

PROYECTO FINAL

# Airbnb en Buenos Aires Y Capital Federal.

Integrantes: Florencia Sun, Federico Mussi.



# Proyecto final

Airbnb en Buenos Aires y Capital Federal



## Índice

<b>Temática a analizar</b>	3
<b>Objetivo del proyecto</b>	3
<b>Alcance</b>	3
<b>Usuario final y nivel de aplicación del análisis</b>	3
<b>Dataset: Definición de tablas</b>	4
<i>Tabla 1: Publicaciones</i>	4
<i>Tabla 2: Tipo_Habitación</i>	4
<i>Tabla 3: Barrio</i>	4
<i>Tabla 4: Ubicación</i>	5
<i>Tabla 5: Dueño</i>	5
<i>Tabla 6: Calificación</i>	5
<b>Diagrama entidad-relación</b>	6
<b>Transformación de datos</b>	6
<b>Columnas, grupos y medidas calculadas</b>	7
En la tabla Barrio:	7
En la tabla dueño:	8
En la tabla publicaciones:	9
<b>Tablas Internas agregadas:</b>	12
<b>Descripción de solapas</b>	13
Proyecto final	13
Inicio	14
Precios	16
Barrios	17
Propietarios	18

## Temática a analizar

La temática a analizar son los alquileres temporarios en la provincia de Buenos Aires y Capital Federal, publicados a través de la plataforma Airbnb. La base de datos cuenta con casi 20.000 registros de publicaciones que incluyen datos de geolocalización, barrio, precio, tipo de habitación, disponibilidad y calificación de las mismas. Para completar el análisis, se cuenta además con el nombre de los dueños y la cantidad de publicaciones que cada uno ha realizado en la plataforma.

## Objetivo del proyecto

El objetivo del presente proyecto es lograr una mayor comprensión sobre la difusión y uso de Airbnb en Capital Federal y Buenos Aires, a través del análisis de la base de datos suministrada por la empresa de forma pública, tratando de entender quienes la usan tanto desde el punto de vista de huéspedes como de los dueños, cuales son los tipos de propiedades, si la finalidad es alquilar para uso turístico de corta duración o estancias temporarias más prolongadas. Otro objetivo es indagar sobre la matriz de precios de las publicaciones de Airbnb y entender cómo se posicionan dichos precios en relación al resto del mercado inmobiliario de CABA, segmentado por barrio y tipo de publicación.

## Alcance

El proyecto toma como muestra una de las localidades más importantes e integradas del país siendo Capital Federal y sus alrededores (Provincia de Buenos Aires). Dicho Dataset se eligió sobre otras localidades y países del mundo debido al número de registros, su diversidad y la facilidad de visualización de sus datos de geolocalización en un mapa.



## Usuario final y nivel de aplicación del análisis

El presente trabajo tiene un propósito de difusión general, donde cualquier persona con motivos de investigación, o con interés por usar o conocer la plataforma puede consultar para profundizar su conocimiento. Una eventual aplicación es la posible comparación en las mismas dimensiones y términos, con respecto a otras ciudades.



## Dataset: Definición de tablas

*Tabla 1: Publicaciones*

CAMPO	TIPO DE CLAVE	TIPO DE CAMPO
Id_Publicación	PK- index	INT
Nombre_publicación		VARCHAR(300)
Id_tipo	FK	INT
Precio		DECIMALES
Min_noches		INT
Id_barrio	FK	INT
Id_ubicación	FK	INT
Id_dueño	FK	INT
Id_calificacion	FK	INT
Disponibilidad		INT

*Tabla 2: Tipo\_Habitación*

CAMPO	TIPO DE CLAVE	TIPO DE CAMPO
Id_tipo	PK- index	INT
Tipo_habitación		VARCHAR(100)

*Tabla 3: Barrio*

CAMPO	TIPO DE CLAVE	TIPO DE CAMPO
Id_barrio	PK- index	INT
Nombre_barrio		VARCHAR(100)

*Tabla 4: Ubicación*

CAMPO	TIPO DE CLAVE	TIPO DE CAMPO
Id_ubicación	PK- index	INT
Latitud		DECIMAL
Longitud		DECIMAL

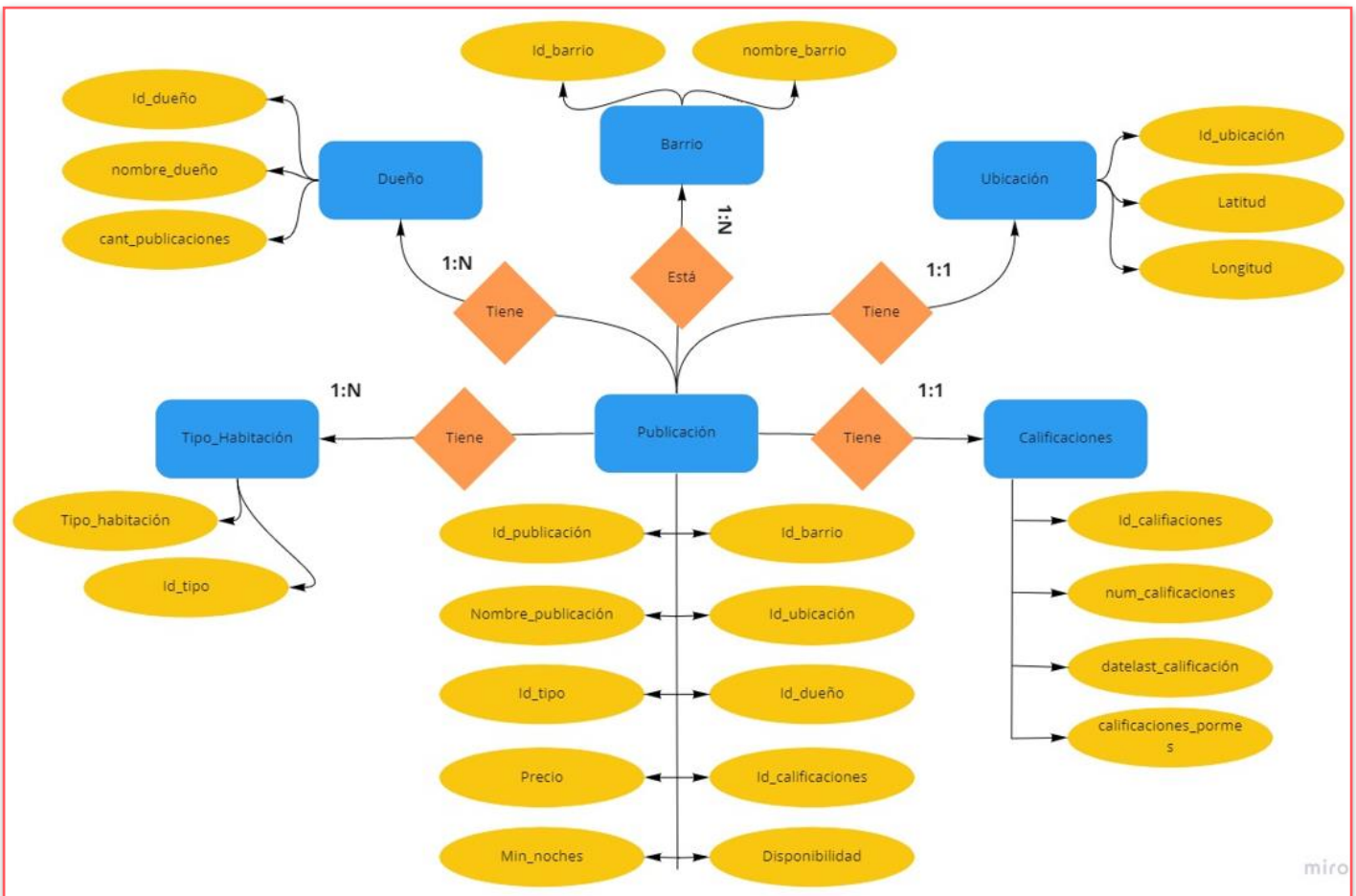
*Tabla 5: Dueño*

CAMPO	TIPO DE CLAVE	TIPO DE CAMPO
Id_dueño	PK- index	INT
Nombre_dueño		VARCHAR(100)
Cantidad_publicaciones		INT

*Tabla 6: Calificación*

CAMPO	TIPO DE CLAVE	TIPO DE CAMPO
Id_calificacion	PK- index	INT
Numero_comentarios		INT
Fecha_comentario		DATETIME
Comentarios_pormes		INT

## Diagrama entidad-relación



## Transformación de datos

- 1- Se cambió la propiedad resumen del campo **Id\_publicacion** en la tabla publicaciones de Suma, a No resumir, transformándolo en una dimensión.
- 2- Se redefinió el campo **Cant\_publicaciones** de la tabla dueño como una medida, al poner Suma en la propiedad Resumen.
- 3- Se reemplazó el valor **Ã±** por ñ en la tabla Barrios.
- 4- Se eliminó un valor defectuoso en la tabla publicaciones.



El proyecto no necesita tablas externas

Si se agregó tablas internas a los fines del Dashboard sin embargo serán explicadas más adelante de este documento en un título especial dedicado a ello.

## Columnas, grupos y medidas calculadas

Realizamos una serie de Columnas y medidas calculadas descriptas a continuación:

### En la tabla Barrio:

- **Precio\_promedio\_Barrio (columna):** Se utilizó para determinar el precio promedio por barrio que sirvió para la realización de diversos gráficos tanto en la solapa de precios como en la solapa de barrios.

```
#Precio_Promedio_Barrio =  
CALCULATE([#Precio_promedio_total],Barrio[Id_barrio])
```

- **FlagPrecio (columna):** Se utilizó para determinar en la solapa barrios, que barrios estaban por encima del precio promedio global y que barrios estaban por debajo. Utilizado luego en la medida “tarjeta”

```
#FlagPrecio = IF(Publicaciones[#Precio_promedio_total]>2812,"Por encima","Por debajo")
```

- **#NumeroDeBarrios (medida):** Se utilizó para contabilizar la cantidad de barrios totales que había. Sirvió para luego calcular promedios de precio por barrio en las solapas nombradas anteriormente.

```
#NumeroDeBarrios = COUNT(Barrio[Id_barrio])
```

- **PrecioConAjuste (columna):** Se realizó a fin de poder calcular, en la solapa de Precios, como se vería afectado el precio de los barrios conforme a la inflación esperada.

```
#PrecioConAjuste = SUM(Publicaciones[Precio])*(1+InflacionEsperada[Valor Ajuste])
```

- **promedio\_barrio (medida):** Originalmente para realizar el cálculo de promedios usamos esta medida, sin embargo, al avanzar del proyecto, entendimos que era necesario convertirlo en una columna (explicada en el inciso 1 de este título). Su utilidad es la misma antes descripta.

```
#promedio_barrio = CALCULATE([#Precio_promedio_total],Barrio[Id_barrio])
```





- **Tarjeta (medida):** Medida realizada con el fin de lograr que la tarjeta que indica si el precio está por debajo o por encima del promedio marque cuando este dentro del valor promedio global (\$5020) diga "En el promedio" caso contrario, dependiendo del valor, indique si está por encima o por debajo.

```
Tarjeta = IF(HASONEVALUE(Barrio[#FlagPrecio]), IF([#Precio_promedio_total]>5020, "Por Encima", "Por Debajo"), "En el Promedio")
```

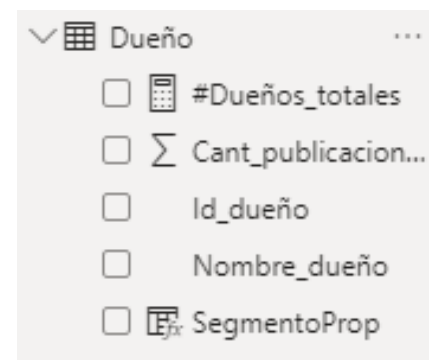
- **Tarjeta barrio (medida):** Medida realizada con el fin de lograr que la tarjeta que indica el barrio diga "Seleccione barrio" en caso de no haber un barrio seleccionado, caso contrario, que indique el nombre del barrio seleccionado.

```
tarjeta barrio = IF(HASONEVALUE(Barrio[Nombre_barrio]), FILTERS(Barrio[Nombre_barrio]), "Seleccione Barrio")
```

#### En la tabla dueño:

- **#Dueños\_totales (medida):** Esta medida fue utilizada para contar la cantidad de dueños totales. Sirvió posteriormente para el KPI utilizado en la solapa de inicio a fin de mostrar la cantidad de personas que utilizan Airbnb para publicar en alquiler sus inmuebles.

```
#Dueños_totales = DISTINCTCOUNT('Dueño '[Id_dueño])
```



- **SegmentoProp (Columna):** Esta columna fue realizada con el fin de poder calcular cuantas propiedades publicadas tenía cada dueño. Fue utilizada en los gráficos de la solapa Propietarios. Nos ayudó a determinar si la mayoría de los propietarios se dedicaban al alquiler de inmuebles o bien contaban con un inmueble fuera del de uso personal/familiar y lo utilizaban como un activo, es decir, generaban dinero con el alquiler.

```
SegmentoProp = IF('Dueño '[Cant_publicaciones] <= 1, "Una Propiedad", "Multiples Propiedades")
```



### En la tabla publicaciones:

- **#Casa (Medida):** fue realizada a fin de calcular la cantidad de publicaciones en donde el tipo de habitación era una casa. La medida fue de utilidad para los gráficos de la solapa barrios en donde se calculó la cantidad de casas totales publicadas en Airbnb así como también se calculó la cantidad de casas publicadas por barrio.

#Casa =

VAR A =

CALCULATE(COUNT(Publicaciones[Id\_tipo]),Publicaciones[Id\_tipo]=1)

VAR B = [#Habitaciones\_totales]

return A/B

- **#H. Compartida (Medida):** Nos remitimos a lo dicho en la medida #Casa del inciso uno de este título solo que en este caso se calcula por tipo de habitación compartida

#H.Compartida =

VAR A = CALCULATE(COUNT(Publicaciones[Id\_tipo]),Publicaciones[Id\_tipo]=3)

VAR B = [#Habitaciones\_totales]

return A/B

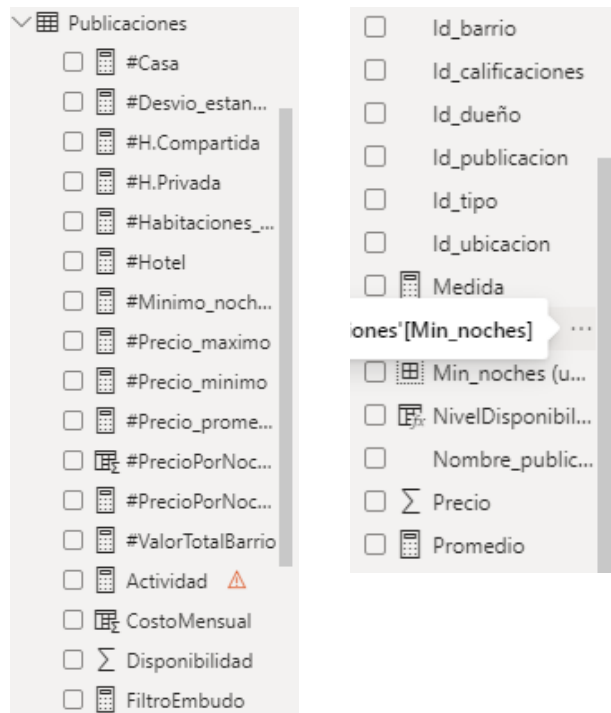
- **#H. Privada (Medida):** Nos remitimos a lo dicho en la medida #Casa del inciso uno de este título solo que en este caso se calcula por tipo de habitación privada.

#H.Privada =

VAR A = CALCULATE(COUNT(Publicaciones[Id\_tipo]),Publicaciones[Id\_tipo]=2)

VAR B = [#Habitaciones\_totales]

return A/B



- **#Hotel(Medida):** Nos remitimos a lo dicho en la medida #Casa del inciso uno de este título solo que en este caso se calcula por tipo de habitación hotel.

#Hotel =

```
VAR A = CALCULATE(COUNT(Publicaciones[Id_tipo]),Publicaciones[Id_tipo]=4)
```

```
VAR B = [#Habitaciones_totales]
```

```
return A/B
```

- **#Habitaciones\_totales (Medida):** medida realizada a fin de contabilizar la cantidad de habitaciones totales. Fue utilizada en el KPI de Total de Habitaciones de la solapa inicio y

```
#Habitaciones_totales = COUNT(Publicaciones[Id_tipo])
```

- **#Minimo\_noche\_promedio (Medida):** medida realizada para calcular el promedio de noches mínimas utilizada en el KPI de la solapa Inicio donde se muestra el promedio de noches mínimas por el total de publicaciones.

```
#Minimo_noche_promedio = AVERAGE(Publicaciones[Min_noches])
```

- **#Precio\_maximo (Medida):** medida realizada para calcular el precio máximo de las publicaciones tanto a nivel global como individual. Fue utilizada en los gráficos de la solapa precio, siendo el KPI de Precio Máximo y la tabla donde estimaba el precio máximo por tipo de habitación

```
#Precio_maximo = MAX(Publicaciones[Precio])
```

- **#Precio\_minimo (Medida):** Nos remitimos a lo dicho anteriormente en el la medida de #Precio\_maximo solo que en este caso se calcula el precio mínimo.

```
#Precio_minimo = MIN(Publicaciones[Precio])
```

- **#Precio\_promedio\_total (Medida):** Medida utilizada para calcular el precio promedio global, utilizada tanto en el KPI sobre el precio promedio en la solapa de inicio, como en el KPI sobre la misma temática en la solapa precios y barrios. Fue utilizada también en los gráficos sobre precio promedio por tipo de habitación.

```
#Precio_promedio_total = AVERAGE(Publicaciones[Precio])
```

- **#PrecioPorNoche (Columna):** Columna realizada a fin de calcular el precio por noche dependiendo del barrio. Fue utilizado en los gráficos de la solapa precios

#PrecioPorNoche = Publicaciones[Precio]/Publicaciones[Min\_noches]

- **#PrecioPorNocheAjustado (Medida):** Medida que fue realizada a fin de calcular el precio promedio por noche conforme a la inflación que pudiera afectar a la economía. Utilizada en los gráficos de la solapa Precios.

#PrecioPorNocheAjustado = AVERAGE(Publicaciones[#PrecioPorNoche])\*([Valor Ajuste])

- **#ValorTotalBarrio (Medida):** Medida que fue realizada a fin de calcular la suma total del precio por cantidad de publicaciones por barrio. Fue utilizada en el KPI sobre sumatoria del precio de la solapa Barrios.

```
#ValorTotalBarrio =  
VAR A = SUM(Publicaciones[Precio])  
Return A
```

- **#CostoMensual (Columna):** Columna que fue realizada a fin de calcular el costo mensual por barrio conforme al precio noche. Fue utilizada en el gráfico sobre Costo mensual por barrio de la solapa Propietarios en donde

CostoMensual = Publicaciones[#PrecioPorNoche]\*30

- **#FiltroEmbudo (Medida):** Medida calculada a fin de contabilizar la cantidad de veces que una publicación determina una cantidad de noches mínimas. La medida fue utilizada en el gráfico de la solapa Propietarios que muestra en una tabla tal información.

```
FiltroEmbudo =  
CALCULATE(SUM(Publicaciones[Min_noches]),FILTER(Publicaciones,Publicaciones[Min_noche  
s]>15))
```

- **Min\_noches – ubicación- (Grupo):** Recuento realizado a fin de calcular el numero de noches mínimas por ubicación. Fue utilizado conto con el Filtro Embudo para realizar el gráfico al cual nos remitimos explicado en la medida #FiltroEmbudo.

- **NivelDisponibilidad (Columna):** Columna realizada a los fines de calcular la disponibilidad de los inmuebles siempre que esta fuera mayor a 3 meses. Se utilizó en el gráfico sobre disponibilidad de las propiedades de la solapa Propietarios a fin de demostrar si los alquileres eran de fines vacacionales o bien con fines de vivienda.

`NivelDisponibilidad = IF(Publicaciones[Disponibilidad] >= 90, "Alta Disponibilidad", "Baja Disponibilidad")`

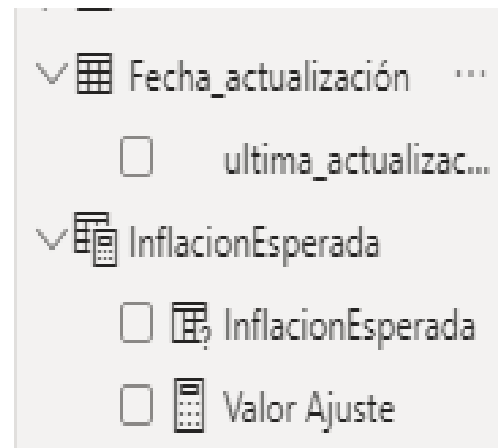
- **#Promedio (Medida):** Medida estática utilizada en el KPI sobre precio promedio global, así como también en el gráfico sobre sumatoria de precio promedio por barrios de la misma solapa. Su funcionalidad es marcar el valor exacto (precio promedio global) a fin de tomarlo como base para calcular si el valor promedio de un barrio está por encima o por debajo del promedio global.

`Promedio = 5020`

### Tablas Internas agregadas:

Al no contar con fechas, no se creó tabla calendario, tampoco se adicionaron tablas externas al modelo, sin embargo se agregaron dos tablas internas

La primera de ella se denomina, **Fecha\_actualización** y fue creada a fin de poder indicar cuando fue la última actualización del Dashboard. El mismo se encuentra expuesto en una tarjeta (contenida en todas las solapas) debajo del título que indican fecha y hora.



La segunda tabla creada denominada InflaciónEsperada contiene un parámetro y una medida.

- Parámetro: `InflacionEsperada = GENERATESERIES(0, 1, 0.1)`
- Medida: `Valor Ajuste = SELECTEDVALUE('InflacionEsperada'[InflacionEsperada])`

Ambos elementos trabajan en conjunto para crear un selector de inflación que pronostique el valor futuro de los alquileres basados en un porcentaje de inflación y fueron utilizados en la solapa precios en los gráficos de la parte inferior.

## Descripción de solapas

### Proyecto final

Es la pantalla de bienvenida que consiste en el título del proyecto, los nombres de los integrantes y un botón que permite acceder a la solapa Inicio.

Cuenta tanto con el logro de la empresa de la cual obtuvimos el Dataset de la cual realizamos el análisis.

Como se verá intentamos mantener los colores acordes al logotipo combinando tonalidades diferentes así como colores contrarios para mantener una cierta armonía en el Dashboard, también mantuvimos un tipo de tipografía denominada "DIN" a lo largo de todo el proyecto.



## ANÁLISIS DE AIRBNB EN LA CIUDAD AUTÓNOMA DE BUENOS AIRES Y PROVINCIA DE BUENOS AIRES.

COMENZAR

FEDERICO MUSSI - FLORENCIA SUN

## Inicio

En la parte superior de la solapa se puede ver que se encuentra el **logo**, el **título del análisis**, así como también una tarjeta, debajo del título, en donde se muestra la **fecha de actualización** con hora incluida.

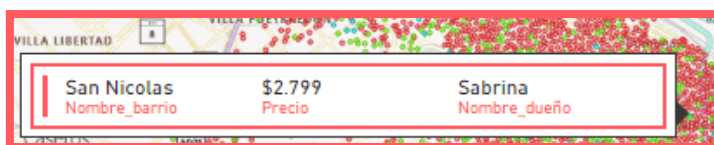
En la parte superior derecha se encuentran los **botones que direccionan a las diferentes solapas** en cuyo interior se encuentran los gráficos relacionados con el Precio, Barrio o Propietarios. Esto se repetirá en todas las demás solapas con la salvedad de que, seleccionada una solapa, siempre aparecerá un botón de inicio que permitirá volver a esta solapa inicial.

A fin de no ser repetitivos, recordamos que esto se repetirá en todas las solapas y por ello no volveremos a incorporarlo en la explicación de las siguientes solapas.

En la parte Izquierda se encontrará con **cuatro Tarjetas** que presentan datos generales siendo:

1. **Total de Dueños**
2. **Precio Promedio**
3. **Total de Habitaciones**
4. **Promedio de Mínimo por noche**

En la posición derecha de la pestaña se encuentra un **mapa** con los datos de **geolocalización** de todos los alquileres presentes en la base de datos. Al posicionar el cursor sobre alguno de los puntos de referencia, un **tooltip** aparece designando datos de la publicación. Es importante que el tooltip fue ocultado.



Como se puede ver arriba **en la leyenda, cada color indica un tipo de habitación**:

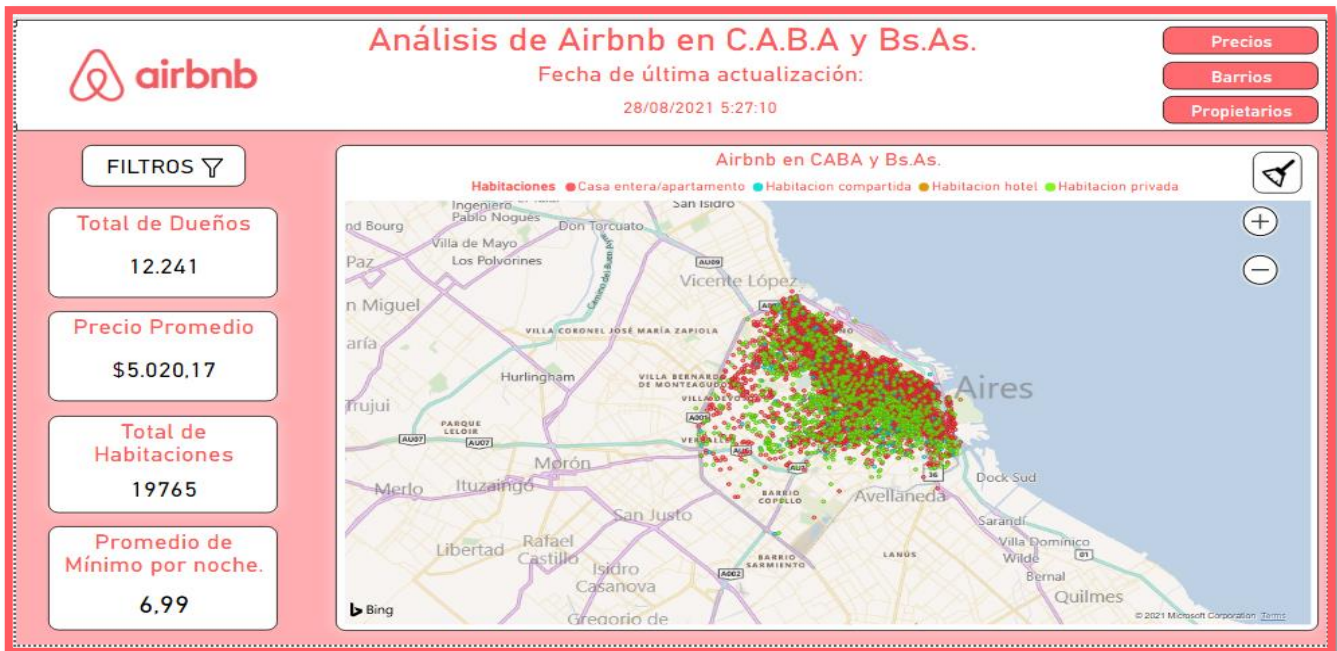
- Rosa: Casas enteras o apartamentos.
- Azul: Habitaciones Compartidas.
- Naranja: Habitaciones de Hotel.
- Verde: Habitaciones Privadas.

El botón **filtros** hace aparecer un panel que permite filtrar los gráficos por **tipo de habitación, barrio, precio, noches mínimas y propietario**. Dentro de este mismo cuadro desplegado se encuentra un **botón que permite cerrar o encoger el panel de filtros** a fin de poder ver la totalidad de la solapa. Es importante destacar que **tanto el panel de filtros como el botón de pliegue se repetirán en todas las solapas** cambiando solo los tipos de filtros.

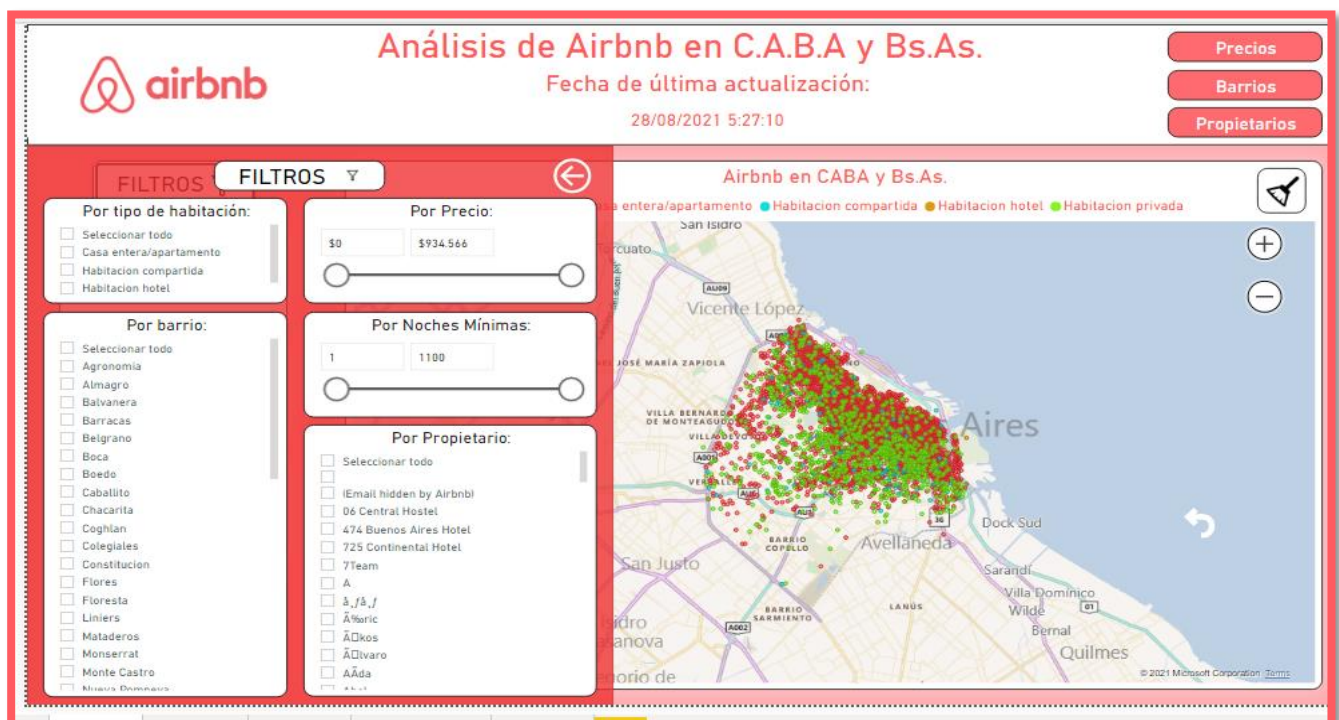
Por último, cuenta con un **botón con icono de escoba** que se encarga de **borrar los filtros** de manera automática sin tener que volver a desplegar el cuadro de filtros. **El mismo se encontrará siempre en el sector derecho de las solapas.**

El objetivo de esta solapa es empezar a interiorizarse con los datos y además presentar la una idea de cómo se concentran las publicaciones en la Ciudad de Buenos Aires y alrededores.

Inicio sin filtros desplegados.



Inicio con filtros desplegados (el cuadro de filtros se repetirá de este modo en todas las solapas con diferencias de los tipos de filtros aplicables)





## Precios

La pestaña precios está diseñada para proveer un insight completo sobre los precios de las publicaciones.

Los primeros gráficos constan de **tres tarjetas**, con la información de precio máximo, mínimo y promedio de cada barrio.

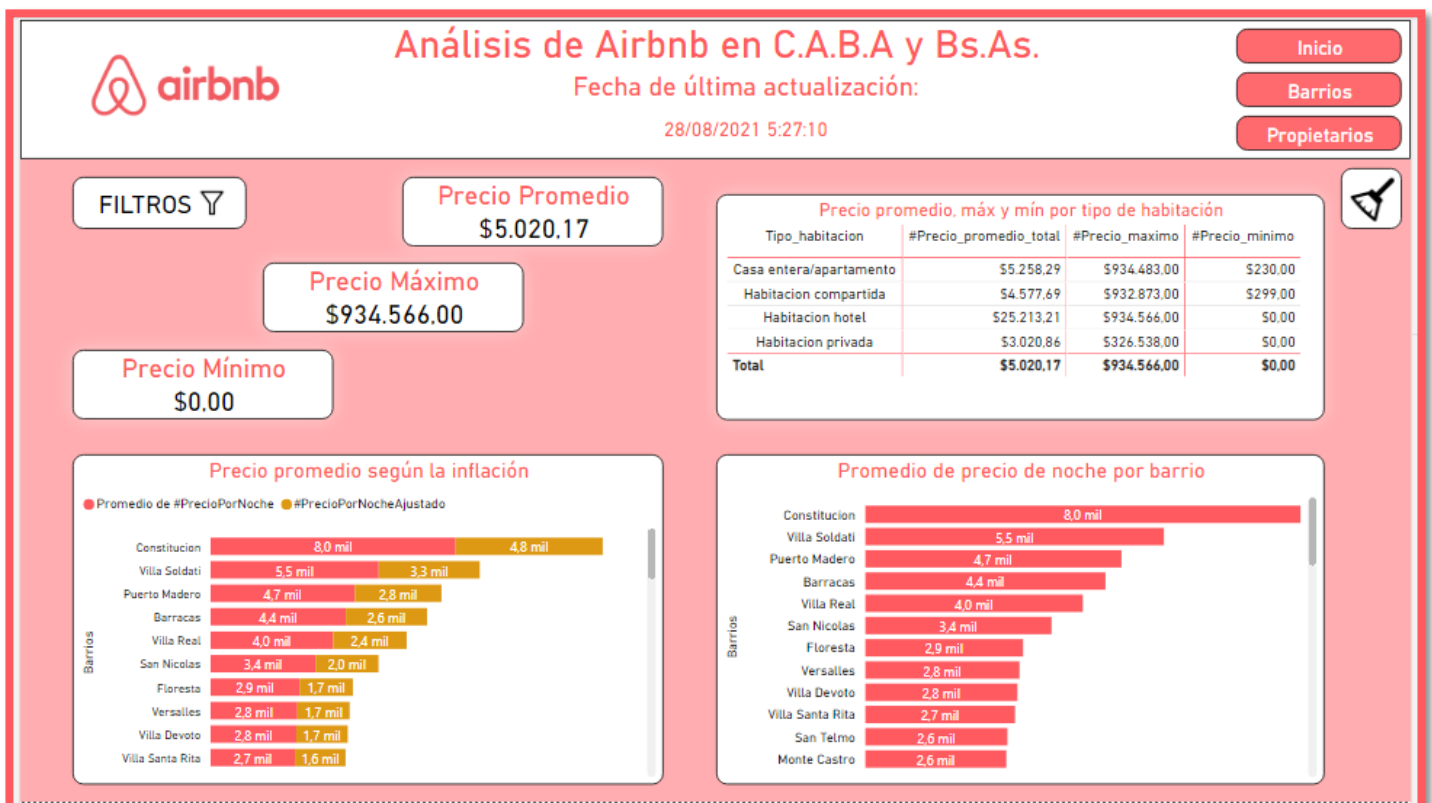
La **tabla** de la derecha desagrega dicha información por tipo de publicación, para realizar una comparación vertical de forma rápida.

La pestaña cuenta con un estimador del precio considerando niveles diferentes de inflación, la información se encuentra resumida en un **gráfico de barras horizontales**. Mediante el uso de un **parámetro** y un **slider** que se encuentra en el menú de filtros, se pueden ajustar diferentes niveles inflacionarios que luego se calcularan y mostraran en el gráfico de barras.

Finalmente, un **gráfico de barras verticales** permite ver y comparar de forma rápida los **precios por noche**.

Así como en las solapas anteriores, cuenta con el **botón de filtro y de borrado de filtros**. Los filtros son por barrio, por tipo de habitación, por precio y por inflación esperada.

### Precios sin filtros desplegados



## Barrios

El primer **KPI** de barrios muestra la sumatoria de todos los alquileres, con el fin de cuantificar el valor total de las propiedades publicadas. La segunda **tarjeta** es un indicador del precio promedio global.

Al medio se encuentra la designación del barrio, segunda tarjeta que indica si el precio promedio de este es superior o inferior al promedio global y una tercera tarjeta que indica el precio promedio de ese barrio en particular.

A la derecha, **un gráfico de anillo**, busca segmentar las publicaciones entre sus cuatro tipos: Casa, Habitación compartida, habitación privada y hotel.

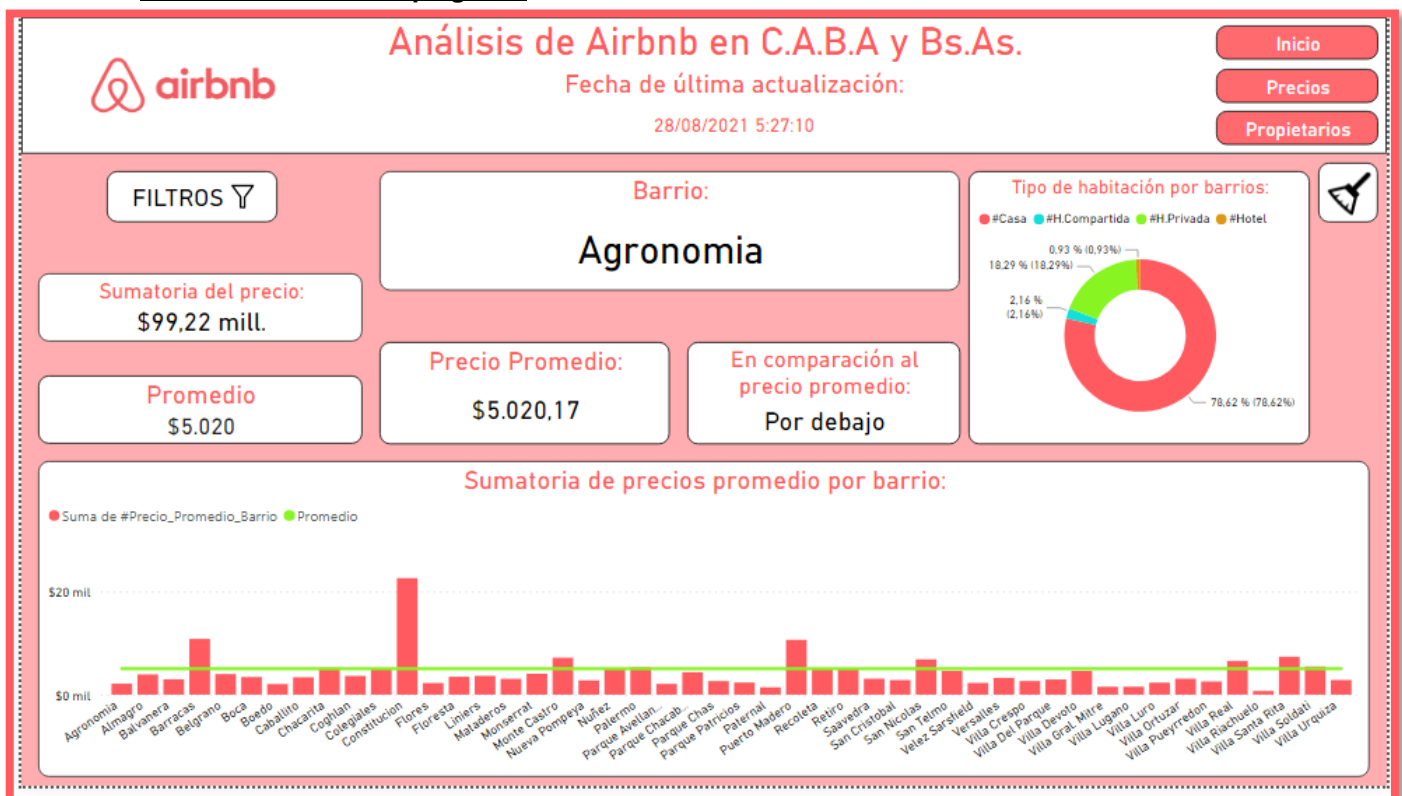
Dentro del gráfico “Sumatoria de precios promedio por barrio” se encuentra el **“tooltip2”** el cual muestra con un KPI que indica si el precio del barrio está por encima o por debajo del promedio global. El tooltip2 fue oculto.



Finalmente, en la parte inferior se encuentra un **gráfico de columnas y líneas**. Este gráfico despliega el precio promedio de cada barrio, y con una línea horizontal el costo promedio total.

Así como en las solapas anteriores, cuenta con el **botón de filtro y de borrado de filtros**. Los filtros son por tipo de habitación, por precio y por barrio.

### Barrios sin filtros desplegados



## Propietarios

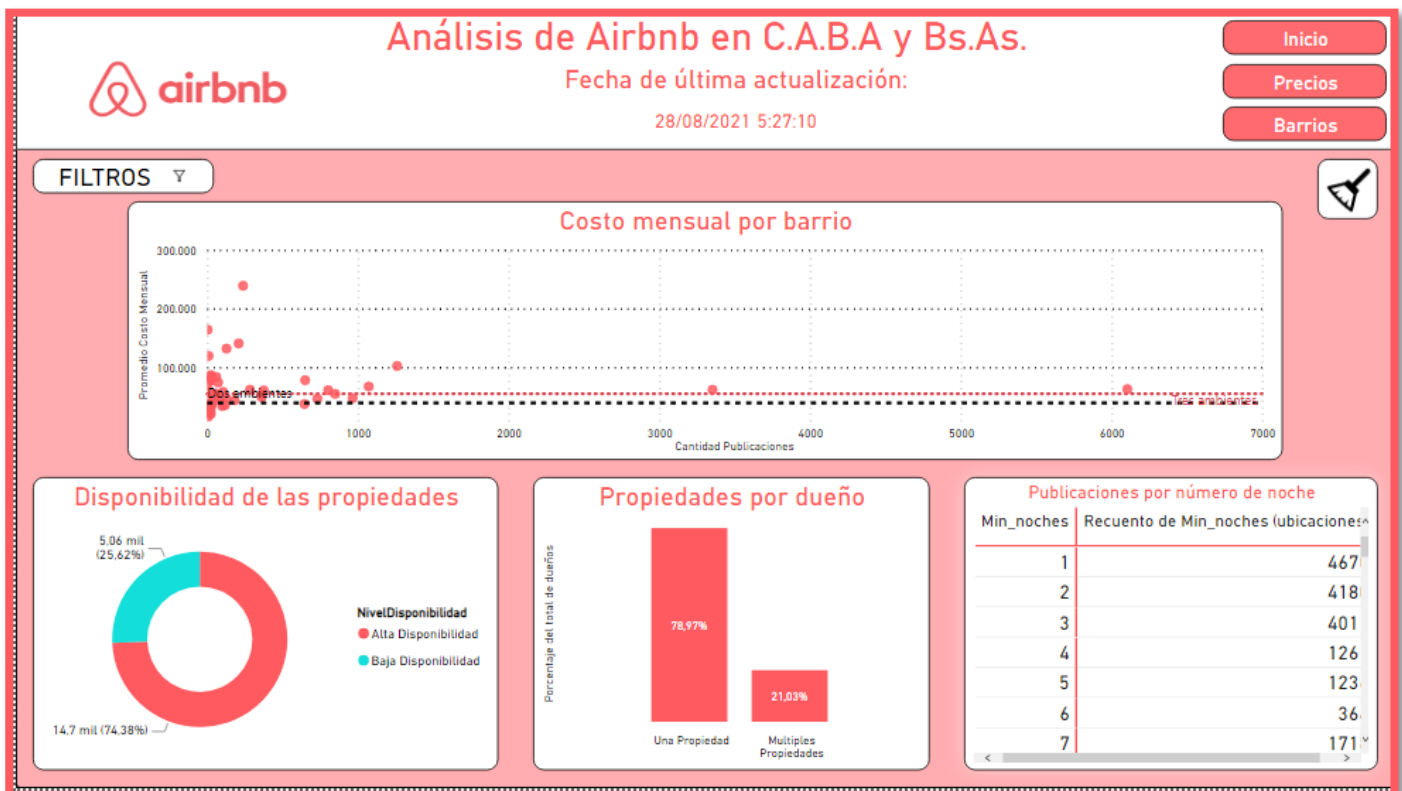
Esta solapa tiene como busca interiorizar en quien es el usuario tipo de Airbnb, indagando en diferentes dimensiones y datos que pueden obtenerse.

La solapa da la primera pauta a través de un **gráfico de dispersión**, que muestra en el eje X la cantidad de publicaciones, y en el eje Y, el costo mensual de las publicaciones. Dos líneas horizontales dan dimensión a los datos, al consignar el costo promedio de alquileres de departamentos de dos y tres ambientes, obtenidos de la página de estadísticas del gobierno de la ciudad.

Un **gráfico de anillos** permite ver el porcentaje de publicaciones con alta y baja disponibilidad. El **grafico de barras** muestra el resultado de una columna calculada que categoriza los dueños por cantidad de publicaciones y finalmente una **tabla** que resume los días mínimos impuestos por cada publicación que más se repiten.

Al igual que las otras pestañas, esta posee un **botón filtro** que permite segmentar más la información mostrada y un **botón de borrado de filtros**. Los filtros son por tipo de habitación y por barrio.

### Propietarios sin filtros desplegados



## Análisis final del trabajo.

Luego de analizar los diversos resultados dados por los gráficos comprendimos la operatividad de Airbnb logrando deducir que Airbnb es utilizado en general por sectores de la sociedad que contando con alguna propiedad fuera de la propiedad que usan como vivienda, en otras palabras, la propiedad “Sobrante” la ponen a alquilar mediante esta aplicación a fin de generar un activo.

Por otra parte, se deduce que **Airbnb no funciona tanto como una aplicación de alquileres con fines turísticos sino más bien, de vivienda** y es que, del análisis de los gráficos de cantidad de propiedades por barrios y disponibilidad mayor a 100 días, se infiere que, en realidad, la mayoría de las propiedades se brindan para un alquiler mayor de 3 meses, así como también, que en su mayoría son casas o habitaciones privadas, habiendo muy pocas habitaciones de hoteles o habitaciones compartidas. Entiendo que, a un nivel práctico, el marketing y acciones de Airbnb debería estar dirigido entonces a sectores de clase media y a personas que estén buscando una vivienda.

Como dato adicional, la cantidad de alquileres que exige un mínimo de noches de 1 implica que Airbnb se especializa en alquileres temporarios.

Por último, hay que remarcar que, dado que Airbnb es utilizado con fines de vivienda, los precios no superan en general al promedio de precios de lo que vale un 2 y 3 ambiente en Capital Federal y alrededores, afirmando lo dicho anteriormente.