



USUARIOS TRANSPORTADOS CON SUBE

PROYECTO FINAL

FELICITAS REPETTO, FERNANDO PARISI, EZEQUIEL MUÑOZ

AÑO:2022

Índice

USUARIOS TRANSPORTADOS CON SUBE	3
1. Tema	3
2. Objetivo	3
3. Alcance	3
4. Datasets	3
5. Diagrama Entidad - Relación	4
6. Listado de tablas	4
7. Listado de columnas por tablas	5
8. Transformación de datos	6
9. Modelo relacional en Power BI	6
10. Medidas calculadas generadas y sus fórmulas	7
11. Segmentaciones elegidas	9
12. Solapas del tablero en Power BI	10
12.1 Portada	10
12.2 Usuarios	11
12.3 Provincias y Líneas	12
12.4 Empresas	13
12.5 Variación MoM	14
13. Conclusión	15

USUARIOS TRANSPORTADOS CON SUBE

PROYECTO FINAL

1. Tema: SUBE – Cantidad de Usuarios por día

Se extrajo información de la página datos.gob.ar publicado por el ministerio de transporte de Argentina, acerca de los viajes realizados por usuarios de SUBE en transporte público como colectivo, tren, subte y lanchas. Los registros que obtuvimos están detallados por tipo de transporte, día que se realizaron los viajes, locación, empresa y línea. Los datos van desde enero de 2020 hasta abril de 2022.

2. Objetivo

Según los datos que se extraigan de SUBE, podremos analizar la cantidad de pasajeros que se transportan por día según el tipo de transporte, en cada Provincia. Se verá cómo varía la cantidad de pasajeros haciendo una comparación mes a mes desde 2020 hasta 2022.

3. Alcance

El proyecto va dirigido a los miembros del Ministerio de Transporte de la Nación que se encargan de la administración de SUBE.

El tablero realizado les permitirá analizar diferentes cuestiones como por ejemplo la evolución de los usuarios transportados en los diferentes territorios de Argentina, y detectar la necesidad de aumentar líneas de transporte donde se encuentren cantidades de pasajeros saturadas y también reducir líneas en aquellos lugares en donde el transporte de personas es muy reducido. El objetivo sería reducir gastos y satisfacer la demanda del transporte público de forma adecuada.

Además, se podrá analizar la participación de cada empresa en el transporte público en general teniendo en claro la cantidad de usuarios que transporta cada una y cuáles son las empresas claves que contribuyen a satisfacer la demanda de transporte.

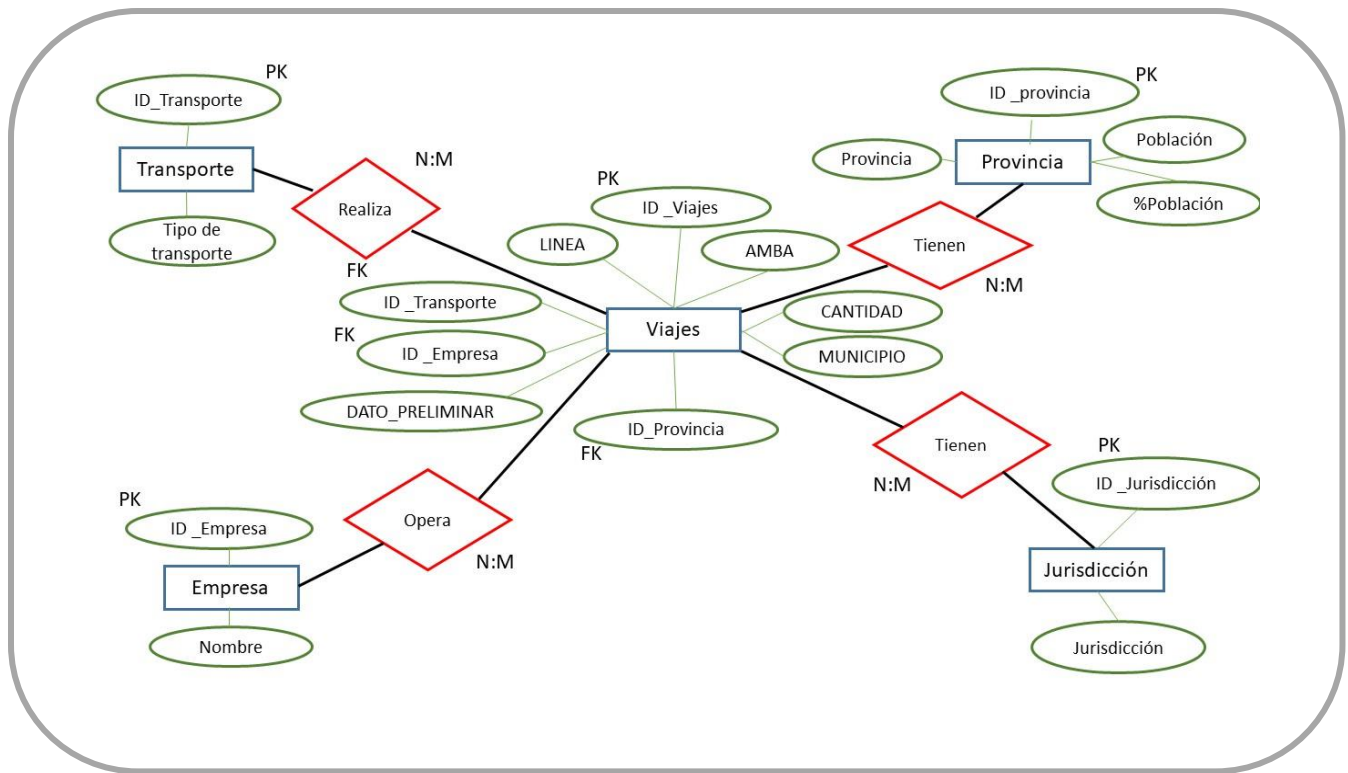
Al poder analizar el transporte de los usuarios por periodos de tiempo, se podrán sacar conclusiones de los meses con mayor demanda y aquellos en que la misma disminuye, accionando así a administrar la frecuencia de cada línea de transporte adecuadamente.

4. Datasets

A continuación, podemos ver el link que nos dirige a la Base de Datos utilizada para este trabajo.

https://docs.google.com/spreadsheets/d/1mDOjZkXj_MNfeW2kL2wX1yOR7eCHayAh/edit?usp=sharing&ouid=104608276796681899864&rtpof=true&sd=true

5. Diagrama Entidad - Relación



6. Listado de tablas:

- **Tabla 1: Viajes** Nos muestra todos los viajes realizados, y para cada uno especifica el día en que se realizó, La empresa que lo realizó, el tipo de transporte, La Línea, La ubicación y la cantidad de pasajeros.
- **Tabla 2: Empresa** Nos muestra el nombre de cada Empresa de Transporte que realiza los viajes y el ID de cada una.
- **Tabla 3: Provincia** Nos muestra el nombre de cada provincia, su ID, así como también la cantidad de población que tiene y que porcentaje representa en relación con la población total del país.
- **Tabla 4: Transporte** Nos muestra el tipo de Transporte que realizan los viajes y el ID de cada uno.
- **Tabla 5: Jurisdicción** Nos muestra el nombre de la Jurisdicción donde se realizaron los viajes y el ID de cada uno.

7. Listado de columnas por tablas

Tabla 1: Viajes

Tipo de clave	Campo	Tipo de Campo
PK	ID_Viajes	Int
-	DIA_TRANSPORTE	Datetime
FK	ID_EMPRESA	Int
-	LINEA	Varchar (255)
-	AMBA	Text
FK	ID_TIPO_TRANSPORTE	Int
FK	ID_JURISDICCION	Int
FK	ID_PROVINCIA	Int
-	MUNICIPIO	Varchar (255)
-	CANTIDAD	Int
-	DATO_PRELIMINAR	Text

Tabla 2: Empresa

Tipo de clave	Campo	Tipo de Campo
-	NOMBRE_EMPRESA	Varchar (255)
PK	ID_EMPRESA	Int

Tabla 3: Provincia

Tipo de clave	Campo	Tipo de Campo
-	PROVINCIA	Text
PK	ID_PROVINCIA	Int
-	POBLACION	Int
-	%POBLACION	Decimal (3,1)

Tabla 4: Transporte

Tipo de clave	Campo	Tipo de Campo
-	TIPO_TRANSPORTE	Text
PK	ID_TIPO_TRANSPORTE	Int

Tabla 5: Jurisdicción

Tipo de clave	Campo	Tipo de Campo
-	JURISDICCION	Varchar (255)
PK	ID_JURISDICCION	Int

8. Transformación de datos

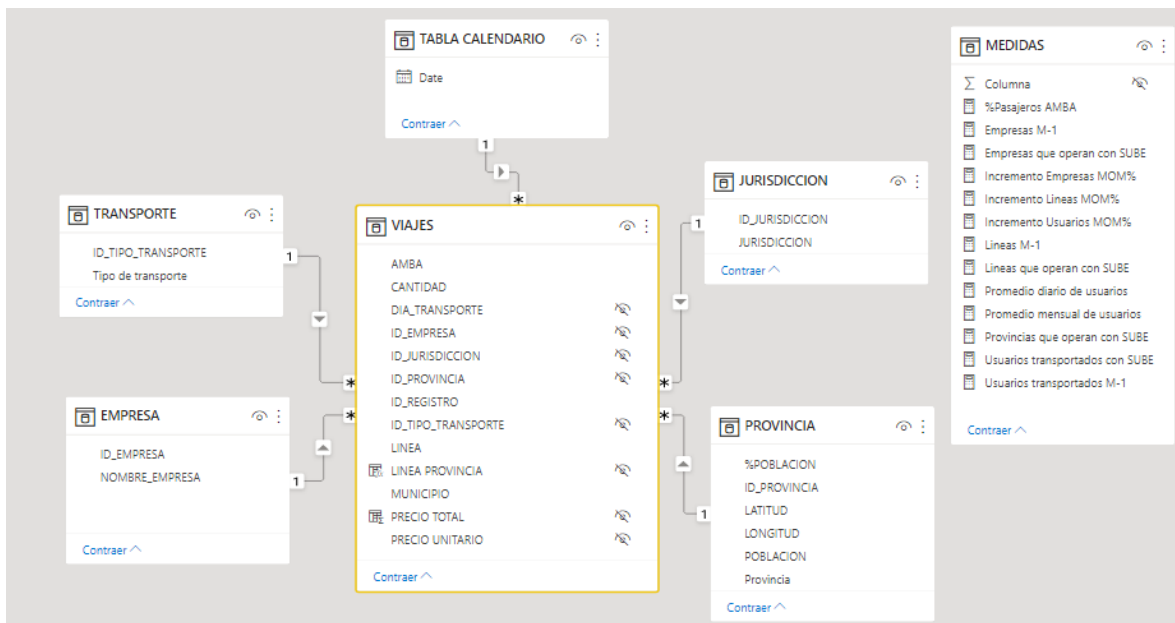
Tabla Viajes

- Columna DATO PRELIMINAR -> Se eliminó columna
- Columna: ID_REGISTRO -> Se cambió de formato alfanumérico a número entero
- Columna: ID_EMPRESA -> Se cambió de formato alfanumérico a número entero
- Columna: AMBA -> Se cambió de formato alfanumérico a texto
- Columna CANTIDAD -> Se cambió de formato alfanumérico a número entero
- Columna: PRECIO UNITARIO -> Se agregó para calcular el precio total

Tabla Provincia

- Columna: %POBLACION -> Se cambió de formato de número decimal a porcentaje
- Columna: ID_PROVINCIA -> Se cambió de formato alfanumérico a número entero
- Columna: POBLACIÓN -> Se cambió de formato alfanumérico a número entero
- Columna: LATITUD -> Se agregó columna
- Columna: LONGITUD -> Se agregó columna
- Fila: Jurisdicción Nacional (ID_PROVINCIA: 1) -> Se eliminó fila

9. Modelo relacional en Power BI



Respecto al modelo entidad-relación que se graficó al principio, se realizaron algunos cambios:

- Se agregó la TABLA CALENDARIO al modelo.
- Se agregaron las columnas LATITUD y LONGITUD en la Tabla PROVINCIA para poder definir correctamente la ubicación de las burbujas en el mapa
- Se colocaron las MEDIDAS en una única tabla por fuera del modelo.
- Se incorporó la columna PRECIO UNITARIO en la Tabla VIAJES, para poder cumplir con la consigna de “Columnas Calculadas”
- Se agregaron las columnas PRECIO TOTAL y LINEA PROVINCIA, para poder cumplir con la consigna de “Columnas Calculadas”

10. Medidas calculadas generadas y sus fórmulas:

- **%Pasajeros AMBA =**

```
VAR Cantidad_Pasajeros_AMBA =
    CALCULATE ( SUM ( VIAJES[CANTIDAD] ), VIAJES[AMBA]="SI" )
RETURN
    (Cantidad_Pasajeros_AMBA / SUM(VIAJES[CANTIDAD]) + 0)
```

- **Incremento Lineas MOM% =**

```
VAR Cantidad_Mes_Actual =
    CALCULATE ( DISTINCTCOUNT(VIAJES[LINEA]) )

VAR Cantidad_Mes_Anterior =
    CALCULATE (
        DISTINCTCOUNT(VIAJES[LINEA]),
        DATEADD ( 'TABLA CALENDARIO'[Date], -1, MONTH )
    )

RETURN
    ( ( Cantidad_Mes_Actual - Cantidad_Mes_Anterior ) /
    Cantidad_Mes_Anterior )
```

- **Incremento Empresas MOM% =**

```
VAR Cantidad_Mes_Actual =
    CALCULATE ( DISTINCTCOUNT(VIAJES[ID_EMPRESA]) )
```

```

VAR Cantidad_Mes_Anterior =
    CALCULATE (
        DISTINCTCOUNT(VIAJES[ID_EMPRESA]),
        DATEADD ( 'TABLA CALENDARIO'[Date], -1, MONTH )
    )
RETURN
    ( ( Cantidad_Mes_Actual - Cantidad_Mes_Anterior ) /
    Cantidad_Mes_Anterior )

```

- **Incremento Usuarios MOM% =**

```

VAR Cantidad_Mes_Actual =
    CALCULATE ( SUM ( VIAJES[CANTIDAD] ) )
VAR Cantidad_Mes_Anterior =
    CALCULATE (
        SUM ( VIAJES[CANTIDAD] ),
        DATEADD ( 'TABLA CALENDARIO'[Date], -1, MONTH )
    )
RETURN
    ( ( Cantidad_Mes_Actual - Cantidad_Mes_Anterior ) /
    Cantidad_Mes_Anterior )

```

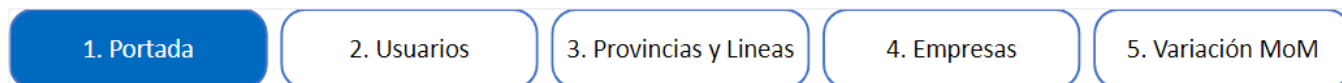
- **Empresas que operan con SUBE =**
`IF(ISBLANK(DISTINCTCOUNT(VIAJES[ID_EMPRESA])),0,DISTINCTCOUNT(VIAJES[ID_EMPRESA]))`
- **Empresas M-1 =**
`CALCULATE(DISTINCTCOUNT(VIAJES[ID_EMPRESA]), DATEADD('TABLA CALENDARIO'[Date], -1, MONTH))`
- **Usuarios transportados con SUBE =**
`IF(ISBLANK(SUM(VIAJES[CANTIDAD])),0,SUM(VIAJES[CANTIDAD]))`

- **Usuarios transportados M-1 =**
`CALCULATE(sum(VIAJES[CANTIDAD]), DATEADD('TABLA CALENDARIO'[Date],-1,MONTH))`
- **Lineas M-1 =**
`CALCULATE(DISTINCTCOUNT(VIAJES[LINEA]), DATEADD('TABLA CALENDARIO'[Date],-1,MONTH))`
- **Lineas que operan con SUBE =**
`IF(ISBLANK(DISTINCTCOUNT(VIAJES[LINEA])),0,DISTINCTCOUNT(VIAJES[L
INEA]))`
- **Provincias que operan con SUBE =**
`CALCULATE(DISTINCTCOUNT(PROVINCIA[ID_PROVINCIA]),PROVINCIA[ID_PRO
VINCIA]>1)`
- **Promedio mensual de usuarios =**
`([Promedio diario de usuarios]*30)`
- **Promedio diario de usuarios =**
`IFERROR([Usuarios transportados con
SUBE]/DISTINCTCOUNT(VIAJES[DIA_TRANSPORTE]), 0)`

11. Segmentaciones elegidas:

- **Botones:**

Los siguientes botones nos permiten cambiar de página o solapa:



Este botón nos permite borrar los filtros y visualizar la totalidad de los datos:



- **Filtros:** En todas las páginas (excepto en la portada), existirán los filtros de Año, Mes, Provincia, AMBA, y Tipo de transporte

Año
 Todas
☐ 2020
☐ 2021
☐ 2022

Mes
 Todas
☐ enero
☐ febrero
☐ marzo
☐ abril
☐ mayo
☐ junio
☐ julio

Provincia
 Todas
☐ BUENOS AIRES
☐ CATAMARCA
☐ CHACO
☐ CHUBUT
☐ CORDOBA
☐ CORRIENTES
☐ ENTRE RÍOS

AMBA
 Todas
☐ NO
☐ SI

Tipo de transporte
 Todas
☐ COLECTIVO
☐ LANCHAS
☐ SUBTE
☐ TREN

12. Solapas del tablero en Power BI

12.1 Portada:



USUARIOS TRANSPORTADOS CON SUBE

Felicitas Repetto, Fernando Parisi, Ezequiel Muñoz

1. Portada

2. Usuarios

3. Provincias y Lineas

4. Empresas

5. Variación MoM

Actualización datos: 30/04/22

Indica el tema del tablero, los integrantes, botones para cambiar de solapas y última actualización de los datos.

12.2 Usuarios:



Se muestran los usuarios transportados con SUBE en Argentina. Para determinar los mismos, en primer lugar, se utilizarán 5 filtros: año, mes, provincia, AMBA y tipo de transporte. A la derecha de estos, existirá un botón que permitirá borrar los filtros.

Para mostrar la cantidad de pasajeros, utilizaremos 3 tarjetas: Usuarios transportados con SUBE, Promedio mensual de usuarios y Promedio diario de usuarios.

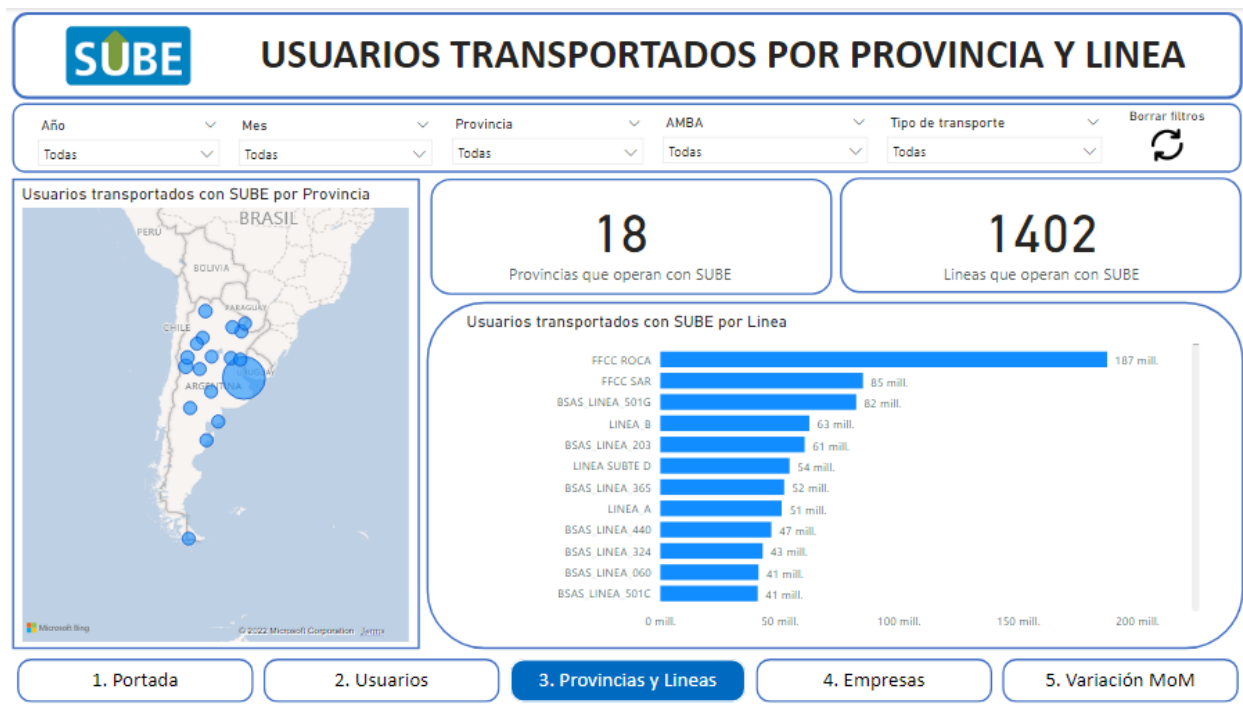
El gráfico de torta indicará la distribución de los pasajeros en función del tipo de transporte: colectivo, tren, subte o lancha. Este último, es ínfimo a comparación de los otros tres.

Finalmente, el gráfico de línea mostrará la evolución temporal (mes a mes) de los pasajeros en el período seleccionado.

Los valores de los indicadores y gráficos dependerán de los filtros aplicados. En el caso de utilizar todo el periodo de tiempo, se considerará de enero de 2020 a abril de 2022.

Debajo del tablero, tendremos los botones necesarios para poder navegar por las diferentes solapas.

12.3 Provincias y Líneas:



Se muestran los usuarios transportados con SUBE por provincia y líneas.

Un indicador mostrará la cantidad de provincias que operan con SUBE y el otro la cantidad de líneas.

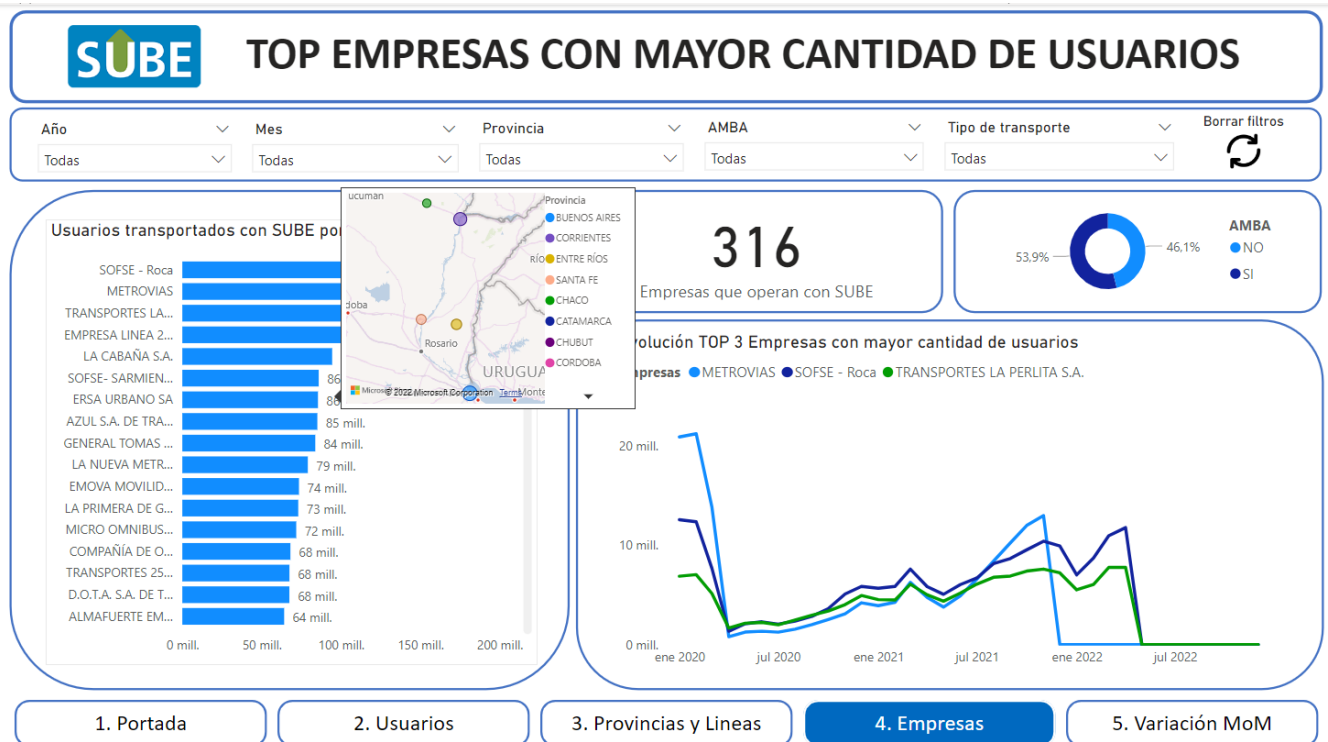
Utilizamos un mapa del territorio analizado, para mostrar la cantidad de usuarios por provincia mediante burbujas, que, al posicionarse sobre ellas, se mostrará dicho valor, y así hacerlo visualmente más dinámico.

Además, incluimos un gráfico de barras para visualizar la cantidad de usuarios que opera cada línea ordenados de mayor a menor. Esto nos da la pauta de las líneas con mayor caudal de usuarios transportados.

Los valores de estos gráficos e indicadores dependerán de los siguientes filtros aplicados: año, mes, provincia, AMBA y tipo de transporte. A la derecha de estos, existirá un botón que permitirá borrar los filtros. En el caso de utilizar todo el periodo de tiempo, se considerará de enero de 2020 a abril de 2022.

Debajo del tablero, tendremos los botones necesarios para poder navegar por las diferentes solapas.

12.4 Empresas:



Se muestran los usuarios transportados con SUBE por empresa.

El indicador mostrará la cantidad de empresas que operan con SUBE. A su derecha, se creó un gráfico de anillo, indicando que porcentaje de empresas pertenecen al AMBA y cuáles no.

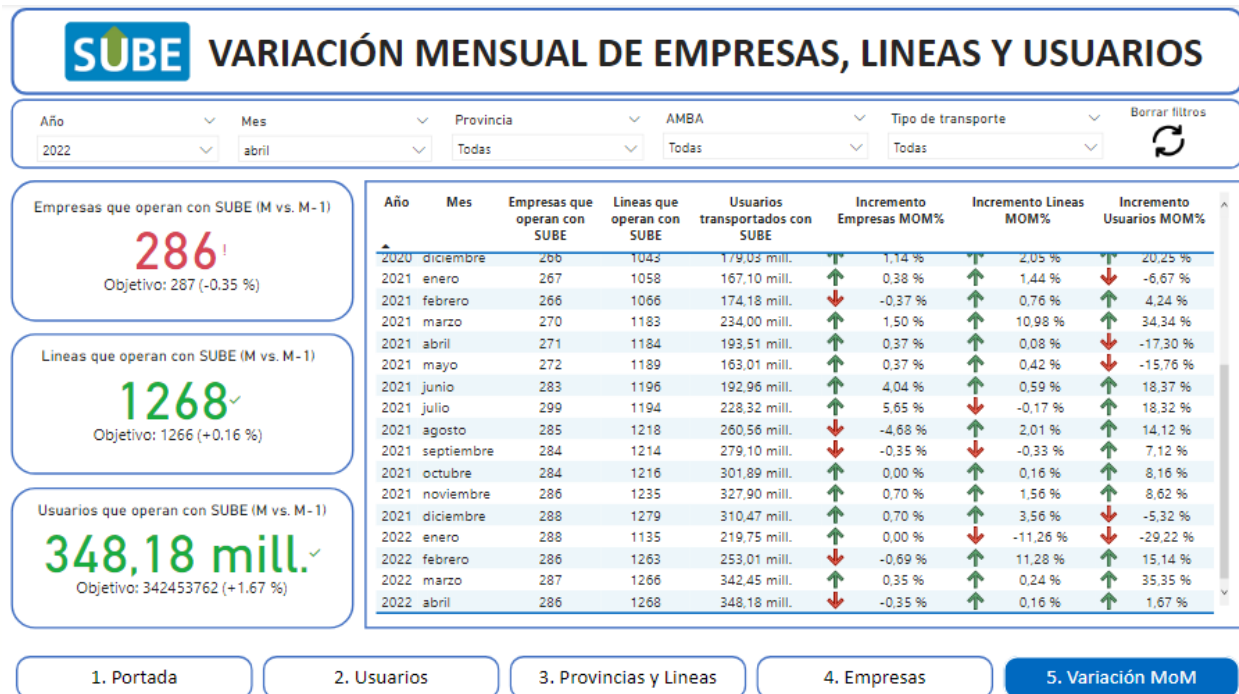
Utilizamos un gráfico de barras para visualizar la cantidad de usuarios que opera cada empresa ordenados de mayor a menor. Esto nos da la pauta de las empresas con mayor caudal de usuarios transportados. Además, este gráfico, tendrá un Tooltip que servirá para mostrar un mapa con la ubicación en la que opera cada empresa.

Además, incluimos un gráfico de líneas que nos muestra la evolución temporal de la cantidad de usuarios transportados por el TOP 3 de empresas más grandes.

Los valores de estos gráficos y del indicador dependerán de los siguientes filtros aplicados: año, mes, provincia, AMBA y tipo de transporte. A la derecha de estos, existirá un botón que permitirá borrar los filtros. En el caso de utilizar todo el periodo de tiempo, se considerará de enero de 2020 a abril de 2022.

Debajo del tablero, tendremos los botones necesarios para poder navegar por las diferentes solapas.

12.5 Variación MoM:



Se muestra la variación de empresas, líneas y usuarios mes a mes.

Incluimos tres KPIs que nos muestran la siguiente información:

- El primero nos muestra las empresas que operan con SUBE comparando el mes corriente con el anterior, y su variación en términos porcentuales.
- El segundo nos muestra las líneas que operan con SUBE comparando el mes corriente con el anterior, y su variación en términos porcentuales.
- El tercero nos muestra la cantidad de usuarios que viajan con SUBE comparando el mes corriente con el anterior, y su variación en términos porcentuales.

En los tres casos el objetivo a cumplir es el mes anterior, y en base a eso, se obtendrá una variación porcentual negativa (roja) o positiva (verde) según corresponda.

A la derecha de los KPIs, insertamos una tabla con los mismos datos con mayor detalle, tuvimos en cuenta un formato condicional que nos indica, una flecha hacia arriba si la variación porcentual es positiva, y una flecha hacia abajo en caso de que sea negativa. Incluimos en la tabla, el periodo analizado, considerando saltos mensuales.

Los valores de estos KPIs y de la tabla dependerán de los siguientes filtros aplicados: año, mes, provincia, AMBA y tipo de transporte. En el caso del período filtrado, habrá que tener en cuenta que se deberá seleccionar un solo año y un solo mes para obtener el valor correcto de los 3 KPIs. A la derecha de estos, existirá un botón que permitirá borrar los filtros. Debajo del tablero, tendremos los botones necesarios para poder navegar por las diferentes solapas.

13. Conclusión

A partir del tablero realizado, se pudo obtener información útil de los usuarios, líneas, y empresas que operan con SUBE, durante los últimos 28 meses, en las distintas provincias de Argentina. Este proyecto fue dirigido a los directivos del Ministerio de Transporte, quienes podrán tomar decisiones en base al mismo, para optimizar el transporte público operado por SUBE, buscando el crecimiento y la eficiencia de las empresas y líneas que operan con este sistema.

Algunas de las conclusiones que se pueden extraer del tablero son las siguientes:

- La gran mayoría de los usuarios transportados con SUBE utiliza como medio de transporte el colectivo.
- El número de pasajeros transportados mediante lancha es insignificante a comparación del resto de los medios de transporte. Además, solo opera en la localidad de Tigre, Buenos Aires.
- Los subtes solo operan en AMBA
- Durante los meses de pandemia del 2020, se puede observar que hubo una caída brusca en la demanda debido a la cuarentena.
- En los meses de enero, febrero y julio, se ve un descenso en el número de pasajeros transportados, debido al receso escolar y vacaciones.
- La mayor cantidad de usuarios se encuentra en la provincia de Buenos Aires.
- Hay 18 provincias que operan con SUBE, faltando 5 que se adhieran al sistema.
- La línea de transporte público más concurrida es el tren Roca, por amplia diferencia, ubicado en Buenos Aires.
- Las empresas que transportan mayor número de pasajeros son: SOFSE (de trenes), Metrovias (de subte) y Transportes La Perlita (de colectivos).
- En abril de 2020, se puede observar una variación negativa en realización al mes anterior, respecto a los usuarios transportados, líneas y empresas que operan con SUBE, debido al inicio de la cuarentena.
- En los meses de marzo, la variación positiva de pasajeros transportados con respecto al mes anterior es muy evidente, ya que es el mes en el cual comienza la actividad escolar y termina el periodo de vacaciones.

La cantidad de conclusiones que se pueden obtener son numerosas debido a dinamismo del tablero, de acuerdo con los filtros aplicados.