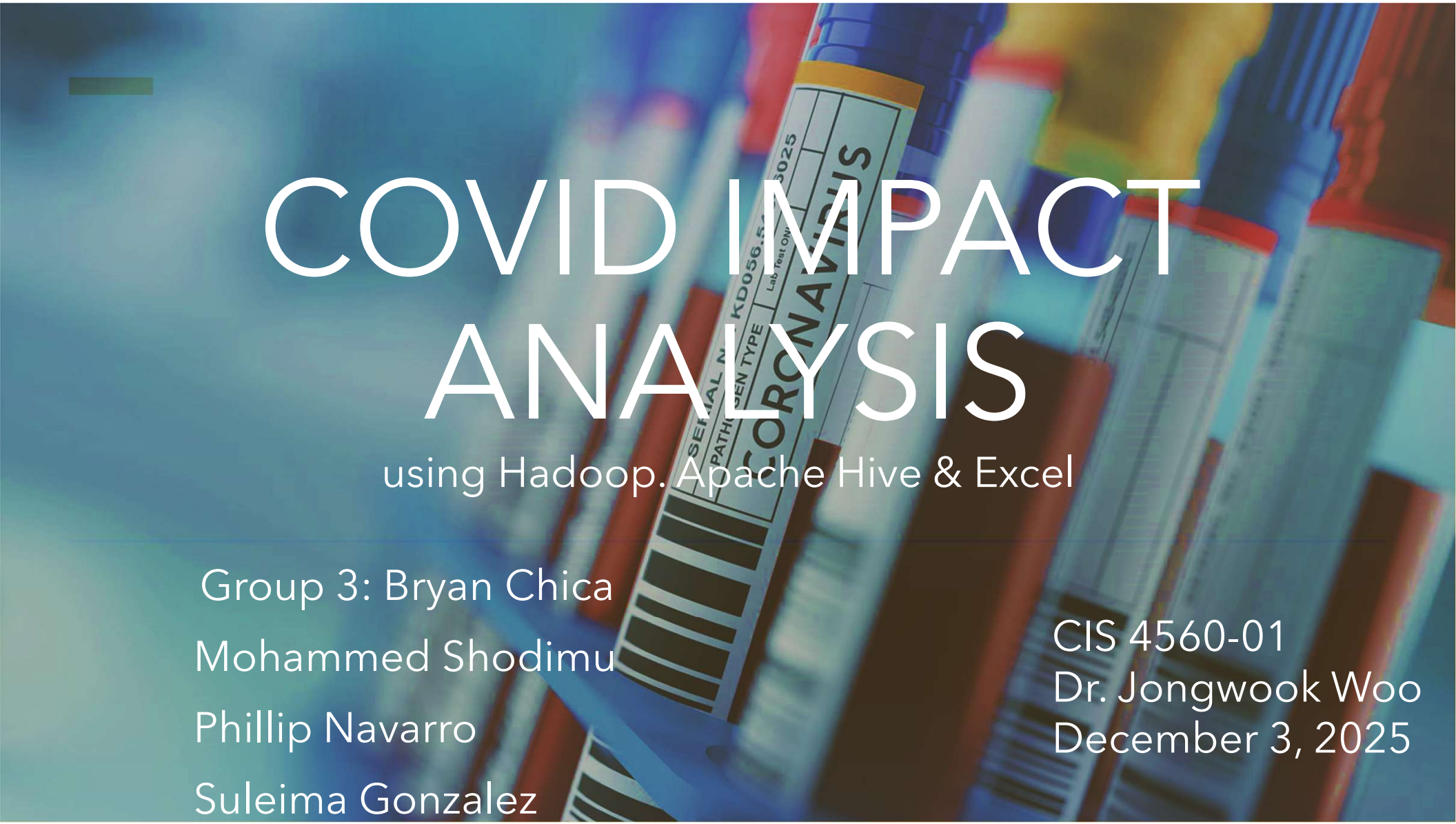


COVID IMPACT ANALYSIS

using Hadoop. Apache Hive & Excel

Group 3: Bryan Chica
Mohammed Shodimu
Phillip Navarro
Suleima Gonzalez

CIS 4560-01
Dr. Jongwook Woo
December 3, 2025

Agenda

- Introduction
- Dataset Overview
- Platform Specifications
- Architecture
- Workflow / Flowchart
- Tempo-Spatial Visualizations
- Analysis & Findings
- Hive Processing
- Github Repository
- Conclusion

Introduction

This Big data project studies the global impact of COVID-19 by analyzing a large Google Public COVID-19 Open dataset. We implemented tempo-spatial analysis to understand how the virus affected people with regards to:

- *Different groups i.e age, gender, geography*
- *Health availability and expenditures*
- *Varying measures i.e. Vaccine doses*

Relevance: Using tools such as Hive, HDFS and Power BI we processed millions of records to identity meaningful patterns and problem areas in COVID's impact on healthcare, economic and social sectors worldwide.

Dataset Specifications

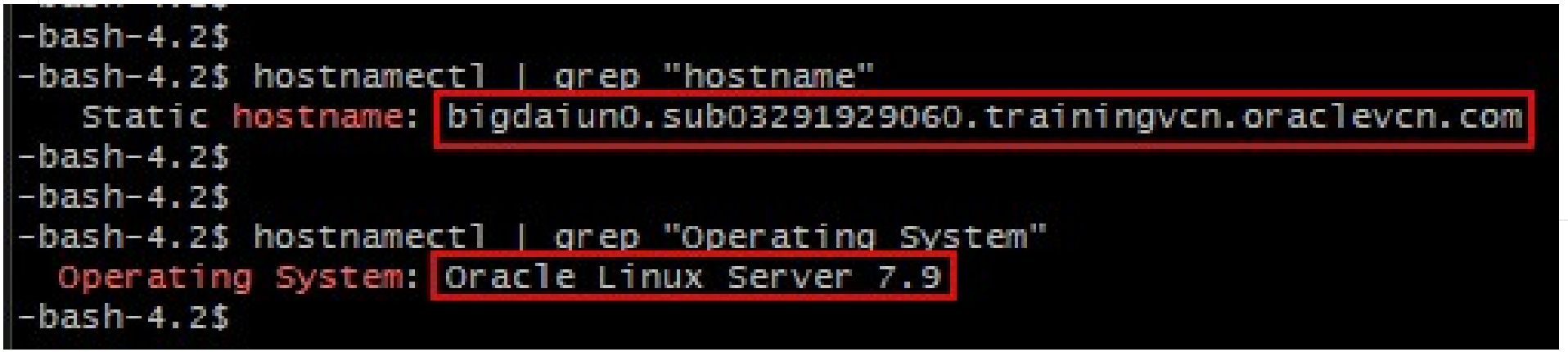
Dataset

- **Source:** Google Public Dataset
- **File name:** aggregated.csv
- **Compressed size:** 1.32 GB
- **Extracted size**: 20.9 GB
- **Records:** 3,000,599
- **Columns:** 708
- **Format:** .csv
- **Data Types:** STRING, FLOAT, INT, BOOLEAN,DATETIME
- **Link:** *https://storage.googleapis.com/covid19-open-data*

Github

Includes

- **Hive .hql scripts**
- **CSV ouputs**
- **Hadoop commands**
- **Workflow documantation**
- **Presentation Slides**
- **Architecture diagrams**
- **Flowchart**
- **Link:** *https://github.com/chica-94/COVID19-Data-Analysis*



Host name & Operating System

- **Hostname:** bigdaiun0.sub03291929060.trainingvcn.oraclevcn.com
- **OS:** Oracle Linux Server 7.9

System Resources

```
-bash-4.2$ lscpu
Architecture:          x86_64
CPU op-mode(s):        32-bit, 64-bit
Byte Order:            Little Endian
CPU(s):                6
On-line CPU(s) list:   0-5
Thread(s) per core:    2
Core(s) per socket:    3
Socket(s):             1
NUMA node(s):          1
Vendor ID:             AuthenticAMD
CPU family:            25
Model:                 1
Model name:            AMD EPYC 7J13 64-Core Processor
Stepping:              1
CPU MHz:               2445.406
```

- **CPU:** AMC EPYX 7J13 64-Core Processor @ 2.45 Ghz
- **Storage:** 120 GB
- **Memory Size:** 32 GB (8 GB swap)

```
-bash-4.2$
-bash-4.2$ df -h ~
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda3       120G  103G   18G  86% /
-bash-4.2$
-bash-4.2$ free -h
              total        used        free      shared  buff/cache   available
Mem:            31G         26G        739M        507M        4.2G        4.1G
Swap:          8.0G        2.6G        5.4G
-bash-4.2$
-bash-4.2$
```

Free storage space

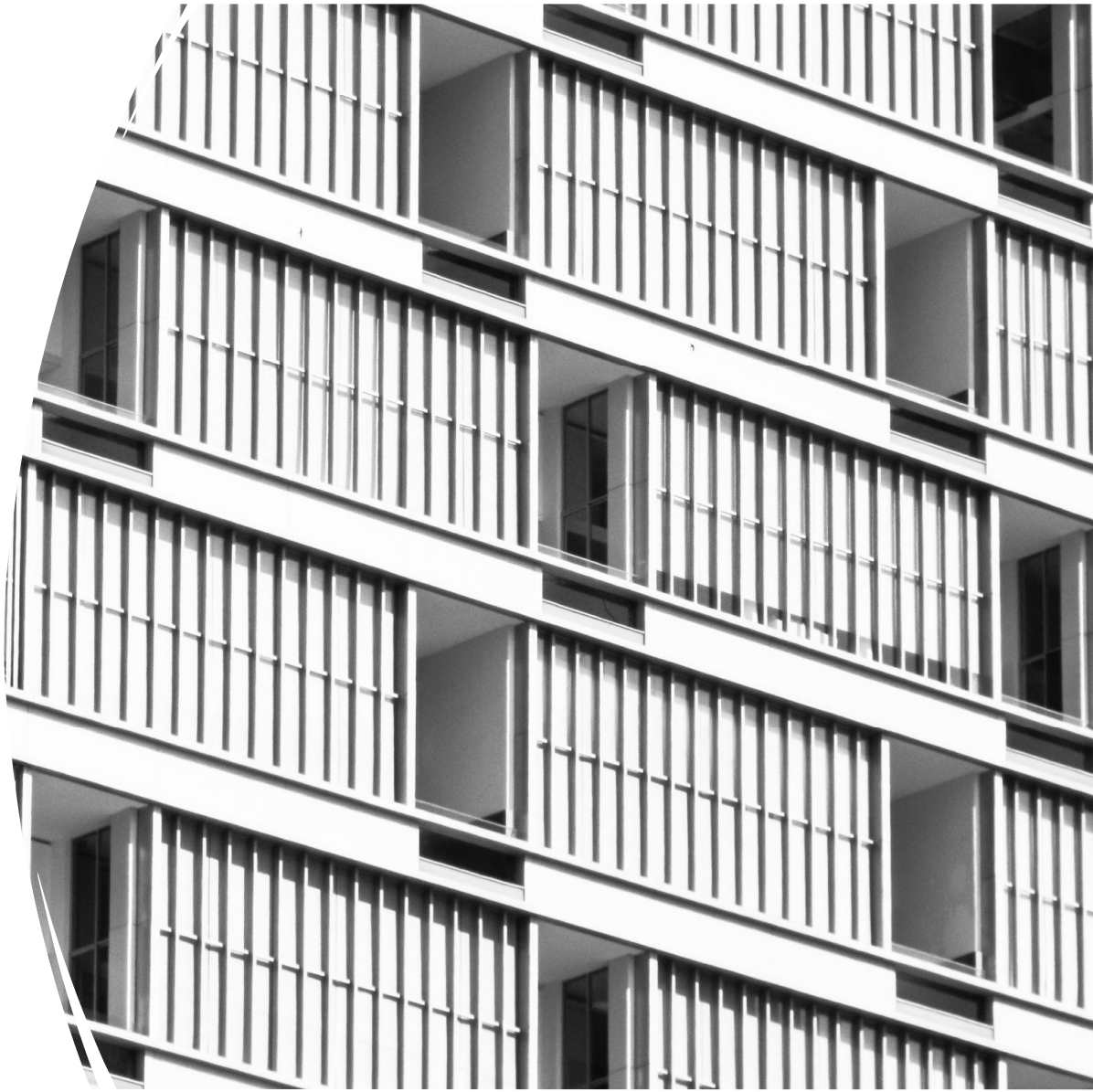
Other Specifications

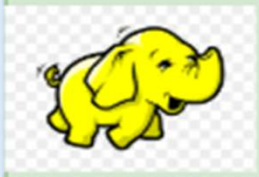
```
-bash-4.2$
-bash-4.2$ hadoop version
Hadoop 3.1.2
Source code repository ssh://git@bitbucket.oci.oraclecorp.com:7999/bdcs/apache_bigtop.git -r 2ce21
64edd68ab34a98a99cea5a1c2037a3a7ca1
Compiled by root on 2024-07-30T16:09Z
Compiled with protoc 2.5.0
From source with checksum eaec305dfb8e514a567cd7fdba59f0
This command was run using /usr/odh/1.1.13/hadoop/hadoop-common-3.1.2.jar
-bash-4.2$
```

```
-bash-4.2$
-bash-4.2$ yarn node -list all
25/11/28 04:15:31 INFO client.RMProxy: Connecting to ResourceManager at bigdaimn0.sub03291929060.t
rainingvcn.oraclevcn.com/10.1.0.251:8050
25/11/28 04:15:32 INFO client.AHSProxy: Connecting to Application History server at bigdaiun0.sub0
3291929060.trainingvcn.oraclevcn.com/10.1.0.73:10200
Total Nodes:3
         Node-Id             Node-State Node-Http-Address       Number-of-Running-Containers
bigdaiwn2.sub03291929060.trainingvcn.oraclevcn.com:45454               RUNNING bigdaiwn2.sub03291
929060.trainingvcn.oraclevcn.com:8042                                        0
bigdaiwn0.sub03291929060.trainingvcn.oraclevcn.com:45454               RUNNING bigdaiwn0.sub03291
929060.trainingvcn.oraclevcn.com:8042                                        0
bigdaiwn1.sub03291929060.trainingvcn.oraclevcn.com:45454               RUNNING bigdaiwn1.sub03291
929060.trainingvcn.oraclevcn.com:8042                                        1
-bash-4.2$
```

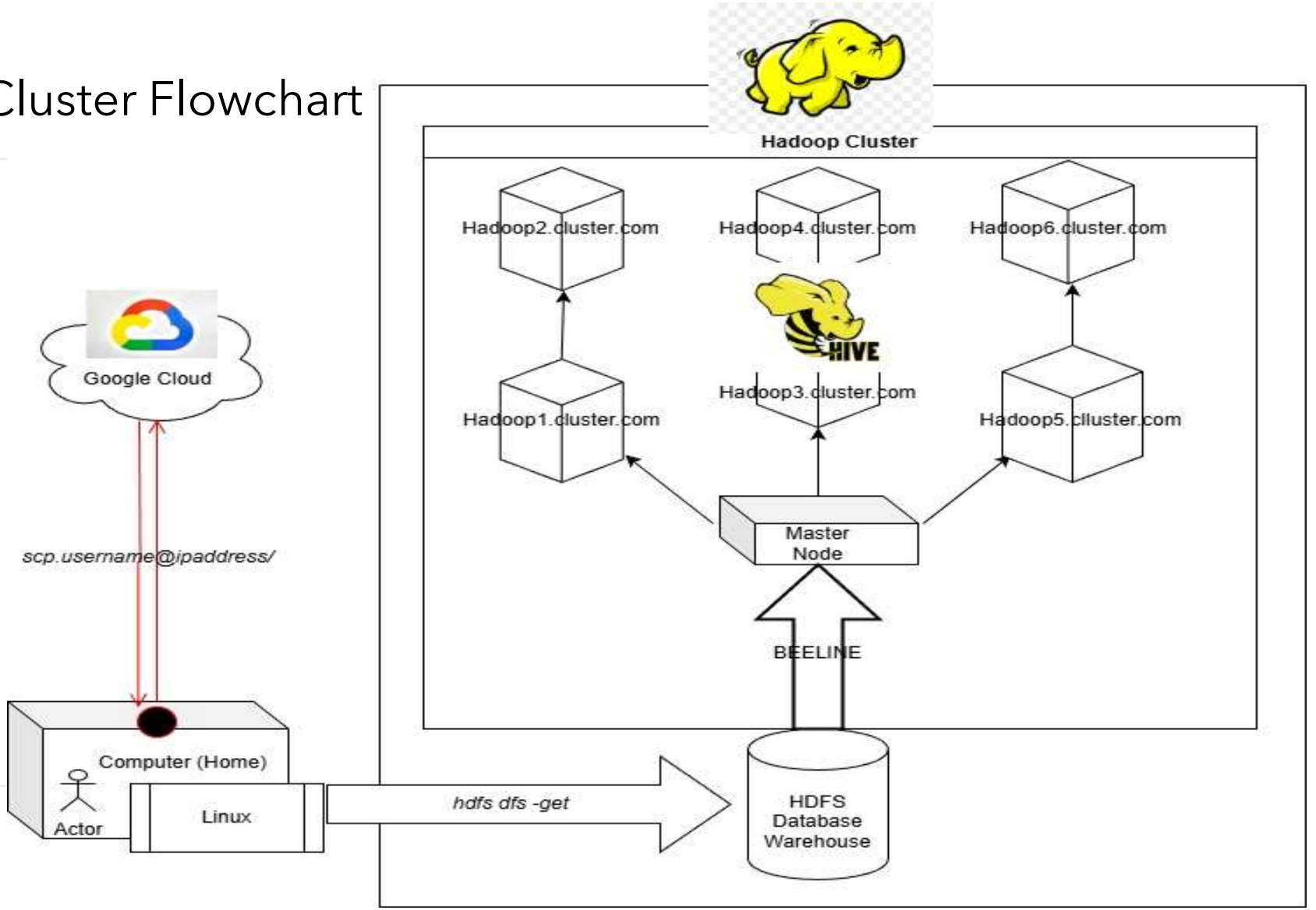
- **Cluster Version:** Hadoop 3.1.2
- **Cluster Number of Nodes:** 3

Architecture

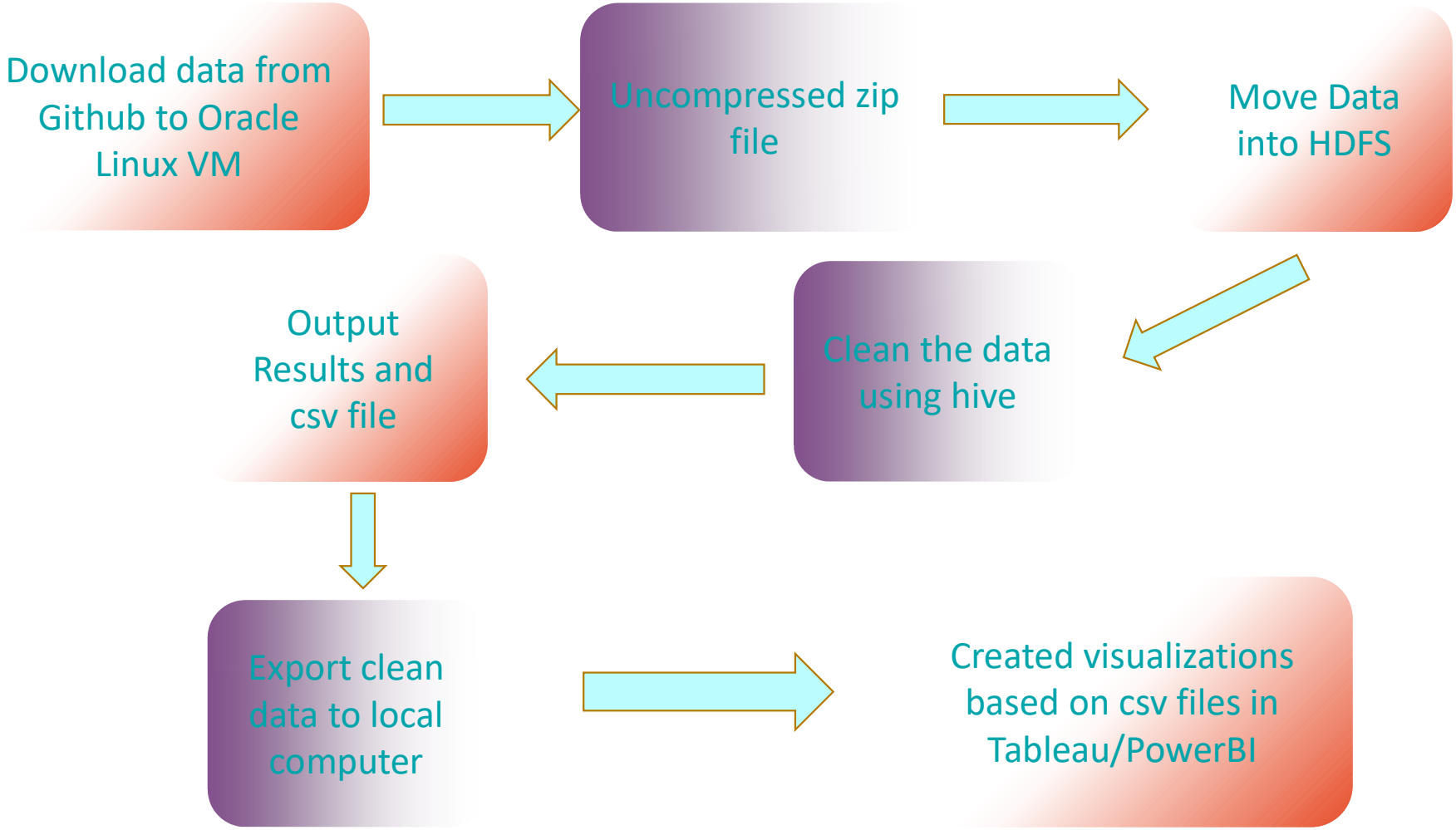


SOURCE	STAGING	DATA WAREHOUSE	QUERY & ANALYSIS	VISUALIZATION	REPOSITORY
					
GOOGLE CLOUD (Source Code / Raw Data)	**LINUX** Open Source O.S. Kernel	**HDFS HADOOP** File System	**APACHE HIVE**	**POWER BI**	**GITHUB**

Hadoop Cluster Flowchart

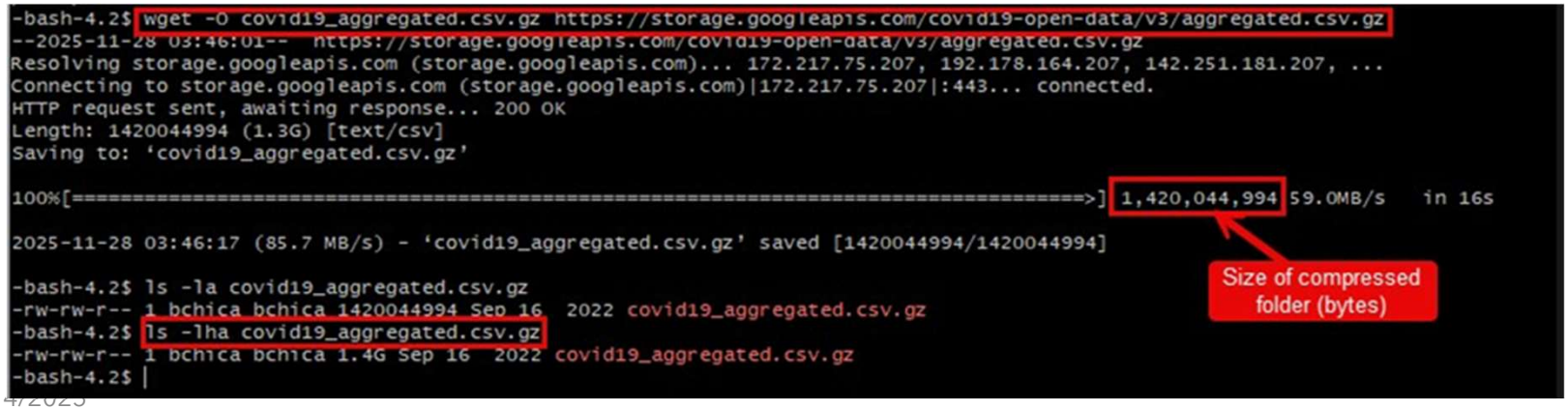


Workflow for Data Analysis



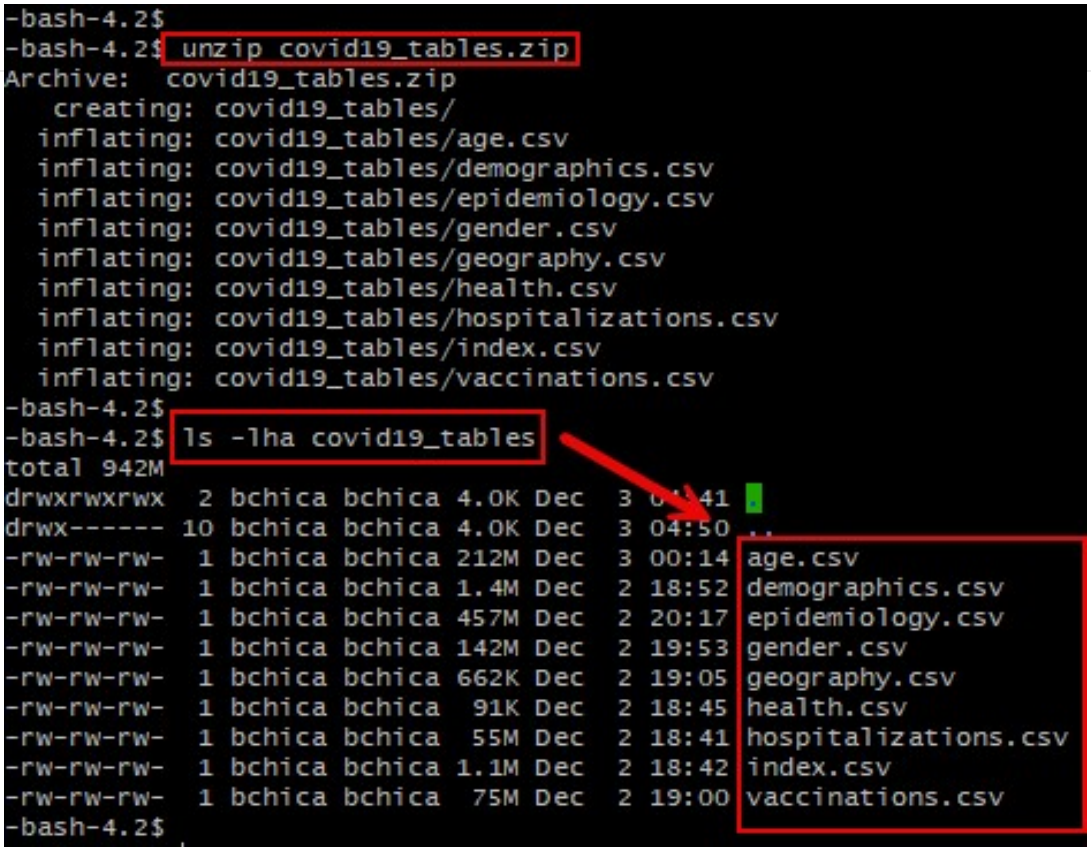
Retrieving the Data

- The complete dataset was pulled from Github
- Wget command is used to download the data to Oracle server
- *wget* *https://github.com/chica-94/COVID19-Data-Analysis/raw/refs/heads/main/covid19_tables.zip*



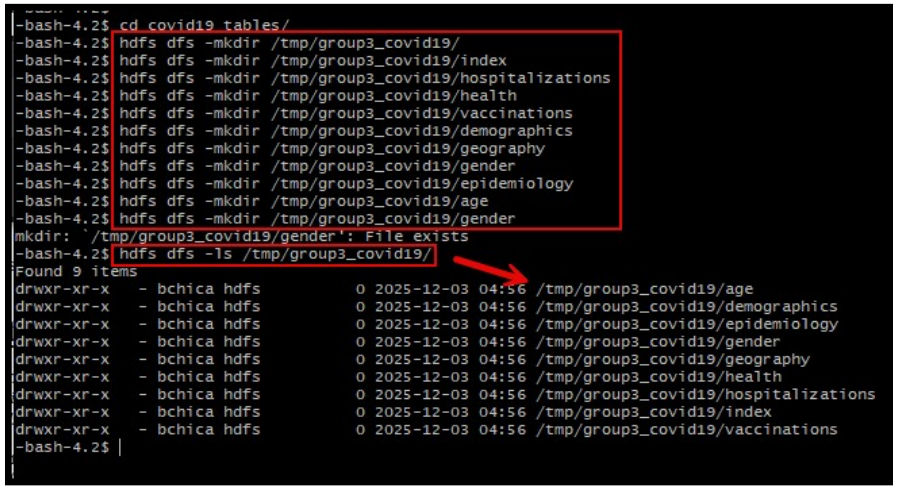
Extracting Dataset

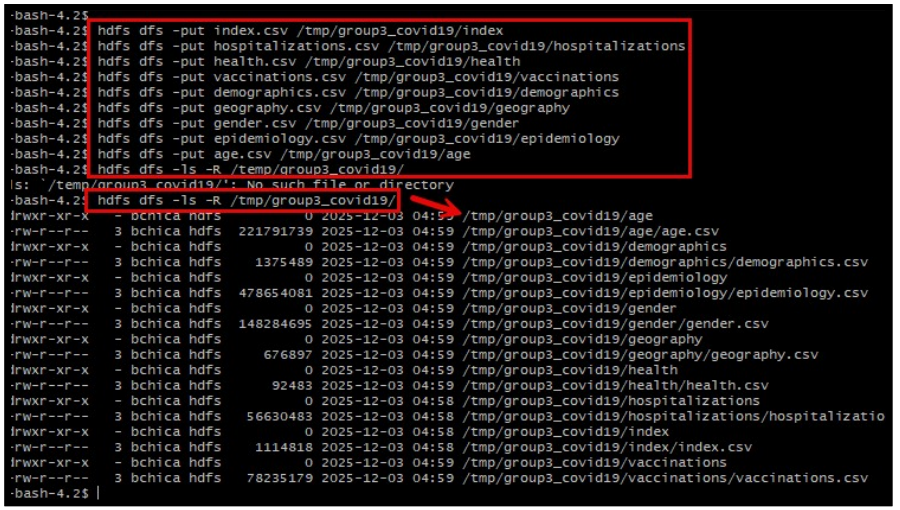
- *Unzip the compressed folder*
- List all tables inside extracted folder



Moving the Data

- Created a new directory in HDFS from the Linux environment to store the data.
- Uploaded the dataset into HDFS using the **-put** command.





Hive Data Processing (Preparing the data)

- Used Hive to clean, filter, and analyze large datasets
- Key Tasks:
  - Removed null/invalid rows
  - Aggregated data by (country, month, years,….)
  - Calculated case fatality rates
  - Generated reduced tables for visualization

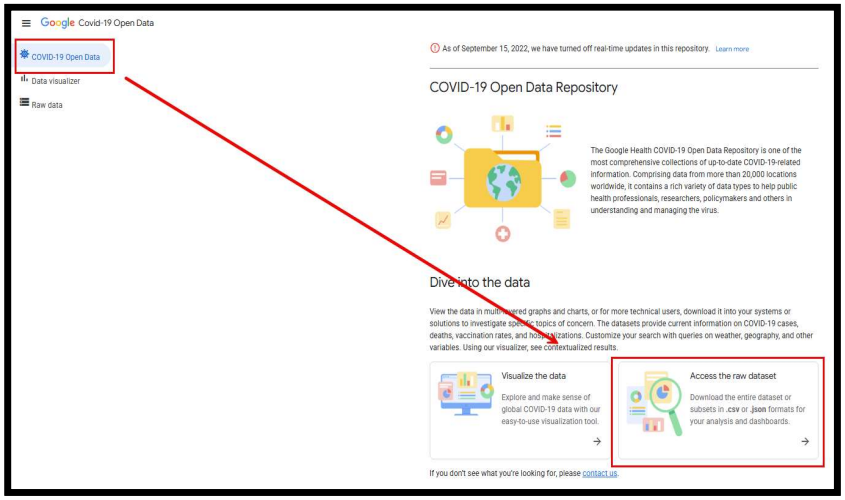
- Cleaned the data using OpenRefine and Hive
  - Final process dataset cleaned and reduced from 20.9 GB (Origin) -> 3.2 GB (MapReduce) -> 2.1 MB (Visualization)

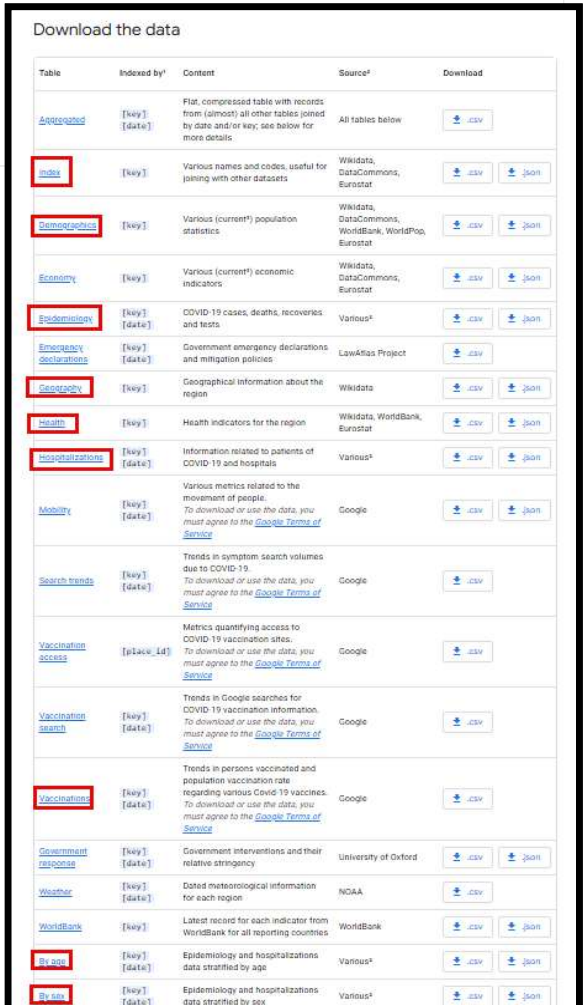
Obtain datasets via Google COVID-19 Open Data

Download the csv files of the following tables:

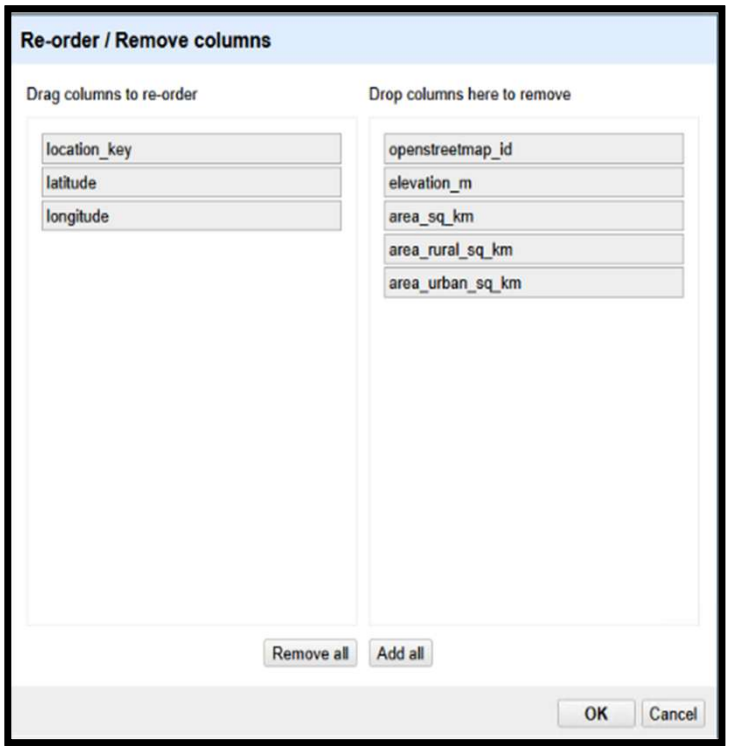
- Index
- Demographics
- Epidemiology
- Geography
- Health
- Hospitalizations
- Vaccinations
- By age
- By sex

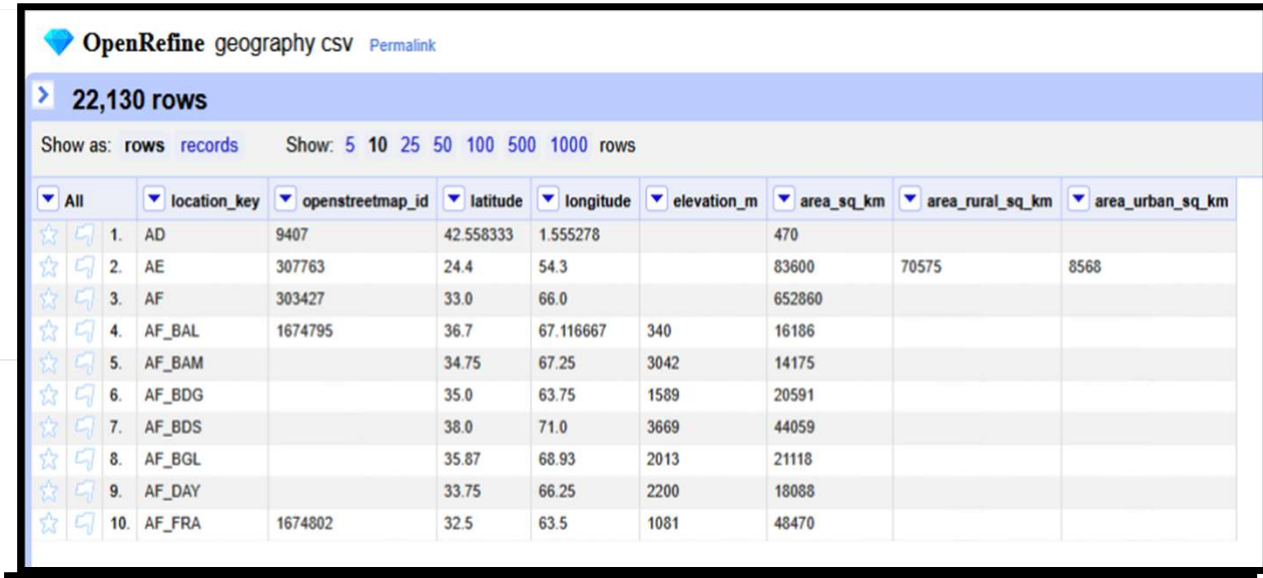
Total size: ~**21 GB**

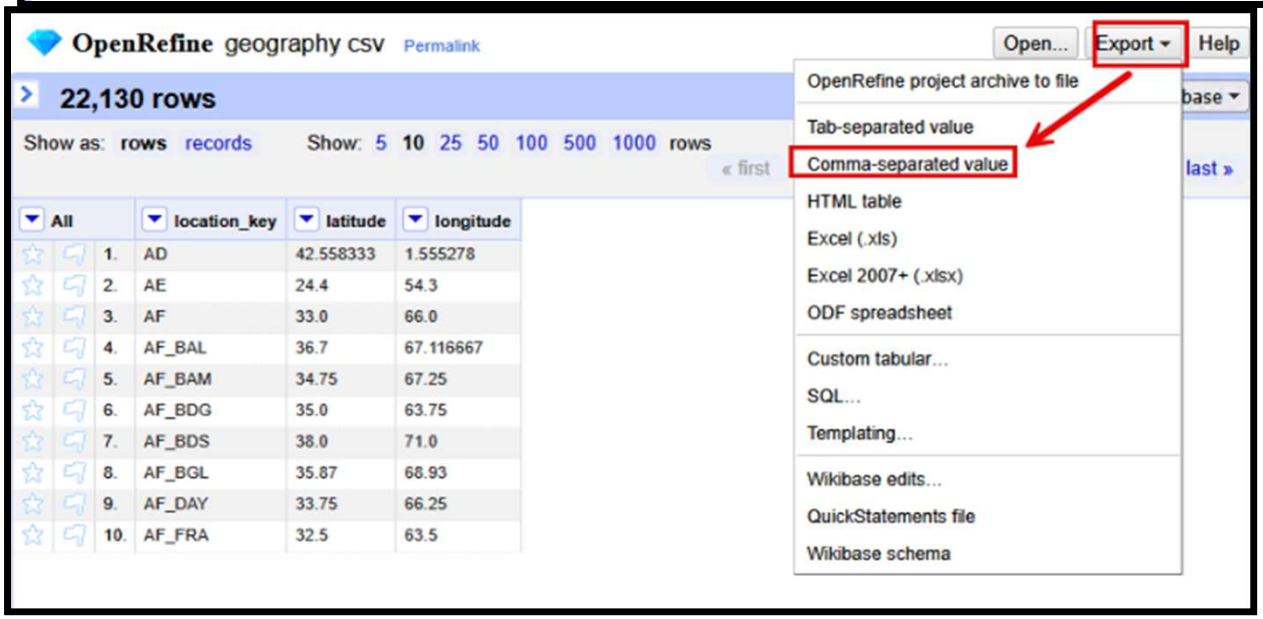




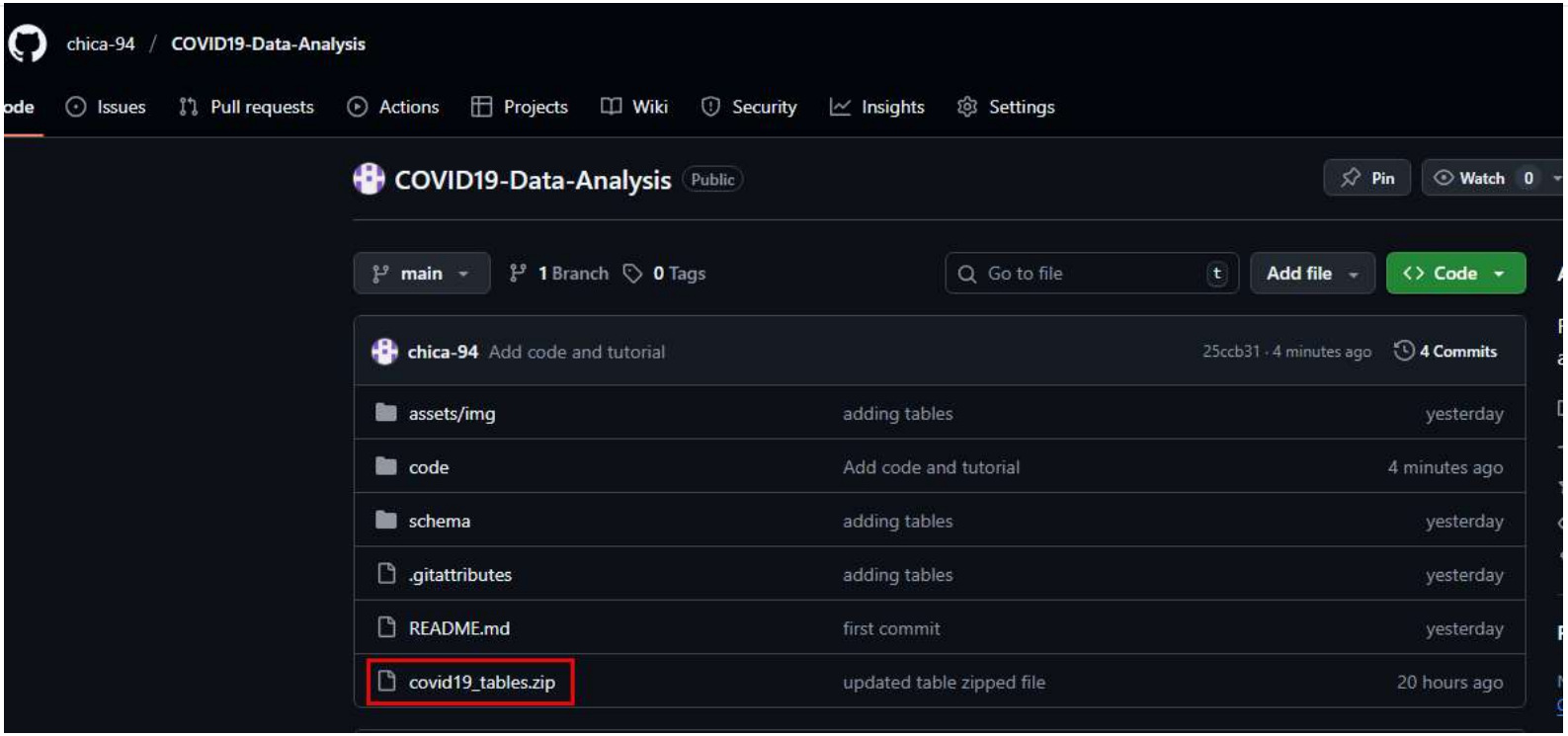
Cleaning up: (OpenRefine)



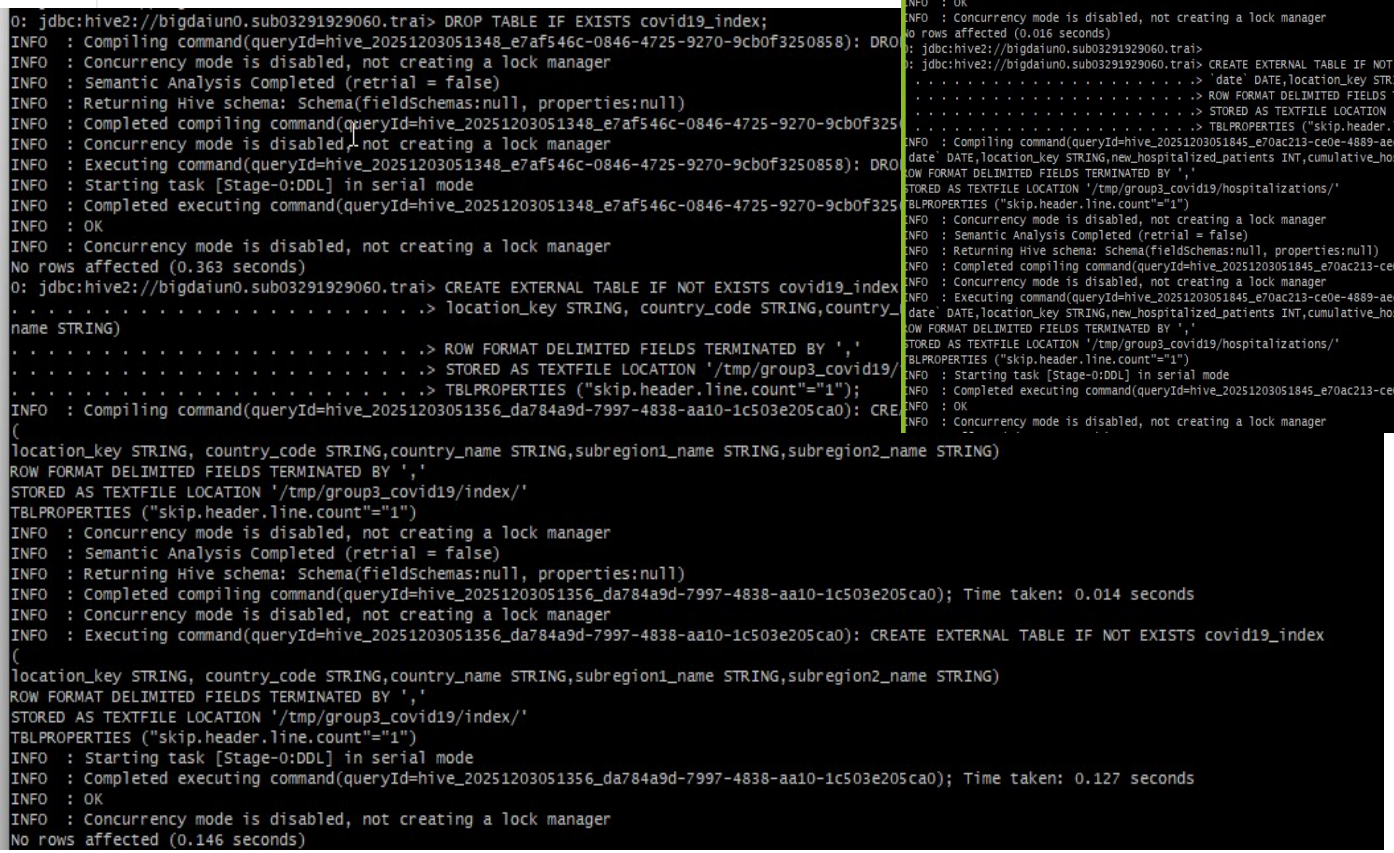


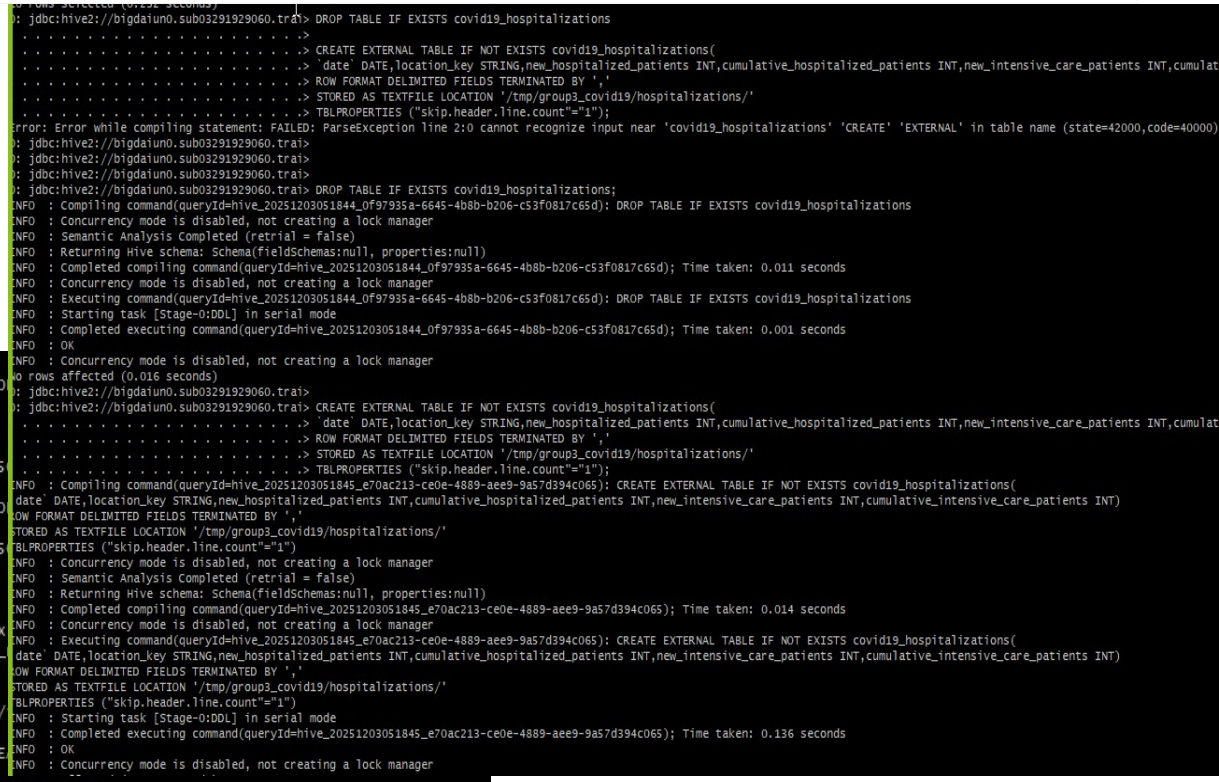


Added cleaned tables to Github repo



Creating Tables: Apache Hive





HiveQL: Select Query (index)

```
No rows affected (0.146 seconds)
0: jdbc:hive2://bigdaiun0.sub03291929060.trai> SELECT * FROM covid19_index LIMIT 10;
INFO  : Compiling command(queryId=hive_20251203051403_41ed71c1-efa3-4a50-9f41-fcb9c6ab137b): SELECT * FROM covid19_index LIMIT 10
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Semantic Analysis Completed (retrial = false)
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:covid19_index.location_key, type:string, comment:null), FieldSchema(name:
covid19_index.country_code, type:string, comment:null), FieldSchema(name:covid19_index.country_name, type:string, comment:null), FieldSchema(n
ame:covid19_index.subregion1_name, type:string, comment:null), FieldSchema(name:covid19_index.subregion2_name, type:string, comment:null)], pr
operties:null)
INFO  : Completed compiling command(queryId=hive_20251203051403_41ed71c1-efa3-4a50-9f41-fcb9c6ab137b); Time taken: 0.218 seconds
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Executing command(queryId=hive_20251203051403_41ed71c1-efa3-4a50-9f41-fcb9c6ab137b): SELECT * FROM covid19_index LIMIT 10
INFO  : Completed executing command(queryId=hive_20251203051403_41ed71c1-efa3-4a50-9f41-fcb9c6ab137b); Time taken: 0.0 seconds
INFO  : OK
INFO  : Concurrency mode is disabled, not creating a lock manager
+-----------------------------+-----------------------------+-----------------------------+-------------------------------+-----------------
--------------+
| covid19_index.location_key  | covid19_index.country_code  | covid19_index.country_name  | covid19_index.subregion1_name  | covid19_index.sub
region2_name  |
+-----------------------------+-----------------------------+-----------------------------+-------------------------------+-----------------
--------------+
| AD                          | AD                          | Andorra                     |                                |
              |
| AE                          | AE                          | United Arab Emirates        |                                |
              |
| AF                          | AF                          | Afghanistan                 |                                |
              |
| AF_BAL                      | AF                          | Afghanistan                 | Balkh                          |
              |
| AF_BAM                      | AF                          | Afghanistan                 | Bamyan                         |
              |
| AF_BDG                      | AF                          | Afghanistan                 | Badghis                        |
              |
| AF_BDS                      | AF                          | Afghanistan                 | Badakhshan                     |
              |
| AF_BGL                      | AF                          | Afghanistan                 | Baghlan                        |
              |
| AF_DAY                      | AF                          | Afghanistan                 | Daykundi                       |
              |
| AF_FRA                      | AF                          | Afghanistan                 | Farah                          |
              |
+-----------------------------+-----------------------------+-----------------------------+-------------------------------+-----------------
--------------+
10 rows selected (0.232 seconds)
```

HiveQL: Select Query (health)

```
0: jdbc:hive2://bigdaiun0.sub03291929060.trai> SELECT * FROM covid19_health LIMIT 10;
INFO  : Compiling command(queryId=hive_20251203052123_6b9cd0dd-a2f7-435d-9a12-0426a98ee287): SELECT * FROM covid19_health LIMIT 10
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Semantic Analysis Completed (retrial = false)
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:covid19_health.location_key, type:string, comment:null), FieldSchema(name:covid19_health.adult_mal
e_mortality_rate, type:double, comment:null), FieldSchema(name:covid19_health.adult_female_mortality_rate, type:double, comment:null), FieldSchema(name:covid19_health.
life_expectancy, type:double, comment:null), FieldSchema(name:covid19_health.diabetes_prevalence, type:double, comment:null), FieldSchema(name:covid19_health.health_ex
penditure_usd, type:double, comment:null), FieldSchema(name:covid19_health.out_of_pocket_health_expenditure_usd, type:double, comment:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20251203052123_6b9cd0dd-a2f7-435d-9a12-0426a98ee287); Time taken: 0.22 seconds
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Executing command(queryId=hive_20251203052123_6b9cd0dd-a2f7-435d-9a12-0426a98ee287): SELECT * FROM covid19_health LIMIT 10
INFO  : Completed executing command(queryId=hive_20251203052123_6b9cd0dd-a2f7-435d-9a12-0426a98ee287); Time taken: 0.0 seconds
INFO  : OK
INFO  : Concurrency mode is disabled, not creating a lock manager
+----------------------------+----------------------------------------+----------------------------------------+--------------------------------+----------
----------------------+----------------------------------------+----------------------------------------------------+
| covid19_health.location_key  | covid19_health.adult_male_mortality_rate  | covid19_health.adult_female_mortality_rate  | covid19_health.life_expectancy  | covid19_he
alth.diabetes_prevalence  | covid19_health.health_expenditure_usd  | covid19_health.out_of_pocket_health_expenditure_usd |
+----------------------------+----------------------------------------+----------------------------------------+--------------------------------+----------
----------------------+----------------------------------------+----------------------------------------------------+
| AD                         | NULL                                   | NULL                                   | NULL                           | 7.7
                      | 4040.786621                            | 1688.12146                                         |
| AE                         | 69.555                                 | 44.863                                 | 77.814                         | 16.3
                      | 1357.017456                            | 256.034485                                         |
| AF                         | 237.554                                | 192.532                                | 64.486                         | 9.2
                      | 67.12265                               | 50.665913                                          |
| AG                         | 126.917                                | 83.136                                 | 76.885                         | 13.1
                      | 673.85968                              | 235.749039                                         |
| AL                         | 93.315                                 | 49.486                                 | 78.9                           | 9.0
                      | NULL                                   | NULL                                               |
| AM                         | 173.428                                | 65.595                                 | 74.945                         | 6.1
                      | 407.635864                             | 343.832977                                         |
| AO                         | 327.044                                | 220.291                                | 60.782                         | 4.5
```

HiveQL: Select Query (geography)

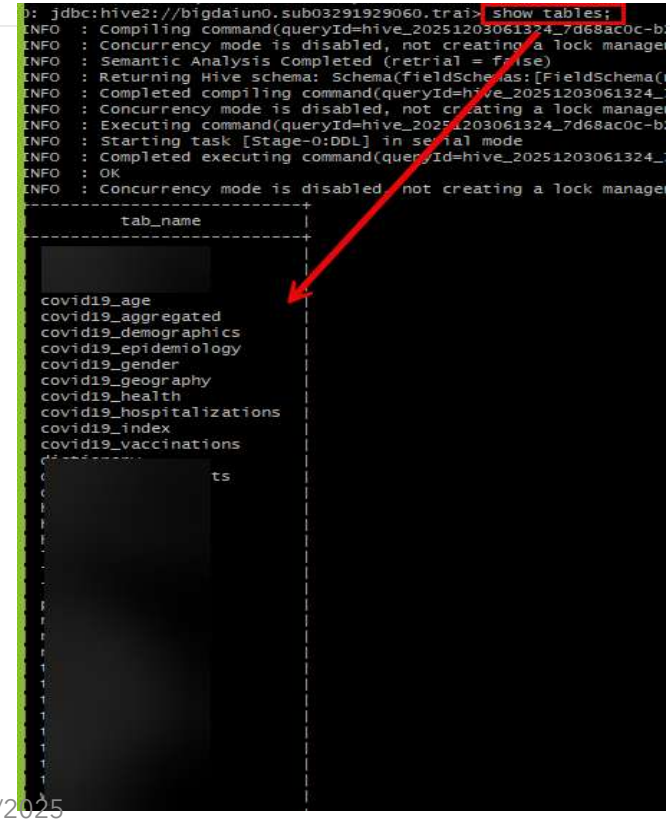
```
0: jdbc:hive2://bigdaiun0.sub03291929060.trai>
0: jdbc:hive2://bigdaiun0.sub03291929060.trai>  SELECT * FROM covid19_geography LIMIT 10;
INFO  : Compiling command(queryId=hive_20251203052620_5b9b836e-9d35-46cf-aed2-97447e3613c2): SELECT * FROM covid19_geography LIMIT 10
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Semantic Analysis Completed (retrial = false)
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:covid19_geography.location_key, type:string, comment:null), FieldSch
itude, type:double, comment:null), FieldSchema(name:covid19_geography.longitude, type:double, comment:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20251203052620_5b9b836e-9d35-46cf-aed2-97447e3613c2); Time taken: 0.232 seconds
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Executing command(queryId=hive_20251203052620_5b9b836e-9d35-46cf-aed2-97447e3613c2): SELECT * FROM covid19_geography LIMIT 10
INFO  : Completed executing command(queryId=hive_20251203052620_5b9b836e-9d35-46cf-aed2-97447e3613c2); Time taken: 0.0 seconds
INFO  : OK
INFO  : Concurrency mode is disabled, not creating a lock manager
+---------------------------------+-----------------------------+-----------------------------+
| covid19_geography.location_key  | covid19_geography.latitude  | covid19_geography.longitude |
+---------------------------------+-----------------------------+-----------------------------+
| AD                              | 42.558333                   | 1.555278                    |
| AE                              | 24.4                        | 54.3                        |
| AF                              | 33.0                        | 66.0                        |
| AF_BAL                          | 36.7                        | 67.116667                   |
| AF_BAM                          | 34.75                       | 67.25                       |
| AF_BDG                          | 35.0                        | 63.75                       |
| AF_BDS                          | 38.0                        | 71.0                        |
| AF_BGL                          | 35.87                       | 68.93                       |
| AF_DAY                          | 33.75                       | 66.25                       |
| AF_FRA                          | 32.5                        | 63.5                        |
+---------------------------------+-----------------------------+-----------------------------+
10 rows selected (0.25 seconds)
```

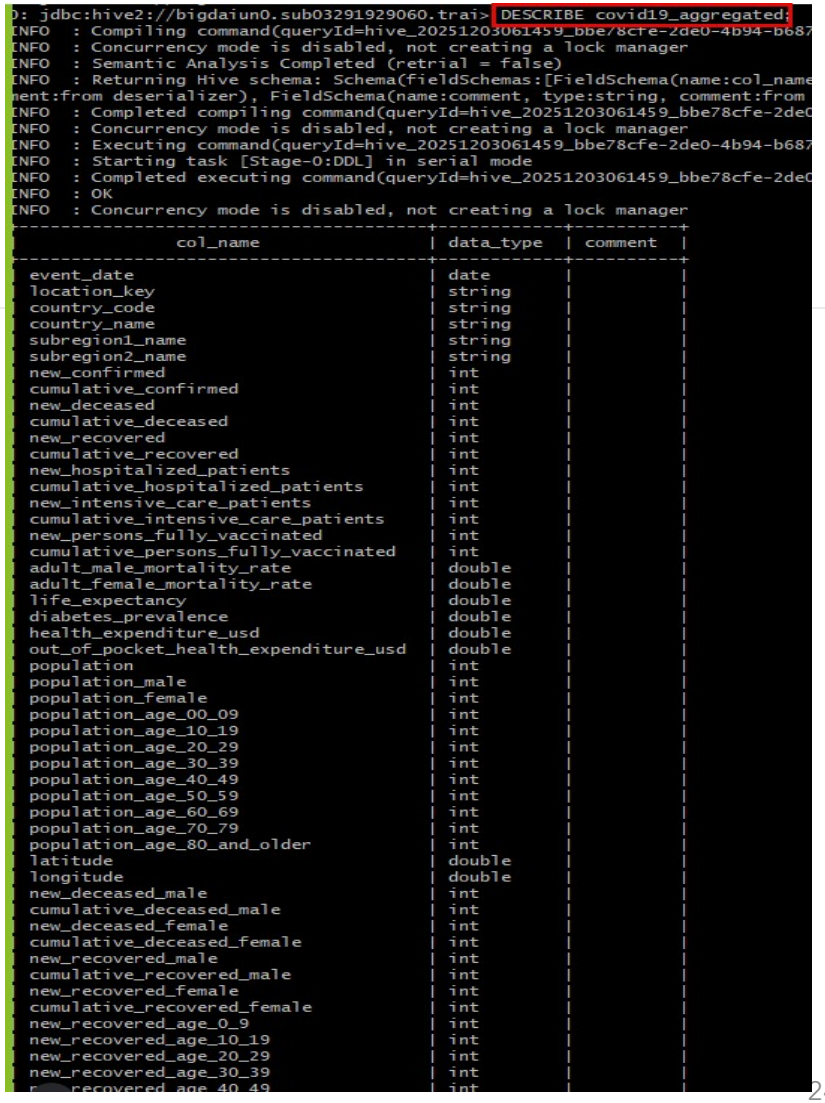
HiveQL: Insert Overwrite (aggregated)

```
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Executing command(queryId=hive_20251203060935_5be0bfcd-67f4-4ee4-b156-90e4dc59ff00): INSERT OVERWRITE TABLE covid19_aggregated
SELECT epi.event_date,epi.location_key,idx.country_code,idx.country_name,idx.subregion1_name,idx.subregion2_name,epi.new_confirmed,epi.cum
deceased,epi.cumulative_deceased,epi.new_recovered,epi.cumulative_recovered,hosp.new_hospitalized_patients,hosp.cumulative_hospitalized_pa
care_patients,hosp.cumulative_intensive_care_patients,vac.new_persons_fully_vaccinated,vac.cumulative_persons_fully_vaccinated,health.adul
th.adult_female_mortality_rate,health.life_expectancy,health.diabetes_prevalence,health.health_expenditure_usd,health.out_of_pocket_health
lation,demo.population_male,demo.population_female,demo.population_age_00_09,demo.population_age_10_19,demo.population_age_20_29,demo.popu
lation_age_40_49,demo.population_age_50_59,demo.population_age_60_69,demo.population_age_70_79,demo.population_age_80_and_older,geo.latitu
w_deceased_male,gender.cumulative_deceased_male,gender.new_deceased_female,gender.cumulative_deceased_female,gender.new_recovered_male,gen
ale,gender.new_recovered_female,gender.cumulative_recovered_female,age.new_recovered_age_0_9,age.new_recovered_age_10_19,age.new_recovered
ed_age_30_39,age.new_recovered_age_40_49,age.new_recovered_age_50_59,age.new_recovered_age_60_69,age.new_recovered_age_70_79,age.new_recov
overed_age_90_99,age.cumulative_recovered_age_0_9,age.cumulative_recovered_age_10_19,age.cumulative_recovered_age_20_29,age.cumulative_rec
ative_recovered_age_40_49,age.cumulative_recovered_age_50_59,age.cumulative_recovered_age_60_69,age.cumulative_recovered_age_70_79,age.cum
9,age.cumulative_recovered_age_90_99
FROM covid19_epidemiology epi
LEFT JOIN covid19_index idx ON epi.location_key = idx.location_key
LEFT JOIN covid19_hospitalizations hosp ON epi.location_key = hosp.location_key AND epi.event_date = hosp.event_date
LEFT JOIN covid19_vaccinations vac ON epi.location_key = vac.location_key AND epi.event_date = vac.event_date
LEFT JOIN covid19_health health ON epi.location_key = health.location_key
LEFT JOIN covid19_demographics demo ON epi.location_key = demo.location_key
LEFT JOIN covid19_geography geo ON epi.location_key = geo.location_key
LEFT JOIN covid19_gender gender ON epi.location_key = gender.location_key AND epi.event_date = gender.event_date
LEFT JOIN covid19_age age ON epi.location_key = age.location_key AND epi.event_date = age.event_date
INFO  : Query ID = hive_20251203060935_5be0bfcd-67f4-4ee4-b156-90e4dc59ff00
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Subscribed to counters: [] for queryId: hive_20251203060935_5be0bfcd-67f4-4ee4-b156-90e4dc59ff00
INFO  : Tez session hasn't been created yet. Opening session
INFO  : Dag name: INSERT OVERWRITE TABLE covi...age.event_date (Stage-1)
INFO  : Setting tez.task.scale.memory.reserve-fraction to 0.30000001192092896
INFO  : Status: Running (Executing on YARN cluster with App id application_1756243379751_1599)

--------------------------------------------------------------------------------
        VERTICES      MODE        STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
--------------------------------------------------------------------------------
Map 5 .......... container     SUCCEEDED      1          1        0        0       0       0
Map 1           container       RUNNING      1          0        1        0       0       0
Map 6           container       RUNNING      1          0        1        0       0       0
Map 7           container       RUNNING      1          0        1        0       0       0
Reducer 2       container        INITED    130          0        0      130       0       0
Map 8 .......... container     SUCCEEDED      1          1        0        0       0       0
Map 9 .......... container     SUCCEEDED      1          1        0        0       0       0
Map 10 ......... container     SUCCEEDED      1          1        0        0       0       0
Reducer 3       container        INITED    226          0        0      226       0       0
Map 11          container       RUNNING      1          0        1        0       0       0
Map 12          container       RUNNING      1          0        1        0       0       0
Reducer 4       container        INITED    818          0        0      818       0       0
--------------------------------------------------------------------------------
VERTICES: 04/12  [>>--------------------------] 0%    ELAPSED TIME: 5.82 s
--------------------------------------------------------------------------------
```

Completed Tables





Get the top 10 countries

```
SELECT country_name,MAX(cumulative_confirmed)
AS total_cases
FROM covid19_aggregated
WHERE cumulative_confirmed IS NOT NULL
GROUP BY country_name
ORDER BY total_cases DESC
LIMIT 10;
```

country_name	total_cases
United States of America	92440495
India	44516479
Brazil	34568833
France	33766090
Germany	32604993
South Korea	24264470
United Kingdom	23554971
Italy	22114423
Russia	20265004
Japan	19868288

Create top 10 table

- Create an empty table called ***covid19_top10_countries***, which will contain the cumulative confirmed cases for each country and its subregions (city & states) from **covid19_aggregated** tables once the data has been formatted.
- It reduced the number of records from 12 million to 10 thousand
  - SELECT COUNT(*) AS covid19_records FROM covid19_aggregated;
  - SELECT COUNT(*) AS covid19_top10_records FROM covid19_top10_countries;
- The size reduced from 3.2GB to 2.1 MB

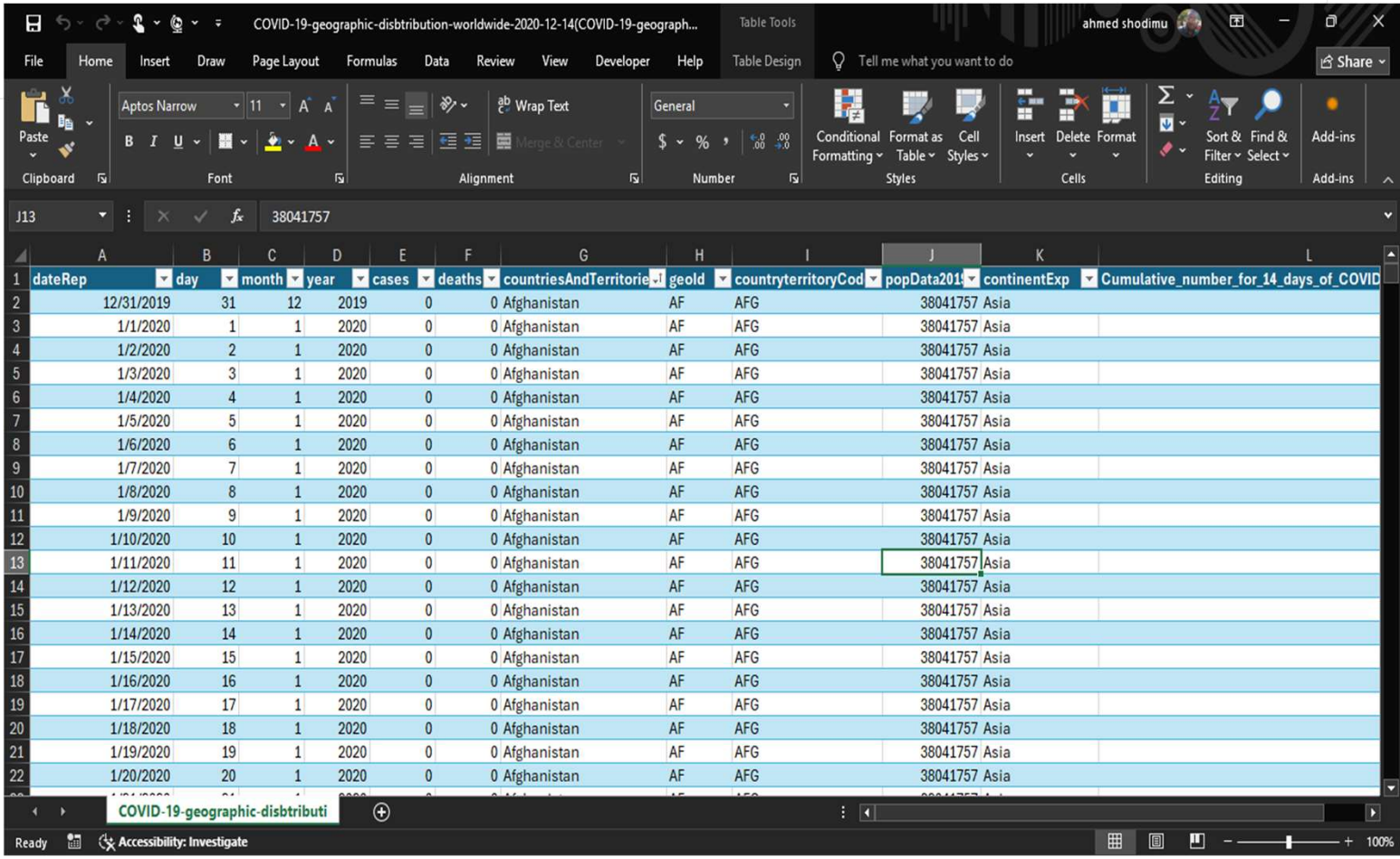
```
+------------------+
| covid19_records  |
+------------------+
| 12525825         |
+------------------+
```

```
+-----------------------+
| covid19_top10_records |
+-----------------------+
| 10660                 |
+-----------------------+
```

```
-bash-4.2$ du -h covid19_*
3.2G    covid19_aggregated.csv
2.1M    covid19_top10_countries.csv
-bash-4.2$
```

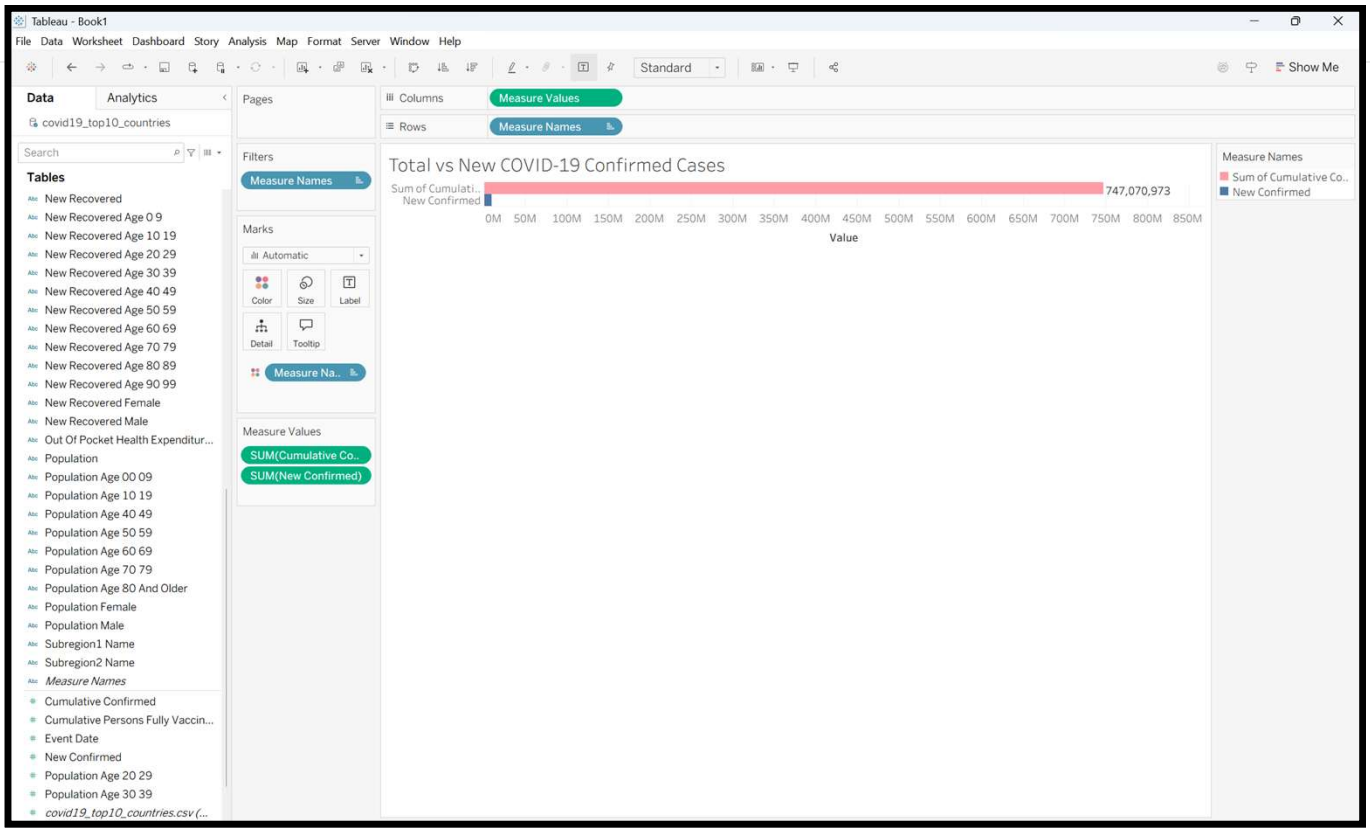
CSV → Visualization

```
bash-4.2$ ls
ge.csv  demographics.csv  epidemiology.csv  gender.csv  geography.csv  health.csv  hospitalizations.csv  index.csv  vaccinations.csv
bash-4.2$ hdfs dfs -getmerge /tmp/group3_covid19/aggregated/000* covid19_aggregated.csv
bash-4.2$ ls covid19_aggregated.csv
ovid19_aggregated.csv
bash-4.2$ ls -lha covid19_aggregated.csv
rw-r--r-- 1 bchica bchica 3.2G Dec  3 06:20 covid19_aggregated.csv
```

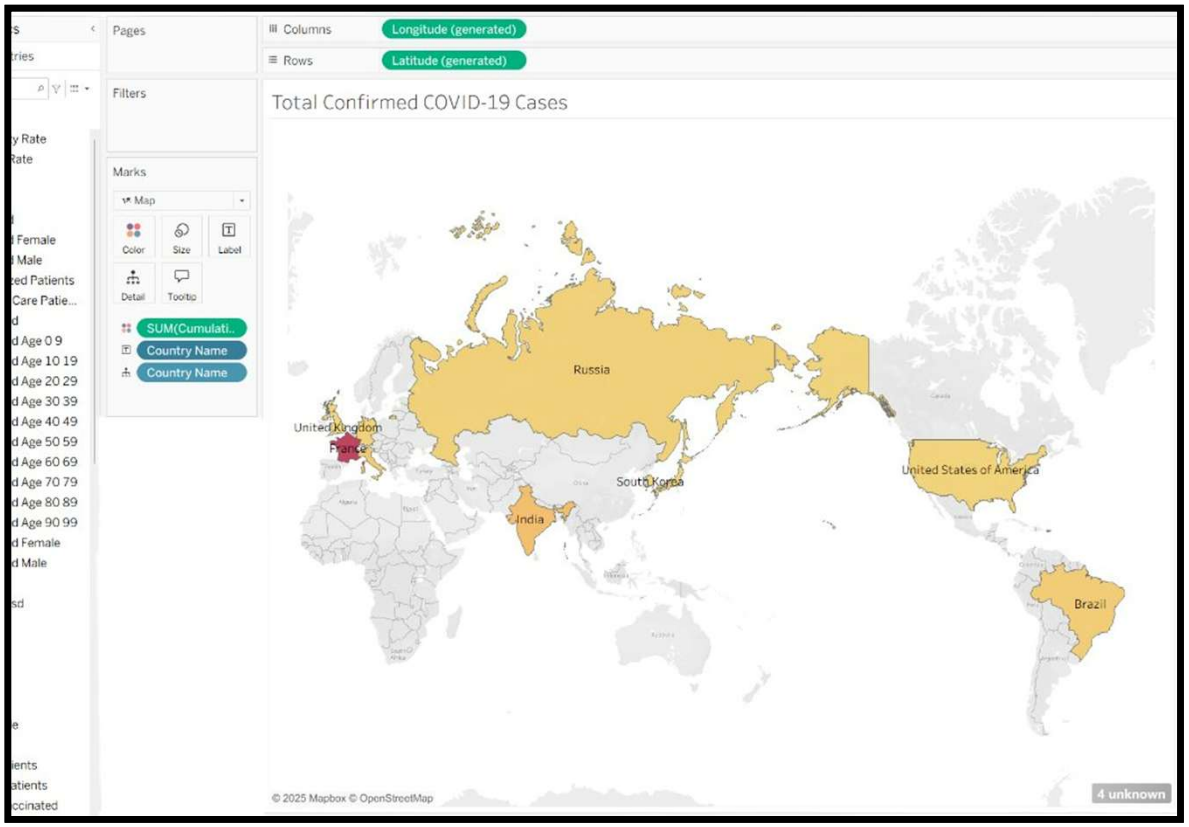


	A	B	C	D	E	F	G	H	I	J	K	L
1	dateRep	day	month	year	cases	deaths	countriesAndTerritorie	geoId	countryterritoryCod	popData201	continentExp	Cumulative_number_for_14_days_of_COVID
2	12/31/2019	31	12	2019	0	0	Afghanistan	AF	AFG	38041757	Asia	
3	1/1/2020	1	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
4	1/2/2020	2	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
5	1/3/2020	3	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
6	1/4/2020	4	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
7	1/5/2020	5	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
8	1/6/2020	6	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
9	1/7/2020	7	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
10	1/8/2020	8	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
11	1/9/2020	9	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
12	1/10/2020	10	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
13	1/11/2020	11	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
14	1/12/2020	12	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
15	1/13/2020	13	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
16	1/14/2020	14	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
17	1/15/2020	15	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
18	1/16/2020	16	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
19	1/17/2020	17	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
20	1/18/2020	18	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
21	1/19/2020	19	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	
22	1/20/2020	20	1	2020	0	0	Afghanistan	AF	AFG	38041757	Asia	

Visualization: Bar Chart

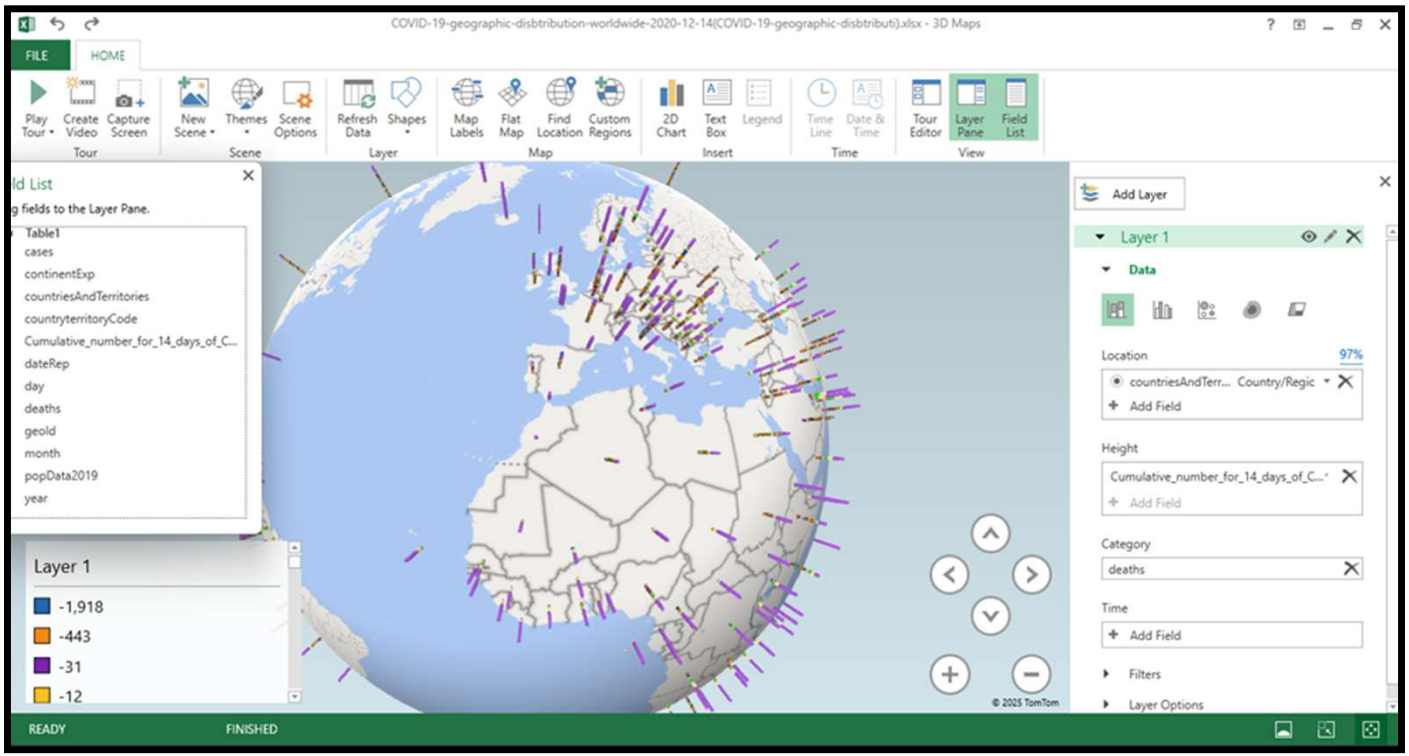


Worldwide map of confirmed COVID-19 Cases



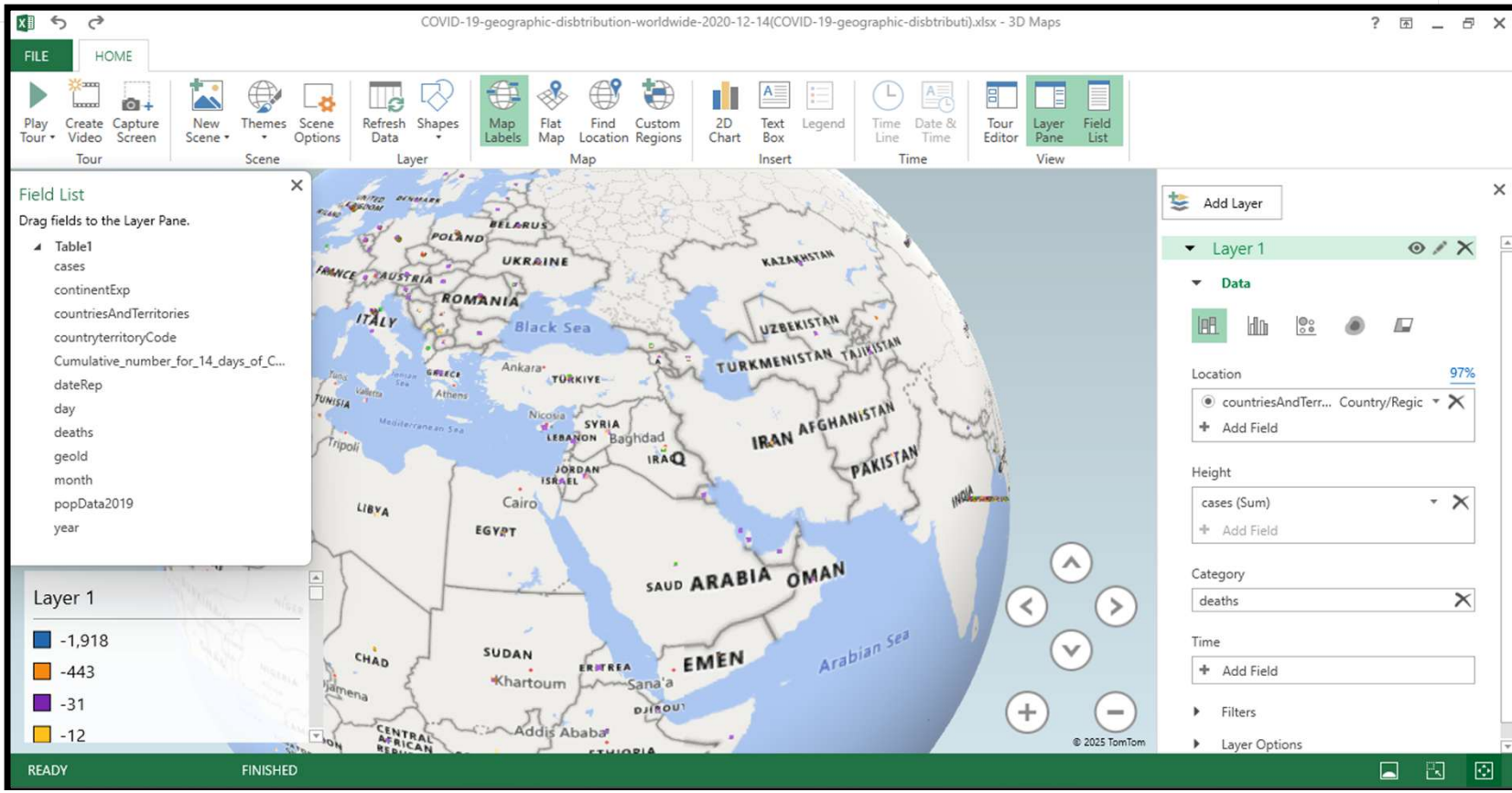
Visualization: Excel 3D Map

- Countries and Territories
- Cumulative number for 14 days per Covid-19 cases per 100,000



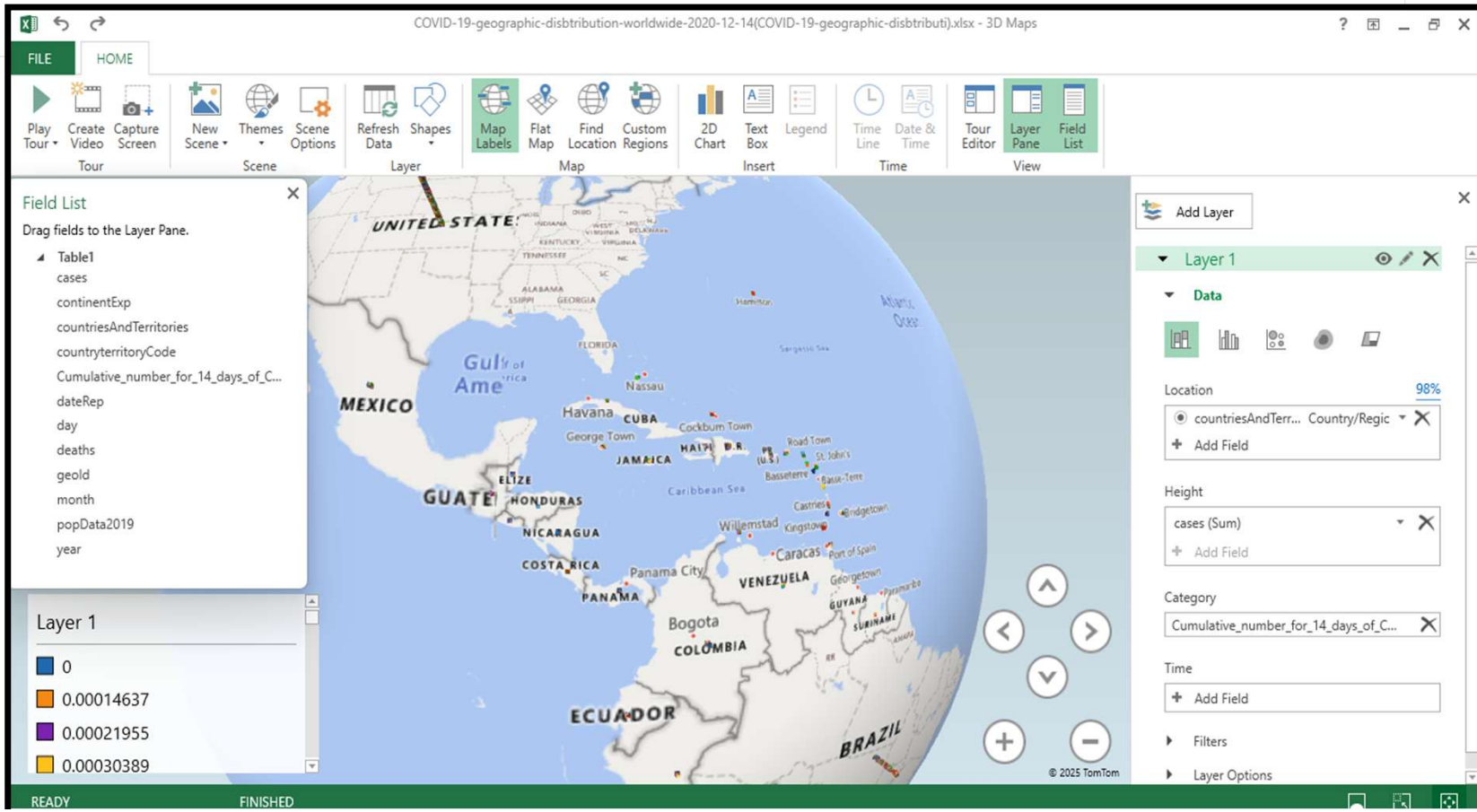
Visualization: Excel 3D Map

- Countries and Territories
- Cases
- Deaths



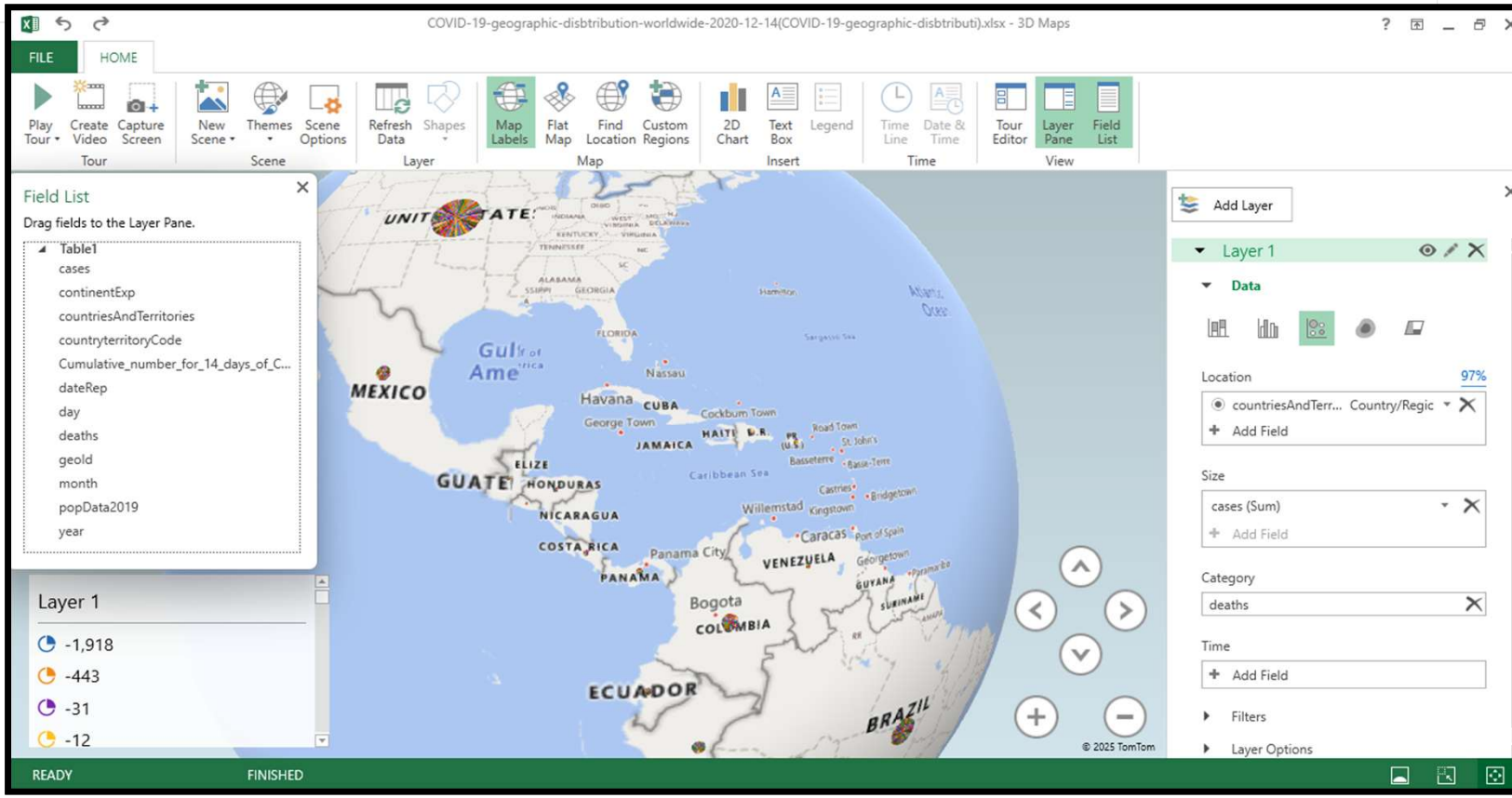
Visualization: Excel 3D Map

- Countries and Territories
- Cases
- Cumulative cases



Visualization: Excel 3D Map (Bubble)

- Countries and territories
- Cases
- Deaths



Key Findings

- COVID-19 showed steep growth trends during major variants
- Countries differed significantly in peak periods and fatality rates
- Higher population density equals higher confirmed case totals
- Earlier lockdowns flattened trends for some regions
- Data confirms seasonal and global wave patterns

Thank you

Any Questions?