

INTELIGENCIA DE NEGOCIOS

INFORME COVID-19

DOCENTE

MAURICIO BEDOYA

ESTUDIANTES

CESAR DAVID ACOSTA

STEVEN MARIN GALLEGO



Introducción

El presente informe se realiza como proyecto final para la asignatura “Inteligencia de Negocios” y como objetivo final aborda varias temáticas y herramientas propuestas en el curso. Es por esta razón que los estudiantes esperamos realizar de manera correcta los análisis pertinentes y con esto cubrir las expectativas y necesidades que el dataset seleccionado contemple.

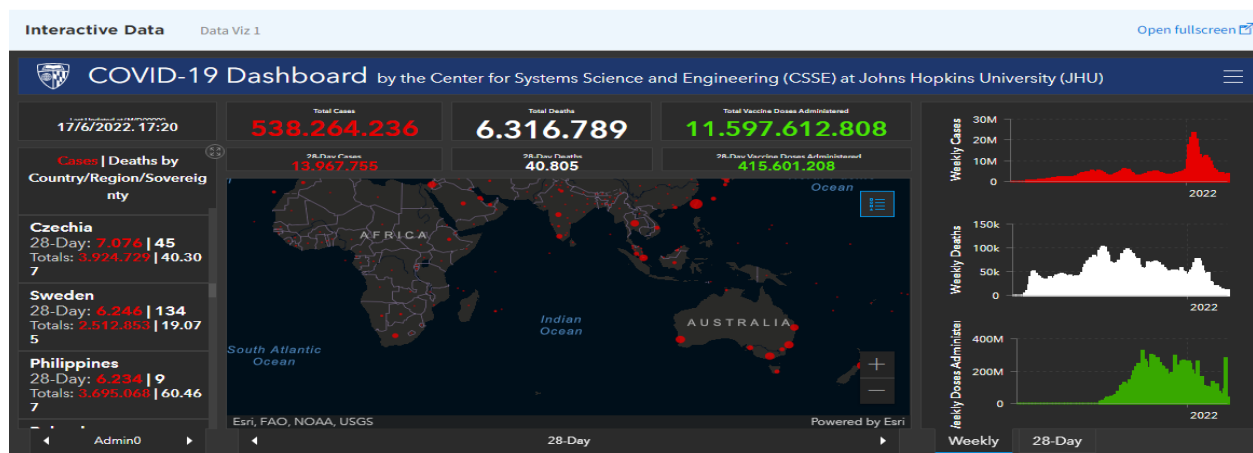
Objetivo general y específico del análisis

El dataset seleccionado para este informe es sobre los resultados post pandemia a nivel mundial, con este queremos analizar el impacto del virus y su alcance, evaluando tres tablas principales con la herramienta POWER BI.

Consideramos que este análisis permitirá identificar cómo fue la incidencia del virus durante el inicio de la pandemia en los países, los picos más altos de incubación y el descenso de los indicadores tras el despliegue de las vacunas y medidas de seguridad adoptadas por el gobierno.

Además a partir de un componente gráfico se podrá mostrar como fue la interacción del virus en cada país según el componente de mortalidad, pacientes contagiados y pacientes recuperados, entregando datos reales que pueden ayudar a la toma de decisiones ante eventuales picos del virus.

Dataset Seleccionados



<https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases>

time_series_covid19_confirmed_global.csv

time_series_covid19_deaths_global.csv

time_series_covid19_recovered_global.csv

Arquitectura propuesta

La arquitectura propuesta es sobre BI usando herramientas de ETL para preparar y consolidar la data usada en el data warehouse.

Herramientas técnicas utilizadas

Usaremos la herramienta sql developer para la creación de un procedimiento que nos permita cargar las tablas y consolidar desde allí la conexión con power BI usando el método de cargue desde ODBC de 64 bits.

Ya con la data cargada en Power BI realizaremos la limpieza final de los datos y consolidaremos las relaciones necesarias de las tablas para lograr los análisis propuestos para esta investigación.

Metodología de desarrollo

La metodología usada para la validación de los datos es por medio de ETL puesto que desde la limpieza inicial de los datasets tendremos que realizar transformaciones a los datos y con estos resultados podremos consolidar nuestro análisis final




KPI

- Contagios
- Mortalidad
- Pacientes recuperados
- Porcentaje de contagios

- Porcentaje de fallecimientos
- Porcentaje de recuperaciones
- Top de países con mayor número de casos
- Rendimiento del virus
- Porcentaje de casos activos

Desarrollo del trabajo

Como paso inicial se descargan los datasets con las 3 tablas desde el repositorio oficial.

	time_series_covid19_confirmed_global.csv	14/06/2022 11:52 a. m.	Archivo de valores...	1.280 KB
	time_series_covid19_deaths_global.csv	14/06/2022 11:51 a. m.	Archivo de valores...	888 KB
	time_series_covid19_recovered_global.csv	14/06/2022 11:52 a. m.	Archivo de valores...	845 KB

Diccionario de Datos Tabla Covid 19

Proceso de importación del archivo al Sql Developer

Vista Previa de Datos

Restaurar Estado

Origen: Archivo Local

Archivo: C:\Users\Anderson\Downloads\archive (1)\covid_19_clean_complete.csv Examinar...

Formato de Archivo

☒ Cabecera Después de Omitir Omitir Filas: 0

Formato: csv ☒ Límite de Vista Previa de Filas: 100

Codificación: Cp1252

Delimitador: ; Terminador de Línea: estándar: CR LF, CR o LF

Cierre a la Izquierda: " Cierre a la Derecha: "

Contenido del Archivo

Province/State	Country/Region	Lat	Long	Date	Confirmed	Deaths	Recovered	Active	WHO Region
	Afghanistan	33.93911	67.709953	2020-01-22	0	0	0	0	Eastern Med...
	Albania	41.1533	20.1683	2020-01-22	0	0	0	0	Europe
	Algeria	28.0339	1.6596	2020-01-22	0	0	0	0	Africa
	Andorra	42.5063	1.5218	2020-01-22	0	0	0	0	Europe
	Angola	-11.2027	17.8739	2020-01-22	0	0	0	0	Africa
	Antigua and...	17.0608	-61.7964	2020-01-22	0	0	0	0	Americas
	Argentina	-38.4161	-63.6167	2020-01-22	0	0	0	0	Americas
	Armenia	40.0691	45.0382	2020-01-22	0	0	0	0	Europe
Australian C...	Australia	-35.4735	149.0124	2020-01-22	0	0	0	0	Western Pa...
New South ...	Australia	-33.8688	151.2093	2020-01-22	0	0	0	0	Western Pa...

Corregimos errores en los nombres de las columnas

Definición de Columna

Para cada columna de la izquierda, define los detalles de la columna en la tabla de la base de datos que se creará para importar los datos.

Vista Previa de Datos

Método de Importación

Seleccionar Columnas

Definición de Columna

Terminar

Columnas de Datos de Origen

- Province/State
- Country/Region
- Lat
- Long
- Date
- Confirmed
- Deaths
- Recovered
- Active
- WHO Region

Columnas de Datos de Destino

Nombre: Province/State

Tipo de Dato: VARCHAR2

Tamaño/Precisión: 128

☒ ¿Nulo? Valor por Defecto:

Comentario:

Datos

Australian Capital Territory
New South Wales
Northern Territory
Queensland
South Australia
Tasmania
Victoria
Western Australia

Estado
El nombre de columna no es válido

Importación Correcta de la tabla

Columnas	Datos	Model	Restricciones	Permisos	Estadísticas	Disparadores	Flashback	Dependencias	Detalles	Particiones	Índices	SQL
Acciones...												
	COLUMN_NAME	DATA_TYPE	NULLABLE	DATA_DEFAULT	COLUMN_ID	COMMENTS						
1	PROVINCESTATE	VARCHAR2(128 BYTE)	Yes	(null)	1	(null)						
2	COUNTRYREGION	VARCHAR2(26 BYTE)	Yes	(null)	2	(null)						
3	LAT	NUMBER(38,0)	Yes	(null)	3	(null)						
4	LONGITUD	NUMBER(38,0)	Yes	(null)	4	(null)						
5	DATES	DATE	Yes	(null)	5	(null)						
6	CONFIRMED	NUMBER(38,0)	Yes	(null)	6	(null)						
7	DEATHS	NUMBER(38,0)	Yes	(null)	7	(null)						
8	RECOVERED	NUMBER(38,0)	Yes	(null)	8	(null)						
9	ACTIVE	NUMBER(38,0)	Yes	(null)	9	(null)						
10	WHOREGION	VARCHAR2(26 BYTE)	Yes	(null)	10	(null)						

Registros de la tabla

PROVINCESTATE	COUNTRYREGION	LAT	LONGITUD	DATES	CONFIRMED	DEATHS	RECOVERED	ACTIVE	WHOREGION
1 (null)	Afghanistan	3393911	67709953	22/01/20	0	0	0	0	0 Eastern Mediterranean
2 (null)	Albania	411533	201683	22/01/20	0	0	0	0	0 Europe
3 (null)	Algeria	280339	16596	22/01/20	0	0	0	0	0 Africa
4 (null)	Andorra	425063	15218	22/01/20	0	0	0	0	0 Europe
5 (null)	Angola	-112027	178739	22/01/20	0	0	0	0	0 Africa
6 (null)	Antigua and Barbuda	170608	-617964	22/01/20	0	0	0	0	0 Americas
7 (null)	Argentina	-384161	-636167	22/01/20	0	0	0	0	0 Americas
8 (null)	Armenia	400691	450382	22/01/20	0	0	0	0	0 Europe
9 Australian Capital Territory	Australia	-354735	1490124	22/01/20	0	0	0	0	0 Western Pacific
10 New South Wales	Australia	-338688	1512093	22/01/20	0	0	0	0	0 Western Pacific
11 Northern Territory	Australia	-124634	1308456	22/01/20	0	0	0	0	0 Western Pacific
12 Queensland	Australia	-274698	1530251	22/01/20	0	0	0	0	0 Western Pacific
13 South Australia	Australia	-349285	1386007	22/01/20	0	0	0	0	0 Western Pacific
14 Tasmania	Australia	-428821	1473272	22/01/20	0	0	0	0	0 Western Pacific
15 Victoria	Australia	-378136	1449631	22/01/20	0	0	0	0	0 Western Pacific
16 Western Australia	Australia	-319505	1158605	22/01/20	0	0	0	0	0 Western Pacific
17 (null)	Austria	475162	145501	22/01/20	0	0	0	0	0 Europe
18 (null)	Azerbaijan	401431	475769	22/01/20	0	0	0	0	0 Europe
19 (null)	Bahamas	25025885	-78035889	22/01/20	0	0	0	0	0 Americas
20 (null)	Bahrain	260275	5055	22/01/20	0	0	0	0	0 Eastern Mediterranean
21 (null)	Bangladesh	23685	903563	22/01/20	0	0	0	0	0 South-East Asia
22 (null)	Barbados	131939	-595432	22/01/20	0	0	0	0	0 Americas
23 (null)	Belarus	537098	279534	22/01/20	0	0	0	0	0 Europe
24 (null)	Belgium	508333	4469936	22/01/20	0	0	0	0	0 Europe
25 (null)	Benin	93077	23158	22/01/20	0	0	0	0	0 Africa

Limpieza de Datos

Hoja de Trabajo

Generador de Consultas

```

alter table covid19
drop column PROVINCESTATE;

alter table covid19
drop column LAT;

alter table covid19
drop column Longitud;

alter table covid19
drop column WHOREGION;

```

Salida de Script

Tarea terminada en 0,797 segundos

Table COVID19 alterado.

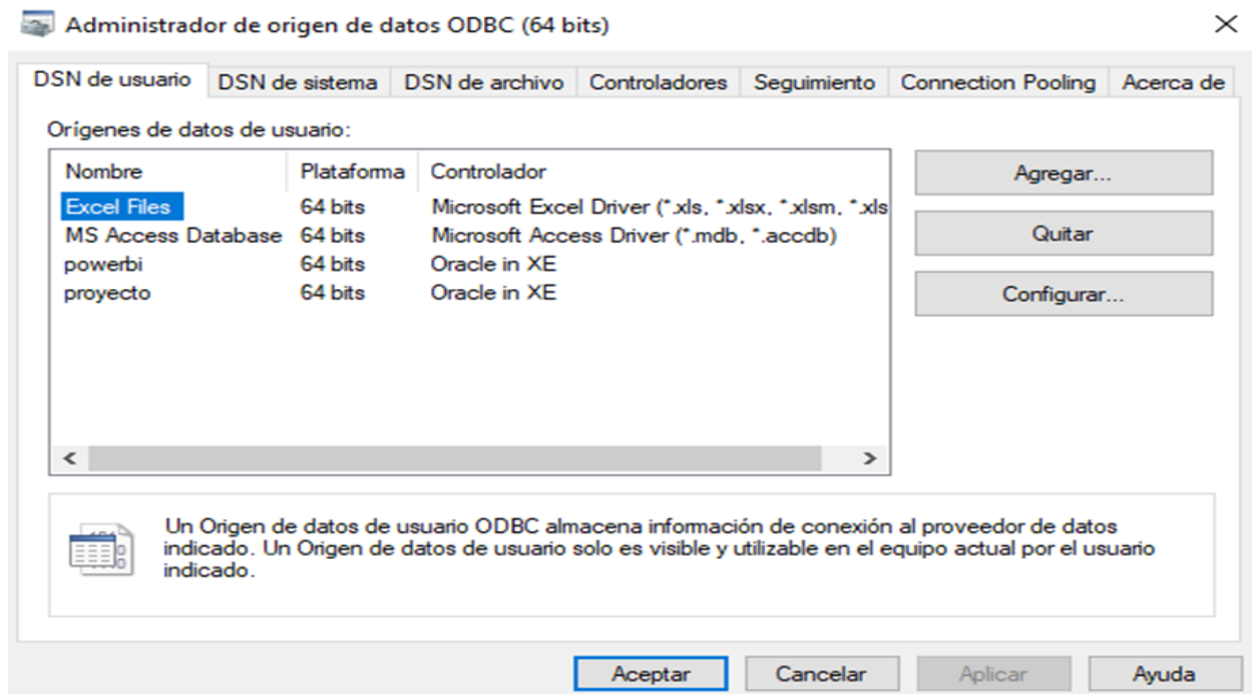
Table COVID19 alterado.

Table COVID19 alterado.

Tabla con las columnas finales

Columnas Datos Model Restricciones Permisos Estadísticas Disparadores Flashback Dependencias Detalles Particiones Índices SQL							
Ordenar... Filtrar:							
COUNTRY/REGION	DATES	CONFIRMED	DEATHS	RECOV...	ACTIVE		
1 Afghanistan	22/01/20	0	0	0	0		
2 Albania	22/01/20	0	0	0	0		
3 Algeria	22/01/20	0	0	0	0		
4 Andorra	22/01/20	0	0	0	0		
5 Angola	22/01/20	0	0	0	0		
6 Antigua and Barbuda	22/01/20	0	0	0	0		
7 Argentina	22/01/20	0	0	0	0		
8 Armenia	22/01/20	0	0	0	0		
9 Australia	22/01/20	0	0	0	0		
10 Australia	22/01/20	0	0	0	0		
11 Australia	22/01/20	0	0	0	0		
12 Australia	22/01/20	0	0	0	0		
13 Australia	22/01/20	0	0	0	0		
14 Australia	22/01/20	0	0	0	0		
15 Australia	22/01/20	0	0	0	0		
16 Australia	22/01/20	0	0	0	0		
17 Austria	22/01/20	0	0	0	0		
18 Azerbaijan	22/01/20	0	0	0	0		
19 Bahamas	22/01/20	0	0	0	0		
20 Bahrain	22/01/20	0	0	0	0		
21 Bangladesh	22/01/20	0	0	0	0		
22 Barbados	22/01/20	0	0	0	0		
23 Belarus	22/01/20	0	0	0	0		
24 Belgium	22/01/20	0	0	0	0		
25 Benin	22/01/20	0	0	0	0		

Conexion ODBC



Agregamos un Nuevo Origen de datos de tipo Oracle in XE

Asignamos el nombre de usuario del espacio de trabajo creado en SqlDeveloper, y el nombre del servidor.

Oracle ODBC Driver Configuration

Data Source Name: proyecto

Description:

TNS Service Name: localhost:1521/xs

User ID: ROBOTIN

Application | Oracle | Workarounds | SQLServer Migration

Enable Result Sets ☒ Enable Query Timeout ☒ Read-Only Connection ☐

Enable Closing Cursors ☐ Enable Thread Safety ☒

Batch Autocommit Mode: Commit only if all statements succeed

Numeric Settings: Use Oracle NLS settings

OK Cancel Help Test Connection

Test de Conexión

Oracle ODBC Driver Configuration

Data Source Name: proyecto

Description:

TNS Service Name: localhost:1521/xs

User ID: ROBOTIN

Application | Oracle | Workarounds | SQLServer Migration

Enable Result Sets ☒ Enable Query Timeout ☒ Read-Only Connection ☐

Enable Closing Cursors ☐ Enable Thread Safety ☒

Batch Autocommit Mode: Commit only if all statements succeed

Numeric Settings: Use Oracle NLS settings

OK Cancel Help Test Connection

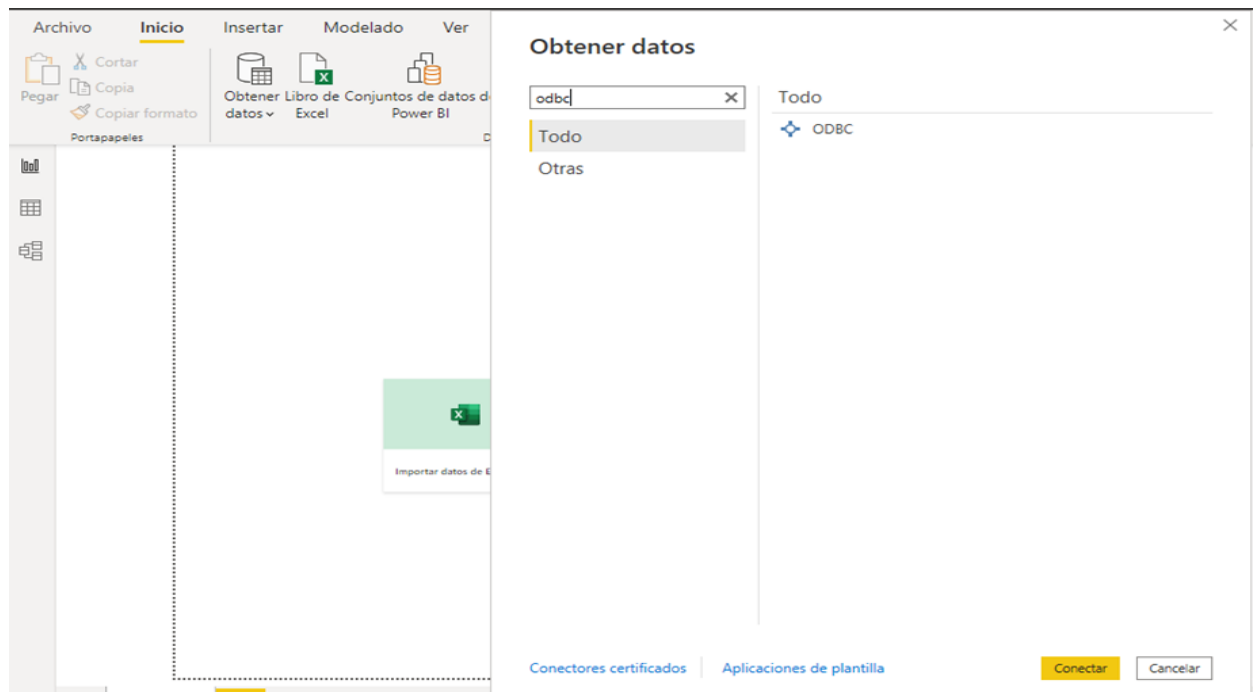
Testing Connection

Connection successful

Aceptar

El siguiente paso es dirigirnos a Power Bi para crear un nuevo proyecto donde incluiremos esta conexión, eso nos permitirá incluir todas las tablas que tenga nuestra conexión declarada previamente en el ODBC.

Seleccionamos desde el botón obtener datos el tipo de conexión que vamos a utilizar para importar nuestra tabla, damos click en conectar.



Seleccionamos la instancia creada y damos clic en aceptar



Seleccionamos el usuario de nuestro espacio de trabajo y escogemos la o las tablas que deseamos importar, si nuestros datos requieren algún cambio previo a la importación debemos dar click en la opción transformar datos, de lo contrario seleccionamos la opción cargar.

Navegador

🔍

Opciones de presentación ▾

📁

ODBC (dsn=proyecto) [5]

▶

📁 APEX_040000

▶

📁 CTXSYS

▶

📁 MDSYS

▶

📁 ROBOTIN [4]

☑

📅 COVID19

☐

📅 DETALLES_TERMINACION_CONTRATO

☐

📅 RECLUTACION_EMPLEADO

☐

📅 TABLA_RENDIMIENTO_SALARIO

▶

📁 XDB

COVID19

📄

COUNTRYREGION	DATES	CONFIRMED	DEATHS	R
Afghanistan	22/01/2020 12:00:00 a. m.	0		
Albania	22/01/2020 12:00:00 a. m.	0		
Algeria	22/01/2020 12:00:00 a. m.	0		
Andorra	22/01/2020 12:00:00 a. m.	0		
Angola	22/01/2020 12:00:00 a. m.	0		
Antigua and Barbuda	22/01/2020 12:00:00 a. m.	0		
Argentina	22/01/2020 12:00:00 a. m.	0		
Armenia	22/01/2020 12:00:00 a. m.	0		
Australia	22/01/2020 12:00:00 a. m.	0		
Australia	22/01/2020 12:00:00 a. m.	0		
Australia	22/01/2020 12:00:00 a. m.	0		
Australia	22/01/2020 12:00:00 a. m.	0		
Australia	22/01/2020 12:00:00 a. m.	0		
Australia	22/01/2020 12:00:00 a. m.	0		
Australia	22/01/2020 12:00:00 a. m.	0		
Australia	22/01/2020 12:00:00 a. m.	0		
Australia	22/01/2020 12:00:00 a. m.	0		
Austria	22/01/2020 12:00:00 a. m.	0		
Azerbaijan	22/01/2020 12:00:00 a. m.	0		
Bahamas	22/01/2020 12:00:00 a. m.	0		
Bahrain	22/01/2020 12:00:00 a. m.	0		
Bangladesh	22/01/2020 12:00:00 a. m.	0		
Barbados	22/01/2020 12:00:00 a. m.	0		
Belarus	22/01/2020 12:00:00 a. m.	0		

<

>

En este último paso todos los registros de la tabla seleccionada quedaron correctamente importados en Power Bi

COUNTRYREGION ▼	DATES ▼	CONFIRMED ▼	DEATHS ▼	RECOVERED ▼	ACTIVE ▼
Afghanistan	22/01/2020 12:00:00 a. m.	0	0	0	0
Albania	22/01/2020 12:00:00 a. m.	0	0	0	0
Algeria	22/01/2020 12:00:00 a. m.	0	0	0	0
Andorra	22/01/2020 12:00:00 a. m.	0	0	0	0
Angola	22/01/2020 12:00:00 a. m.	0	0	0	0
Antigua and Barbuda	22/01/2020 12:00:00 a. m.	0	0	0	0
Argentina	22/01/2020 12:00:00 a. m.	0	0	0	0
Armenia	22/01/2020 12:00:00 a. m.	0	0	0	0
Australia	22/01/2020 12:00:00 a. m.	0	0	0	0
Australia	22/01/2020 12:00:00 a. m.	0	0	0	0
Australia	22/01/2020 12:00:00 a. m.	0	0	0	0
Australia	22/01/2020 12:00:00 a. m.	0	0	0	0
Australia	22/01/2020 12:00:00 a. m.	0	0	0	0
Australia	22/01/2020 12:00:00 a. m.	0	0	0	0
Australia	22/01/2020 12:00:00 a. m.	0	0	0	0
Australia	22/01/2020 12:00:00 a. m.	0	0	0	0
Australia	22/01/2020 12:00:00 a. m.	0	0	0	0
Austria	22/01/2020 12:00:00 a. m.	0	0	0	0
Azerbaijan	22/01/2020 12:00:00 a. m.	0	0	0	0
Bahamas	22/01/2020 12:00:00 a. m.	0	0	0	0
Bahrain	22/01/2020 12:00:00 a. m.	0	0	0	0
Bangladesh	22/01/2020 12:00:00 a. m.	0	0	0	0
Barbados	22/01/2020 12:00:00 a. m.	0	0	0	0
Belarus	22/01/2020 12:00:00 a. m.	0	0	0	0

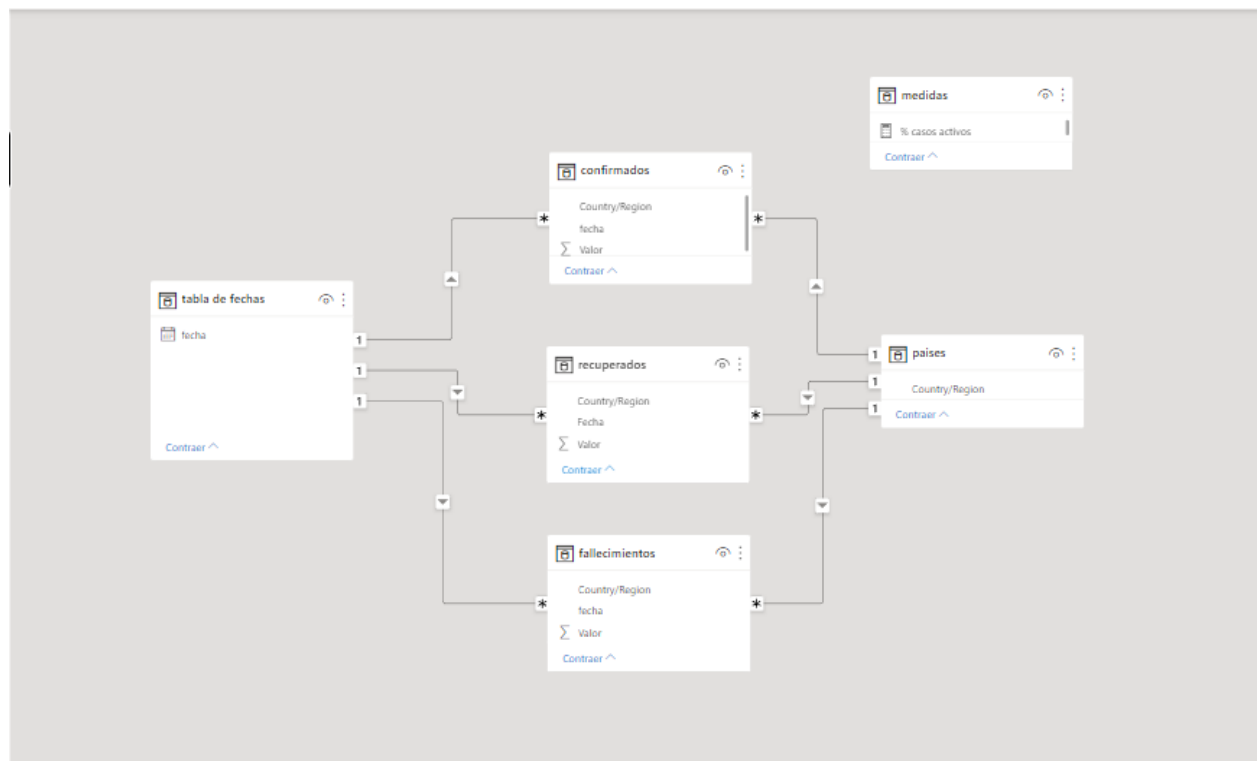
Validación de datos en POWER BI

The screenshot displays the Power BI Desktop interface with the 'Transformar' (Transform) ribbon selected. The main view shows a data table with columns: Country/Region, Valor, and fecha. The table contains 21 rows of data for Afghanistan, with dates ranging from 22/01/2020 to 11/02/2020. The 'Valor' column shows a mix of valid (Válido), error (Error), and empty (Vacío) values.

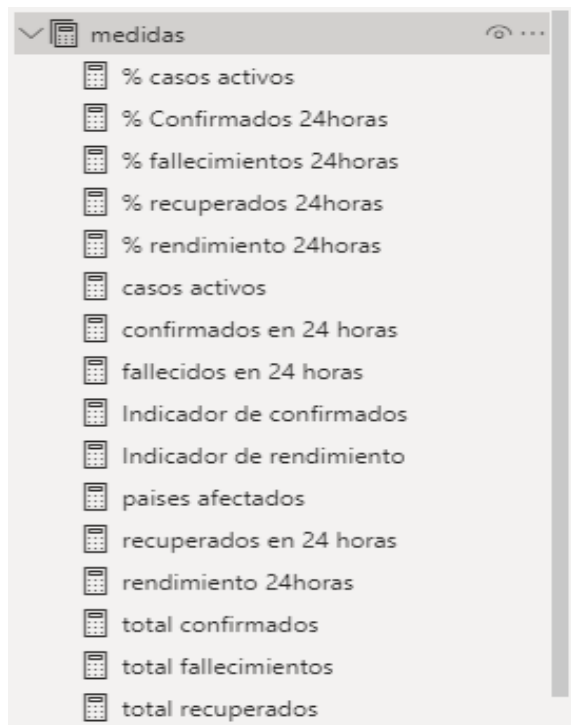
On the right side, the 'Configuración de la consulta' (Query Settings) pane is open, showing the 'PROPIEDADES' (Properties) and 'PASOS APLICADOS' (Applied Steps) sections. The 'PASOS APLICADOS' section lists various transformations applied to the data, including 'Columnas quitadas1' (Columns removed 1).

At the bottom of the interface, a status bar indicates '3 COLUMNAS, 999+ FILAS' and 'Generación de perfiles de columnas basada en las 1000 primeras filas'.

Creación de relaciones entre tablas



Creación de medidas para análisis de datos



Medidas para toma de datos

Total Confirmados = Acumulado de confirmados en el ultimo dia

Total Fallecidos = Acumulado de fallecidos en el último dia

Total Recuperados = Acumulado de recuperados en el último dia

Confirmados en 24 horas: (Total confirmados - Total confirmados dia anterior)

Fallecidos en 24 horas: (Total fallecidos - Total fallecidos dia anterior)

Recuperados en 24 horas: (Total recuperados - Total recuperados dia anterior)

% Confirmados 24h= Confirmados 24h / Total confirmados

% Fallecidos 24h= Fallecidos 24h / Total fallecidos

% Recuperados 24h= Recuperados 24h / Total recuperados

Casos Activos = Total confirmados - Total fallecidos - Total recuperados

%Casos Activos = Casos Activos / Total Recuperados

Países Afectados = Recuento de países con algún caso confirmado

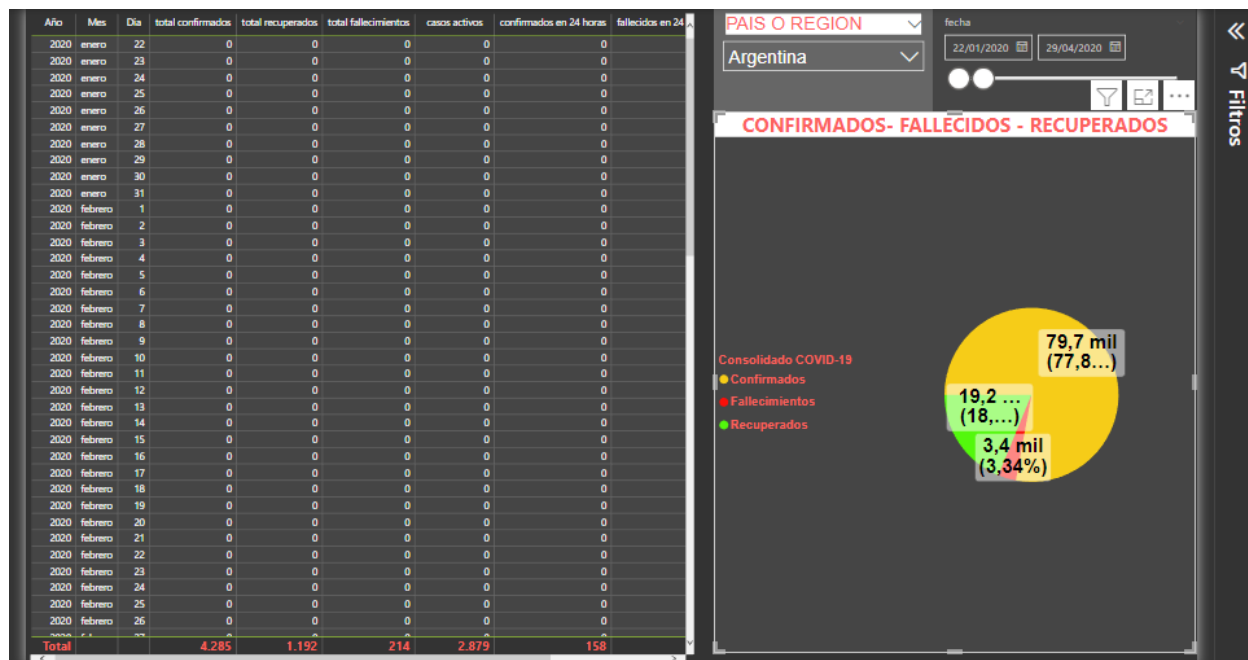
Rendimiento 24h = Recuperados 24h - Confirmados 24h - Fallecidos 24h

% Rendimiento 24h= % Recuperados 24h -% Confirmados 24h - % Fallecidos

Elaboracion del centro interactivo de mando de los datos



Tabla de datos



Centro de mando

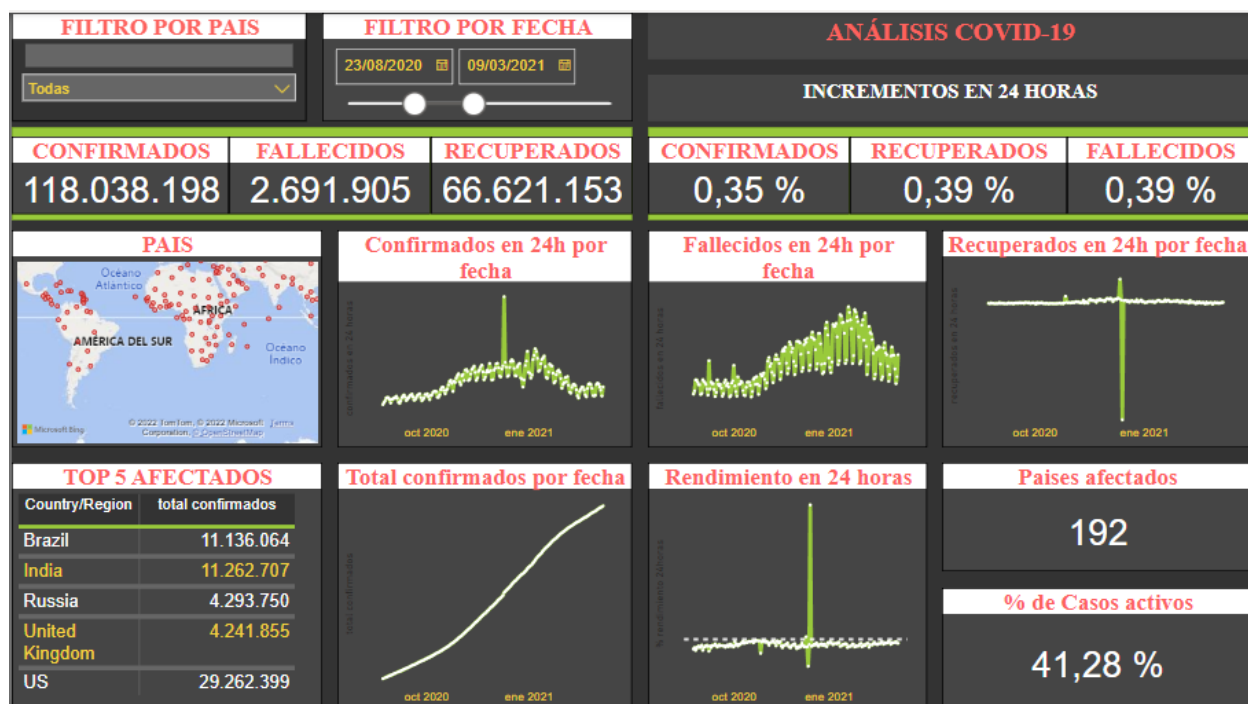


Tabla de Objetivos definidos



Medidas utilizadas

Medidas para toma de datos	≡ Filtros
Total Confirmados = Acumulado de confirmados en el ultimo dia	
Total Fallecidos = Acumulado de fallecidos en el último dia	
Total Recuperados = Acumulado de recuperados en el último dia	
Confirmados en 24 horas: (Total confirmados - Total confirmados dia anterior)	
Fallecidos en 24 horas: (Total fallecidos - Total fallecidos dia anterior)	
Recuperados en 24 horas: (Total recuperados - Total recuperados dia anterior)	
% Confirmados 24h= Confirmados 24h / Total confirmados	
% Fallecidos 24h= Fallecidos 24h / Total fallecidos	
% Recuperados 24h= Recuperados 24h / Total recuperados	
Casos Activos = Total confirmados - Total fallecidos - Total recuperados	
%Casos Activos = Casos Activos / Total Recuperados	
Países Afectados = Recuento de países con algún caso confirmado	
Rendimiento 24h = Recuperados 24h - Confirmados 24h - Fallecidos 24h	
% Rendimiento 24h= % Recuperados 24h - % Confirmados 24h - % Fallecidos	
Indicador confirmados = (% Confirmados en 24 horas + % Confirmados 24 horas dia anterior + % Confirmados 24 horas 2 dias)/3	

Conclusiones

La importancia de las medidas creadas, en este caso las que nos indican los confirmados, fallecidos y recuperados en 24 horas, es que con estos datos podemos hacer seguimiento del incremento del virus y validar en qué fecha del mes, hubo una tendencia al alza con respecto a los contagios, estas medidas reflejan el incremento y la mortalidad que la pandemia generó en cada país.

En el caso de la data estadística a nivel global nuestro dash permite evaluar los países más afectados por la pandemia y con esto podemos asumir ciertas situaciones que pudieron influir en los contagios.

Por ejemplo la cercanía del país evaluado con respecto al país emisor de la pandemia y quizás las malas prácticas de seguridad o el manejo inadecuado de las medidas de prevención, lo cual genera el incremento en los contagios y la mortalidad en el país.

Para hacer esta prueba se tomó como referente el continente europeo durante los 3 primeros meses de la pandemia, con lo cual pudimos confirmar que este tuvo una tendencia más alta de contagios con respecto al continente americano, puesto que hubo más países afectados y además los contagios y la mortalidad siempre mantuvo una tendencia al incremento durante las fechas evaluadas.

Como propuesta de evaluación se propone crear un dash donde a partir de la data recolectada y luego de definir 4 fases de revisión con indicadores de rendimiento, podremos validar en qué países y de acuerdo a rangos de fechas establecidos, que países dieron mejor manejo a la problemática generada con esta situación, en el caso de encontrarse resultados por fuera de los indicadores se propone repetir la prueba.