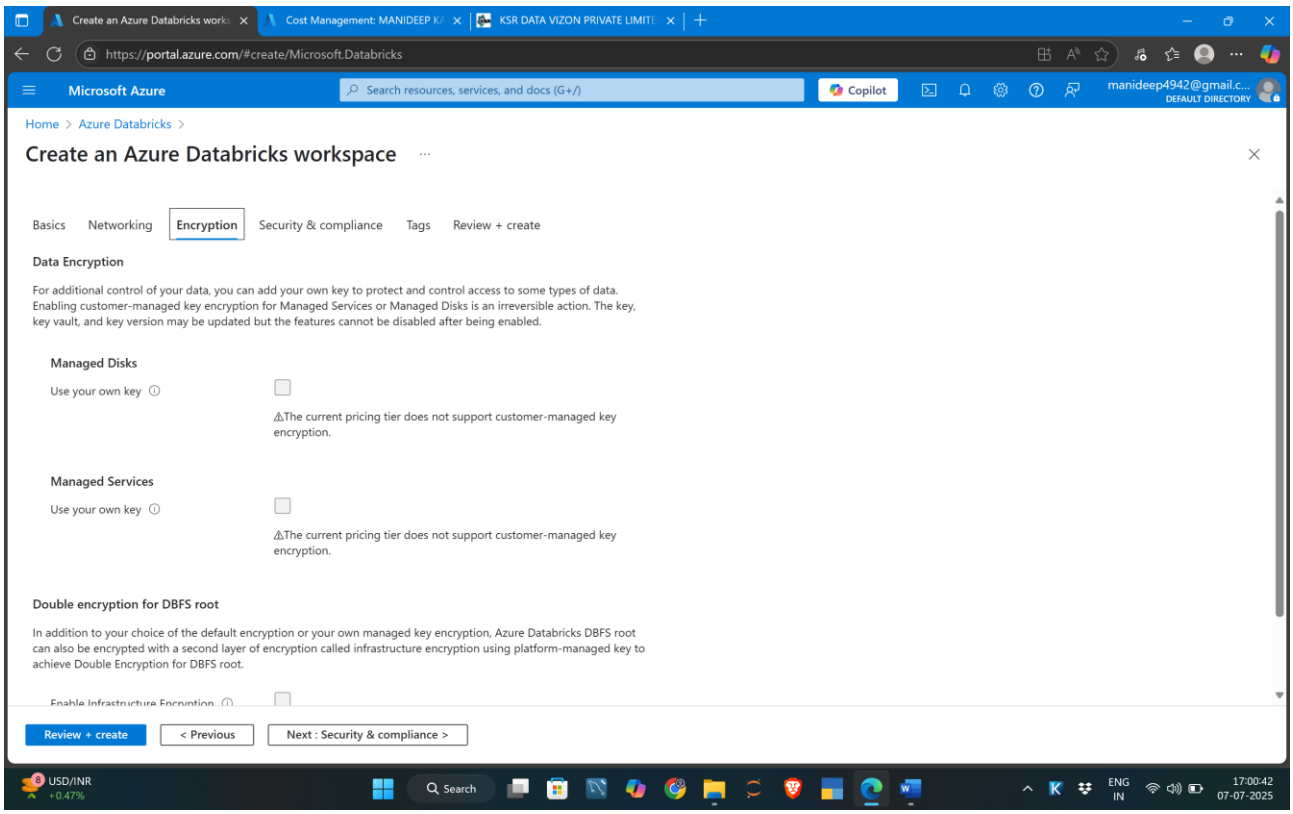
WE NEED TO CREATE THESR SERVICE SO ALEADY WE KNOW THE HOW TO CREATE SOME SERIVCES BUT NOW I AM CREATIGN DATA BRICKS SERVICES.

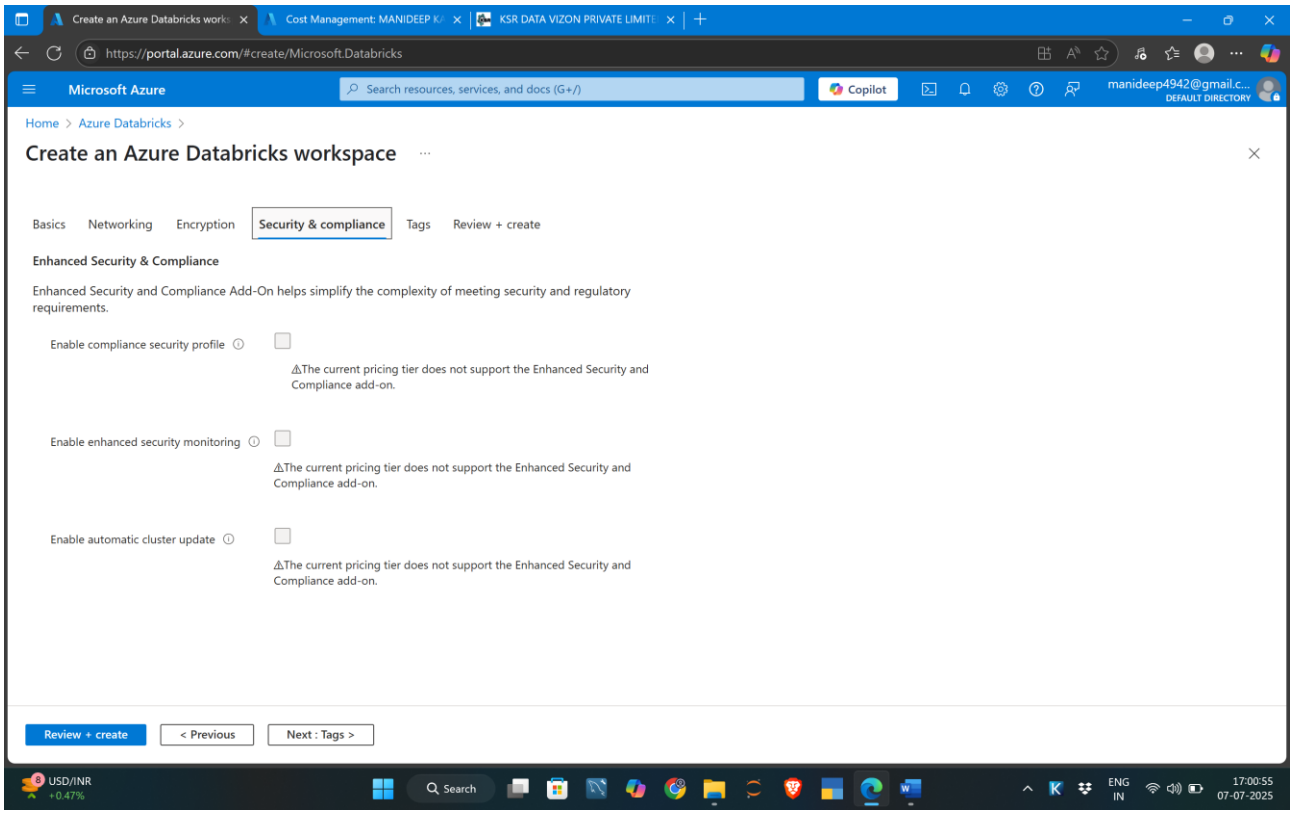


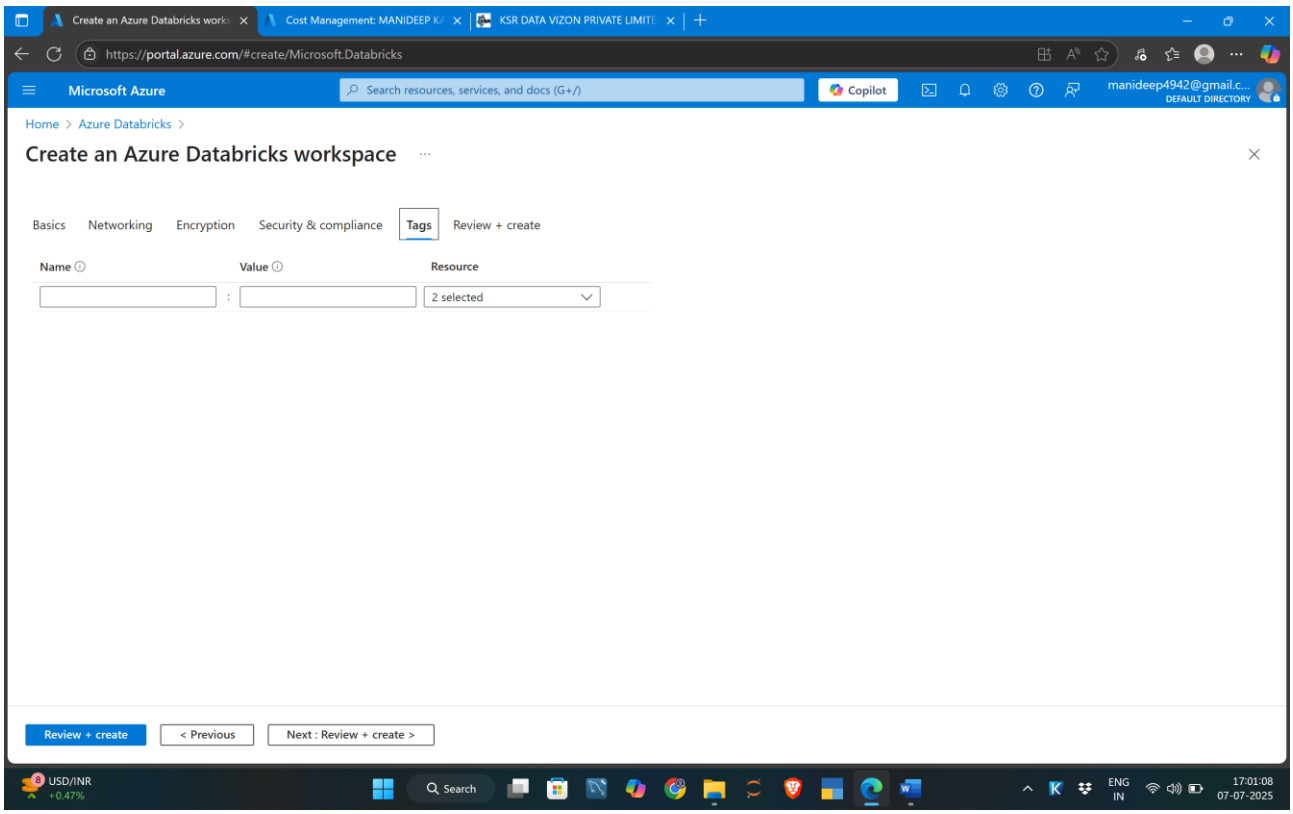


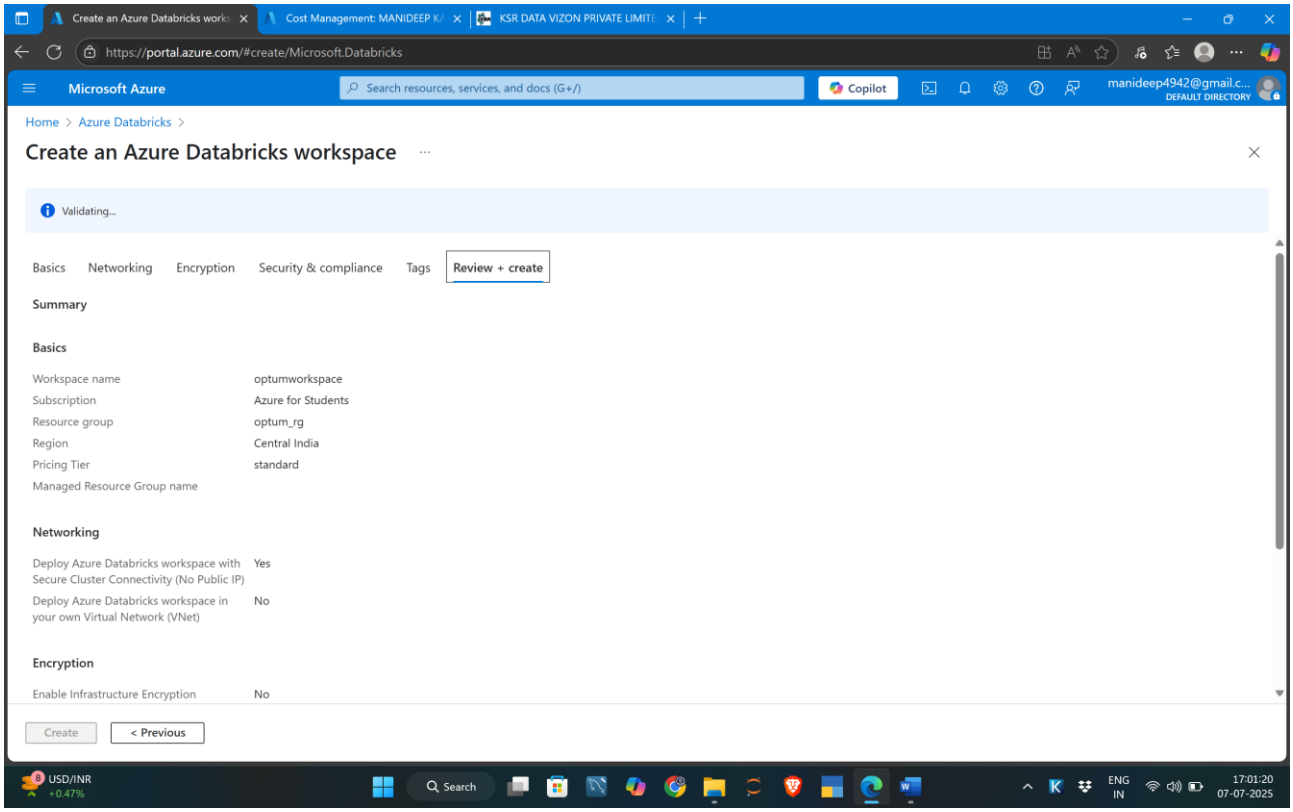

Create an Azure Databricks works
Cost Management: MANIDEEP K
KSR DATA VIZON PRIVATE LIMITE
https://portal.azure.com/#create/Microsoft.Databricks
Microsoft Azure
Search resources, services, and docs (G+/)
Copilot
manideep4942@gmail.c...
DEFAULT DIRECTORY
Home > Azure Databricks >
Create an Azure Databricks workspace
Basics Networking Encryption Security & compliance Tags Review + create
Project Details
Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.
Subscription *
Azure for Students
Resource group *
(New) optum_rg
Create new
Instance Details
Workspace name *
optumworkspace
Region *
Central India
Pricing Tier *
Standard (Apache Spark, Secure with Microsoft Entra ID)
Managed Resource Group name
Enter name for managed resource group
Review + create
< Previous
Next : Networking >
USD/INR
+0.47%
Search
ENG
IN
17:00:15
07-07-2025


Create an Azure Databricks works
Cost Management: MANIDEEP K
KSR DATA VIZON PRIVATE LIMITE
https://portal.azure.com/#create/Microsoft.Databricks
Microsoft Azure
Search resources, services, and docs (G+/)
Copilot
manideep4942@gmail.c...
DEFAULT DIRECTORY
Home > Azure Databricks >
Create an Azure Databricks workspace
Basics Networking Encryption Security & compliance Tags Review + create
Deploy Azure Databricks workspace with Secure Cluster Connectivity (No Public IP)
Yes
No
Deploy Azure Databricks workspace in your own Virtual Network (VNet)
Yes
No
Review + create
< Previous
Next : Encryption >
USD/INR
+0.47%
Search
ENG
IN
17:00:28
07-07-2025

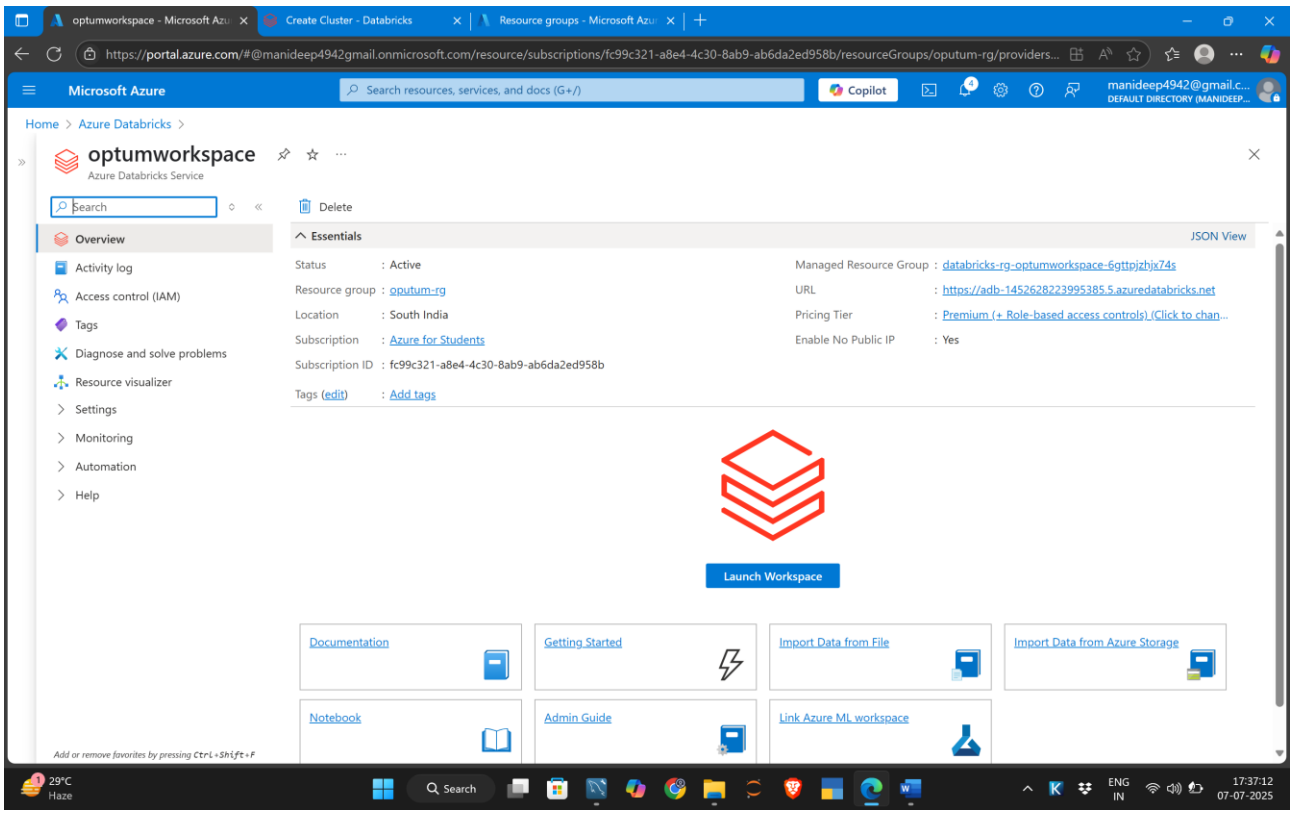

Create an Azure Databricks works
Cost Management: MANIDEEP K/
KSR DATA VIZON PRIVATE LIMITE
https://portal.azure.com/#create/Microsoft.Databricks
Microsoft Azure
Search resources, services, and docs (G+/)
Copilot
manideep4942@gmail.c...
DEFAULT DIRECTORY
Home > Azure Databricks >
Create an Azure Databricks workspace
Basics Networking Encryption Security & compliance Tags Review + create
Data Encryption
For additional control of your data, you can add your own key to protect and control access to some types of data. Enabling customer-managed key encryption for Managed Services or Managed Disks is an irreversible action. The key, key vault, and key version may be updated but the features cannot be disabled after being enabled.
Managed Disks
Use your own key
⚠The current pricing tier does not support customer-managed key encryption.
Managed Services
Use your own key
⚠The current pricing tier does not support customer-managed key encryption.
Double encryption for DBFS root
In addition to your choice of the default encryption or your own managed key encryption, Azure Databricks DBFS root can also be encrypted with a second layer of encryption called infrastructure encryption using platform-managed key to achieve Double Encryption for DBFS root.
Review + create
< Previous
Next : Security & compliance >
USD/INR
+0.47%
Search
ENG
IN
17:00:42
07-07-2025


Create an Azure Databricks works
Cost Management: MANIDEEP K/
KSR DATA VIZON PRIVATE LIMITE
https://portal.azure.com/#create/Microsoft.Databricks
Microsoft Azure
Search resources, services, and docs (G+/)
Copilot
manideep4942@gmail.c...
DEFAULT DIRECTORY
Home > Azure Databricks >
Create an Azure Databricks workspace
Basics Networking Encryption Security & compliance Tags Review + create
Enhanced Security & Compliance
Enhanced Security and Compliance Add-On helps simplify the complexity of meeting security and regulatory requirements.
Enable compliance security profile
⚠The current pricing tier does not support the Enhanced Security and Compliance add-on.
Enable enhanced security monitoring
⚠The current pricing tier does not support the Enhanced Security and Compliance add-on.
Enable automatic cluster update
⚠The current pricing tier does not support the Enhanced Security and Compliance add-on.
Review + create
< Previous
Next : Tags >
USD/INR
+0.47%
Search
ENG
IN
17:00:55
07-07-2025



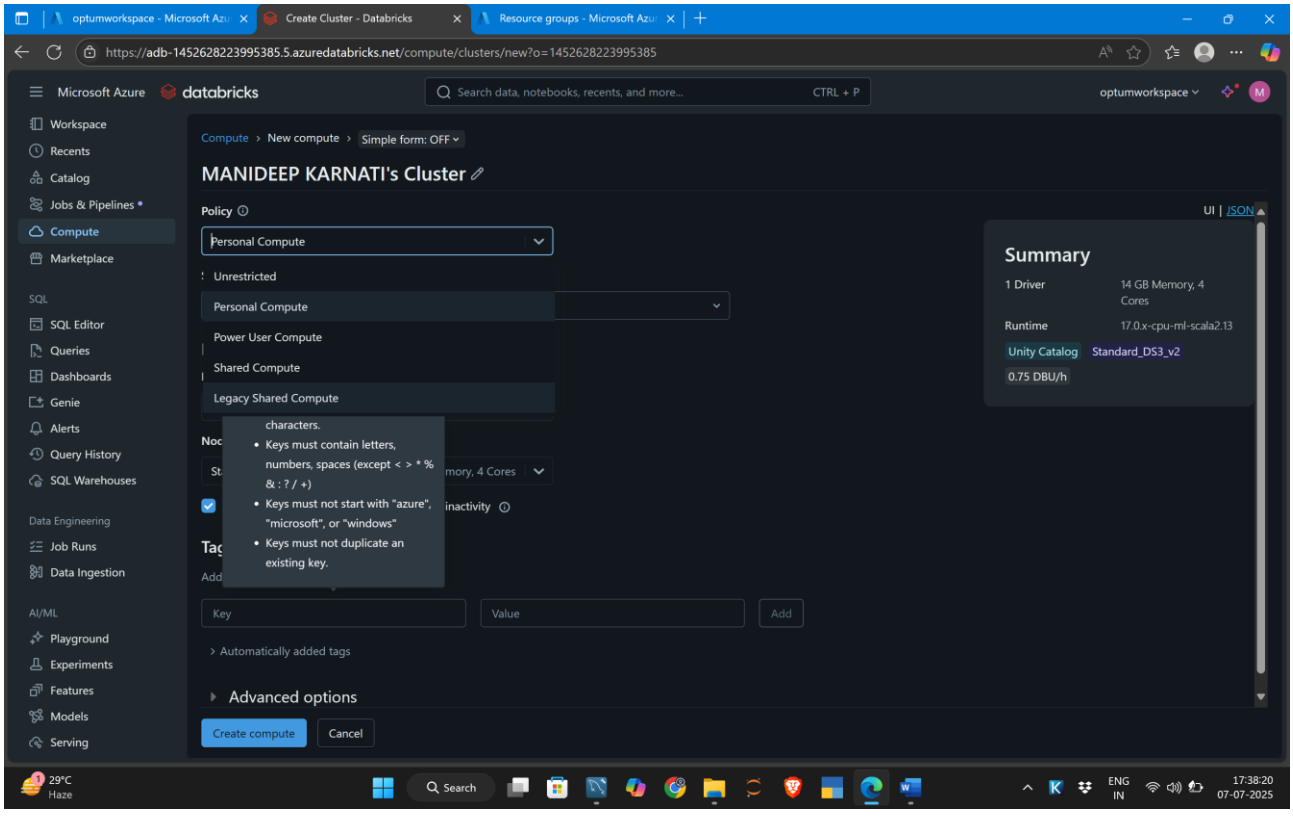


Here it is now we created



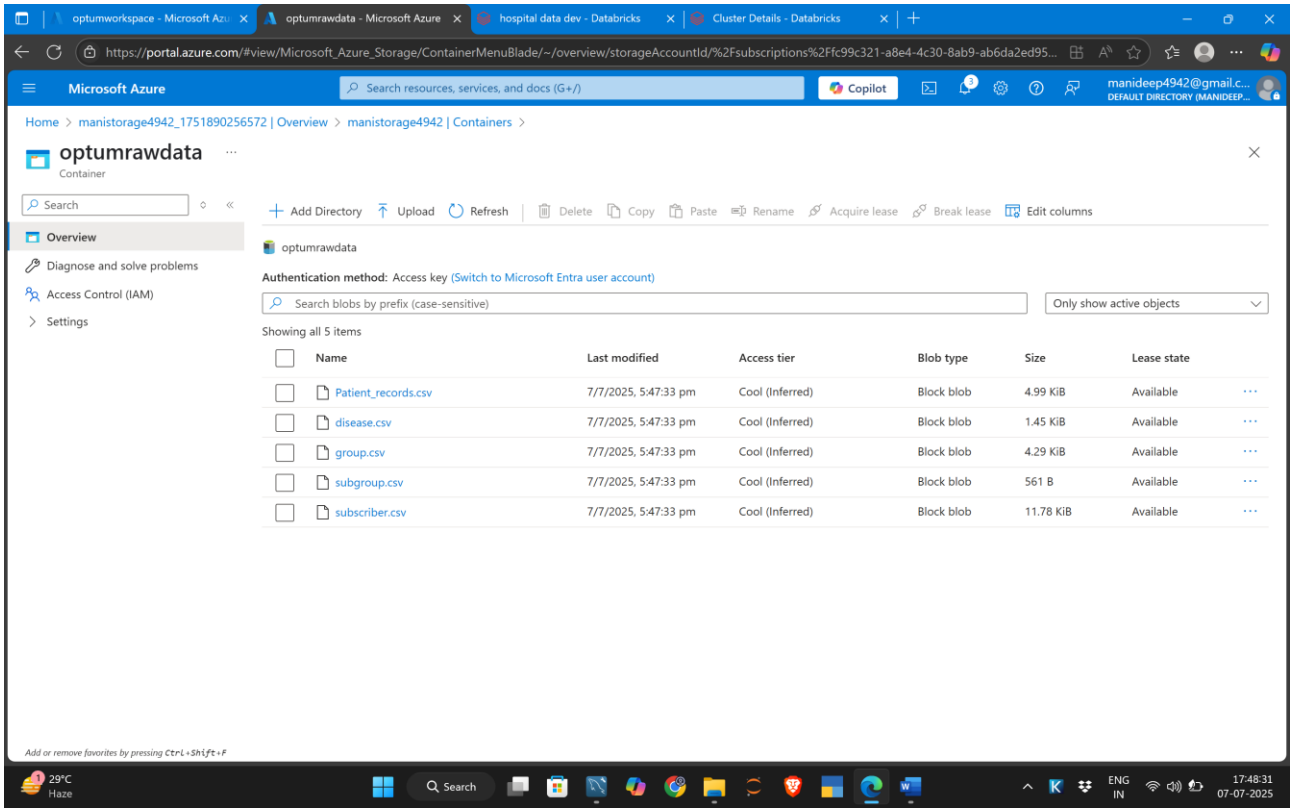
Now create a compute engine to run the queries .

Select the personal computer



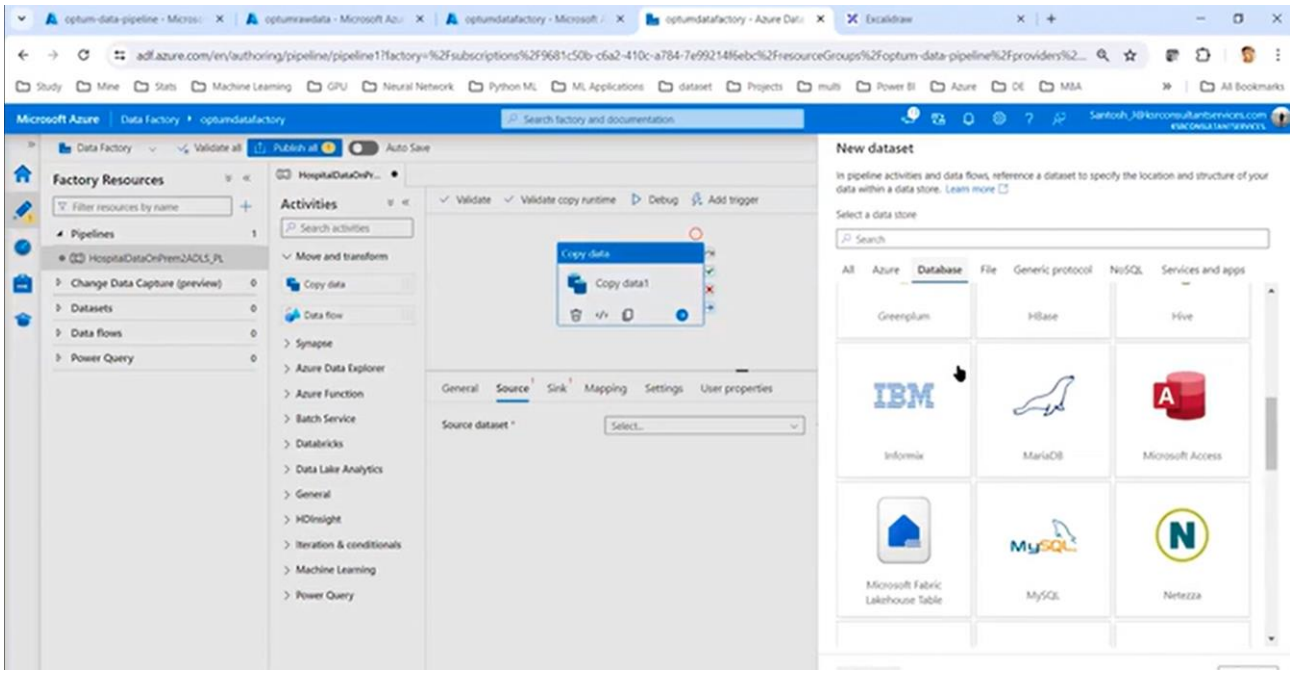
And create the cluster for the notebook

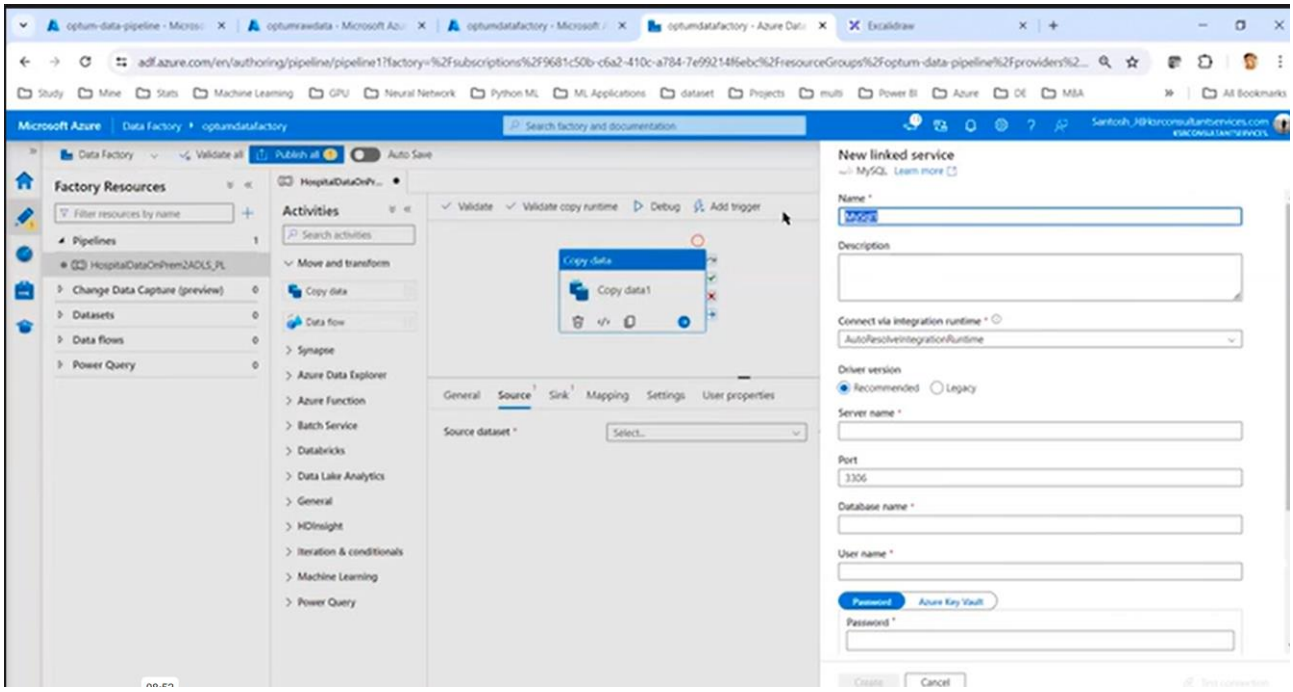
And not create the adls storage and upload these files into this

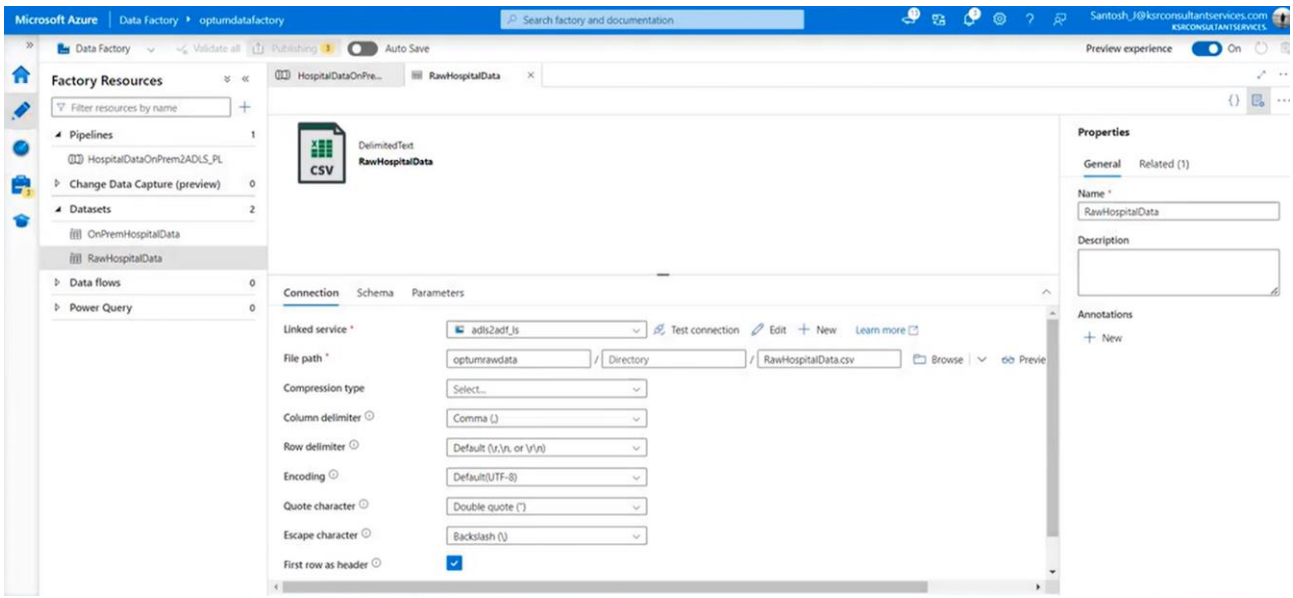


IN THIS WE USE THE DATA BRICKS BECAUSE FOR COPY ACTIVITY , WE CAN ALSO DO FROM DATABRICKS ALSO BUT IT WILL CHARGE LIKE ANY THING .

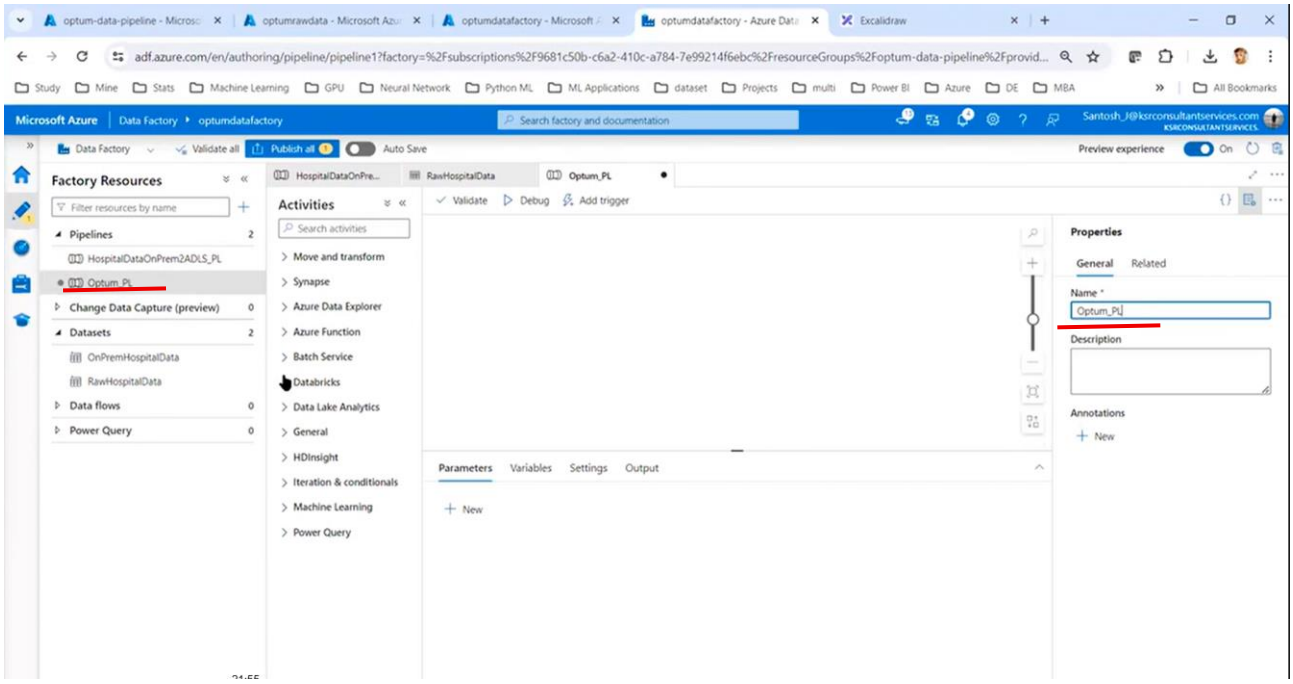
THE FIRST PIPELINE WE NEED TO CREATE (ON-PREM TO ADLS)(HOSPITAL DATASET)



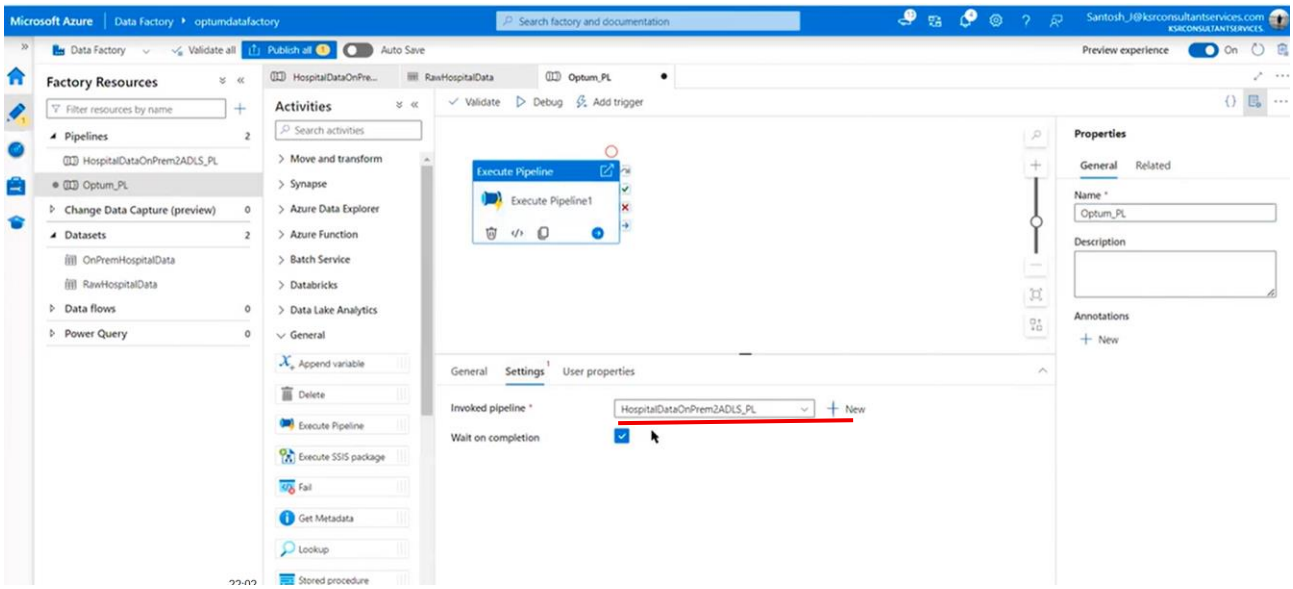




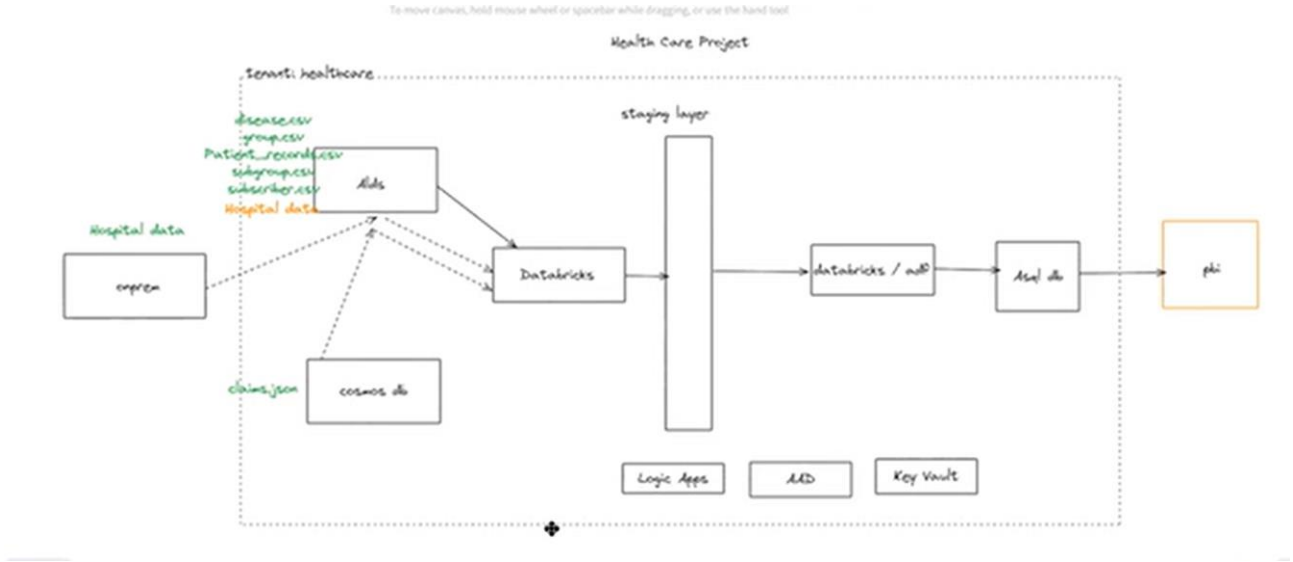
CREATE THE INPUT AND OUTPUT DATA SET TO GET DATA FROM ONPRE TO BLOB.



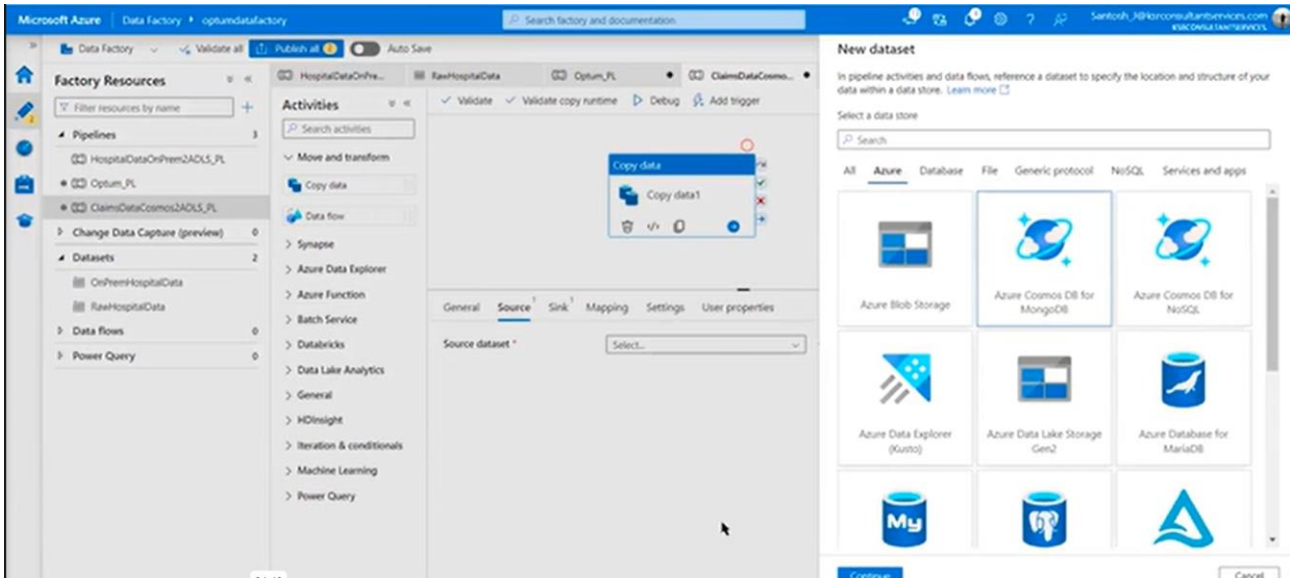
CREATE ONE MAIN PIPELINE

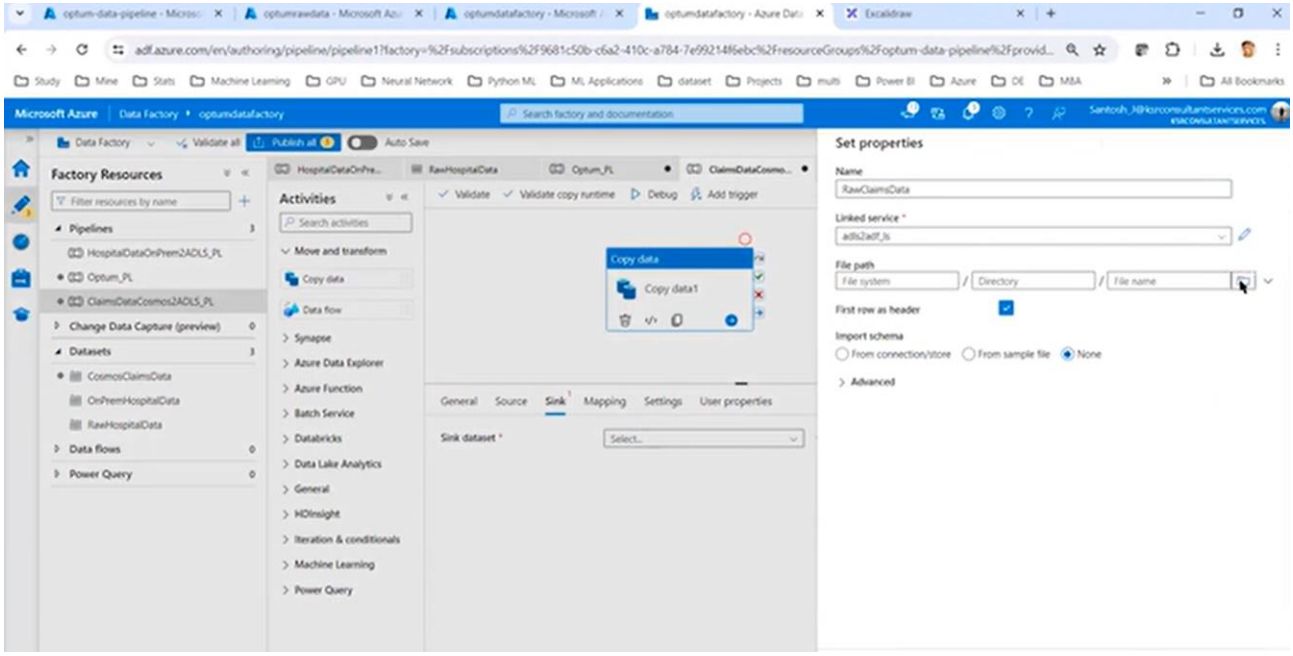


SIMILARY JSON ALSO I LL NEED TO BRING TO THE ADLS

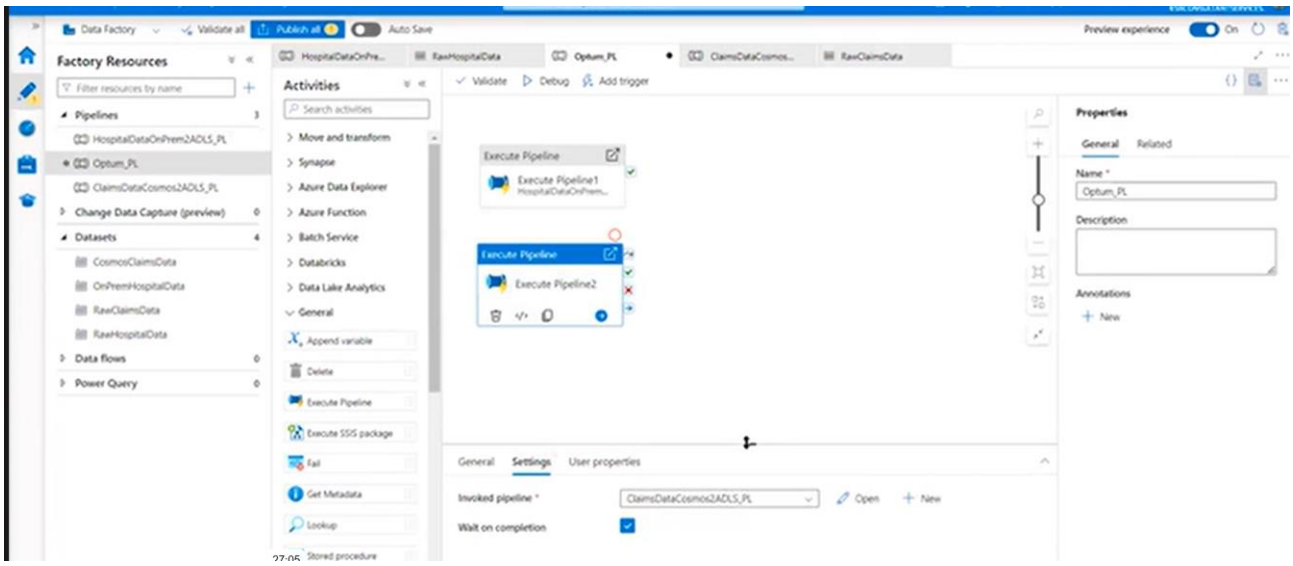


CREATE ANOTHER PIPELINE FOR JSON(CLIMS DATA)



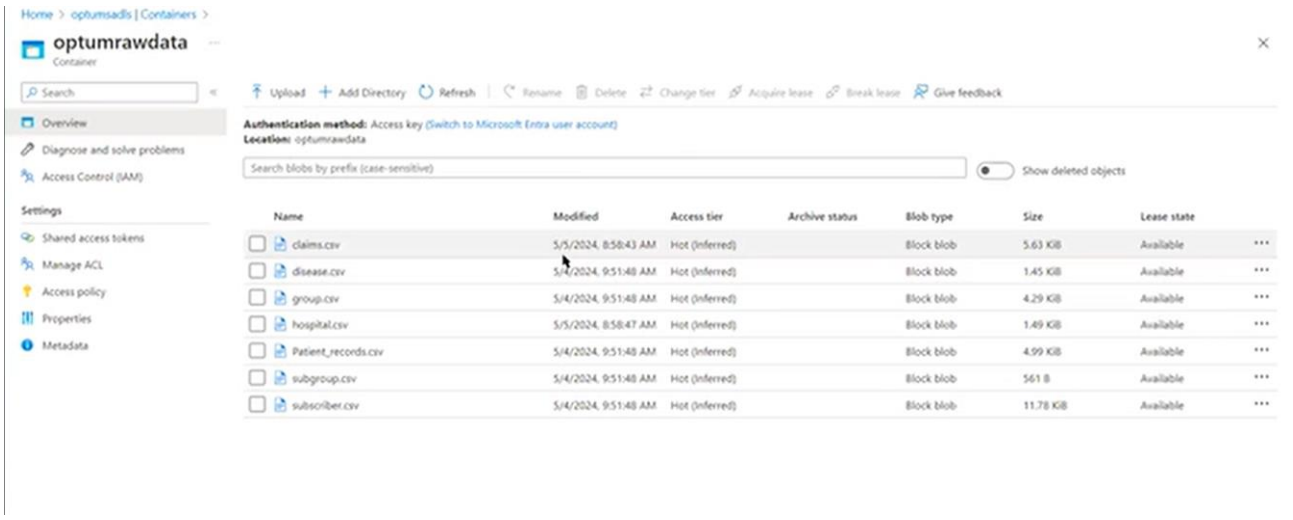


SINK THE DATA IN CSV FORMAT.

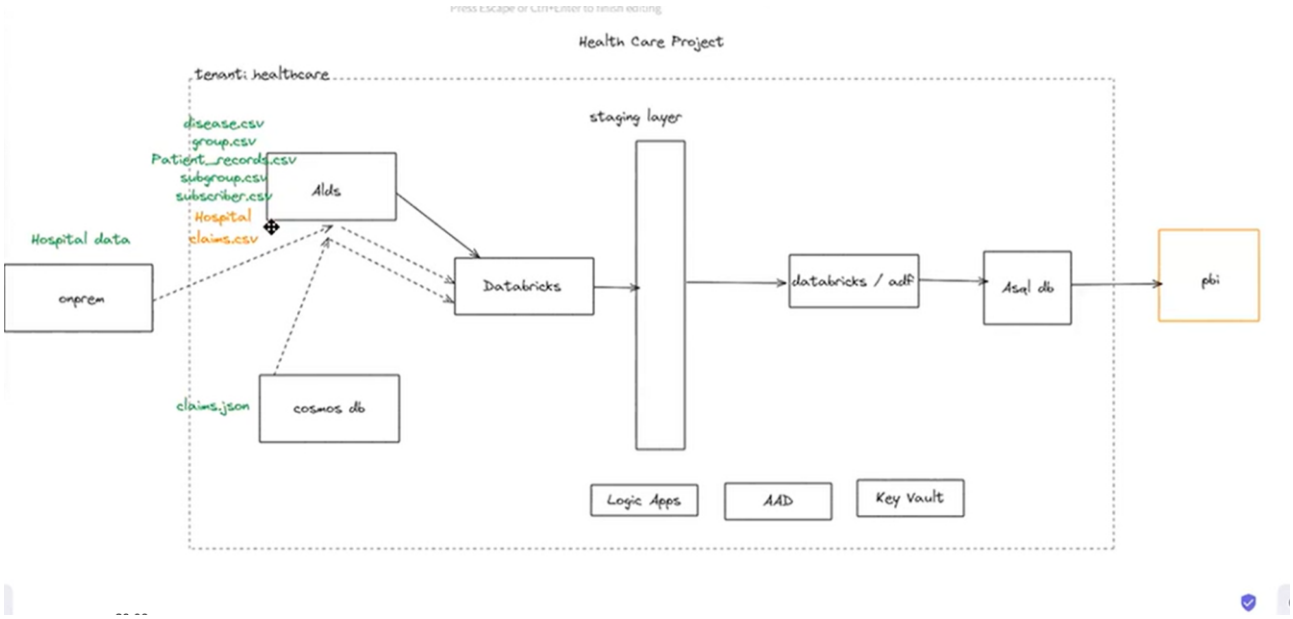


AGAIN ADD TO THE ONPREM_PL TO EXECUTE PARALLEL ACTIVITYS

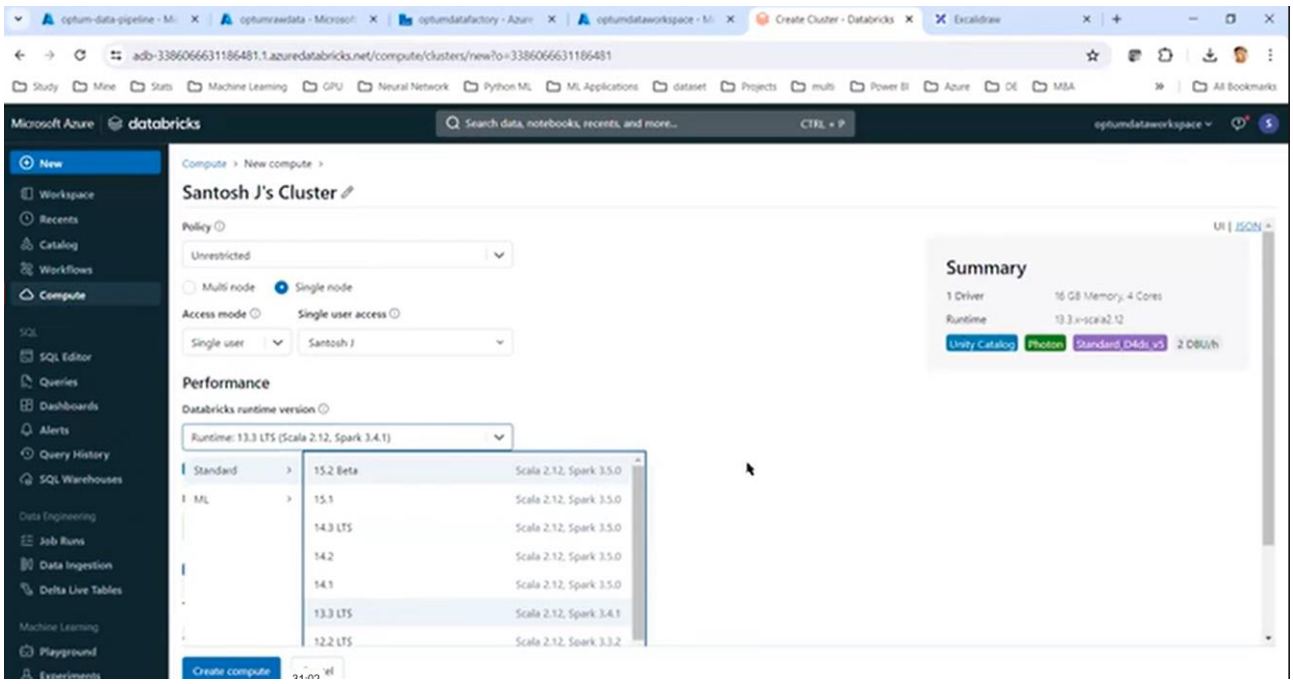
NOW GO AND TRIGGER IT



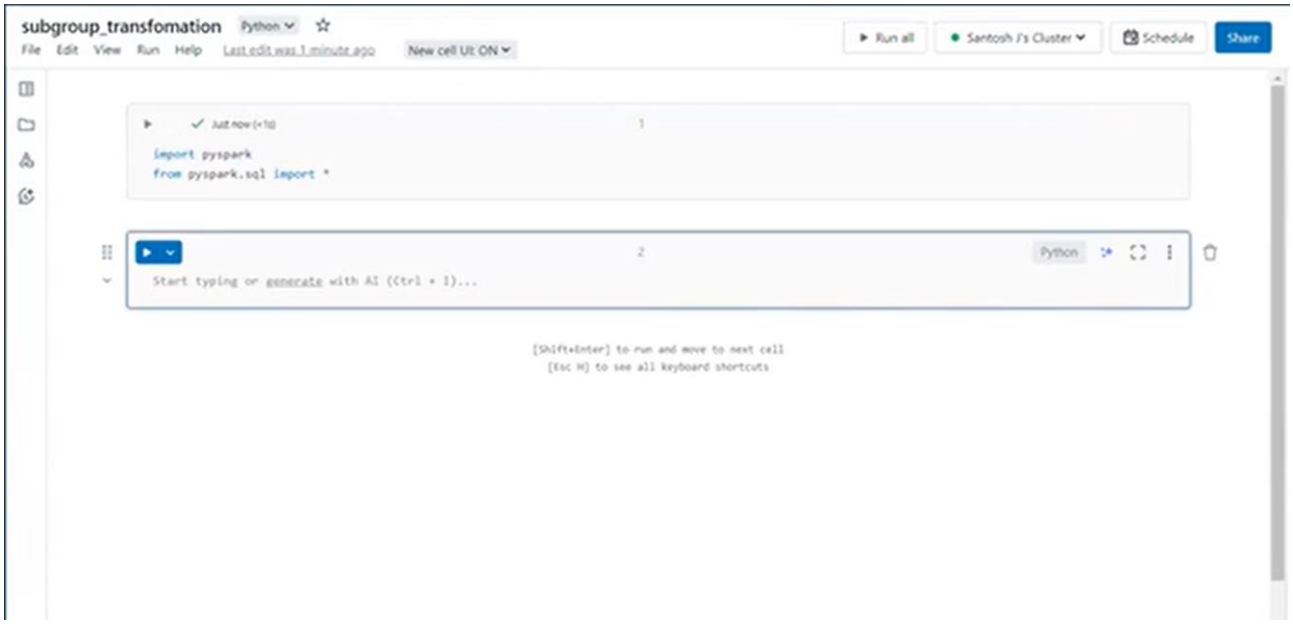
NOW ALL THE DATA AS BEEN CAME INTO ADLS



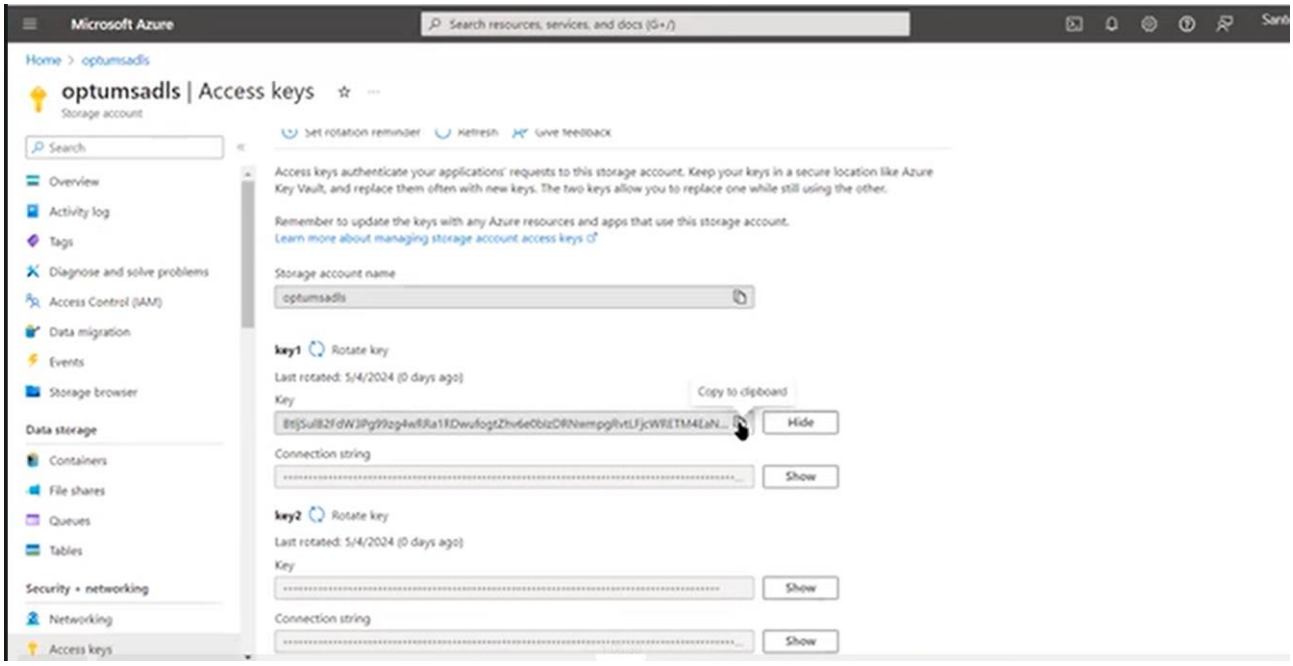
NOW LETS GO THE DATABRICKS----------→



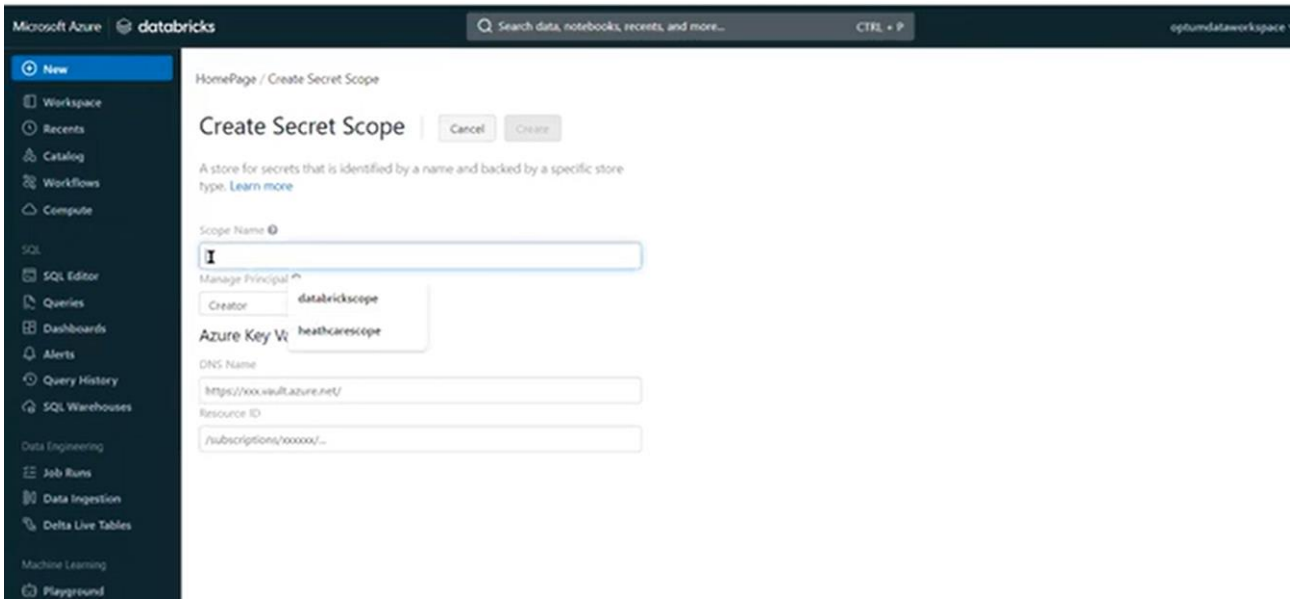
CREATE CLUSTER

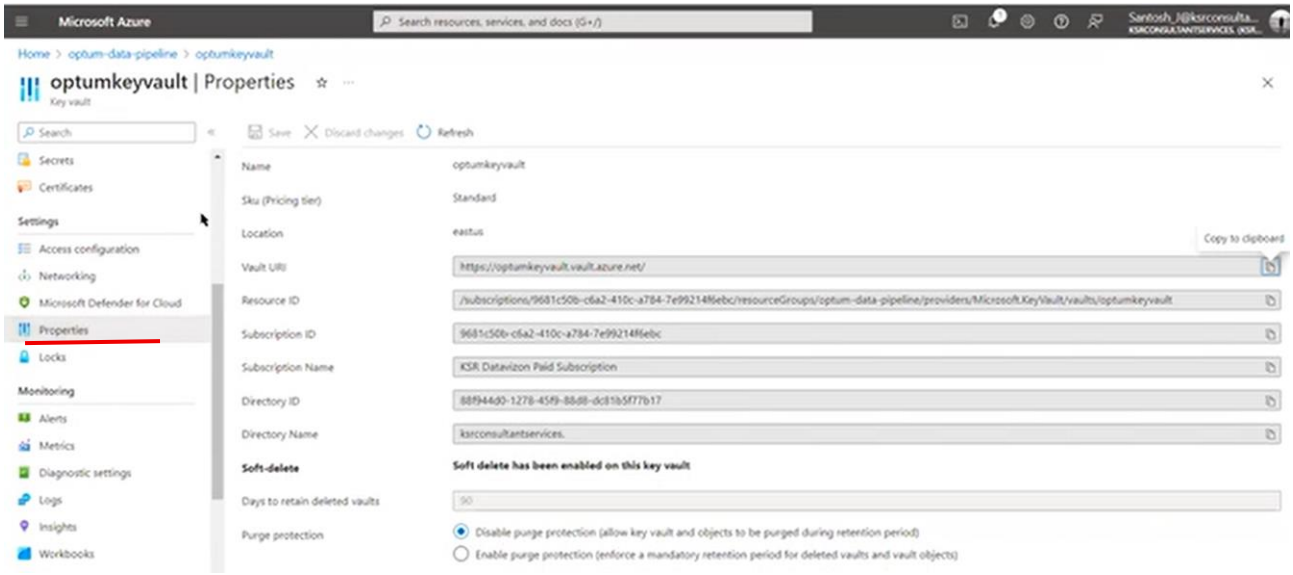


NEXT WE NEED TO DO CREATE A KEY VALUATE FOR DB



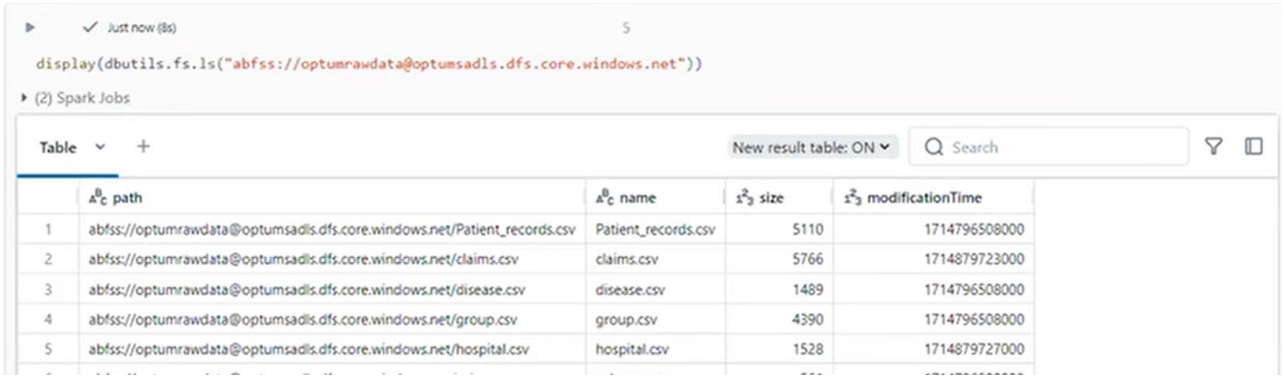
NOW CREATE THE SCOPE IN DATA BRICKS





NOW LIST SCOPE

NOW WE ARE GOING TO USE SUB GROUP DATASET



Just now (8s) 5

```
display(dbutils.fs.ls("abfss://optumrawdata@optumsadls.dfs.core.windows.net"))
```

(2) Spark Jobs

Table + New result table: ON Search

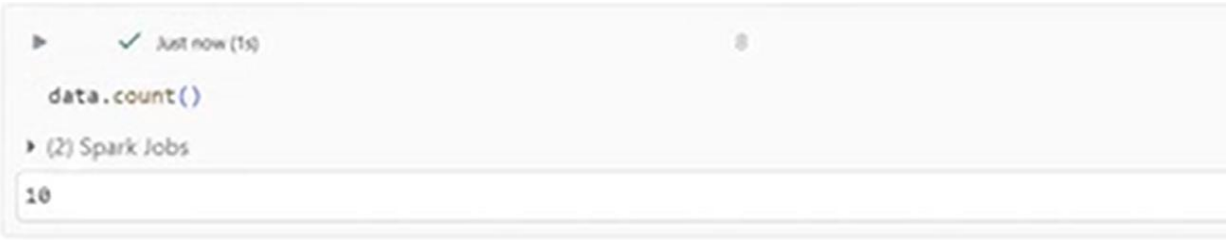
	path	name	size	modificationTime
1	abfss://optumrawdata@optumsadls.dfs.core.windows.net/Patient_records.csv	Patient_records.csv	5110	1714796508000
2	abfss://optumrawdata@optumsadls.dfs.core.windows.net/claims.csv	claims.csv	5766	1714879723000
3	abfss://optumrawdata@optumsadls.dfs.core.windows.net/disease.csv	disease.csv	1489	1714796508000
4	abfss://optumrawdata@optumsadls.dfs.core.windows.net/group.csv	group.csv	4390	1714796508000
5	abfss://optumrawdata@optumsadls.dfs.core.windows.net/hospital.csv	hospital.csv	1528	1714879727000

Just now (3s) 6

```
data = spark.read.csv("abfss://optumrawdata@optumsadls.dfs.core.windows.net/subgroup.csv")
```

(1) Spark Jobs

data: pyspark.sql.dataframe.DataFrame = [_c0: string, _c1: string ... 2 more fields]



Just now (1s) 8

```
data.count()
```

(2) Spark Jobs

```
10
```

Just now (<1s) 9

```
#drop duplicates
data = data.dropDuplicates()
```

data: pyspark.sql.dataframe.DataFrame = [subgrp_sk: string, subgrp_name: string ... 2 more fields]

Just now (1s) 10

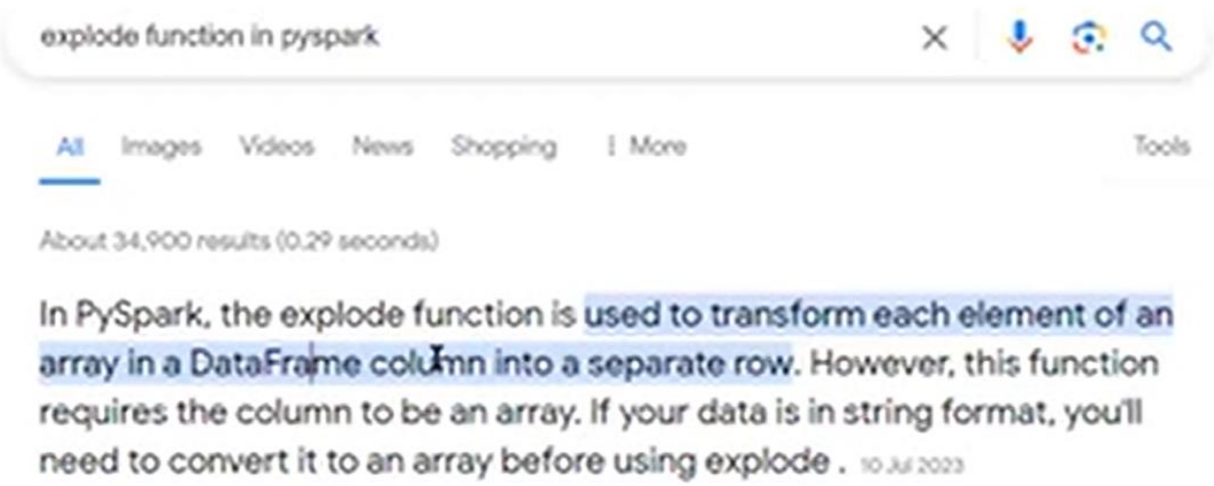
```
data.count()
```

(3) Spark Jobs

```
10
```

Interrupt 0001 10 Python

```
data.select([count(when(isnan(c) | col(c).isNull(), c)).alias(c) for c in data.columns]).show()
```


explode function in pyspark
All
Images
Videos
News
Shopping
More
Tools
About 34,900 results (0.29 seconds)
In PySpark, the explode function is used to transform each element of an array in a DataFrame column into a separate row. However, this function requires the column to be an array. If your data is in string format, you'll need to convert it to an array before using explode . 10 Jul 2023


Waiting
import pyspark
from pyspark.sql import *
from pyspark.sql.functions import *


Just now (<1s)
11
data = data.withColumn("subgrp_id", split(data["subgrp_id"], ","))
data: pyspark.sql.dataframe.DataFrame = [subgrp_sk: string, subgrp_name: string ... 2 more fields]

```
data.show(5,False)
```

(2) Spark Jobs

```
+---------+------------------+-------------+----------------------------------------+
|subgrp_sk|subgrp_name       |monthly_premium|subgrp_id                             |
+---------+------------------+-------------+----------------------------------------+
|S104     |Therapy           |1500         |[GRP103, GRP113, GRP123, GRP133, GRP143]|
|S105     |Allergies         |2300         |[GRP153, GRP104, GRP114, GRP124]        |
|S103     |Physiology        |2000         |[GRP122, GRP108, GRP138, GRP148]        |
|S102     |Accident          |1000         |[GRP110, GRP150, GRP136]                |
|S101     |Deficiency Diseases|3000        |[GRP101, GRP105]                        |
+---------+------------------+-------------+----------------------------------------+
only showing top 5 rows
```

NOW IT AS BEEN CHANGED TO ARRAY

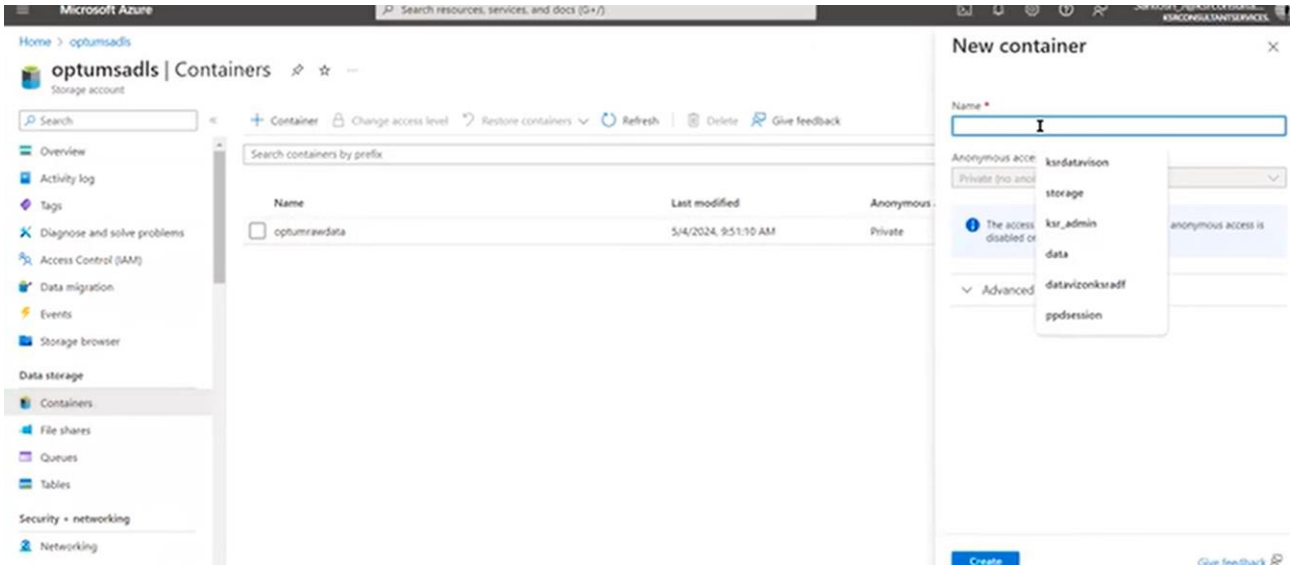
```
data = data.withColumn("subgrp_id", explode(data['subgrp_id']))
```

data: pyspark.sql.dataframe.DataFrame = [subgrp_sk: string, subgrp_name: string ... 2 more fields]

Here the output of the data

```
+---------+------------------+-------------+---------+
|subgrp_sk|       subgrp_name|monthly_premium|subgrp_id|
+---------+------------------+-------------+---------+
|     S101|Deficiency Diseases|         3000|   GRP101|
|     S101|Deficiency Diseases|         3000|   GRP105|
|     S102|          Accident|         1000|   GRP110|
|     S102|          Accident|         1000|   GRP150|
|     S102|          Accident|         1000|   GRP136|
|     S103|         Physiology|         2000|   GRP122|
|     S103|         Physiology|         2000|   GRP108|
|     S103|         Physiology|         2000|   GRP138|
|     S103|         Physiology|         2000|   GRP148|
|     S104|           Therapy|         1500|   GRP103|
|     S104|           Therapy|         1500|   GRP113|
|     S104|           Therapy|         1500|   GRP123|
|     S104|           Therapy|         1500|   GRP133|
|     S104|           Therapy|         1500|   GRP143|
|     S105|         Allergies|         2300|   GRP153|
|     S105|         Allergies|         2300|   GRP104|
|     S105|         Allergies|         2300|   GRP114|
|     S105|         Allergies|         2300|   GRP124|
```

Create an another folder in ADLS TO STAGGING THE DATA.



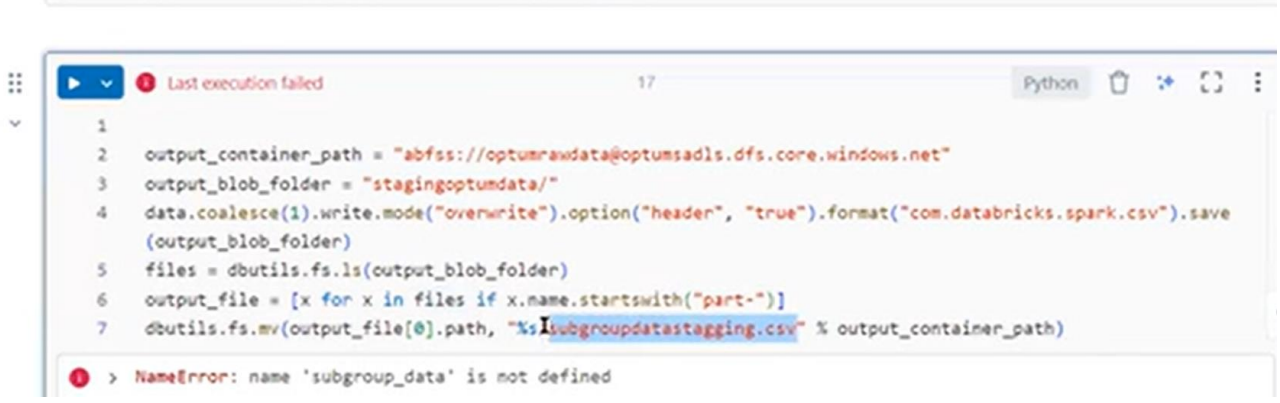
THIS IS BOILER PLATE FOR TO PUSH THE DATA INTO STAGING DATAINTO STAGING FOLDER.

IN 1ST LINE WE NEED TO GIVE PATHA AND GIVE THE CONTAINER NAME IN WHICH FOLDER IT SHOULD BE STORE

2ND LINE WE GIVE THE FOLDER NAME

INSTEAD OF SUBGROUP_DATA WE WRITE THE DATA

BECAUSE WE HAVE SOTRED THE DATEFRAME INTO DATA VARIABLE

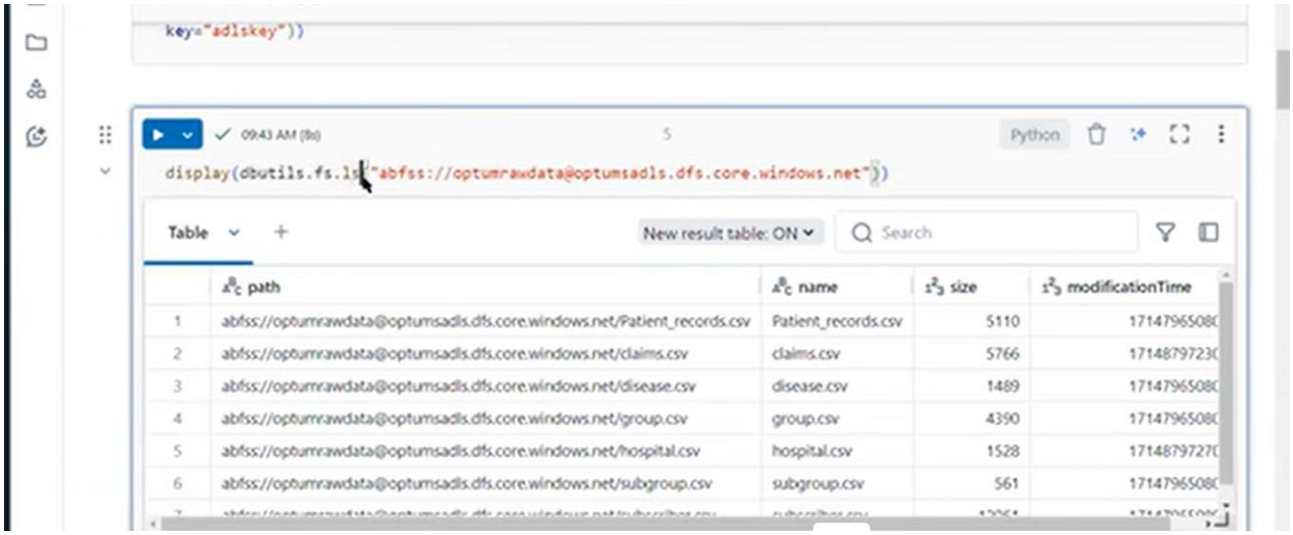


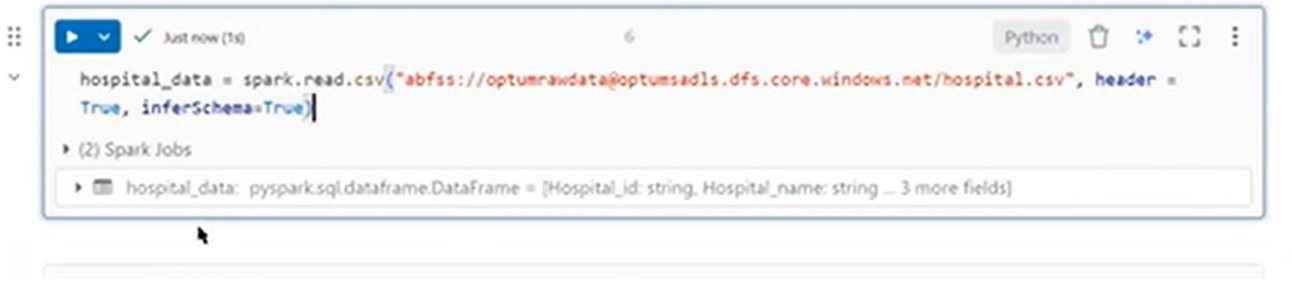
IN THE 7TH LINE WE NEED TO GIVE OUR OWN NAME WHICH FILE NAME YOU NEED

WE GET AS TRUE

SO THIS FILE IS ALL IN DEVELOPING FACE SO WE NEED TO CREATE AN ANOTHER FILE AS PRODUCTION DATA FILE.

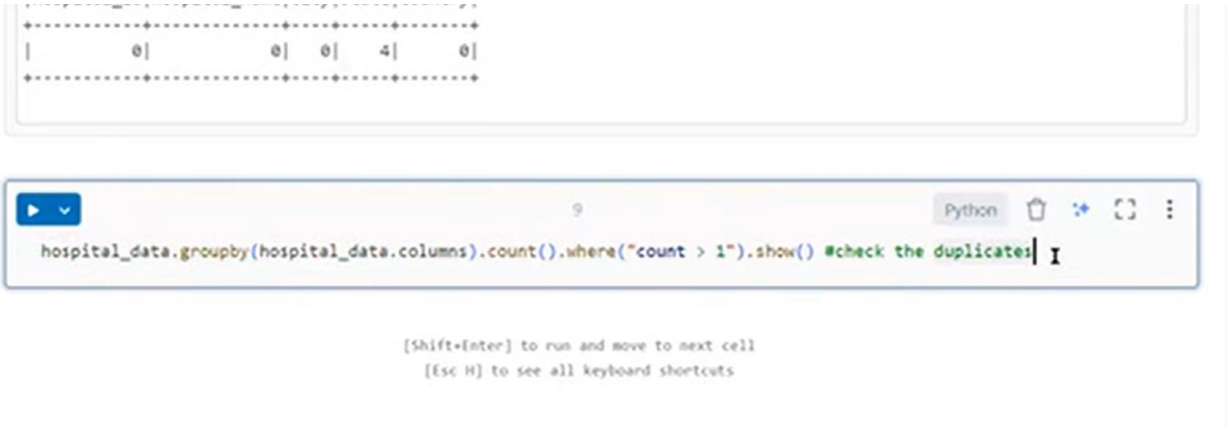
SO NOW DO THE HOSPITAL DATA



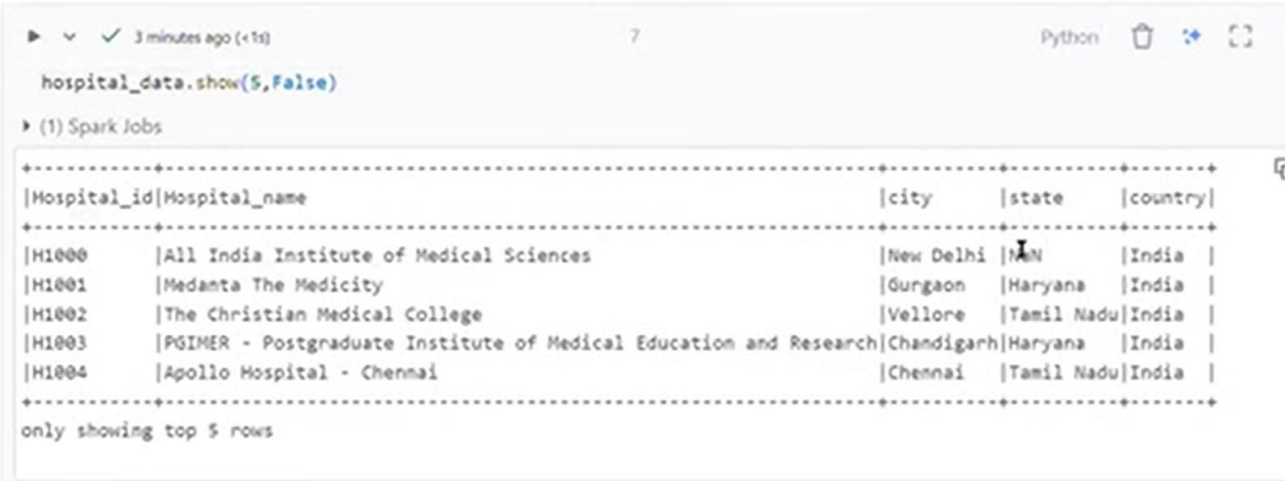


Drop null values

Check the duplicates



We have null in state column



▸ (2) Spark Jobs

```
+-----------+--------------------------------------------------------------+----------+----------+-------+
|Hospital_id|Hospital_name                                                 |city      |state     |country|
+-----------+--------------------------------------------------------------+----------+----------+-------+
|H1004      |Apollo Hospital - Chennai                                     |Chennai   |Tamil Nadu|India  |
|H1002      |The Christian Medical College                                 |Vellore   |Tamil Nadu|India  |
|H1003      |PGIMER - Postgraduate Institute of Medical Education and Research|Chandigarh|Haryana   |India  |
|H1000      |All India Institute of Medical Sciences                       |New Delhi |UT        |India  |
|H1001      |Medanta The Medicity                                          |Gurgaon   |Haryana   |India  |
+-----------+--------------------------------------------------------------+----------+----------+-------+
only showing top 5 rows
```

▶ ⌄ ✓ Just now (1s) 15

```
hospital_data.select("city").distinct().show(10)
```

▸ (2) Spark Jobs

```
+-----------+
|       city|
+-----------+
|    Chennai|
|     Mumbai|
|    Gurgaon|
|    Vellore|
|      Delhi|
| Chandigarh|
|  Bengaluru|
|  New Delhi|
|  Hyderabad|
+-----------+
```

replace column values in pyspark ✕

All Videos Shopping Images News ⋮ More

About 1,30,000 results (0.25 seconds)

Apache Spark
https://spark.apache.org › api › python › reference › api

pyspark.sql.DataFrame.replace

DataFrame.**replace**¶ ... Returns a new DataFrame **replacing** a **value** with another **value**. DataFrame.**replace**() and DataFrameNaFunctions.**replace**() are aliases of each ...

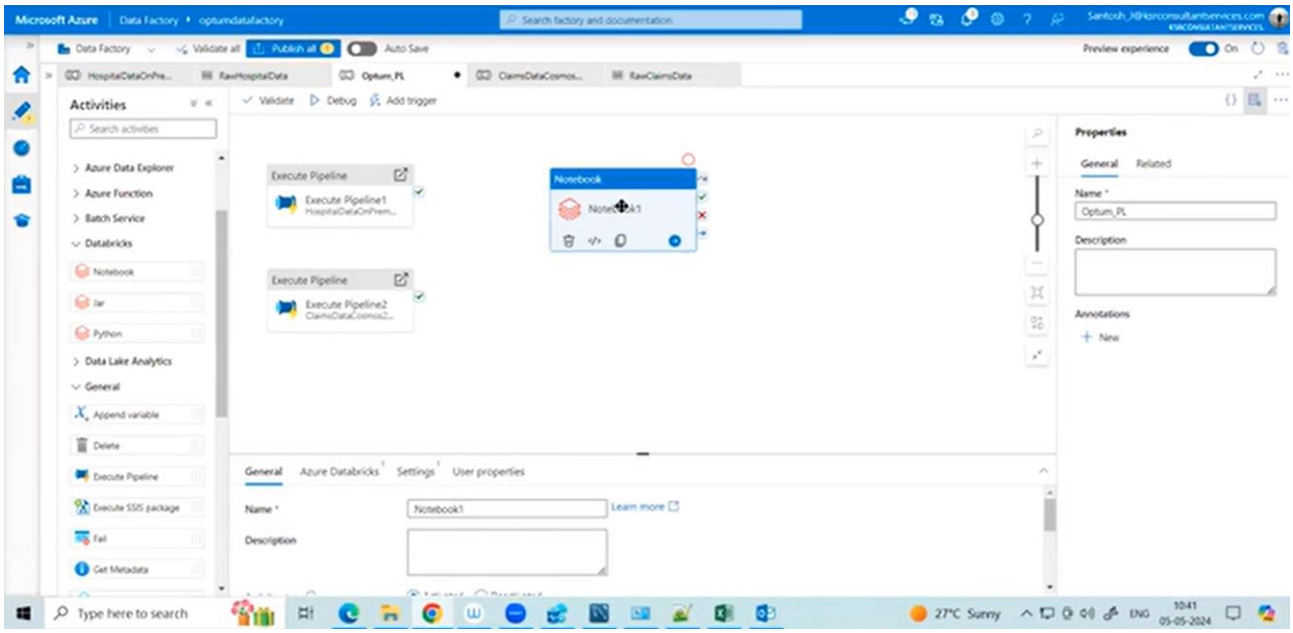
Stack Overflow
https://stackoverflow.com › questions › how-to-condit...

How to conditionally replace value in a column based on ...

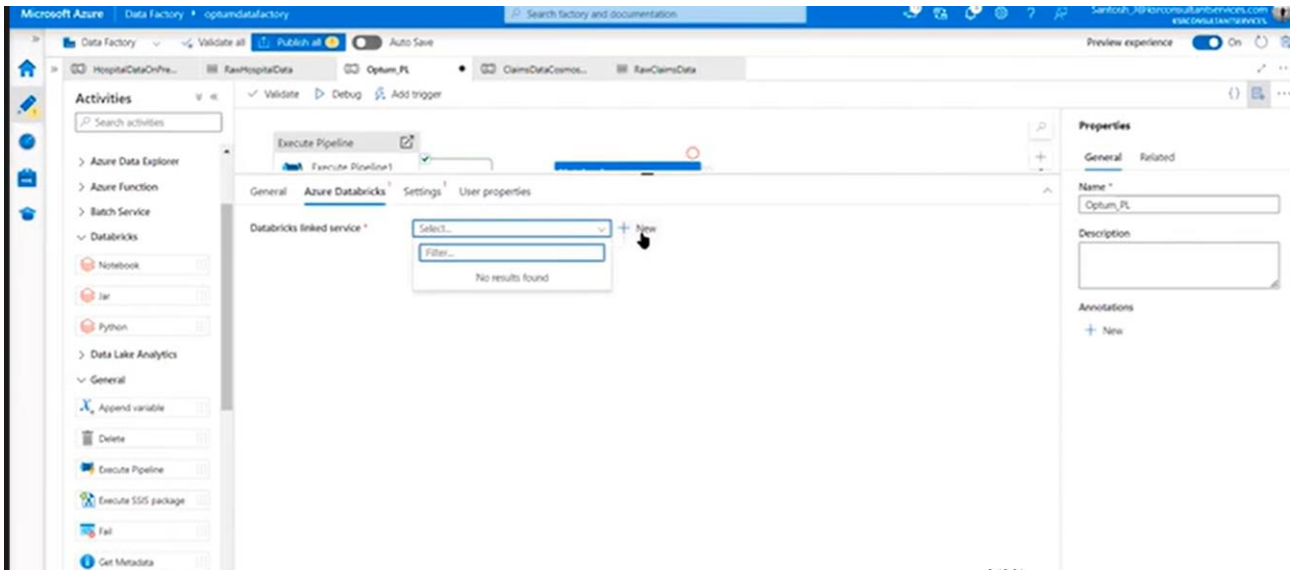
Send it to staging folder

Now do it in dev to production

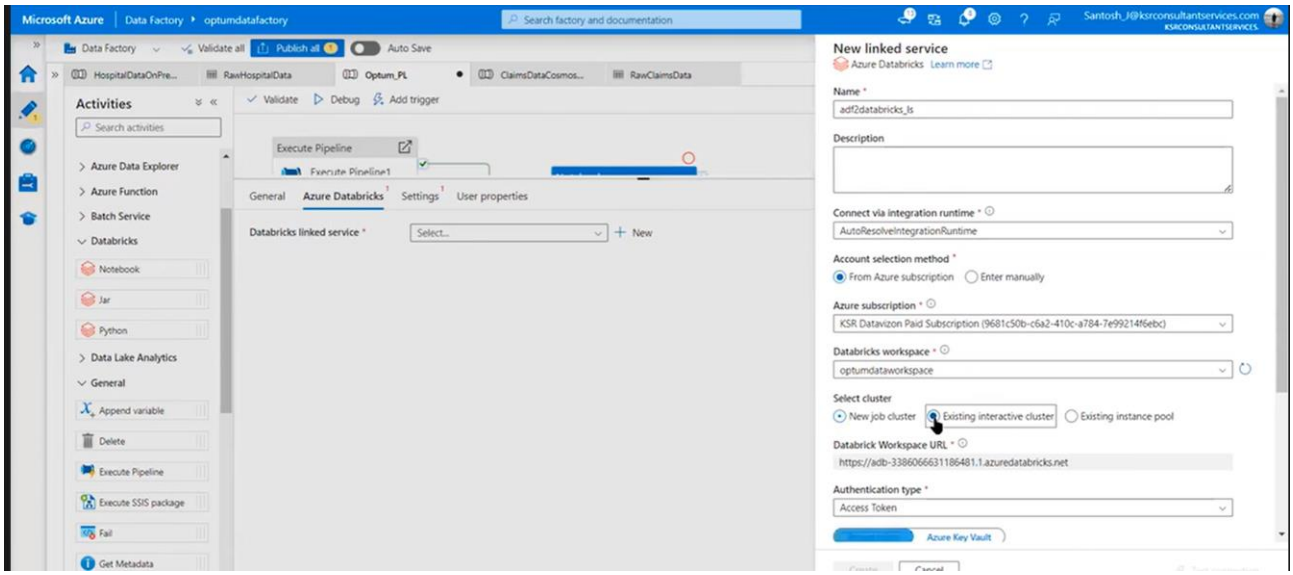
NOW GO TO DATA FACTORY U NEED TO RUN THE PRODUCTION NOTE BOOK FILES IN DATAFACTORY



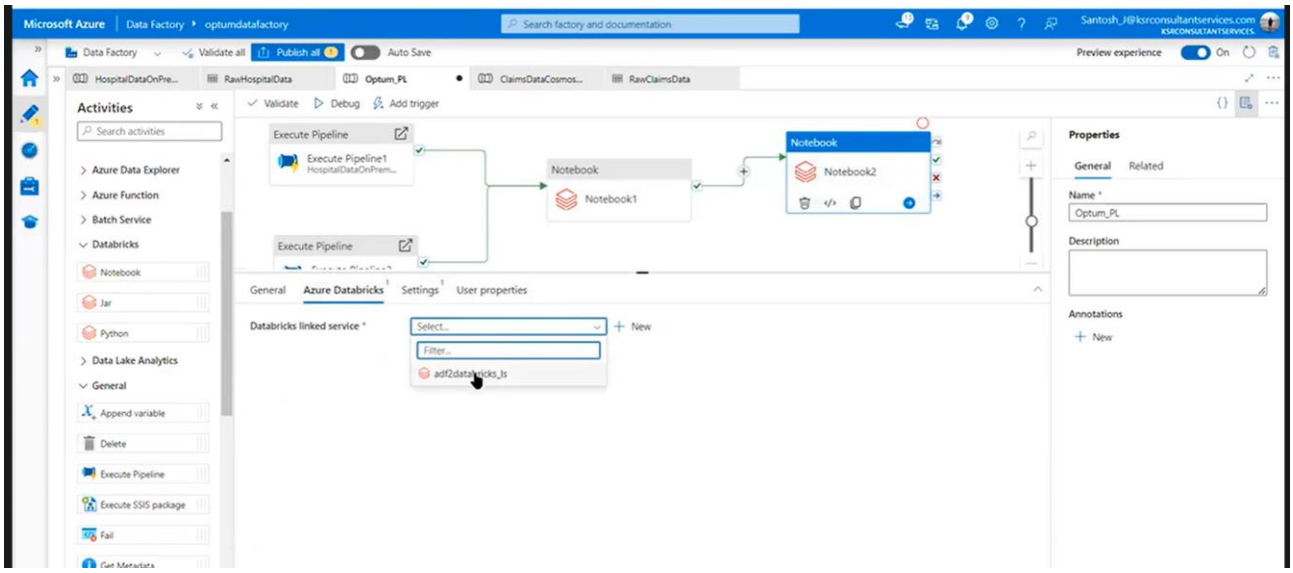
WE CAN CALL THE NOTE BOOK HERE ALSO



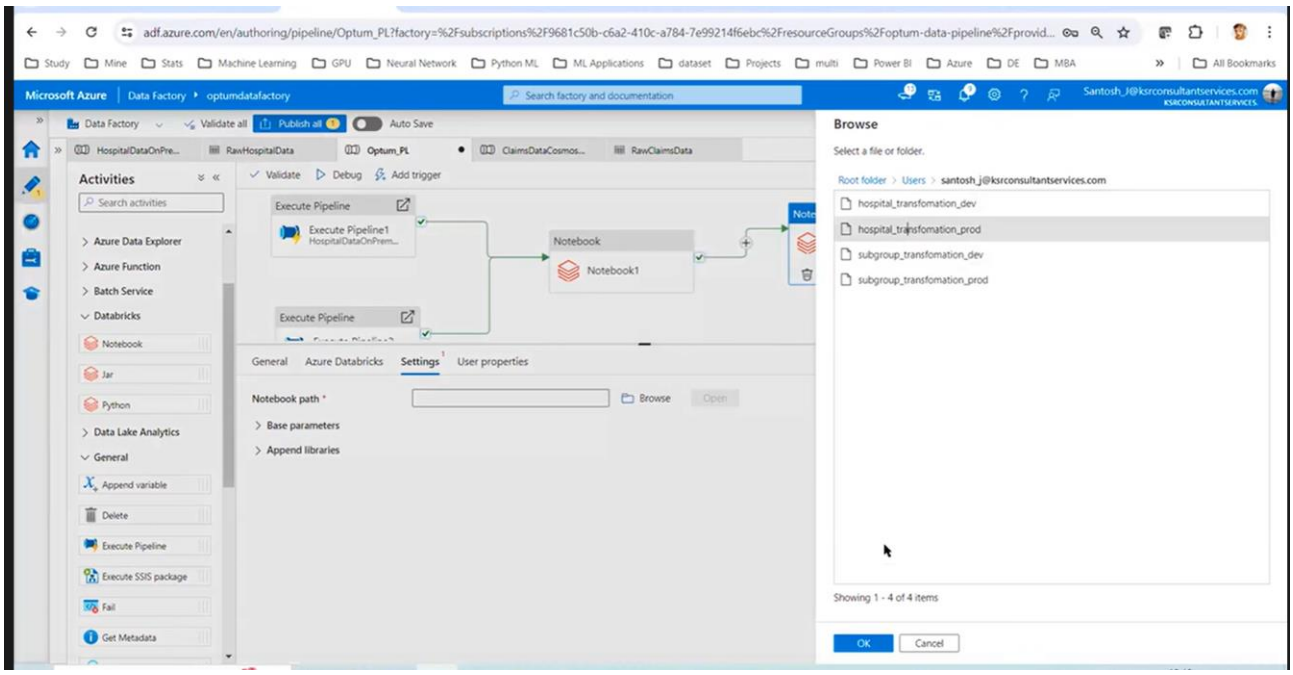
CREATE THE LINK SERVICE FOR THE ADF TO DATA BRICKS LIKE THIS



IN CLUSTER WE NEED TO SLECT CLUSTER AS EXISTING NOW A NEW CLUSTER.

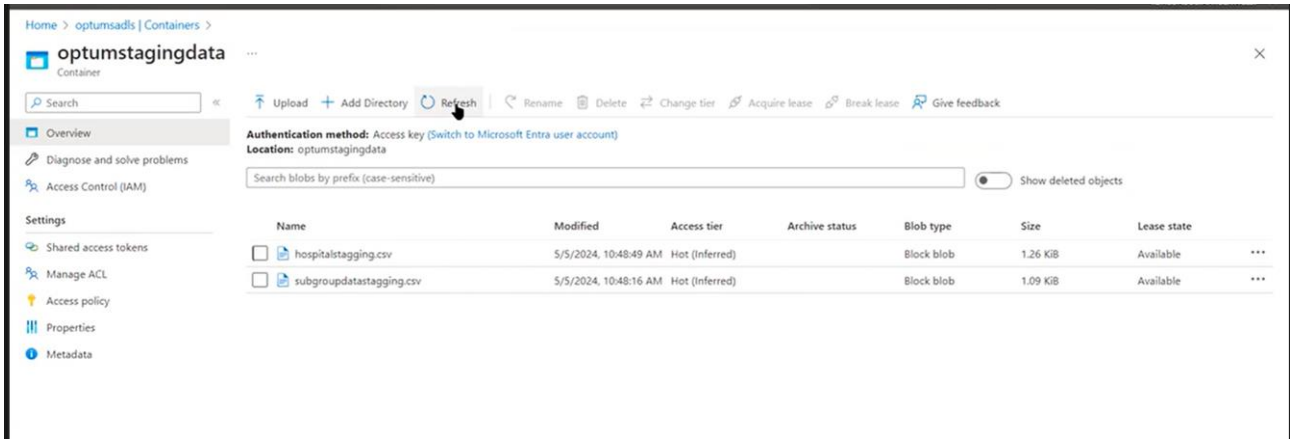


NEXT HOSPITAL DATA NOTEBOOK WE NEED TO SELCT IN ADF

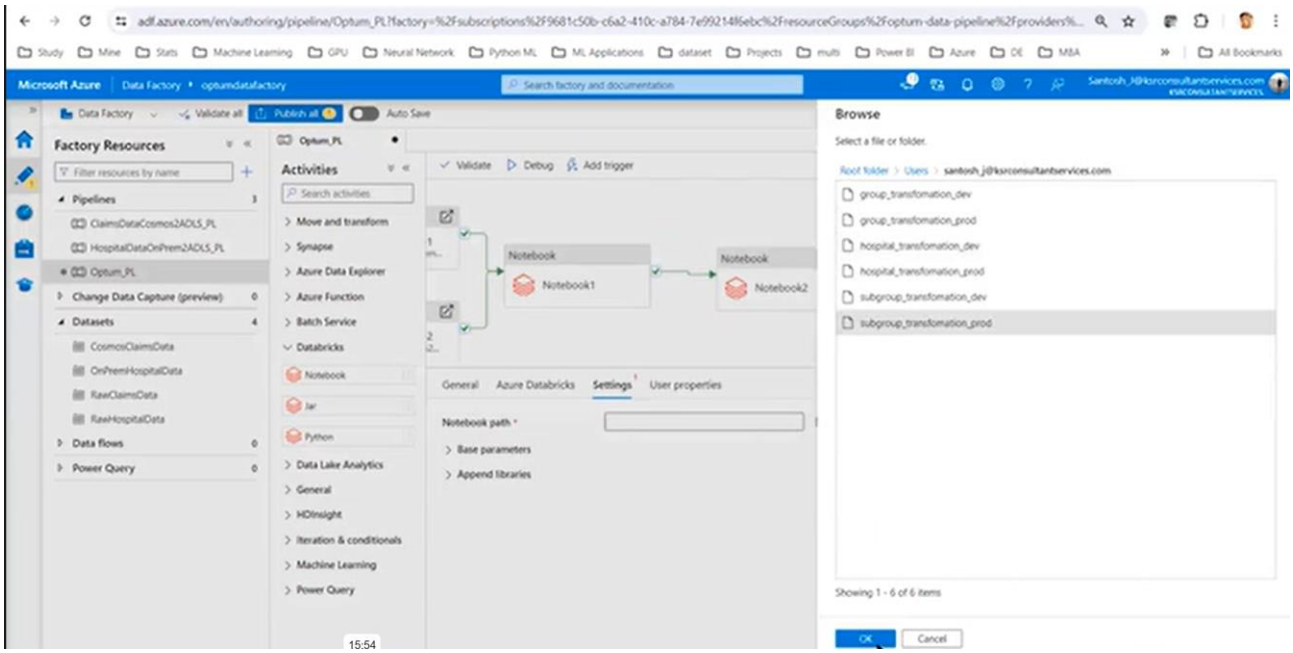


JUST DO THE PUBLISH

NOW DELET ALL THE STAGING DATA RUN THIS PIPELINE



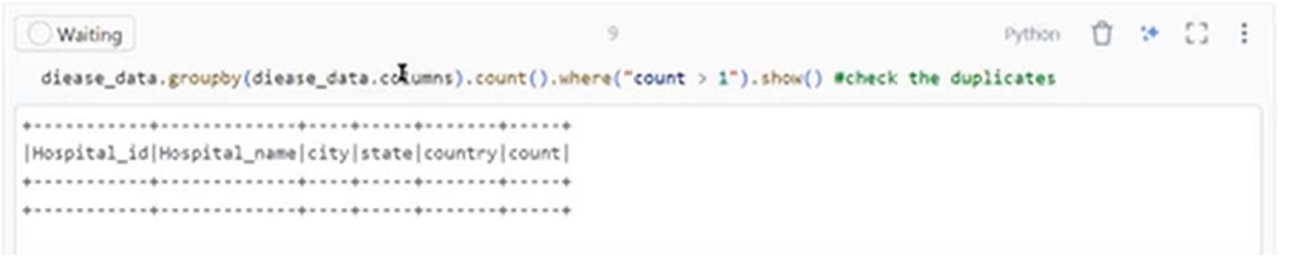
NOW LETS DO THE SUBGROUP DATA

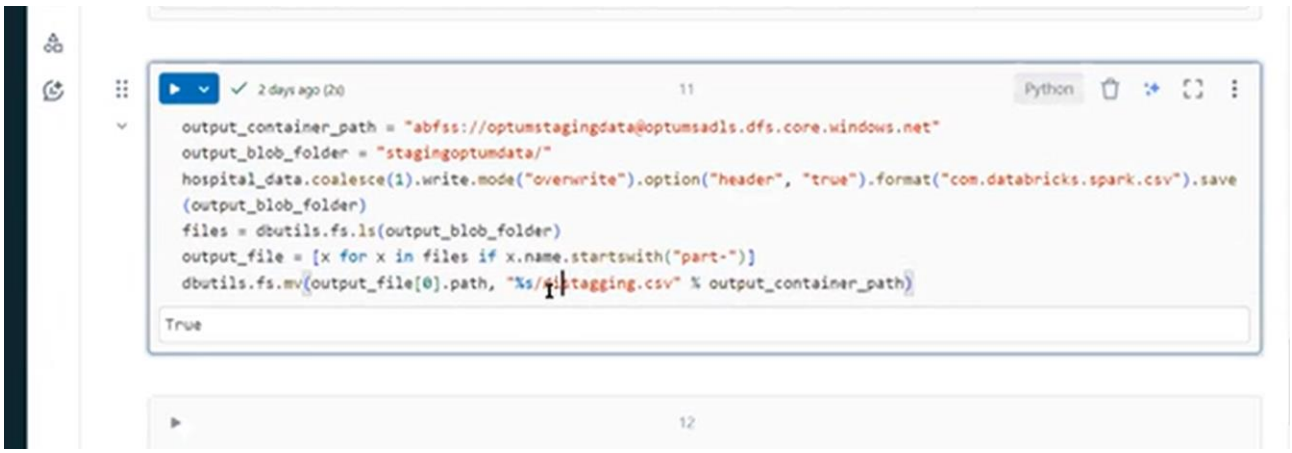


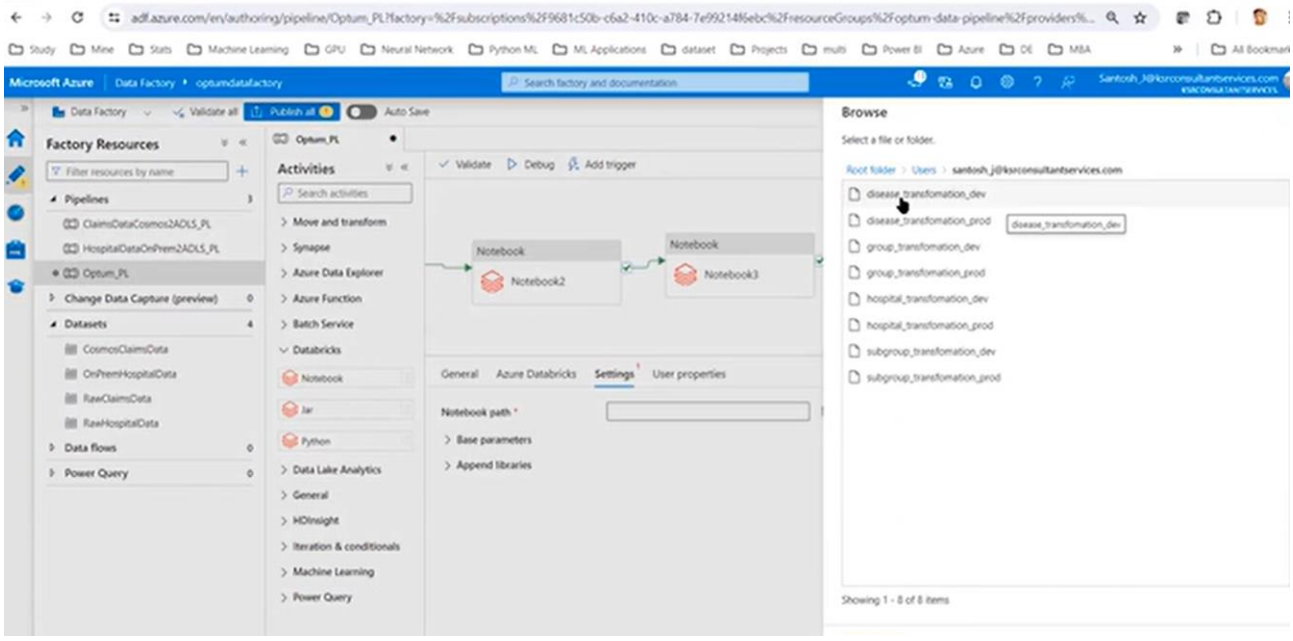
NOW DO THE DISEAS TRANSFORMATION

CHECK THE NULL VALUES

CHECK THE DUPLICATES







NOW CONNECT ANOTHER NOTEBOOK INTO DATABRICKS

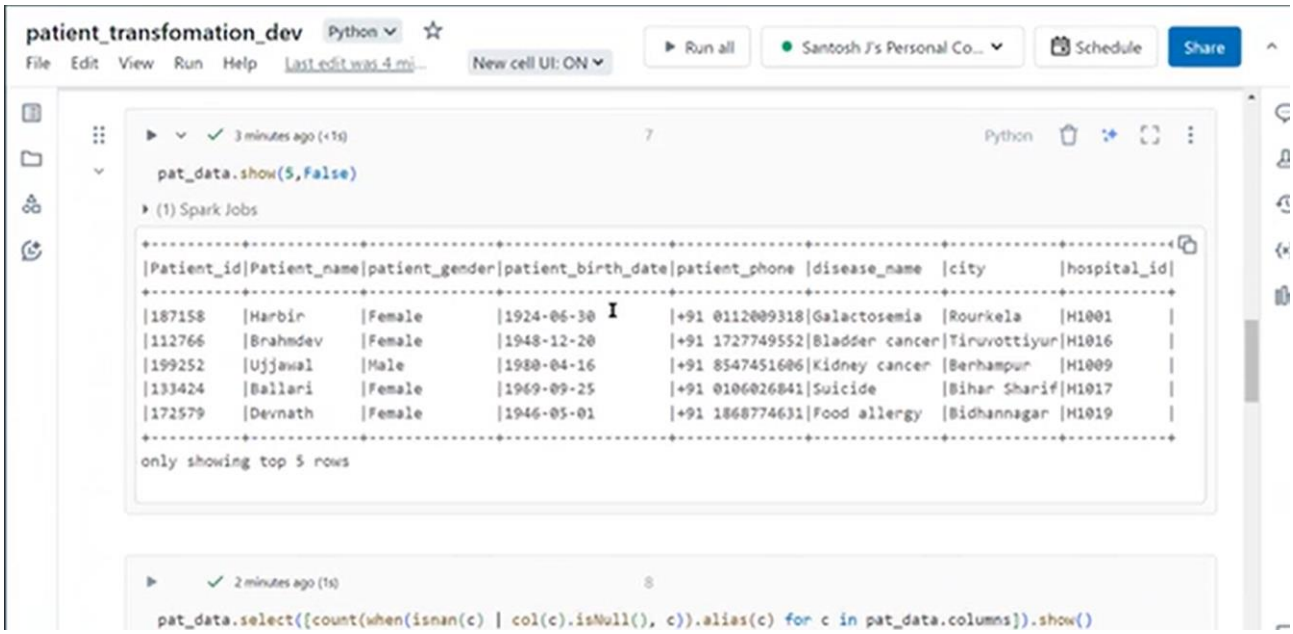
LEST TAKE THE PATIENT DATA

CHECK THE DUPLICATES

FILL THE NULL

```
11
pat_data = pat_data.fillna({"Patient_name":"Guest/NA"})
```

WE DO SOME ONE CHANE



WE DON'T NEED THE PATIENT DOB WE NEED A AGE OF A PATIENT

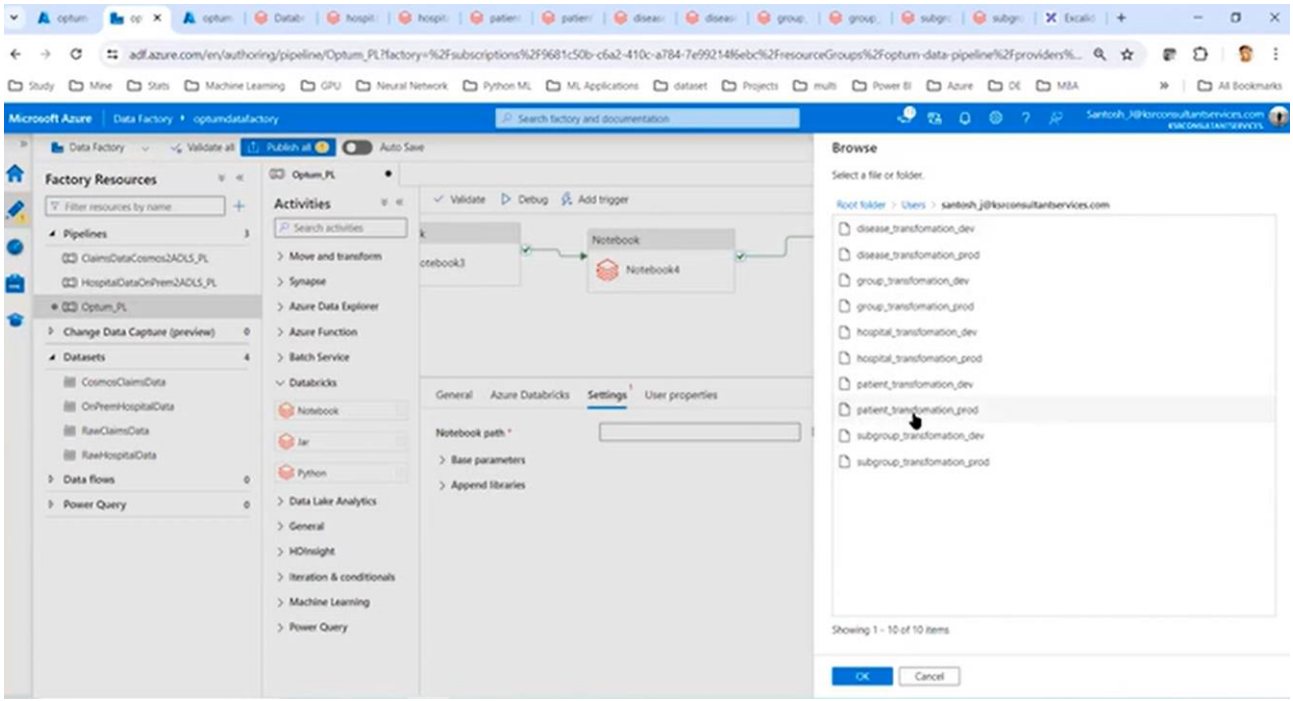
```
12
Python
pat_data = pat_data.withColumn("Patient_age", (months_between(current_date(), col('patient_birth_date'))/12).cast
("int"))
```

```
14
pat_data = pat_data.drop(col("patient_birth_date"))
```

```
output_container_path = "abfss://optumstagingdata@optumsadls.dfs.core.windows.net"
output_blob_folder = "stagingoptumdata/"
hospital_data.coalesce(1).write.mode("overwrite").option("header", "true").format("com.databricks.spark.csv").save
(output_blob_folder)
files = dbutils.fs.ls(output_blob_folder)
output_file = [x for x in files if x.name.startswith("part-")]
dbutils.fs.mv(output_file[0].path, "%s/patientstagging.csv" % output_container_path)
```

```
True
```

NOW JUST TAKE A CLONE AND DO CREATE A PRODUCTION FILE FOR DATA FACTORY

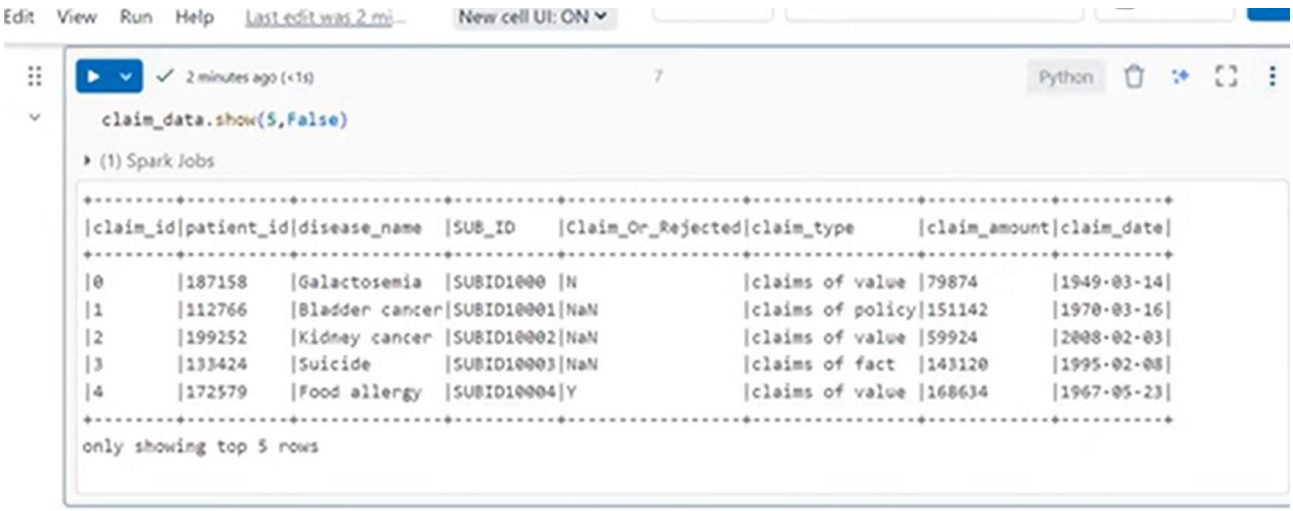


NOW TAKE A CLAIM DATASET

```
claim_data.select([count(when(isnan(c) | col(c).isNull(), c)).alias(c) for c in claim_data.columns]).show()
```

```
+-------+---------------+--------+------+--------+--------+----+
|country|premium_written|zip_code|grp_id|grp_name|grp_type|city|
+-------+---------------+--------+------+--------+--------+----+
|      0|              0|       0|     0|       0|       0|   0|
```

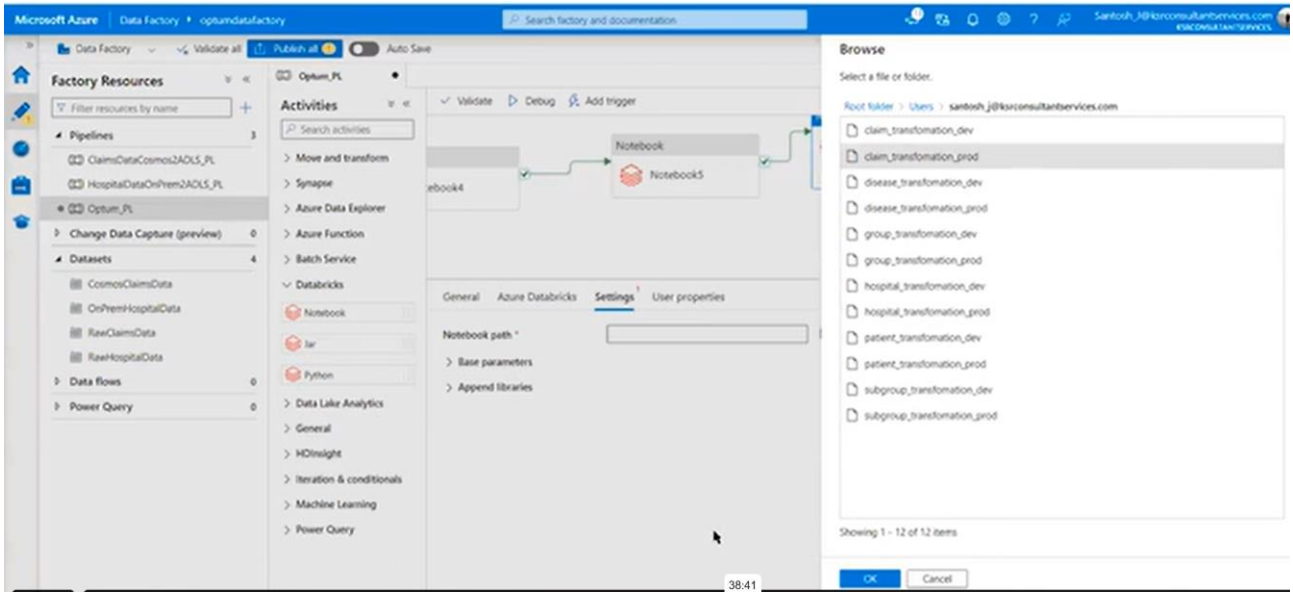
NOW WE NEED TO REMOVE THIS

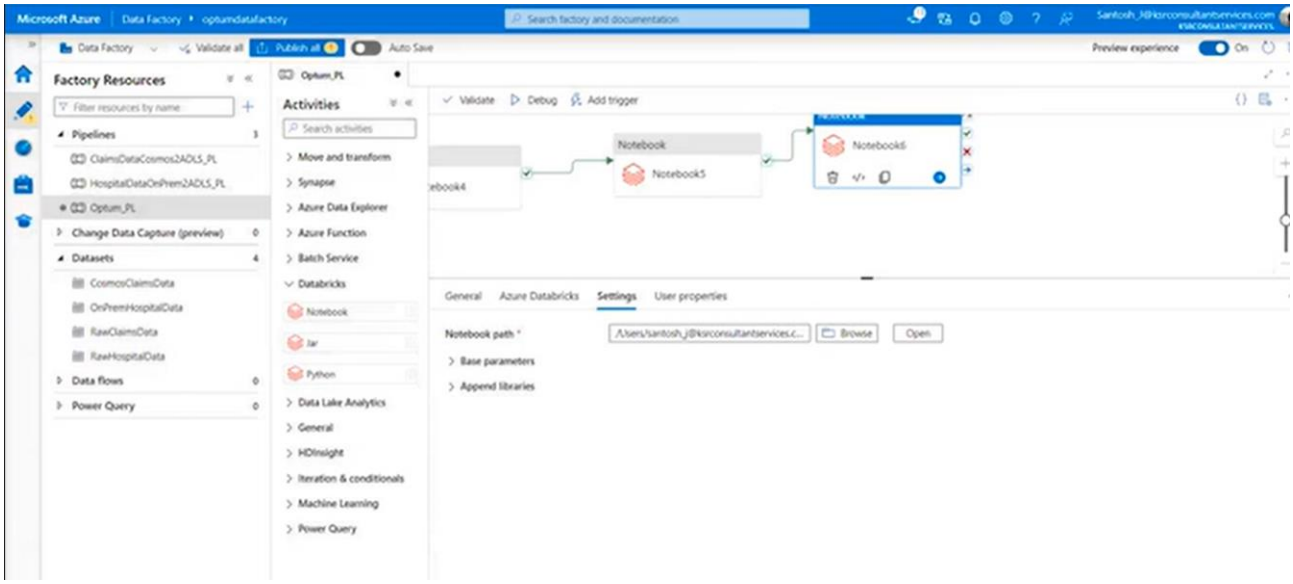


```
output_container_path = "abfss://optumstagingdata@optumsadls.dfs.core.windows.net"
output_blob_folder = "stagingoptumdata/"
hospital_data.coalesce(1).write.mode("overwrite").option("header", "true").format("com.databricks.spark.csv").save
(output_blob_folder)
files = dbutils.fs.ls(output_blob_folder)
output_file = [x for x in files if x.name.startswith("part-")]
dbutils.fs.mv(output_file[0].path, "%s/claimstagging.csv" % output_container_path)
```

```
True
```

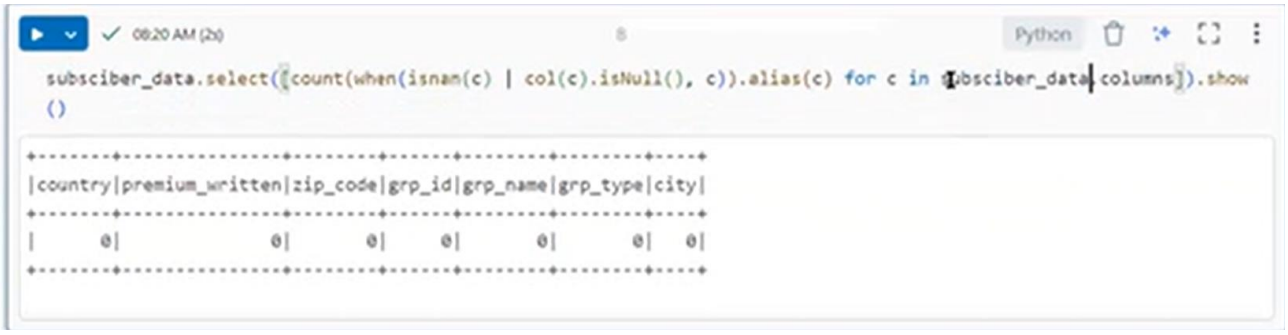
MAKE A PRODUCTIN FILE AND ADD INTO THE ADF





SUBSCRIBER DATA SET


08:19 AM (5s)
6
Python
subsciber_data = spark.read.csv("abfss://optumrawdata@optumsadls.dfs.core.windows.net/subscriber.csv", header = True, inferSchema=True)
group_data: pyspark.sql.dataframe.DataFrame = [country: string, premium_written: integer ... 5 more fields]


08:20 AM (2s)
8
Python
subsciber_data.select([count(when(isnan(c) | col(c).isNull(), c)).alias(c) for c in subsciber_data.columns]).show()
+-------+--------------+--------+------+--------+--------+----+
|country|premium_written|zip_code|grp_id|grp_name|grp_type|city|
+-------+--------------+--------+------+--------+--------+----+
| 0| 0| 0| 0| 0| 0| 0|
+-------+--------------+--------+------+--------+--------+----+

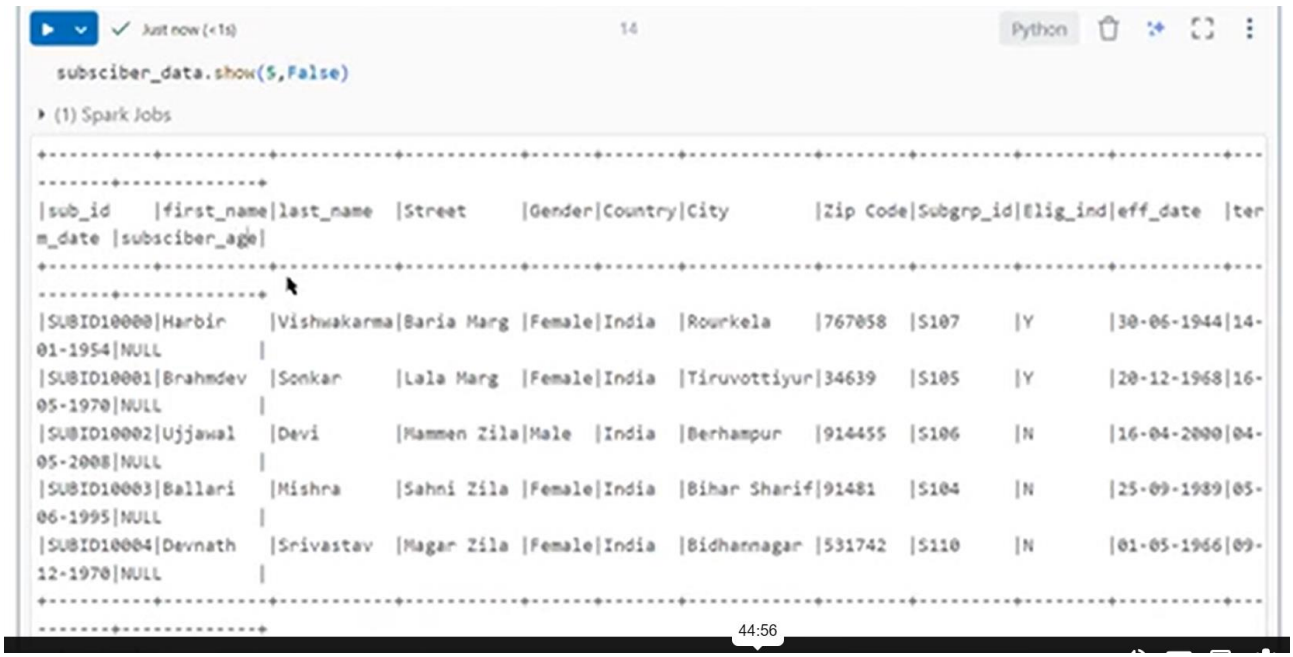

Just now (1s)
9
Python
subsciber_data.groupby(subsciber_data.columns).count().where("count > 1").show() #check the duplicates
(2) Spark Jobs
|sub_id|first_name|last_name|Street|Birth_date|Gender|Phone|Country|City|Zip Code|Subgrp_id|Elig_ind|eff_date|term_date|count|


Just now (<1s)
10
subsciber_data = subsciber_data.dropDuplicates()
subsciber_data: pyspark.sql.dataframe.DataFrame = [sub_id: string, first_name: string ... 12
+ Code
+ Text


Just now (<1s)
12
subsciber_data = subsciber_data.fillna({"Elig_ind":"N", "first_name":"Guest/NA"})
subsciber_data: pyspark.sql.dataframe.DataFrame = [sub_id: string, first_name: string ... 12 more fiel

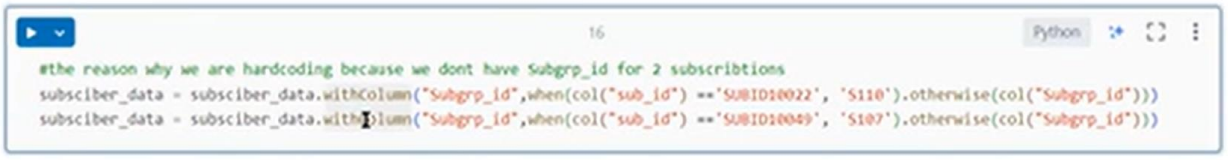
ILL CHECK THE MISSING VALUE COUNT

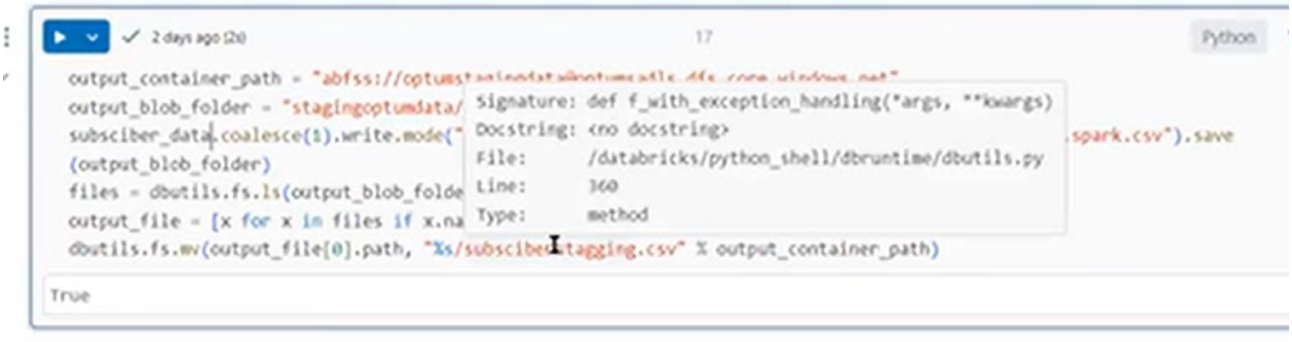
SO THE SUBSCRIBERS ARE SHOWING NULL



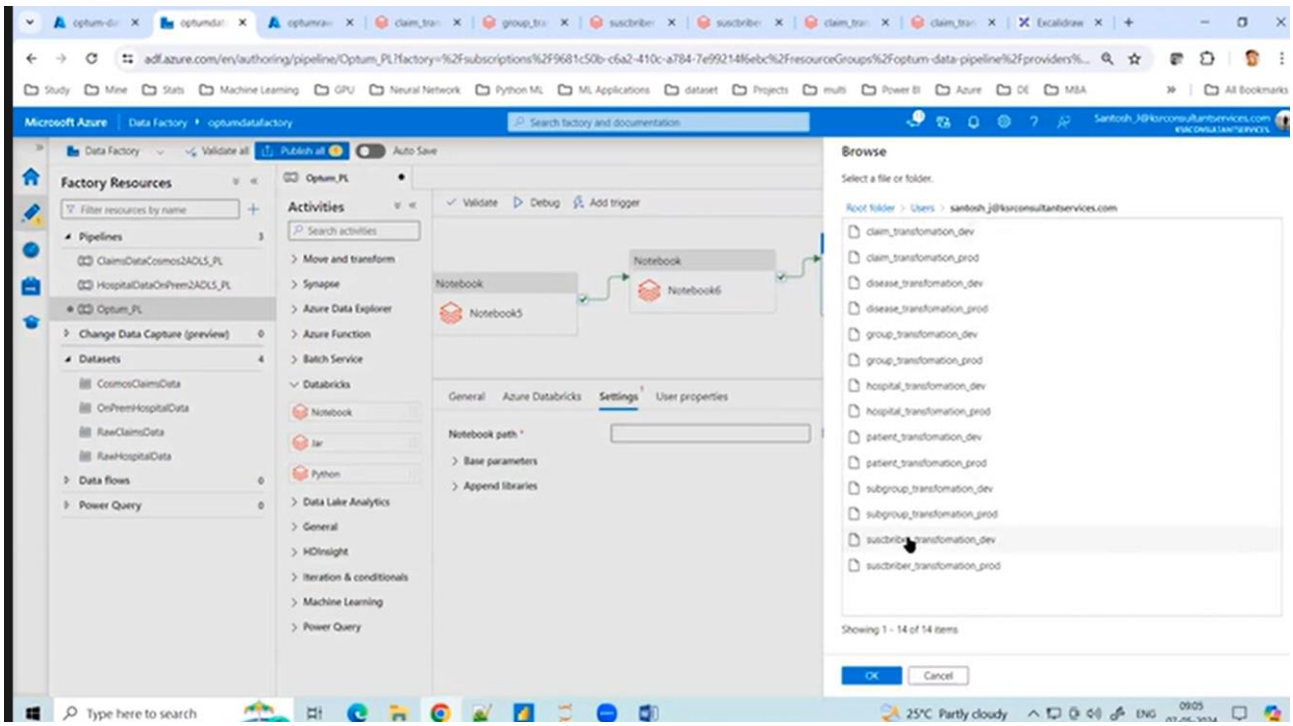
DO IT FOR ANTOHR PATIENT

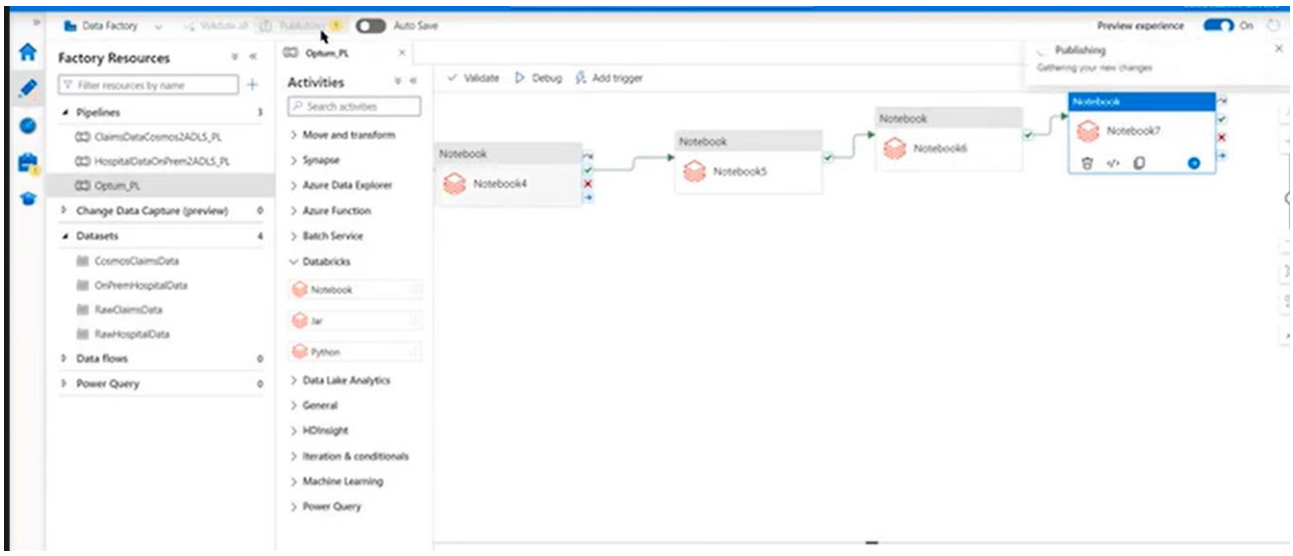
NOW WE NEED DO THE REVERSE ENGINEER



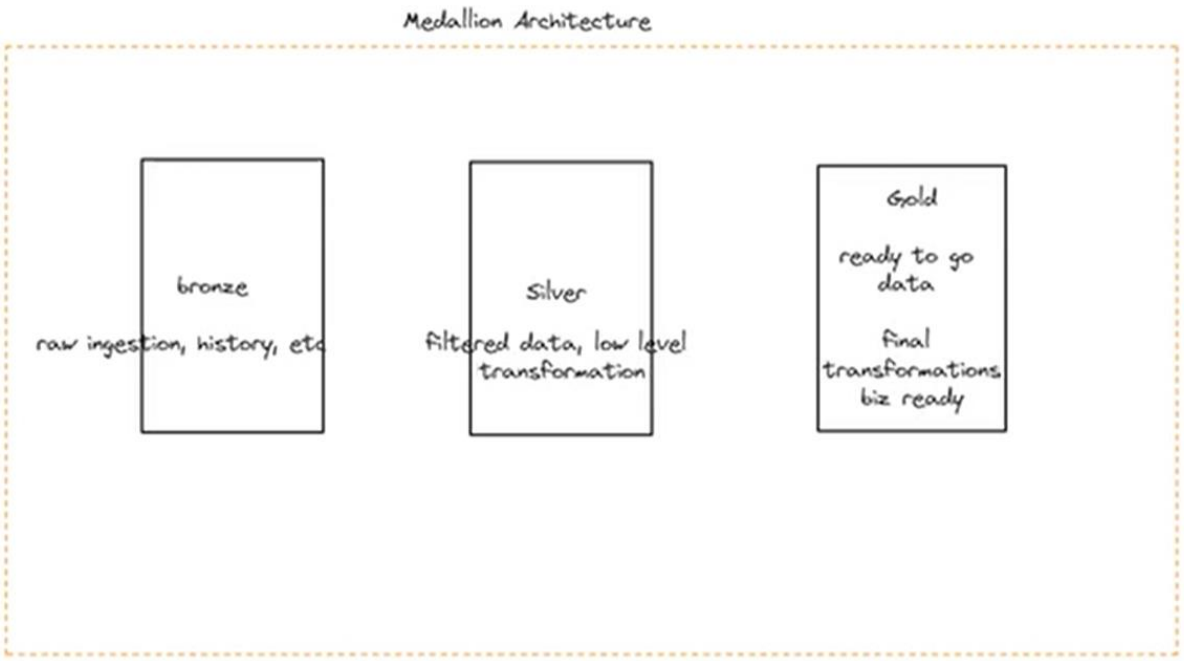


NOW DO THE PRODUCTION NOTEBOOK INTO ADF

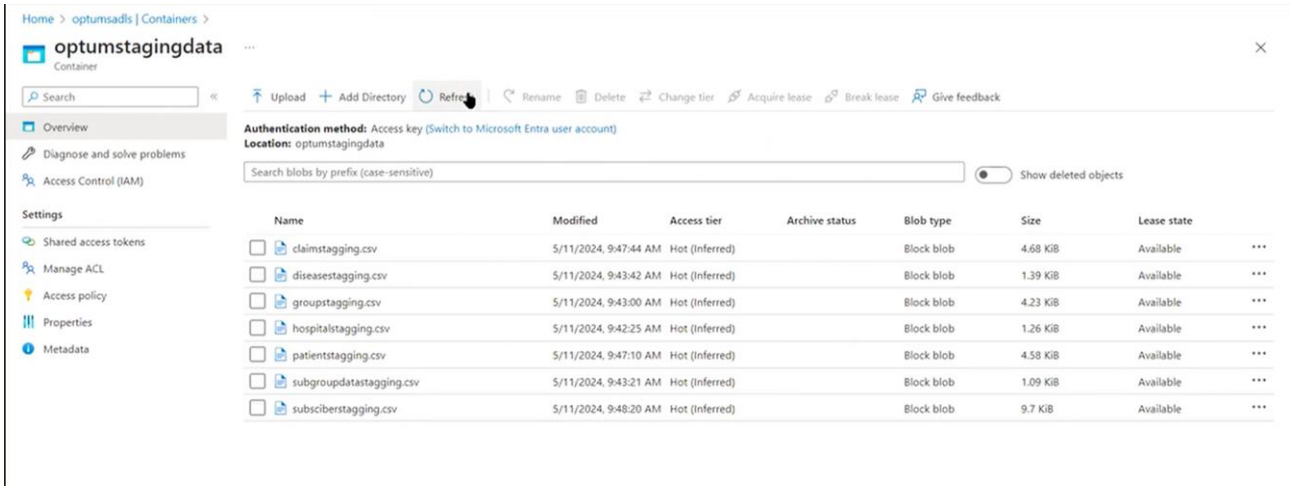




JUST PUBLISH IT



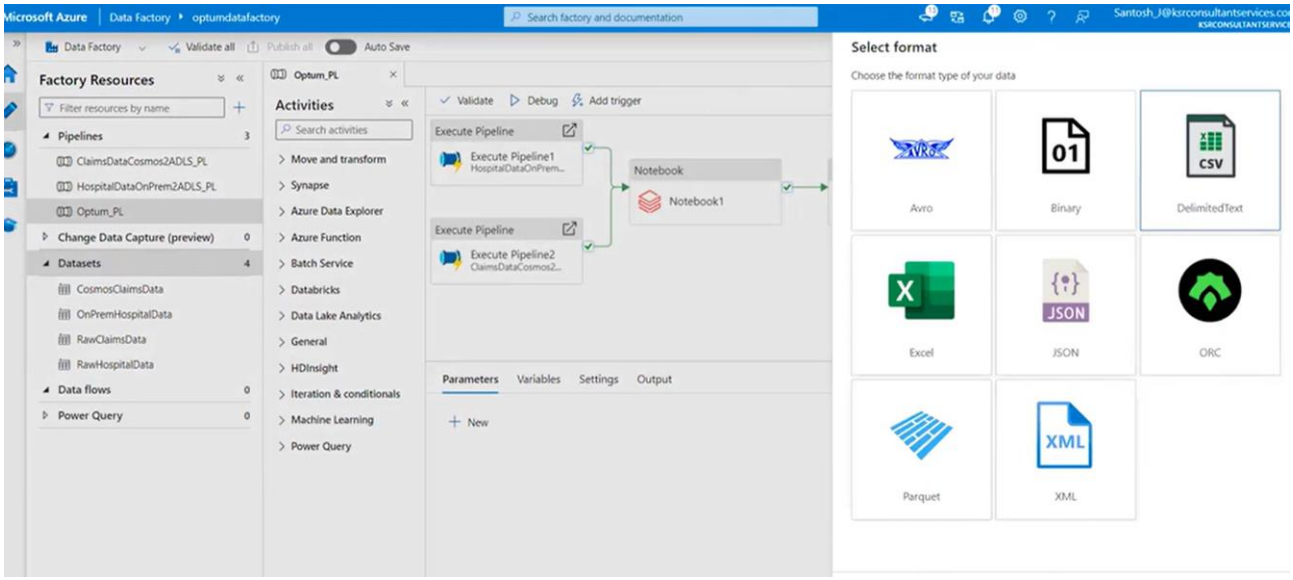
OUR STAGING DATA IS READY

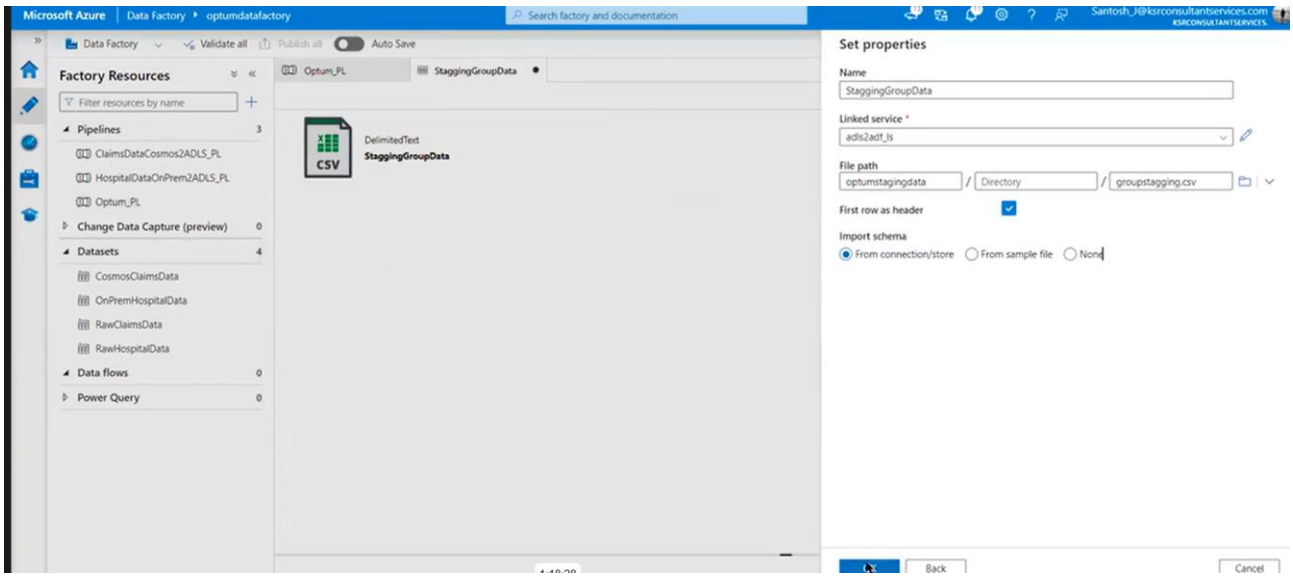


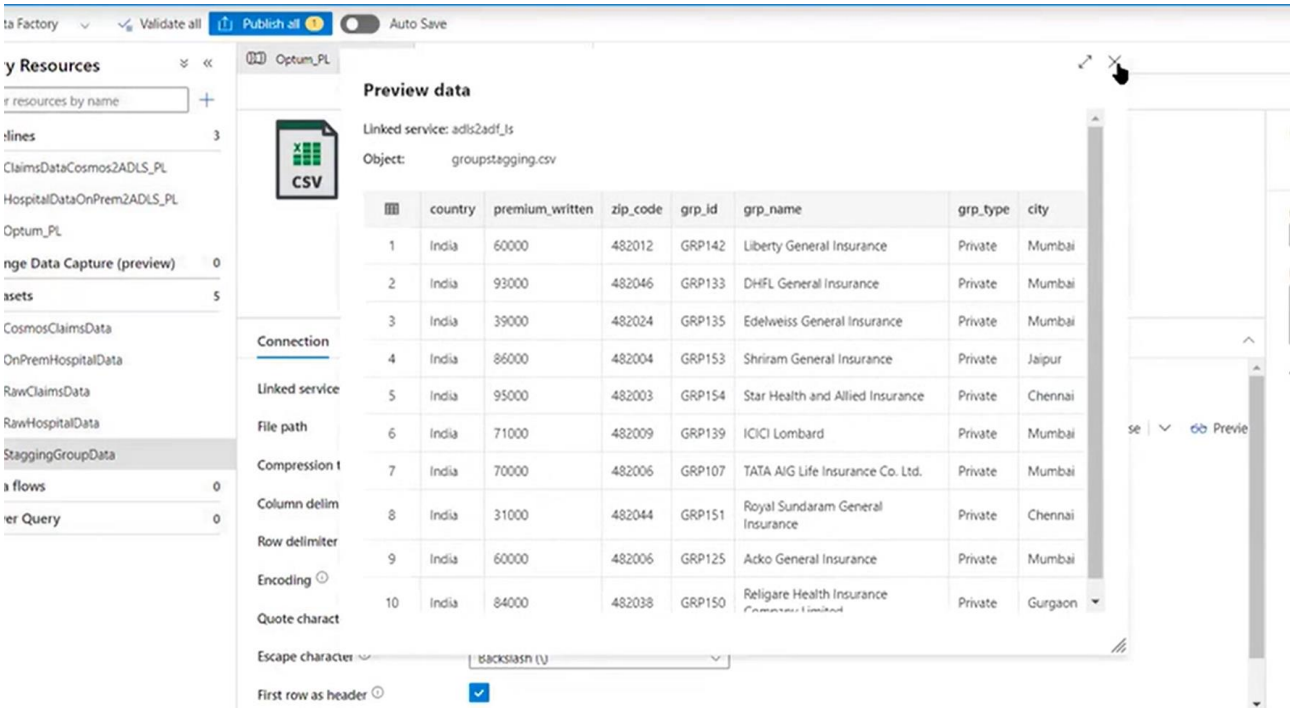
NOW ITS YOU WISH TO DO THE FINAL TRANSFOMATION EITEHR USE ADF ARE NOTEBOOK

I WOULD PREFER ADF BECAUSE IT IS EASY TO JOIN

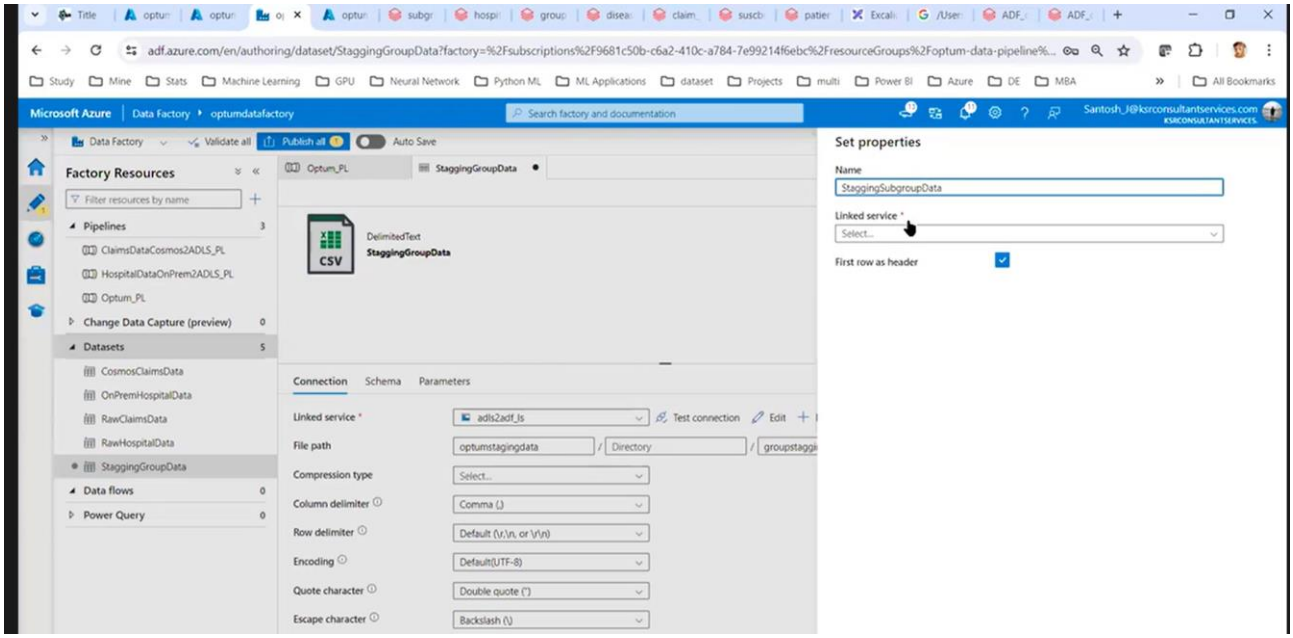
NOW THE DATA IS COMIN GFROM ADLS

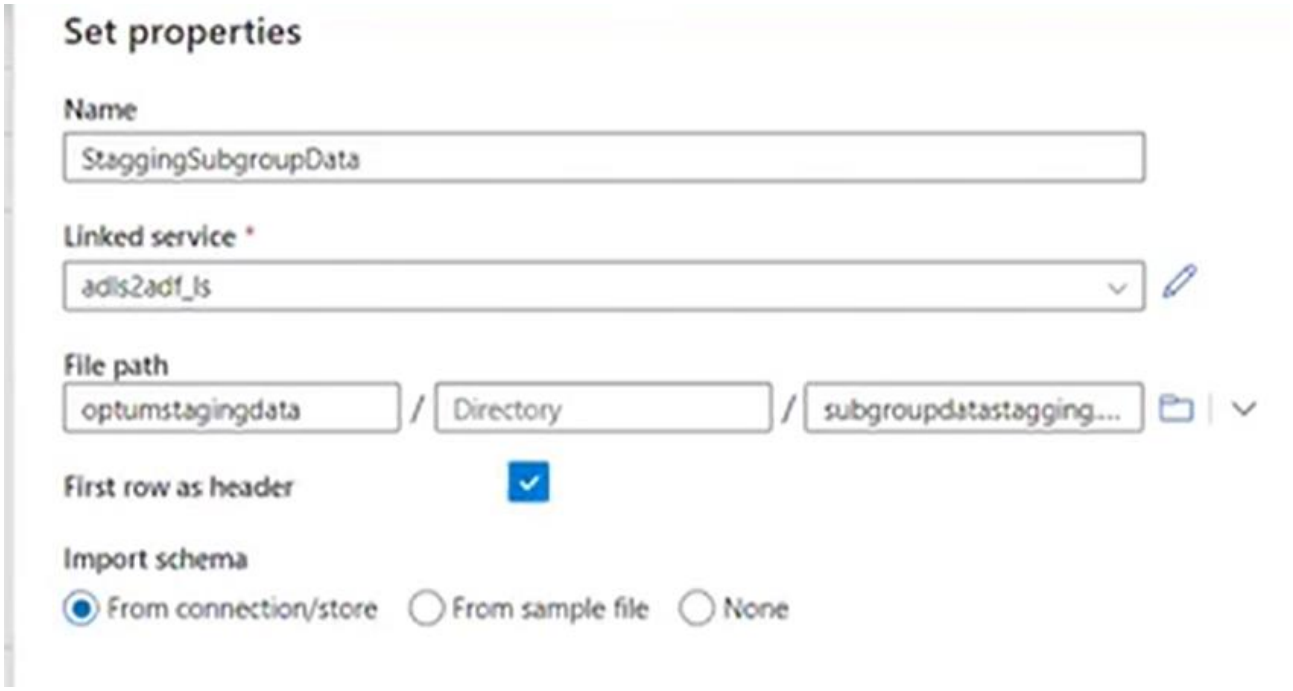




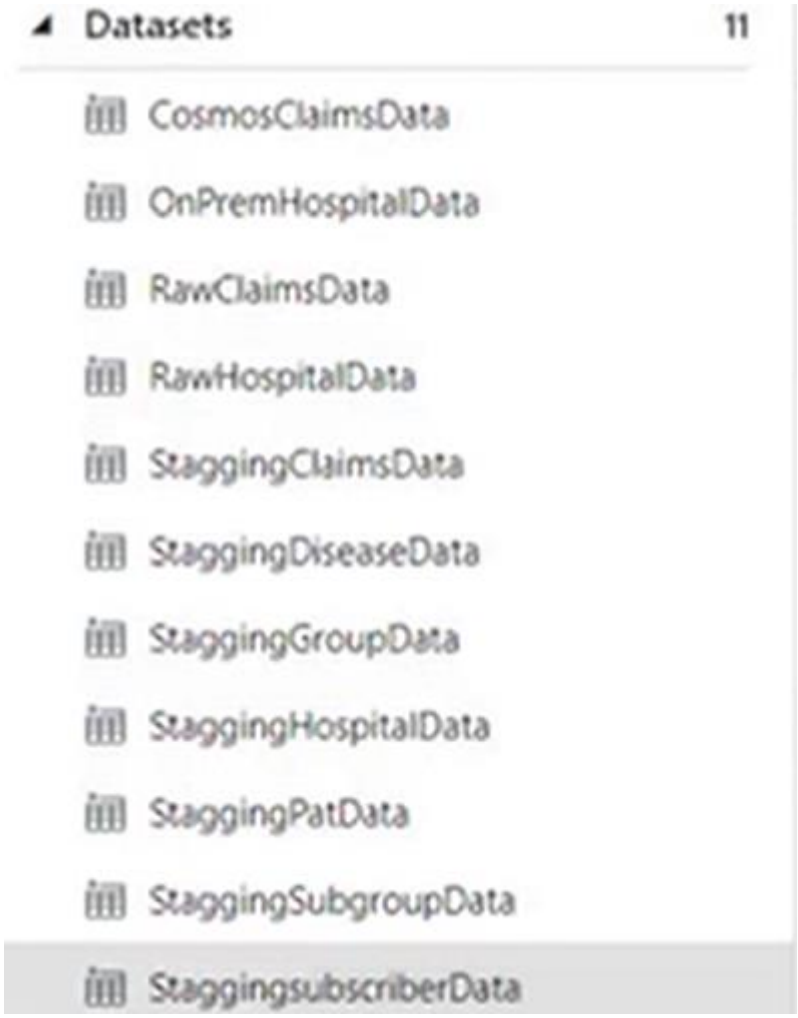


TO JOIN WE NEE A DATA SET



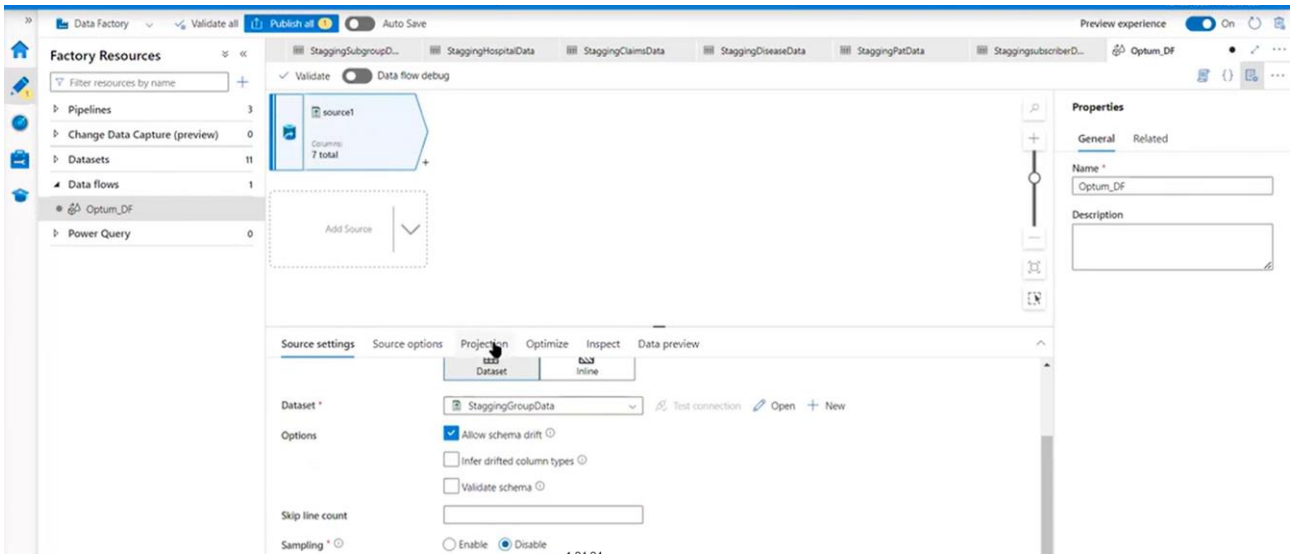


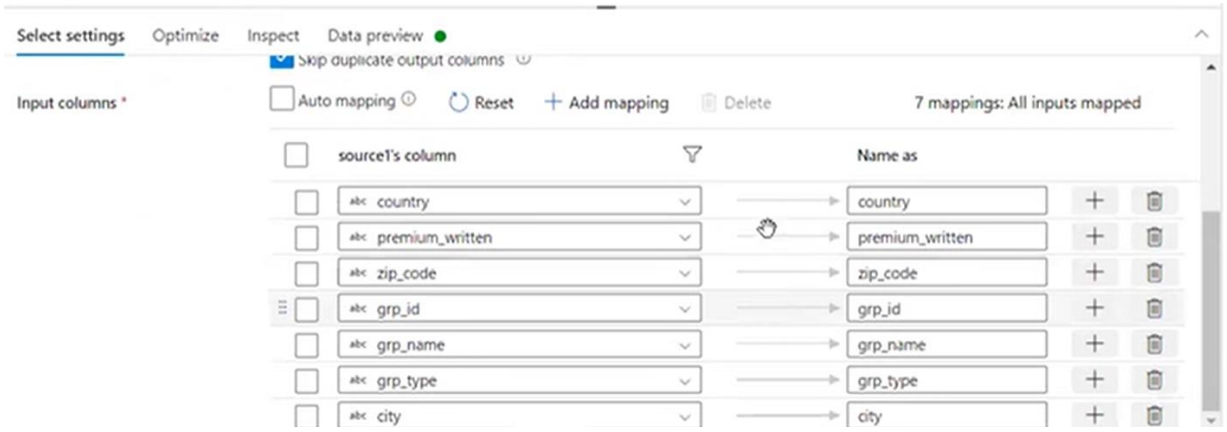
LIKE WE NEED TO CREATE DATA SET FOR ALL STAGINGDATASETS



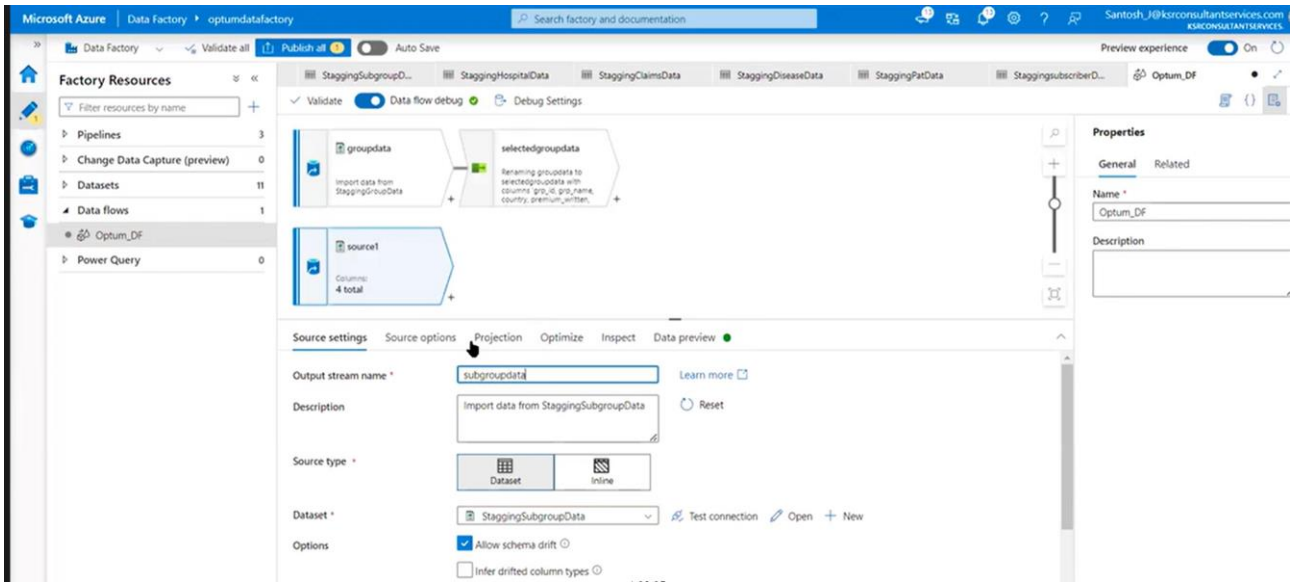
HERE IT IS DATASETS

LETS GO TO THE DATAFLOW

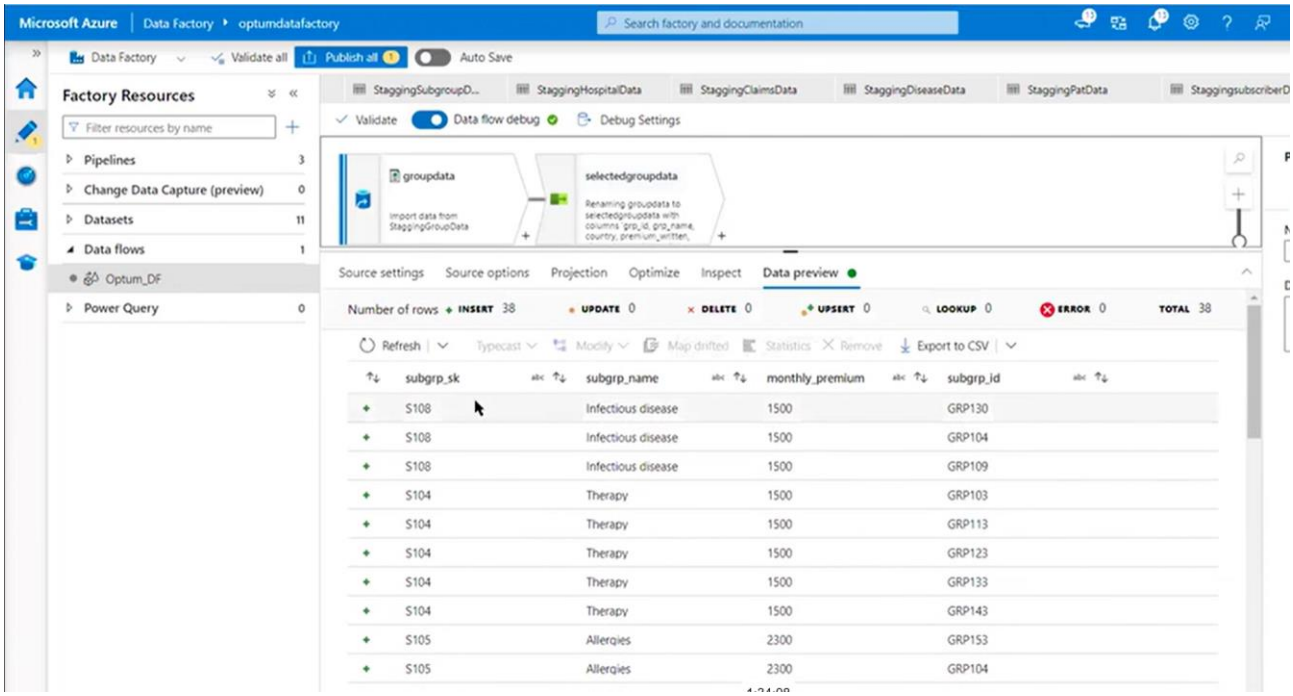


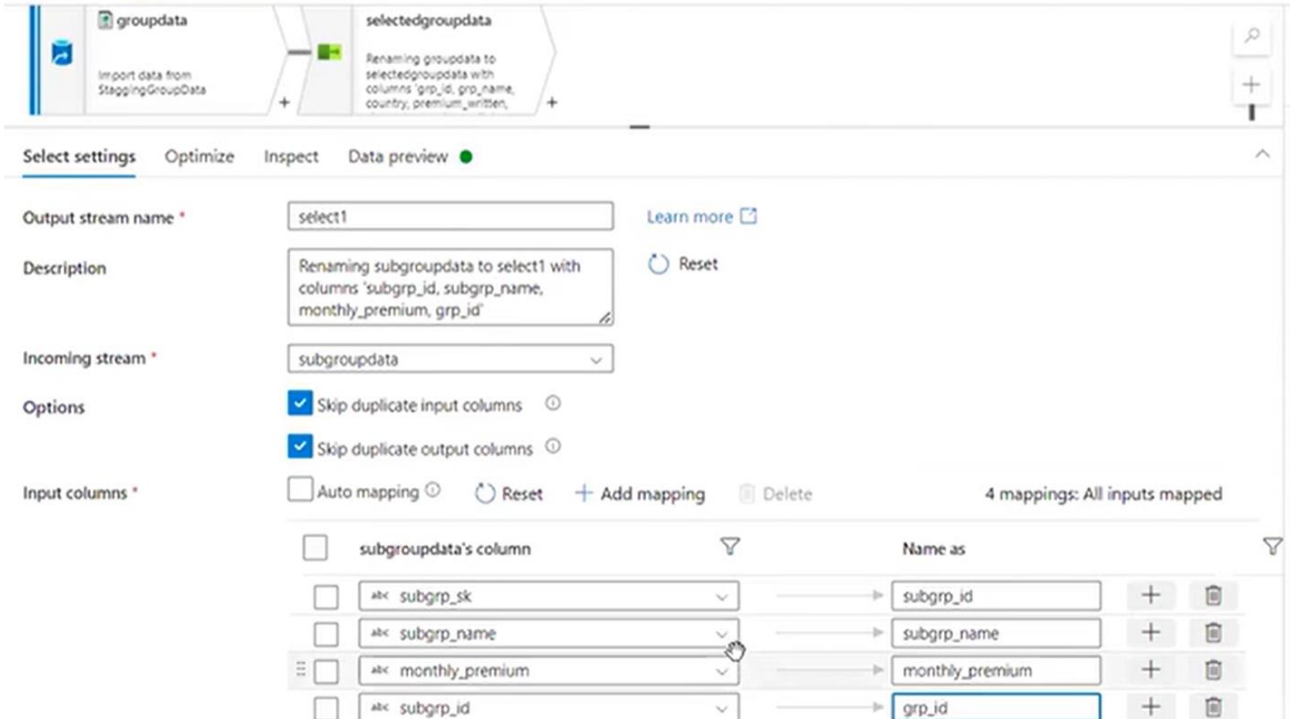


MAKE THE GROUP ID FIRST

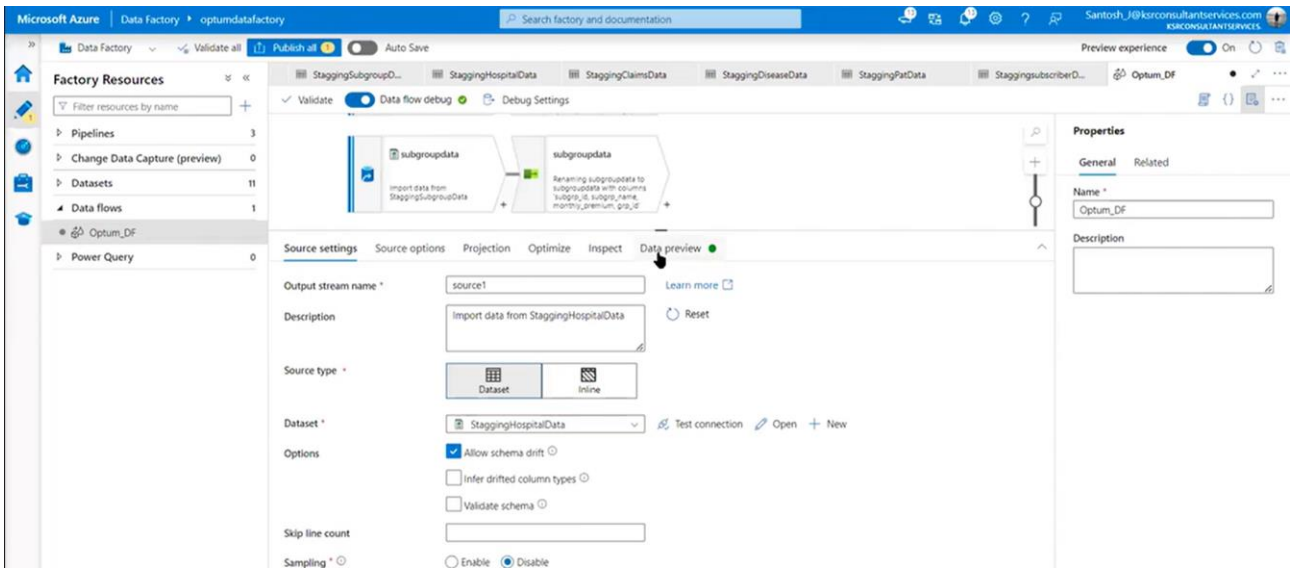


NEED TO EDIT THE COLUMN NAME



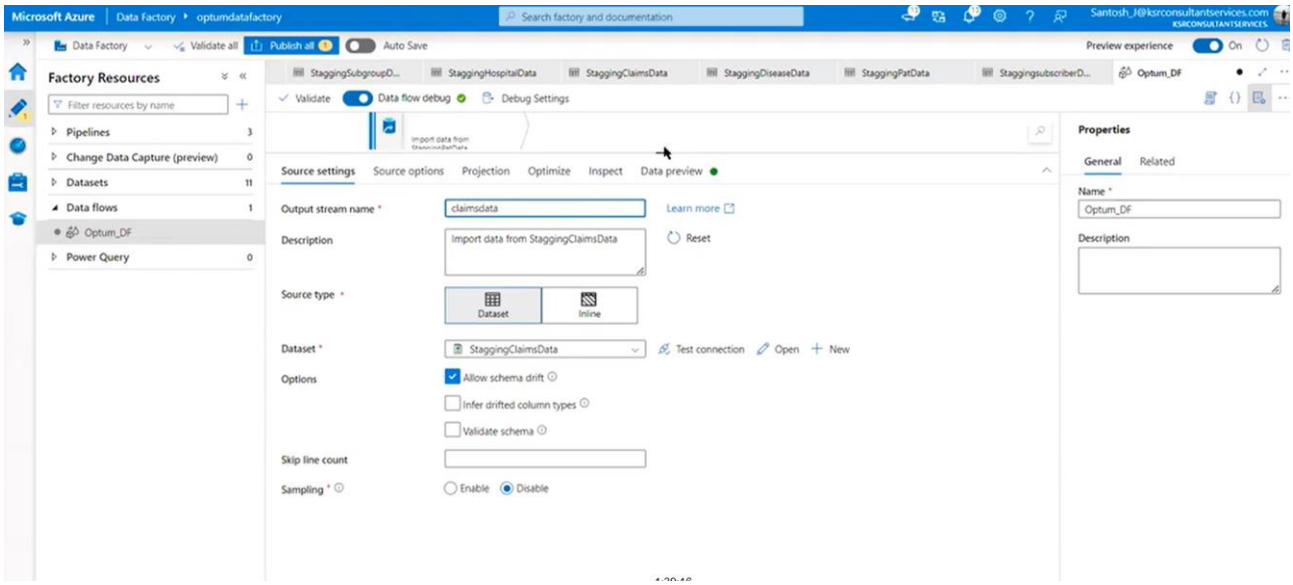


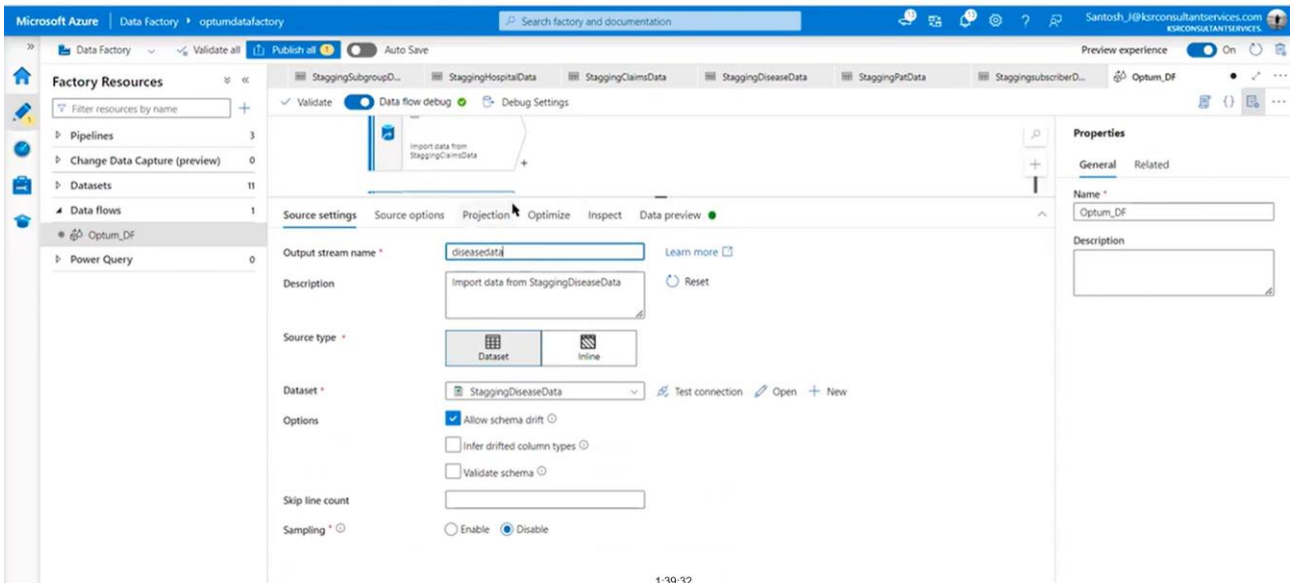
NOW DO ANOTHER HOSPITAL DATA

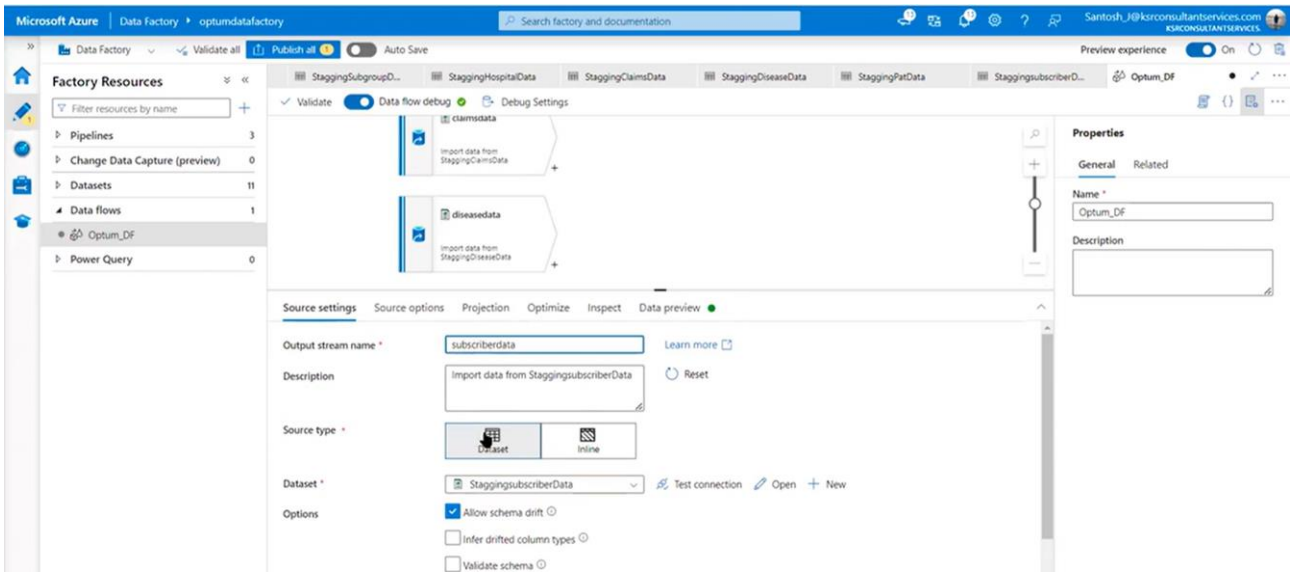


THIS IS AND ASSIGNMENT FOR U

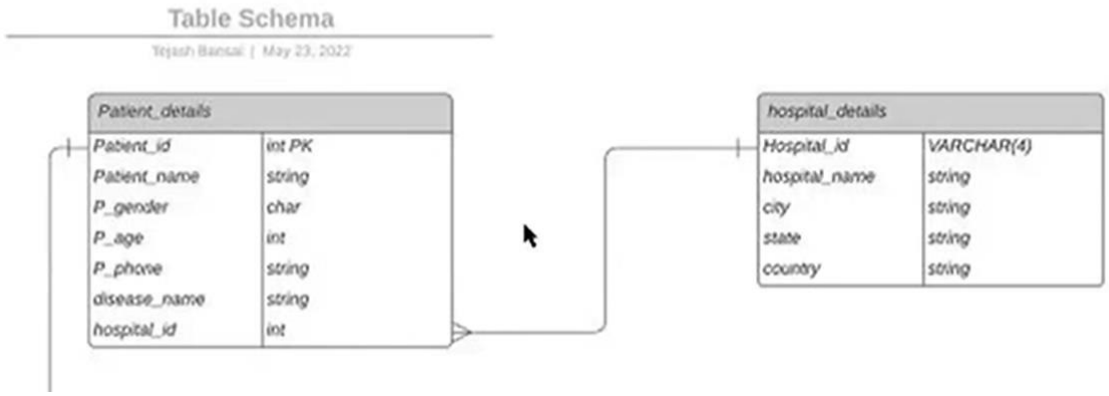
NOW AGAIN CLAIMS DATA SOURCE



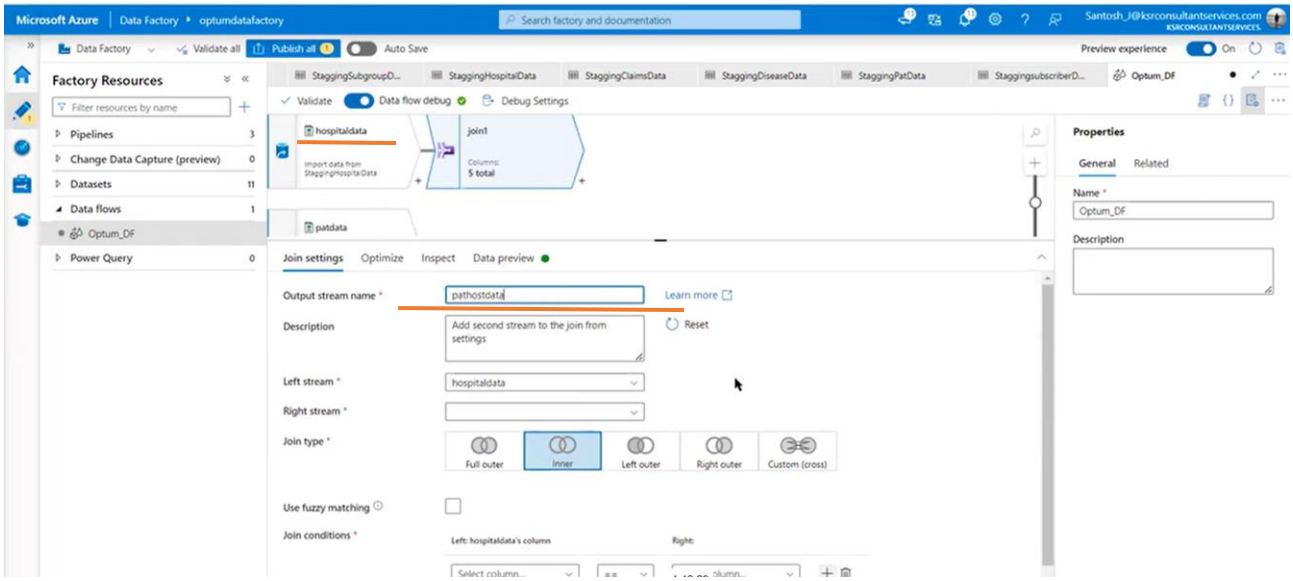


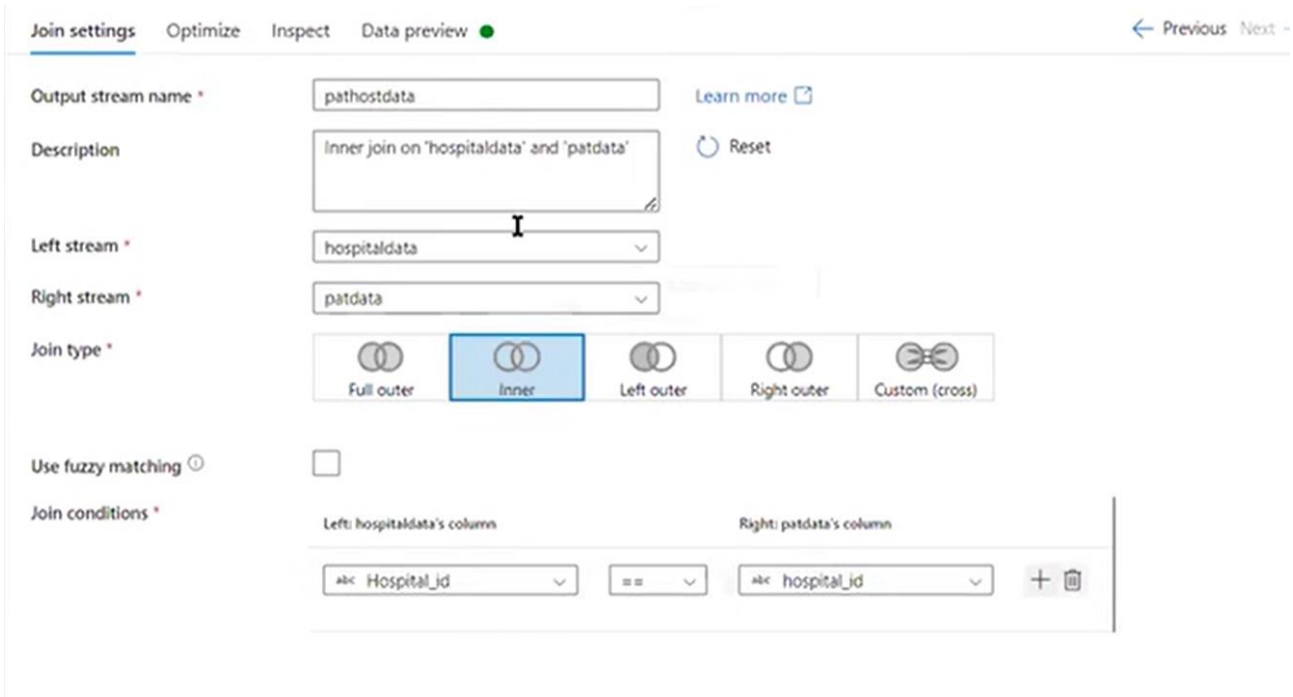


OBSERVE THIS

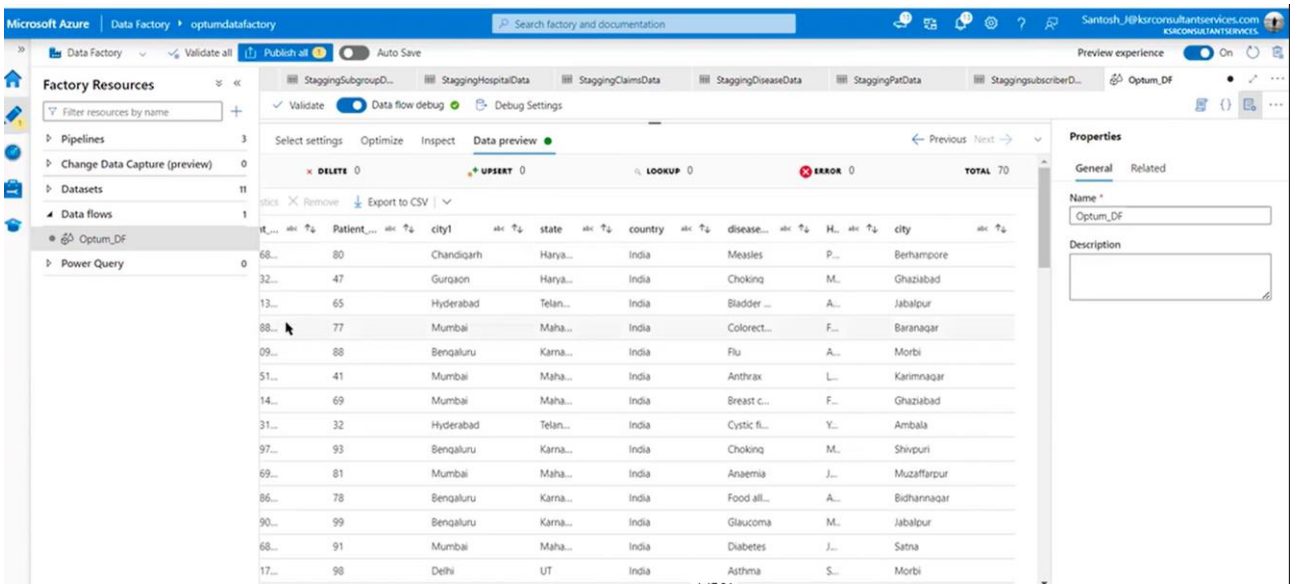


IN ONE HSPTL HAVE MANY PATIENTS



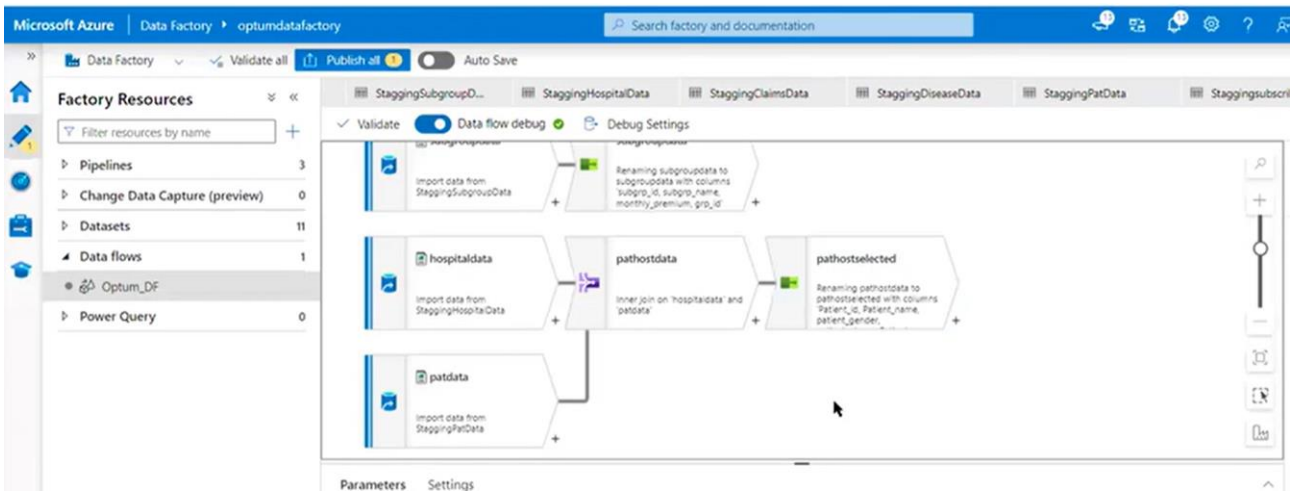


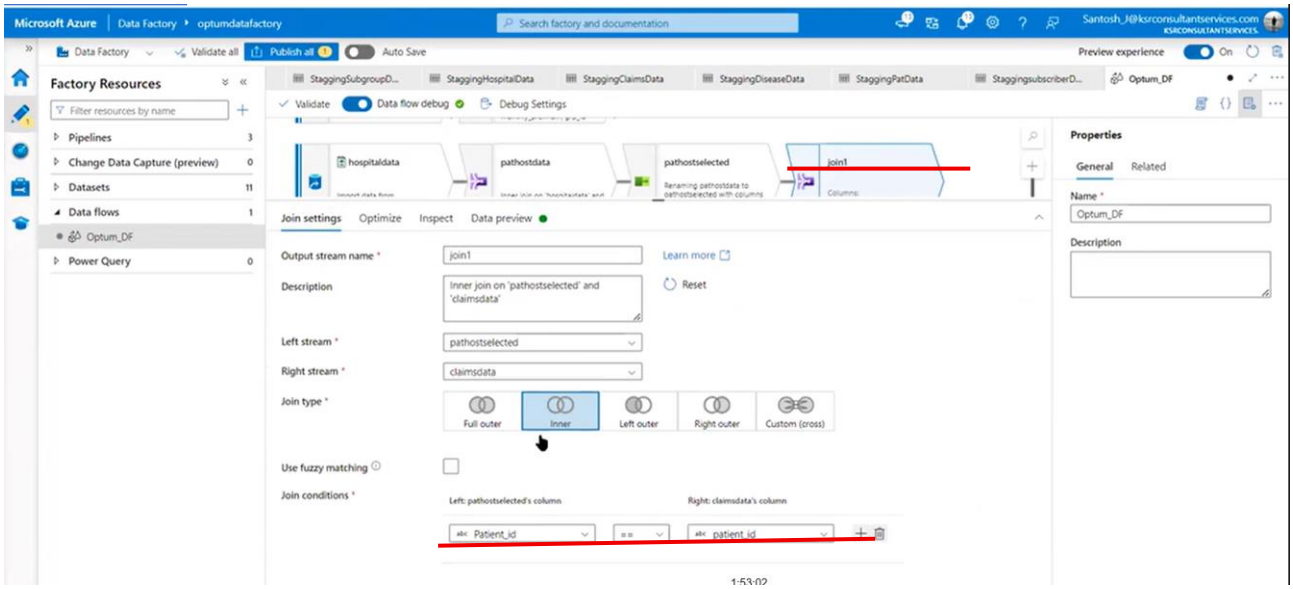
WWE CITY IN 2 TIMES

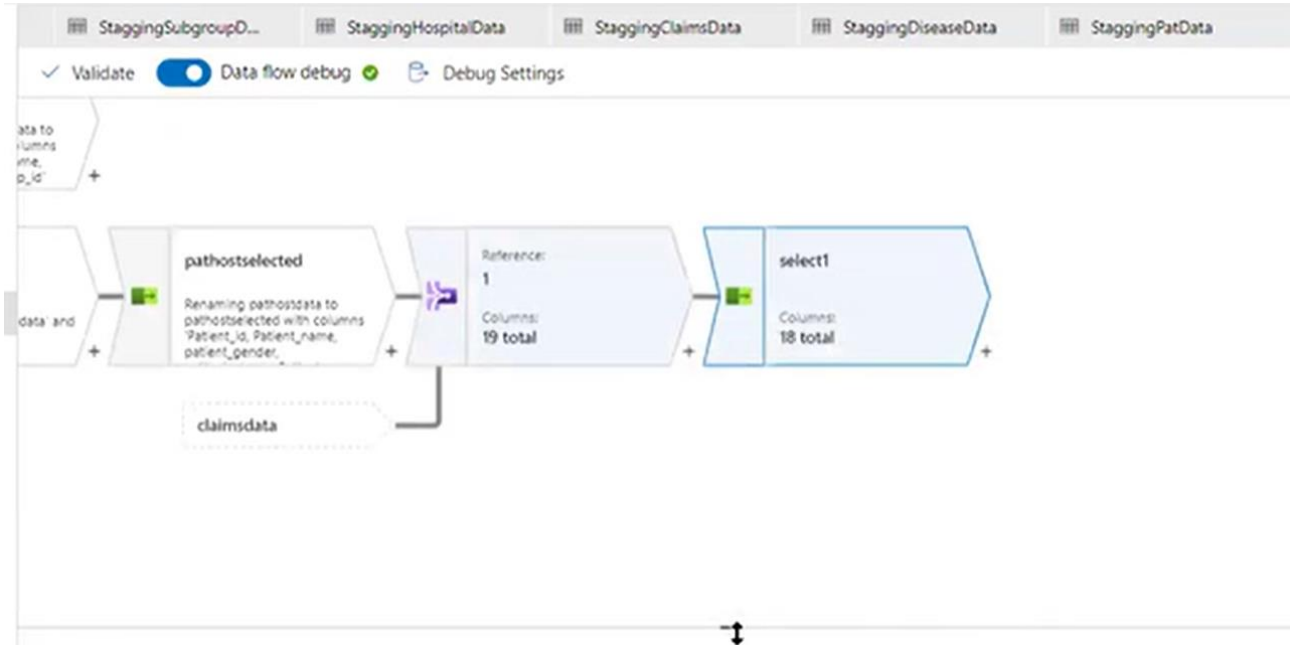


NAME IT CLEAR ONE IS HSPTL CITY ,ONE AS PATIENT CITY

IF BOTH AS SAME VALUE I WILL REMOVE



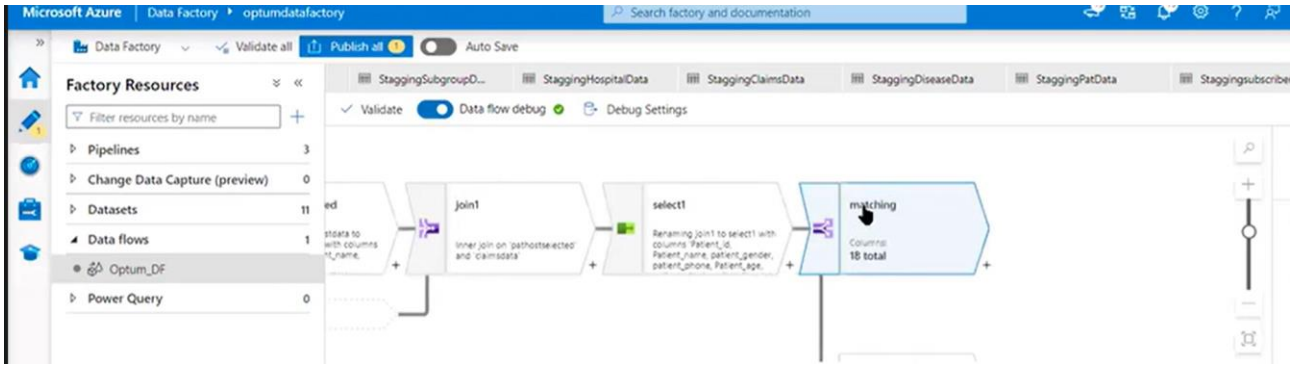


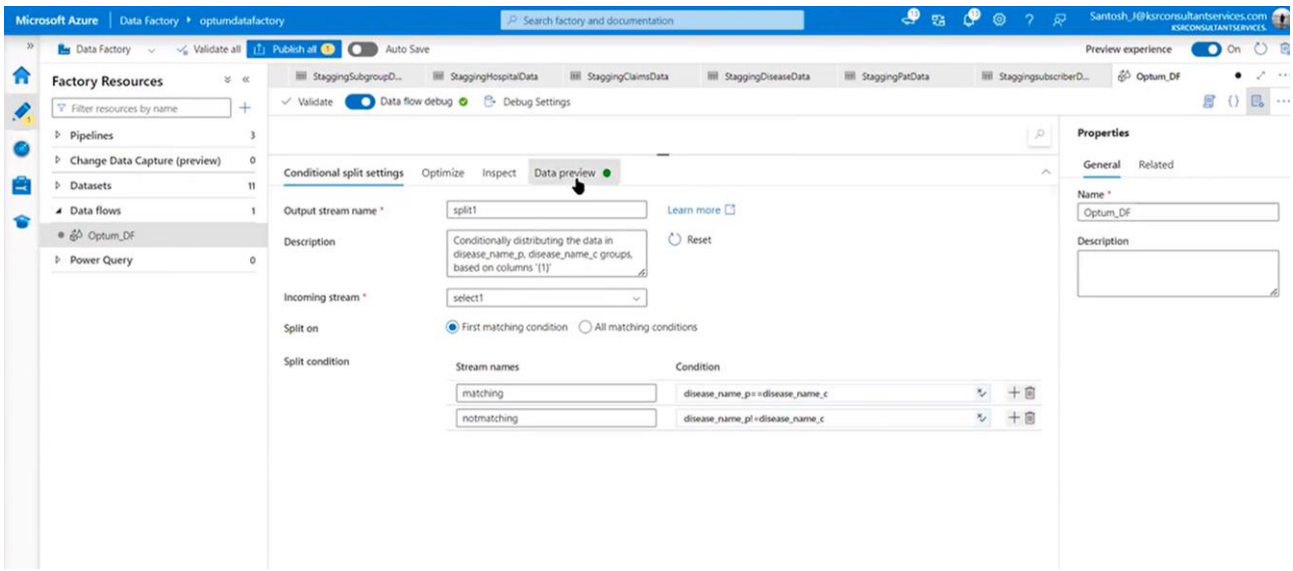


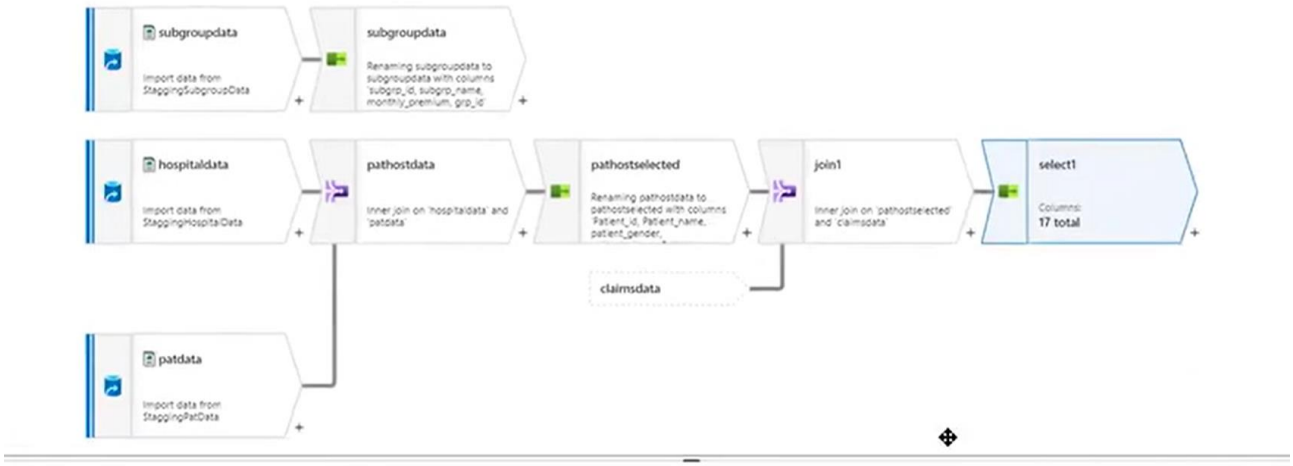
CHECK ALL THE COLUMNS

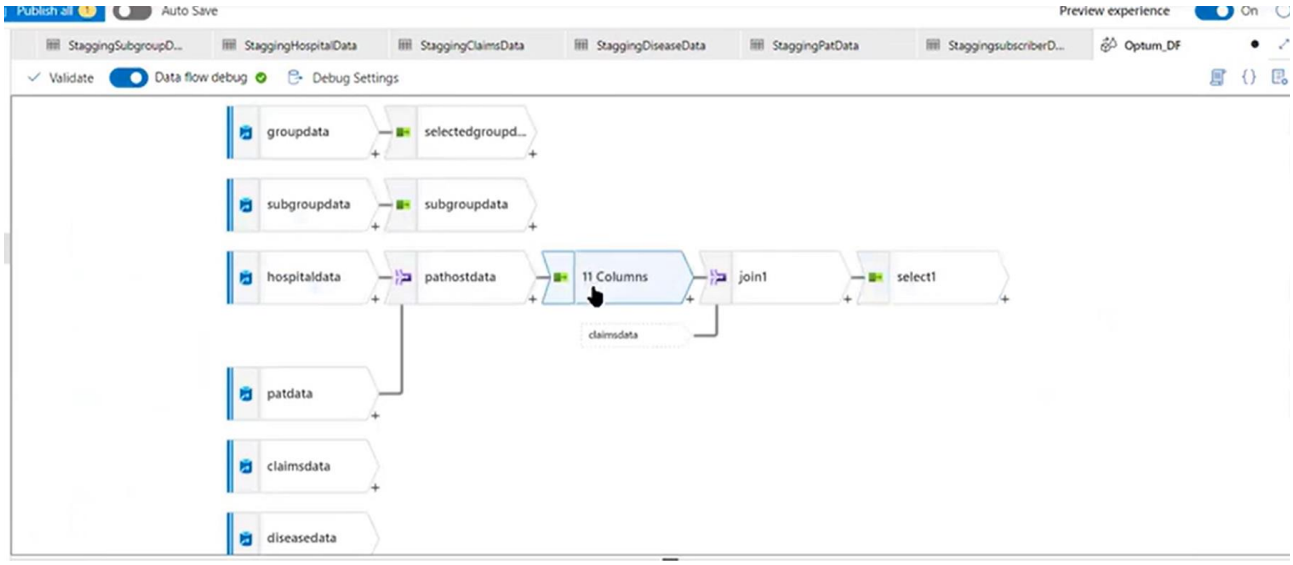
CHECK WHILE IT AS SAME

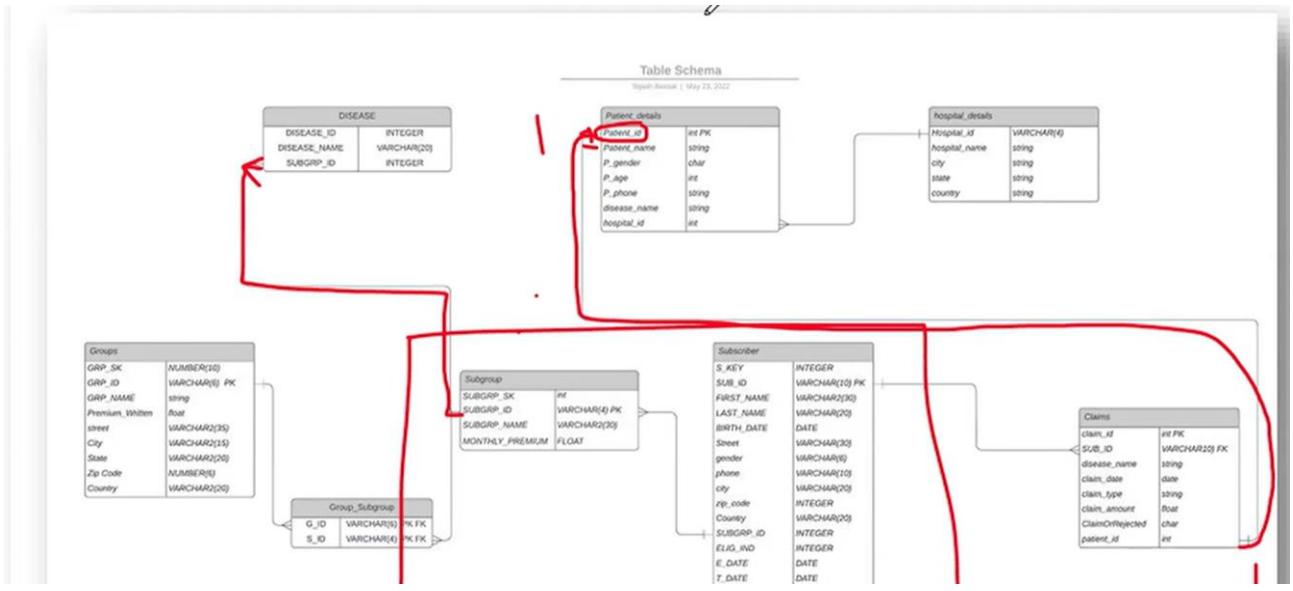
WITHOUT DOING MANUALLY WE CAN USE THE CONDITIONAL ACTIVITY





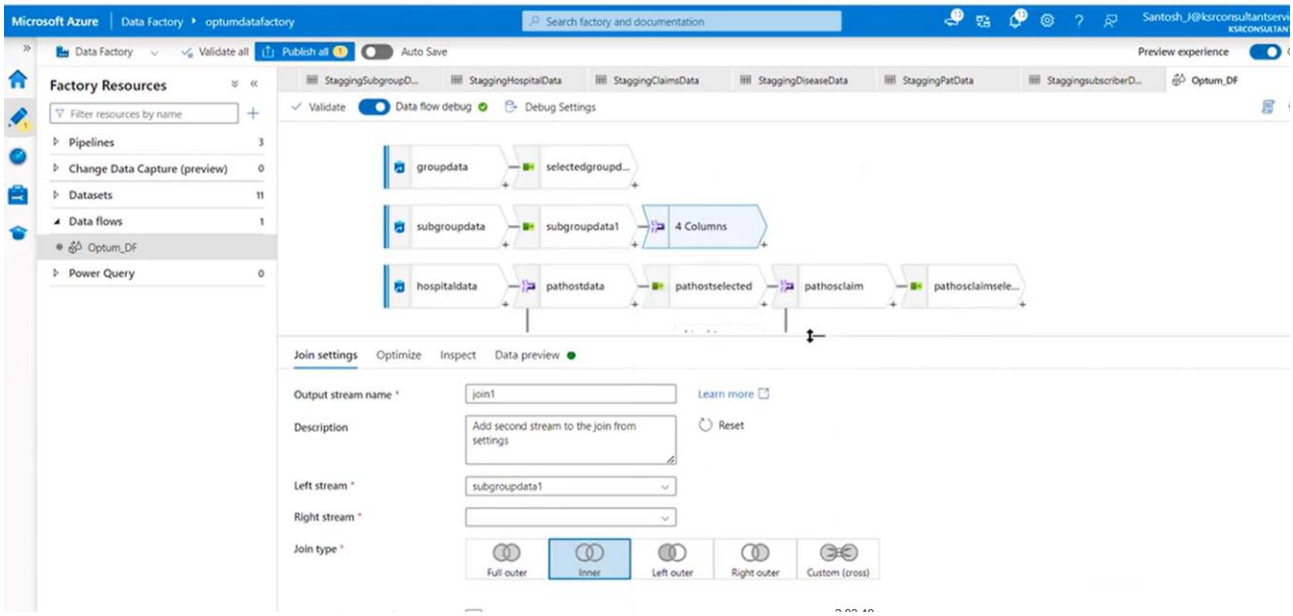


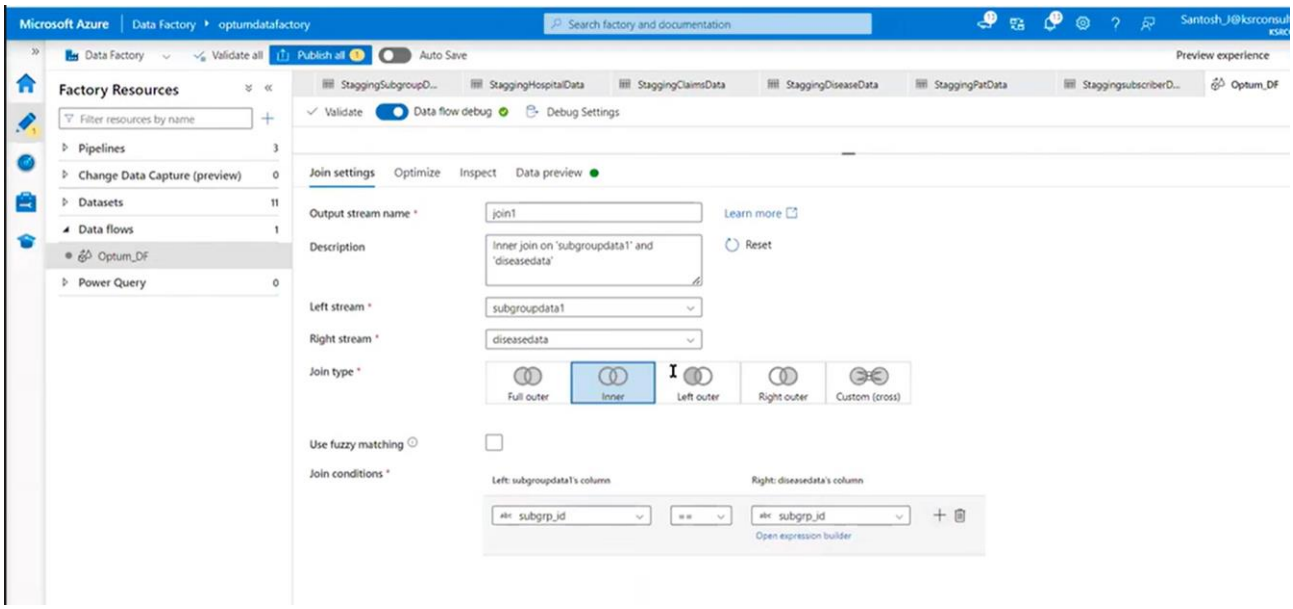




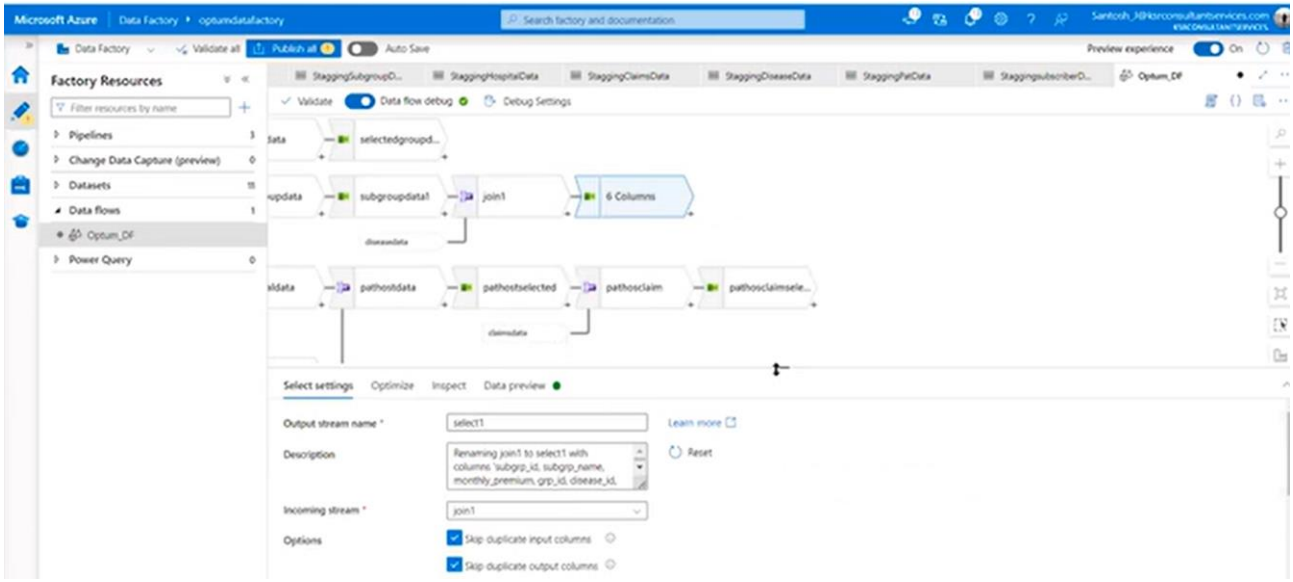
NOW LETS GO TO DISEAS DATA SOURCE

LETS JOIN THIS WITH SUBGROUP DATA

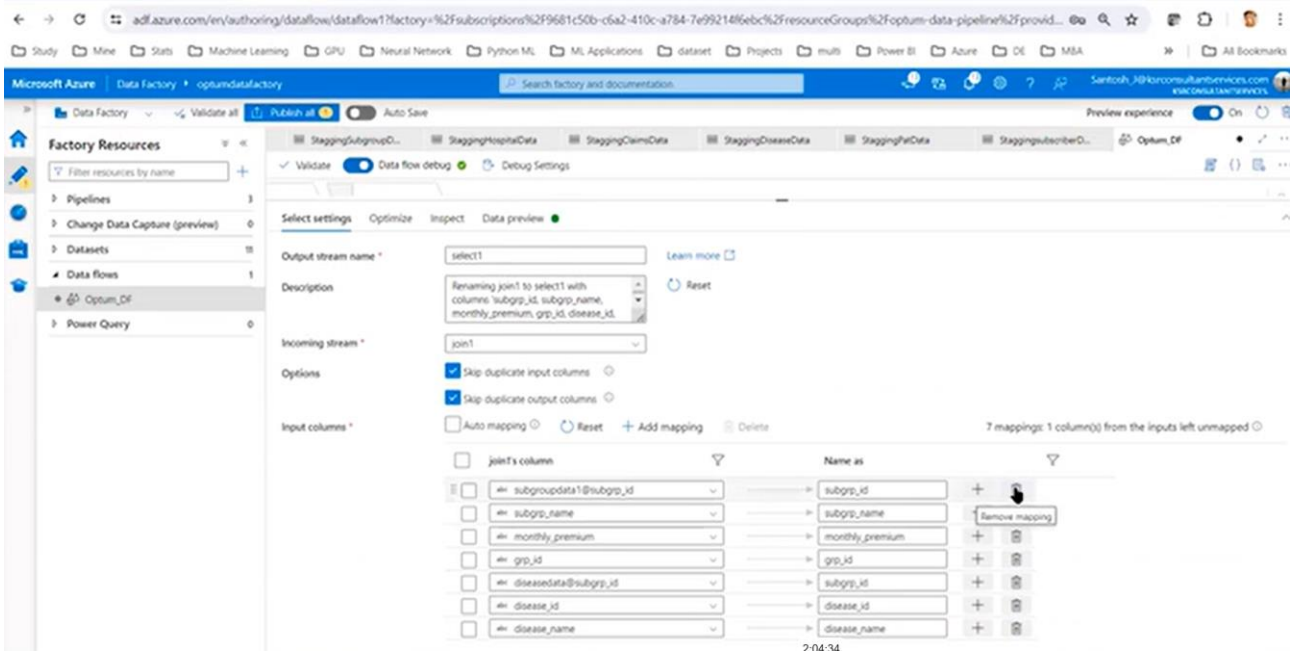




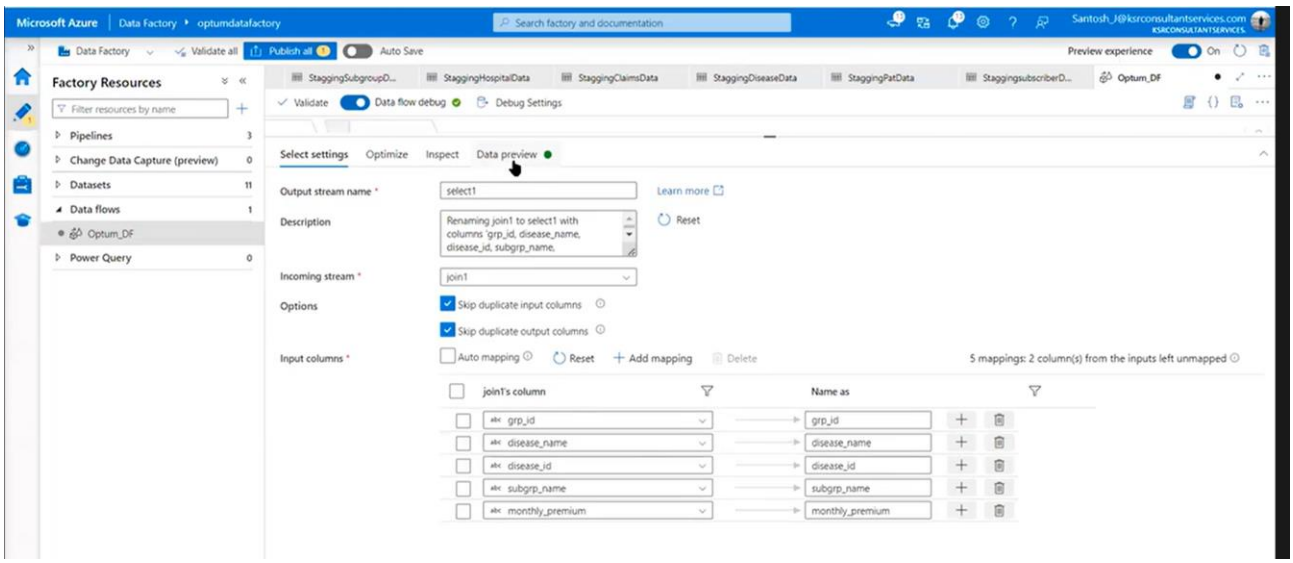
We get the datapreview and we got the groupid in both tables so we need to remove this group id

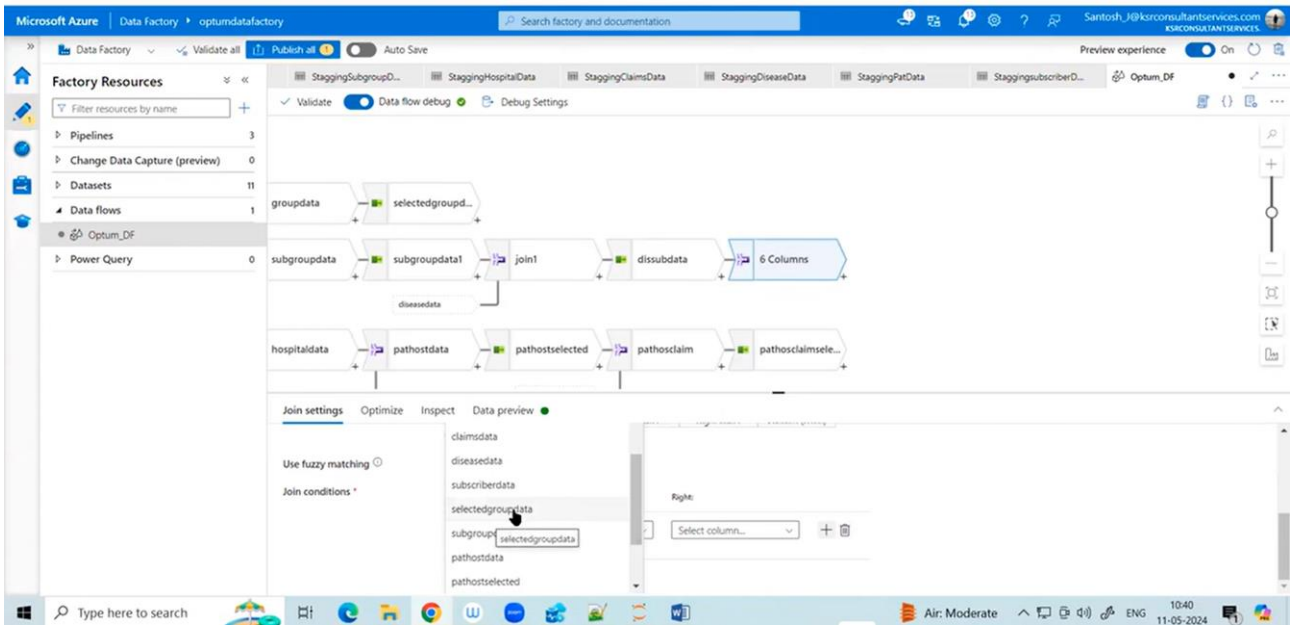


We use the selected activity

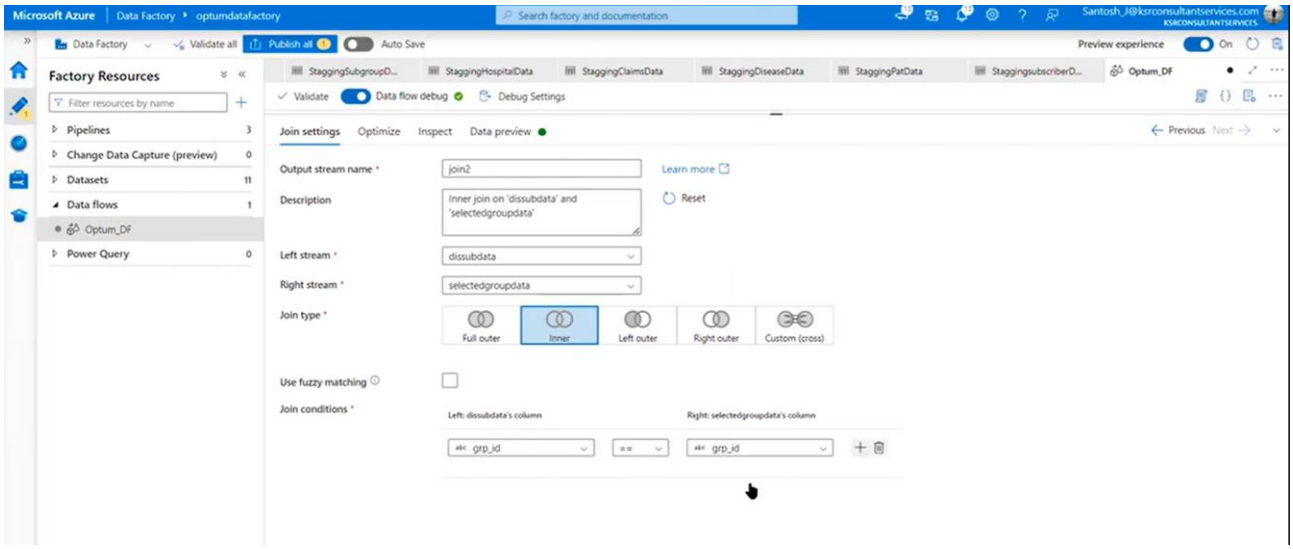


Arrenge like this

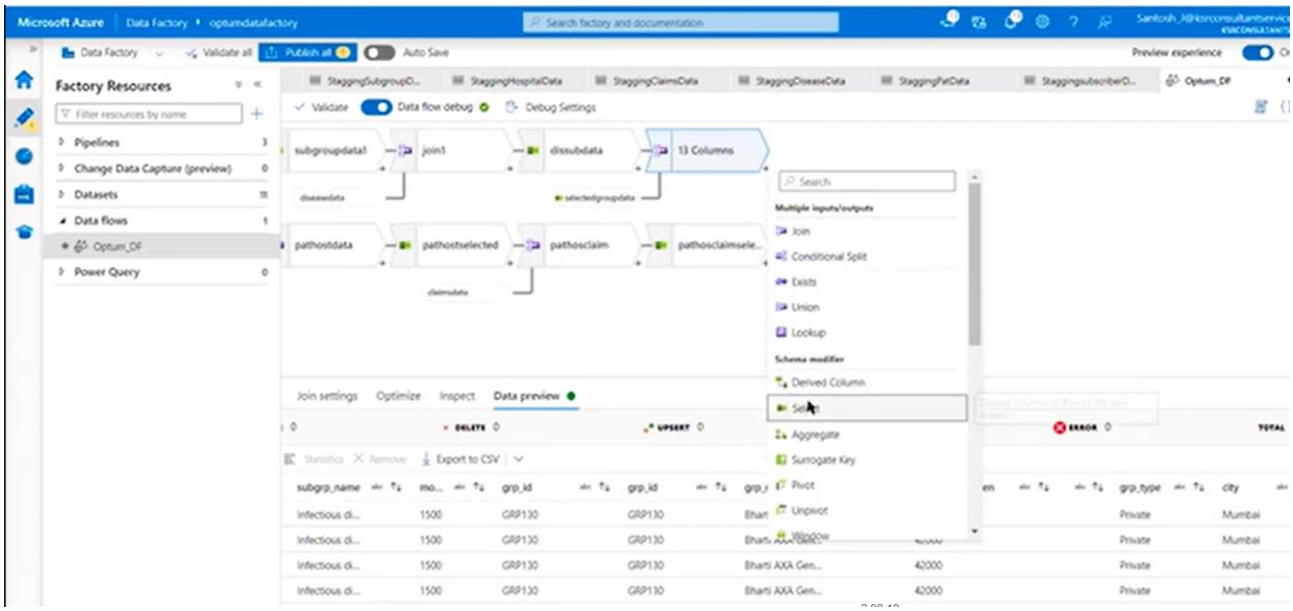




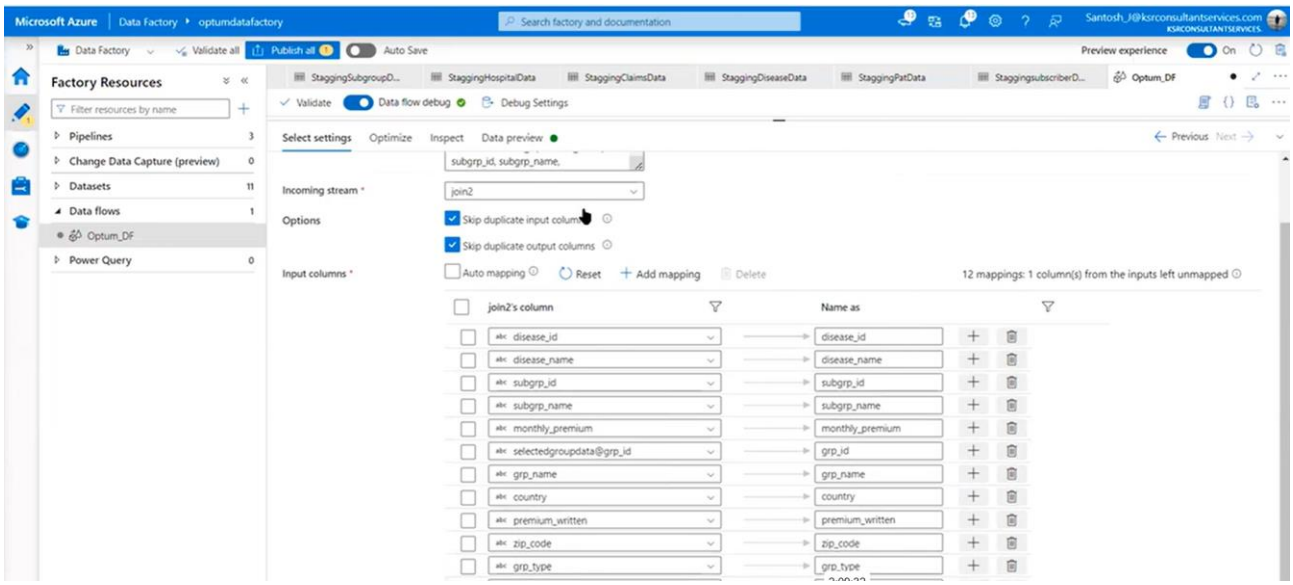
JOIN IT WITH GROUP ID

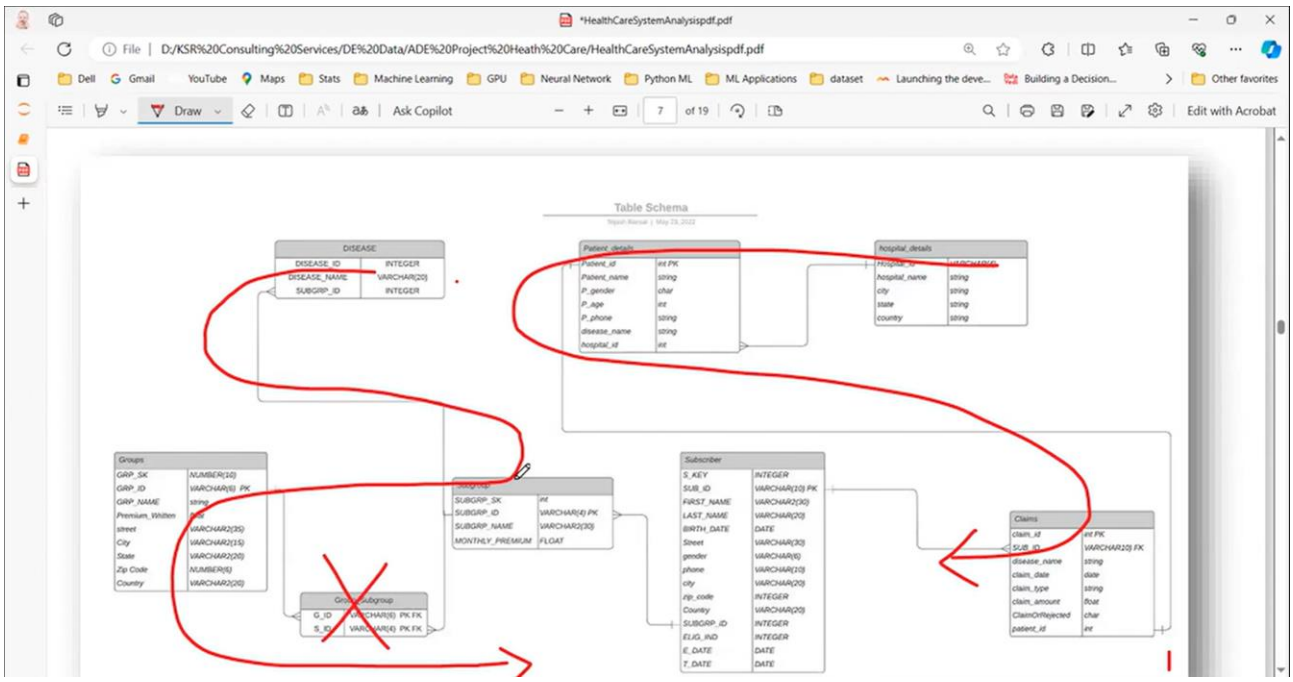


NOW USE THE SELECTED COLUMNS AND REMOVE SOME COLUMNS



THIS WE WILL KEEP LIKE THIS

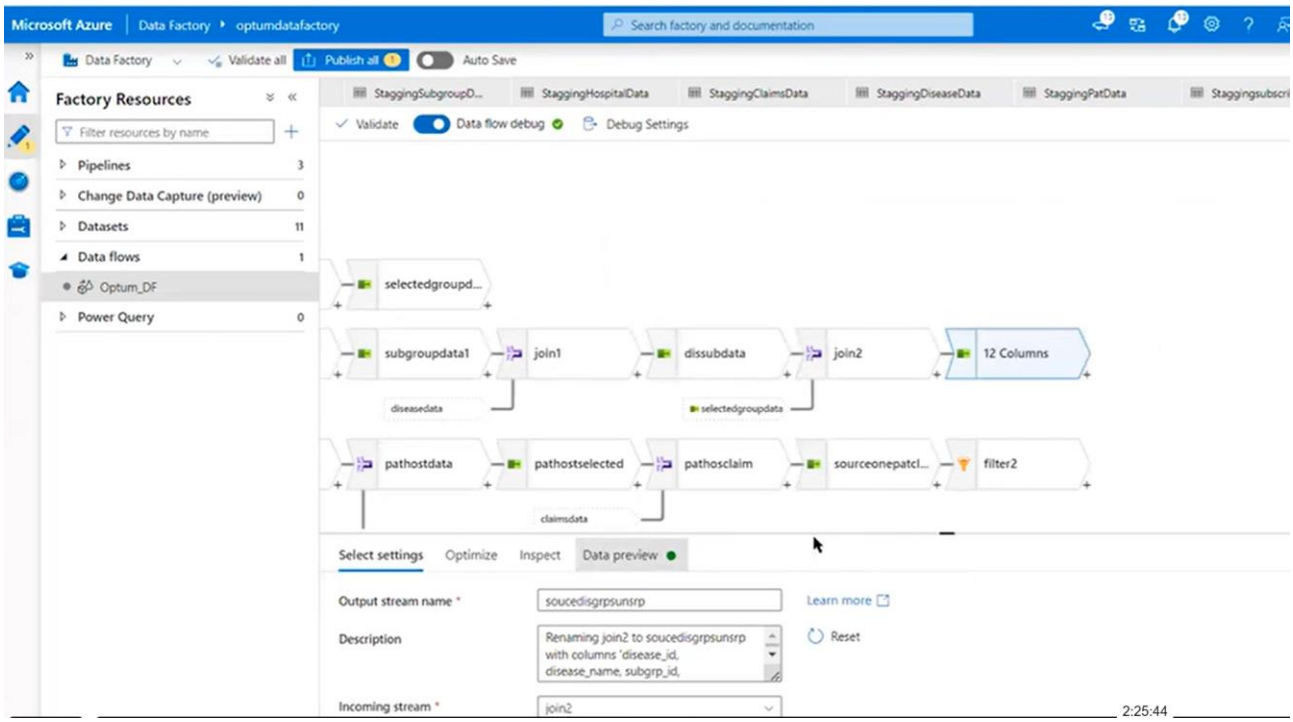




WE HAVE DONE LIKE THIS

NOW LETS GO TO THE SUBSCRIBER DATA.

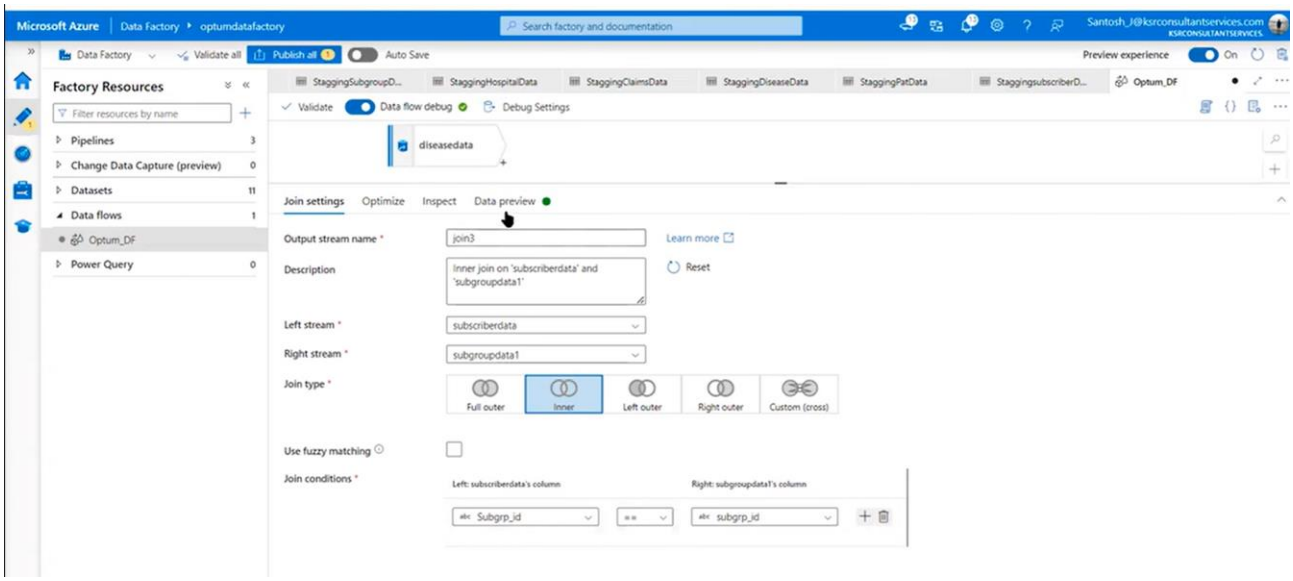
IN SUBSCRIVER DATA WE GET THE SUBGROUP DATA FROM SUBGROUPDATASET



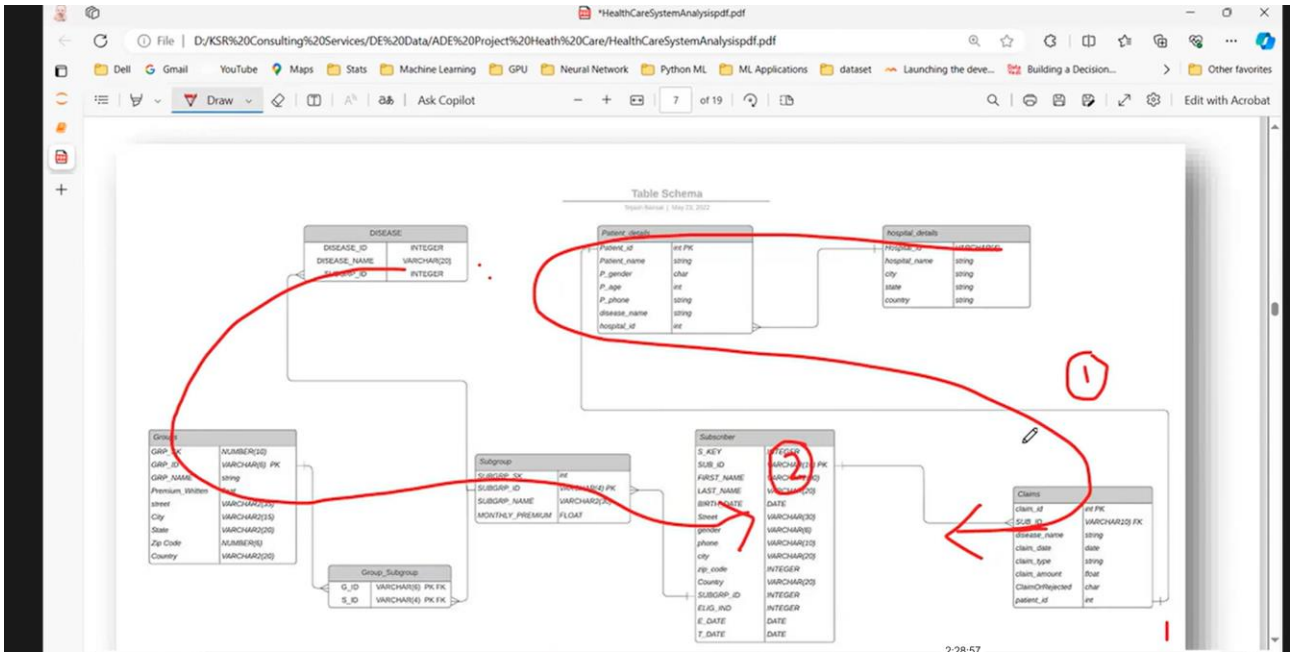
12 COL IS A SUBGROUP ID

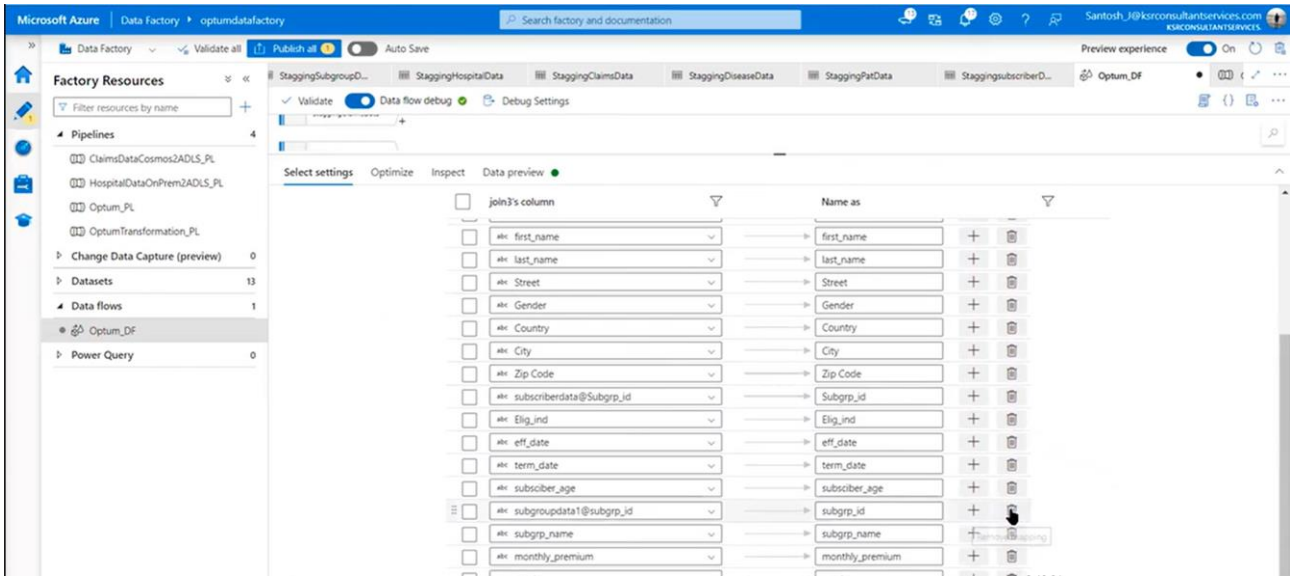
GET THE SUGGROUP ID INFORMATION TO SUBSCRIBER DATA

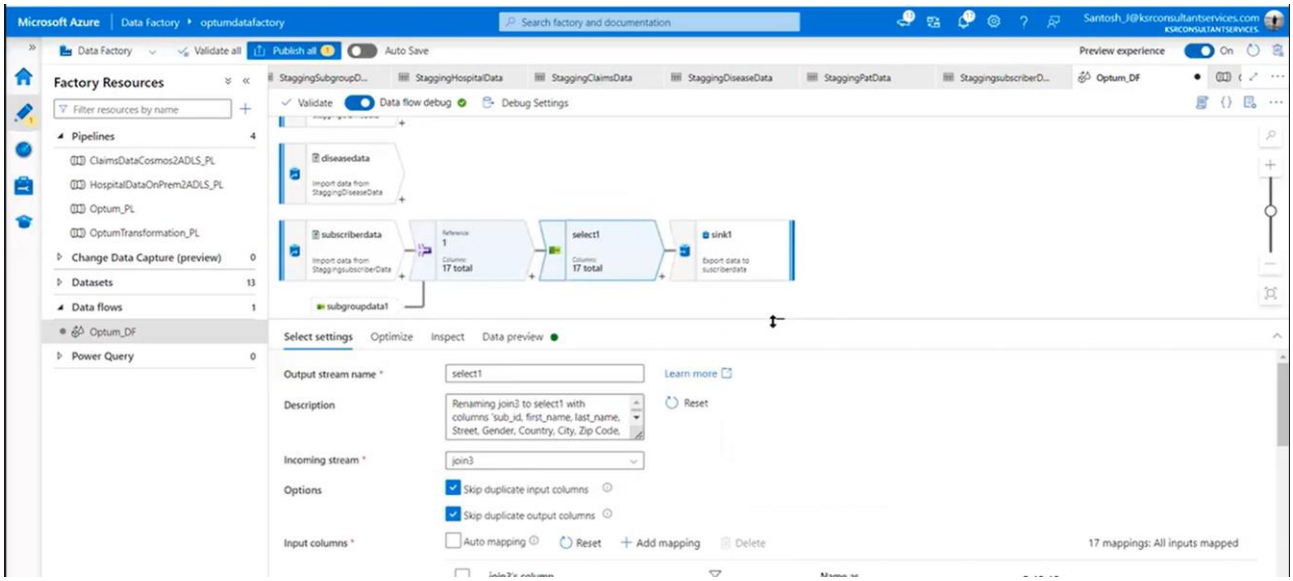
USE JOIN

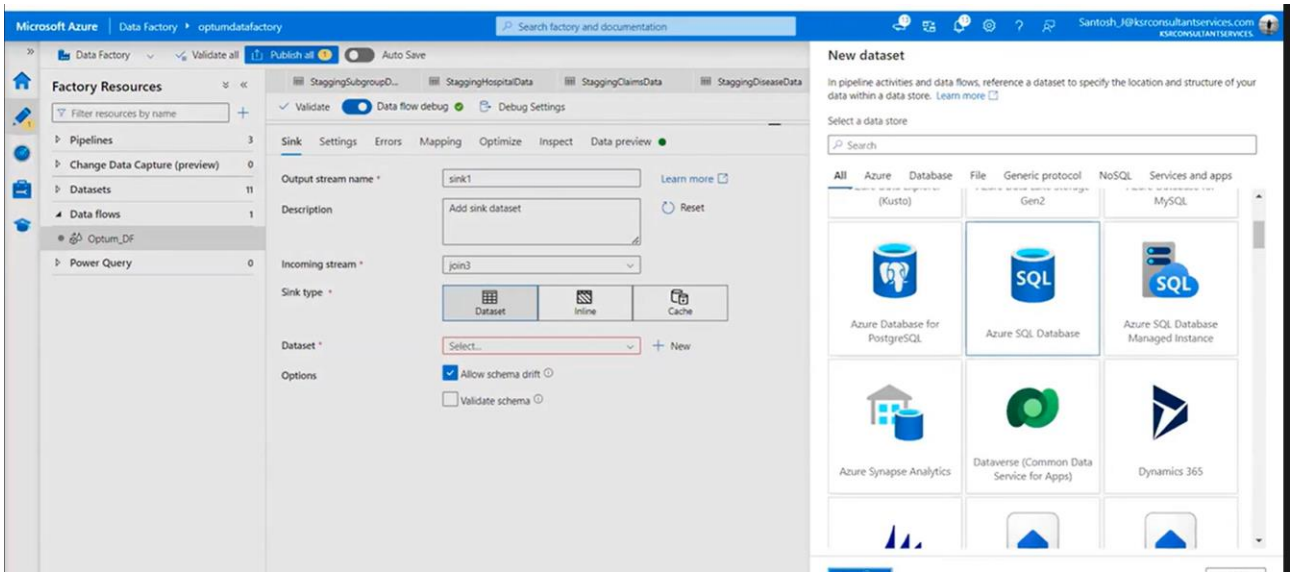


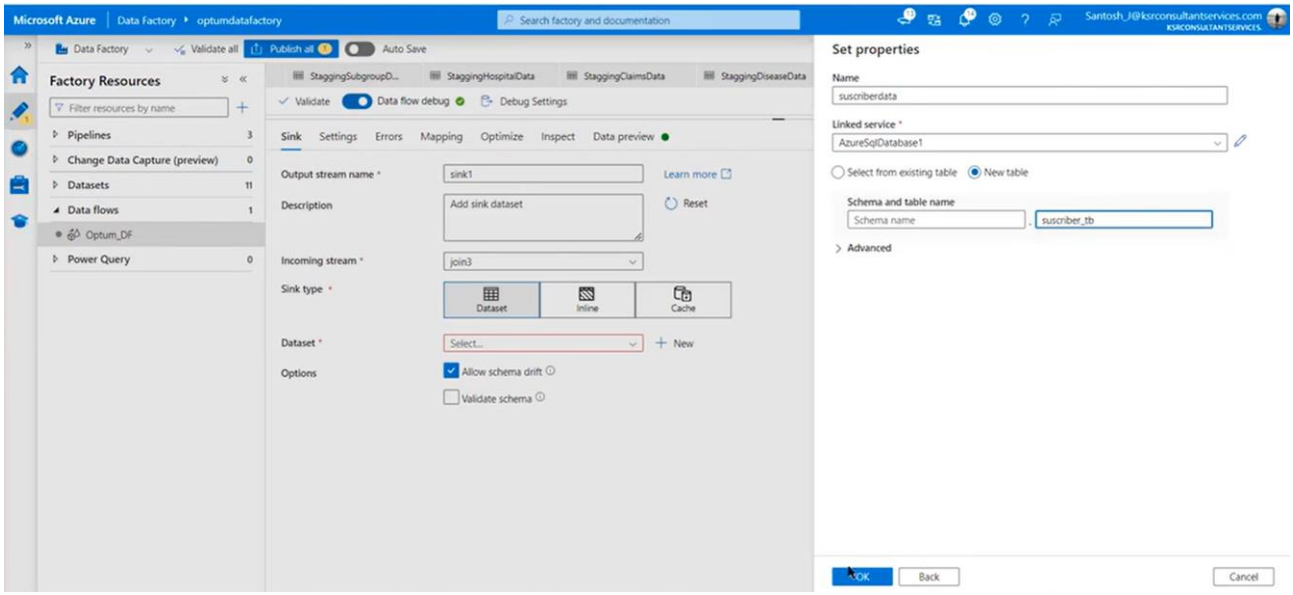
WE HAVE DONE THE 2 TABLES



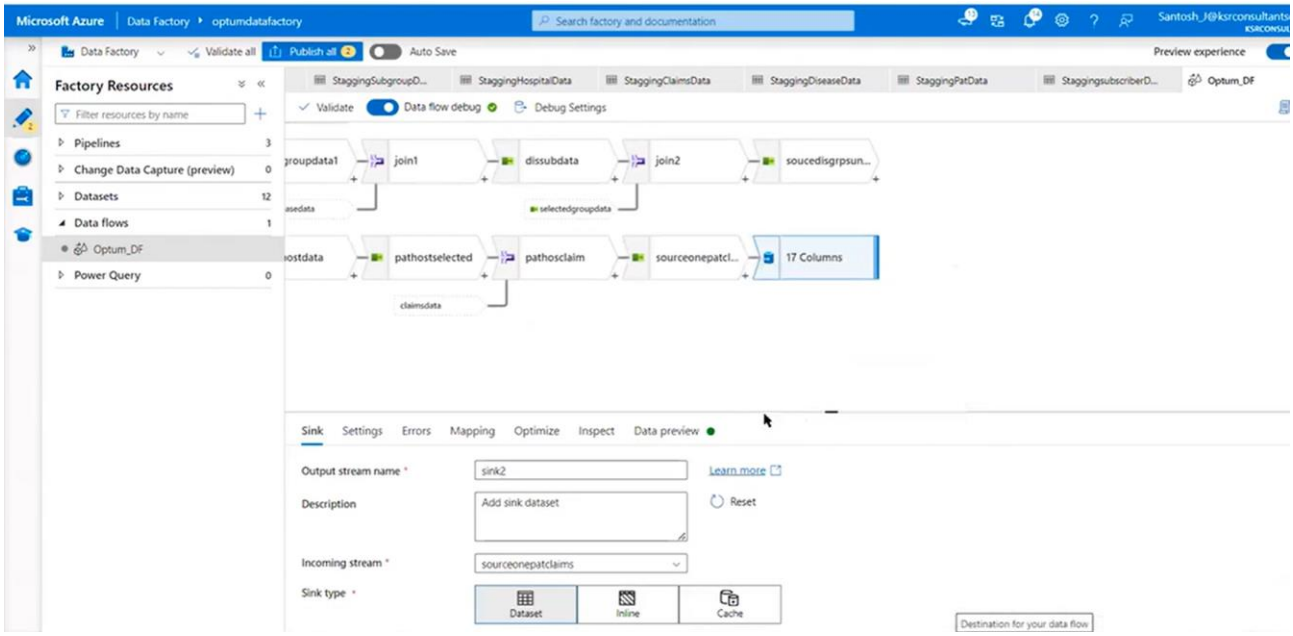


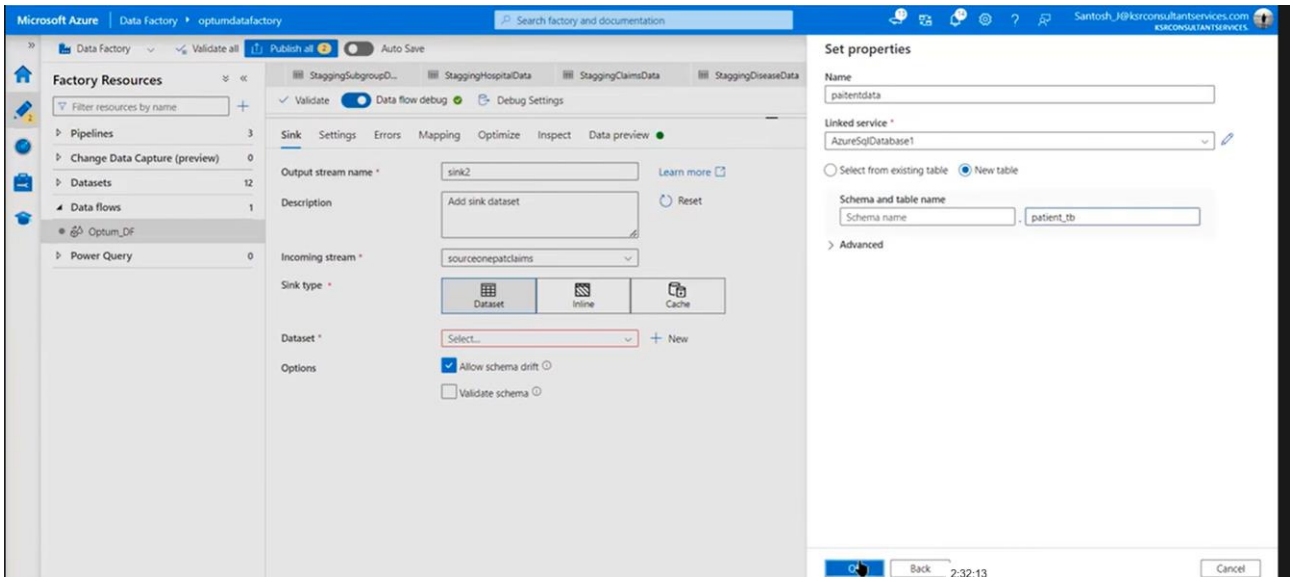




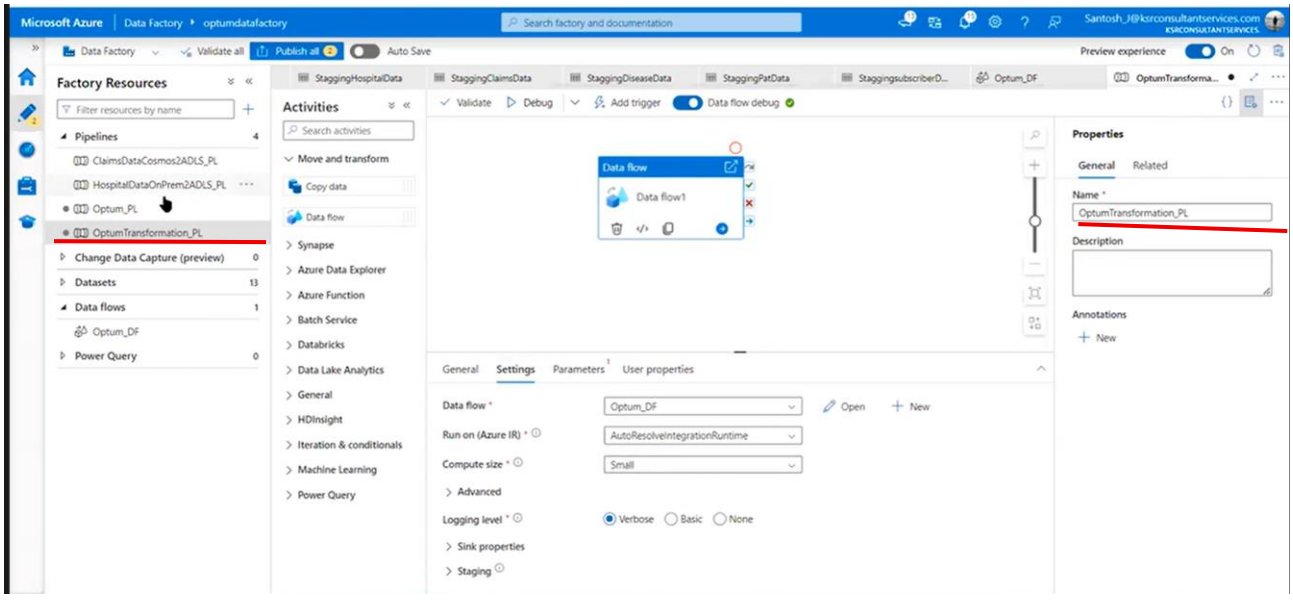


NOW ADD THE ANOTEHR SINK

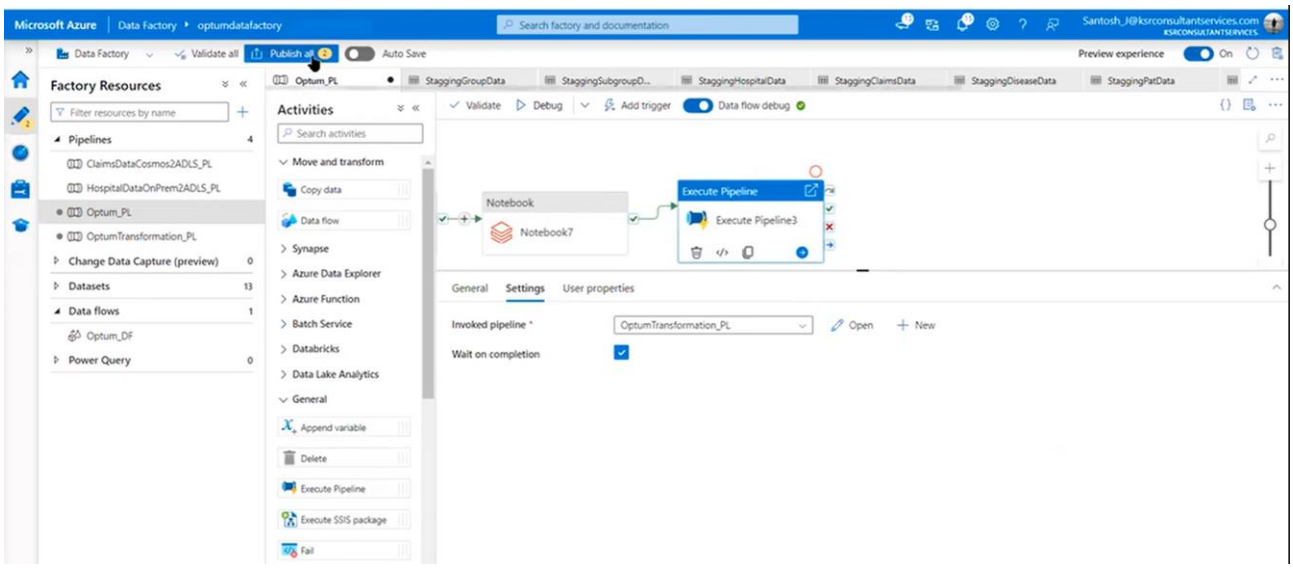




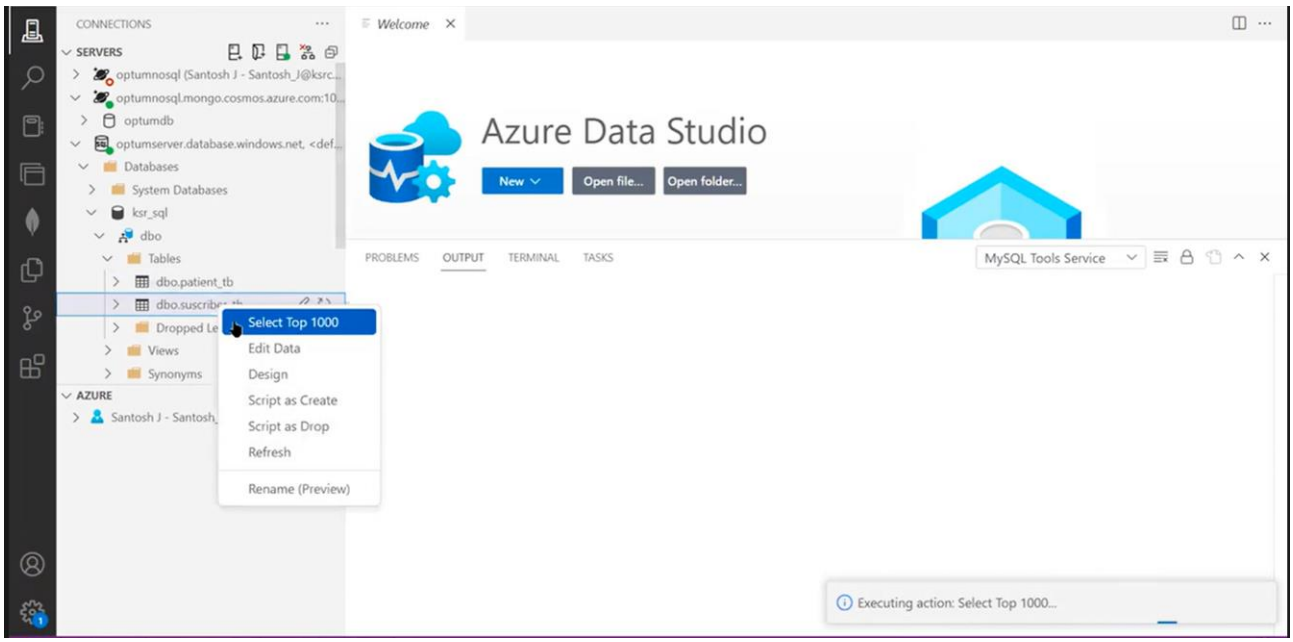
CREATE AN ANOTHER PIPELINE AS OPTUM_TRANSFORMATION_PL

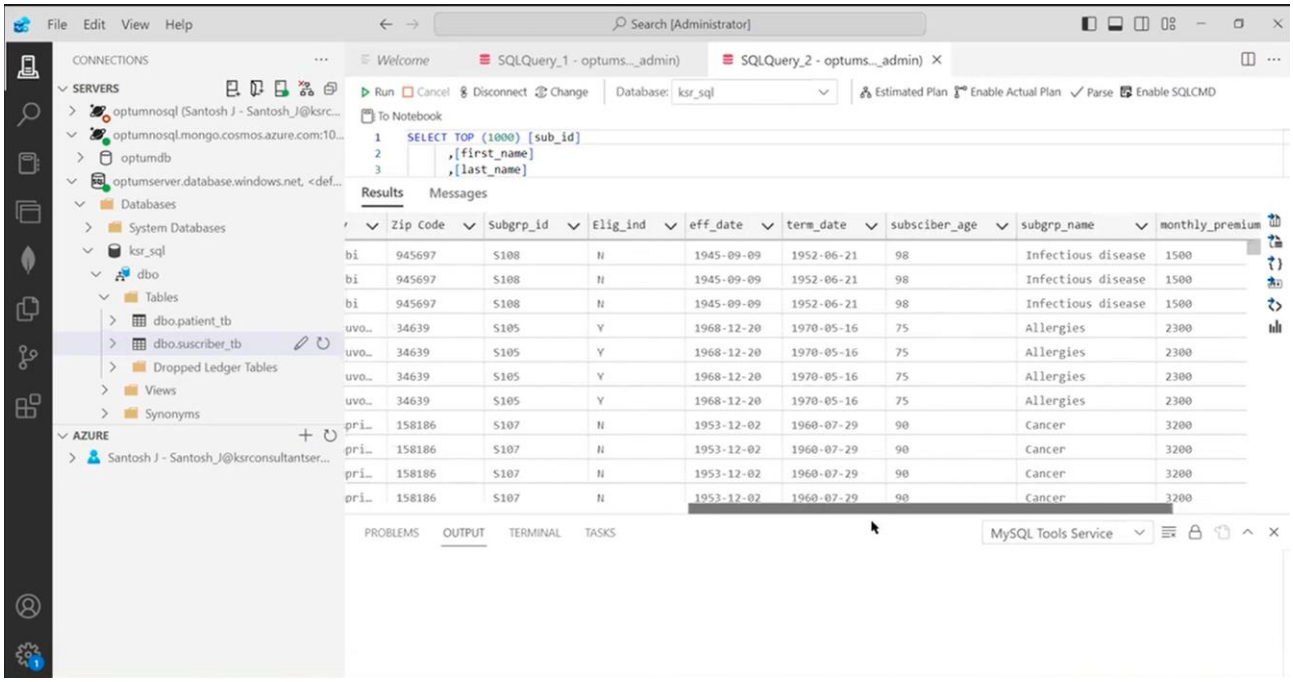


GO TO MAINPIPELINE AND USE EXECUTE PIPELINE



LESTS GO AND EXECUTE





IN THIE WE NEED TO SOLVE THIS QUESITON IN THD SQL DATA BASE

Use Cases?
which disease having maximum number of claims.
Find those Subscribers having age less than 30 and they subscribe any subgroup
Find out which group has maximum subgroups.
Find out hospital which serve most number of patients
Find out which subgroups subscribe most number of times
Find out total number of claims which were rejected
From where most claims are coming (city)
Which groups of policies subscriber subscribe mostly Government or private
Average monthly premium subscriber pay to insurance company.
Find out Which group is most profitable
List all the patients below age of 18 who admit for cancer
List patients who have cashless insura0nce and have total charges greater than or equal for Rs. 50,000.
List female patients over the age of 40 that have undergone knee surgery in the past year