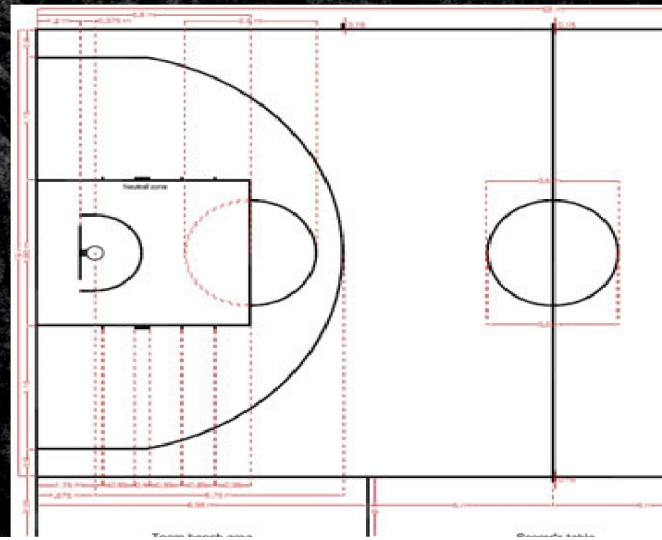


DATA ANALYTICS

# EN BUSCA DE TIRADORES DE 3 PUNTOS



Ramiro Caria



## **INDICE**

***TEMATICA, HIPOTESIS, DATASETS..... 3***

***GLOSARIO, DIAGRAMA ENTIDAD RELACIÓN, LISTADO DE TABLAS.....4***

***TABLAS Y SUS DATOS.....4,5***

***OBJETIVOS DEL PROYECTO, ALCANCE, USUARIO FINAL, SEGMENTACIONES ELEGIDAS, HERRAMIENTAS UTILIZADAS..... 6***

***LISTADO DE TABLAS, DER FINAL, MODELO RELACIONAL.....7***

***LISTADO DE TABLAS Y TRANSFORMACIONES REALIZADAS..... 8,9***

***VISUALIZACIÓN Y EXPLICACIÓN DE CADA SOLAPA..... 10, 11, 12***

***FUTURAS LÍNEAS..... 13***

## PRIMERA ENTREGA DEL PROYECTO FINAL

### TEMATICA:

PARA ESTE PROYECTO, HAREMOS UN ANALISIS ACERCA DE LAS ESTADISTICAS SIMPLES Y AVANZADAS DE LA TEMPORADA 2019-20 DE LA NBA, LA MEJOR LIGA DEL MUNDO DE BALONCESTO.

DENTRO DE LAS ESTADISTICAS, OBSERVAREMOS DATOS COMPLETOS COMO PUEDE SER LOS PUNTOS POR PARTIDO, REBOTES, ASISTENCIAS, ENTRE OTROS, DE TODOS LOS JUGADORES QUE JUGARON EN ESTA TEMPORADA

### HIPOTESIS:

LA CONCLUSIÓN QUE BUSCAMOS OBTENER ES LA SIGUIENTE; NOS PONDREMOS EN EL LUGAR DEL GERENTE PRINCIPAL DE UNA FRANQUICIA DE LA NBA DURANTE EL PERÍODO DE TRASPASOS.

LA SITUACIÓN ES LA SIGUIENTE, NUESTRO ENTRENADOR, LUEGO DE SUFRIR EN ESTE APARTADO DEL JUEGO DURANTE LOS PLAYOFFS, NOS PIDE QUE EN EL MERCADO DE PASES CONSIGAMOS UN SG O UN SF CON UN PORCENTAJE DE 3 PUNTOS SUPERIOR AL 37.5% INTENTANDO MAS DE 4 TRIPLES POR PARTIDO, QUE JUEGUE MAS DE 15 MINUTOS, MENOR A 28 AÑOS, CON AL MENOS 1 STL POR PARTIDO Y QUE HAYA JUGADO MAS DE 45 PARTIDOS EN TEMPORADA REGULAR. ESTAS ESTADISTICAS SERAN EXTRAÍDAS DE LOS MESE OCTUBRE, NOVIEMBRE Y DICIEMBRE DE 2019

### DATASETS

ESTADISTICAS DE LOS JUGADORES.

[STATS 2020.xlsx](#)

Estadísticas de los partidos:

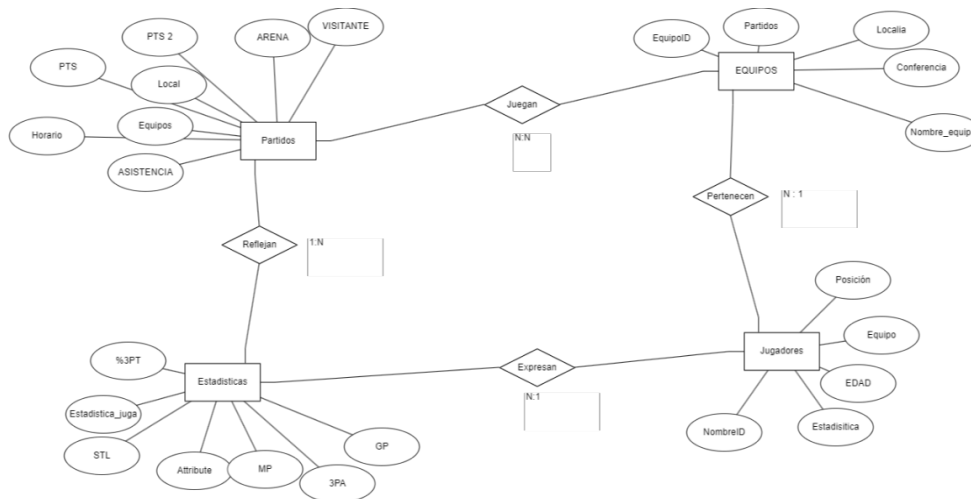
[Partidos data.xlsx](#)

### GLOSARIO

- **POS** : Posición
- **GP** : juegos jugados
- **MIN** : Minutos Por Juego
- **PTS** : Puntos por juego
- **FGM** : Promedio de goles de campo anotados
- **FGA** : Promedio de tiros de campo intentados
- **FG%** : Porcentaje de goles de campo
- **15:00** : Promedio de tiros de campo de 3 puntos anotados
- **3PA** : Promedio de tiros de campo de 3 puntos intentados
- **3P%** : Porcentaje de tiros de campo de 3 puntos
- **FTM** : Promedio de tiros libres anotados
- **FTA** : Promedio de tiros libres intentados
- **FT%** : Porcentaje de tiros libres
- **REB** : Rebotes Por Partido
- **AST** : asistencias por partido
- **STL** : Robos por juego

- **BLK** : Bloques por juego
- **TO** : Pérdidas de balón por partido
- **DD2** : Doble Doble
- **TD3** : Triple Doble
- **PER** : Índice de eficiencia del jugador
- **Franquicia**: equipo de la nba
- **PLAYOFFS**: fase de eliminación

En el siguiente Diagrama de entidad relación podremos ver la relación que hay entre las entidades pertinentes y nuestra fuente principal, la mejor liga de básquet del mundo, la NBA.



## LISTADO DE TABLAS

Las tablas que se mostraran a continuación pertenecen a la primera parte del proyecto, es decir, aquellas hechas antes que el dashboard.

- Partidos
- Estadísticas
- Jugadores
- Equipos

PK: NombreID

### Tabla equipos

Equipos		
Campo	Tipo de clave	Tipo de campo
Partidos	Fk	Int
Nombre_equipo	Fk	Text(n)
Conferencia	-	Text(n)
Localía	Fk	Varchar(n)
EquiposID	FK	Text(n)



**Tabla Jugadores**

Jugadores		
Campo	Tipo de clave	Tipo de campo
NombreID(PK)	PK	Text(n)
Estadísticas (FK)	FK	Varchar(n)
Equipo_ID(FK)	FK	Text(n)
Edad	-	INT
Posición	-	Text(n)

**Tabla Partidos**

Partidos		
Campo	Tipo de clave	Tipo de campo
Equipos	Fk	Text(n)
Local	-	Text(n)
Visitante	-	Text(n)
Horario	-	Varchar(n)
Asistencia	-	Decimal
Arena	-	Text(n)
PTS	-	Int
PTS 2	-	Int

**Tabla Estadísticas**

Estadísticas		
Campo	Tipo de clave	Tipo de campo
Estadística_juga	Fk	Int
MP	-	Varchar(n)
GP	-	Varchar(n)
STL	-	Varchar(n)
%3PT	-	Varchar(n)
3PA	-	Varchar(n)

### **Objetivos del proyecto**

El objetivo final del proyecto será que, a partir de las filtraciones requeridas, lograr llegar a la mejor opción para suplir las necesidades del equipo.

### **Alcance**

Las estadísticas que se tomarán en cuenta serán de partidos entre las fechas 23/10/2019 a 31/12/2019.

A su vez, el dashboard realizado esta hecho con la intención de alcanzar el nivel operativo, es decir, que a través de la especificación de los datos utilizados y de las filtraciones que nos posibilitan distintas visualizaciones del proyecto, poder tomar acción y llegar a las conclusiones indicadas por los altos cargos de la franquicia.

### **Usuario final**

Oficinas operativas del equipo que se encargan de la toma de decisiones en el período de traspasos.

### **Segmentaciones elegidas**

POSICIÓN → SG o SF

3P % → > o = que .375

3PA → > o = 4 por partido

Minutos jugados → > o = 15 min por partido

Edad → < o = 28 años

Robos → > o = 1 robos por partido

Partidos jugados → > o = 45 partidos en temporada regular.

En el dashboard, se ven y se puede filtrar a través de las siguientes visualizaciones



### **Herramientas utilizadas**

Para la realización del trabajo se utilizaron las siguientes herramientas:

**Excel:** Observación y corrección de distintos datases utilizados

**ERD PLUS:** Para la realización del diagrama entidad relación.

**Power BI:** Para la realización del dashboard

### **Listado de tablas**

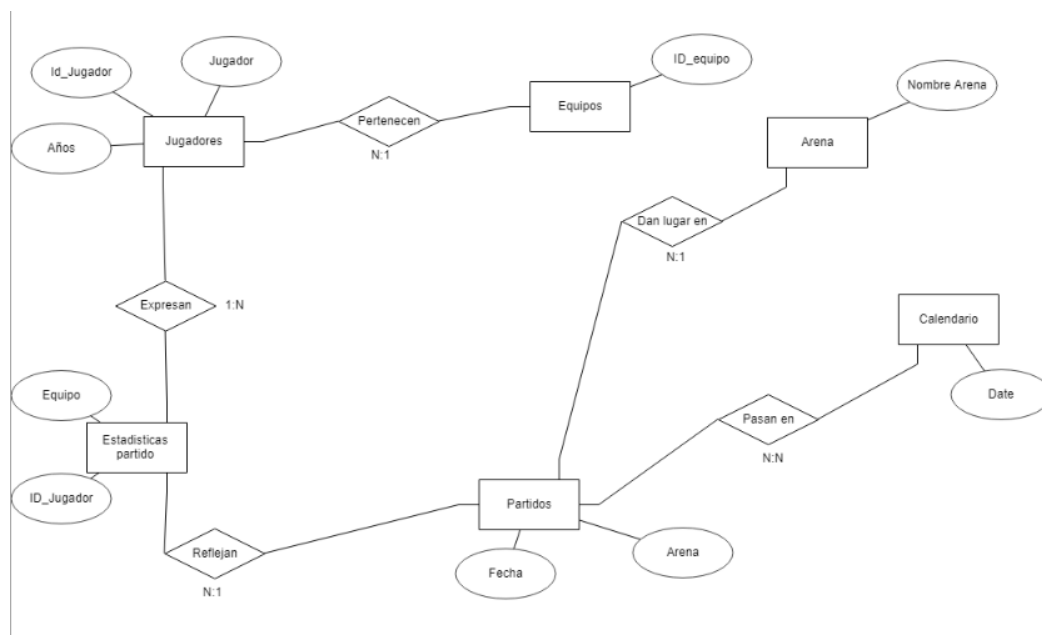
Ya mostramos las tablas que estaban diseñadas al principio del proyecto, pero estas no coinciden con las de luego el dataset utilizado.

Dataset utilizado: [STATS 2020.xlsx](#)

Dentro de la tabla NBA STATS POR PARTIDO 2020, la primera transformación grande realizada fue que los datos venían en formato CSV, por lo que tuve que llevar a cabo la adaptación necesaria.

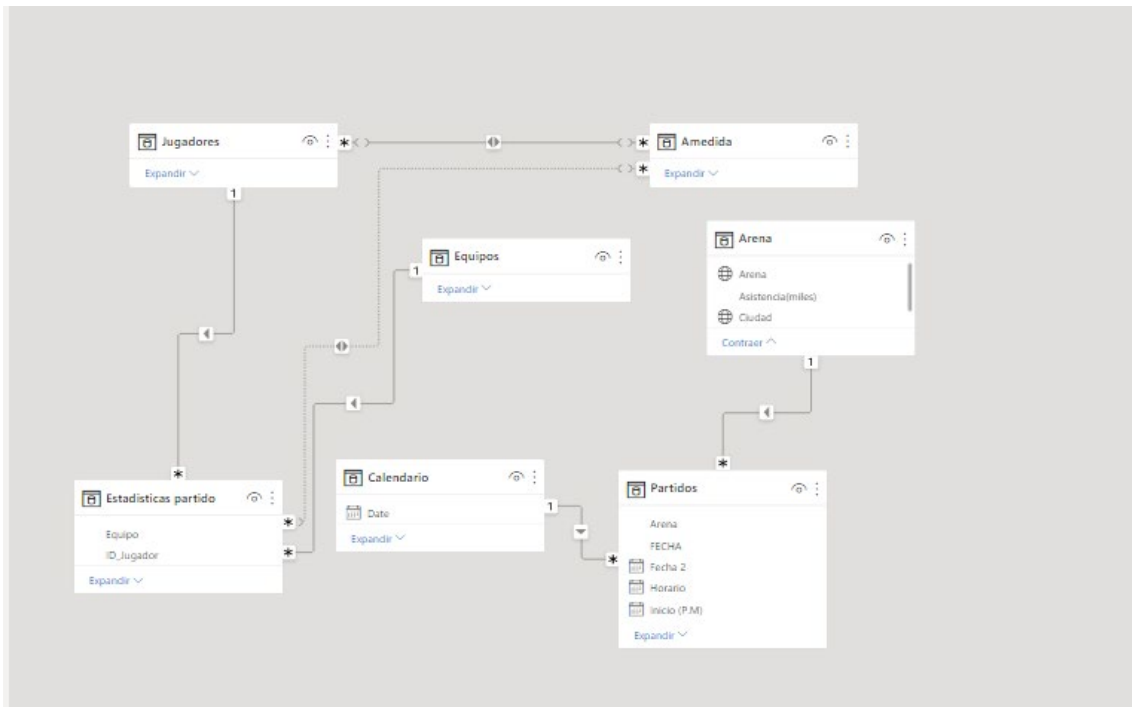
Por otro lado, en la tabla partidos y resultados, la única transformación hecha en Excel fue hacer la tabla una tabla con encabezados. Todas las demás transformaciones hechas fueron dentro de Power BI y Power Query.

### DER final del trabajo



A grandes rasgos, este termino siendo el DER final del proyecto. Igualmente, a continuación, mostrare como quedo el modelo relacional hecho en POWER BI.

### Modelo relacional



### Breve definición de la información que hay en cada tabla y de las transformaciones que hubo que hacer:

**Jugadores:** En esta tabla se especificarán el nombre del jugador, el Id y la edad. Aquí tuve que dividir la tabla proveniente del dataset anteriormente adjuntado ya que necesitaba agrupar los datos de otra forma para hacer más simple y ordenado el análisis.

Por otro lado, para esta tabla, tuve que eliminar los valores duplicados ya que cuando hice la conexión, había jugadores que se repetían debido a que habían cambiado de equipo durante la temporada. Esto generaba una fila para un equipo y otra fila para otro.

30	Keita Bates-Diop	PF-SF	24	DAL
30	Keita Bates-Diop	FP	24	MÍN.
30	Keita Bates-Diop	SF	24	DEN
31	Nicolas Batum	CF	31	CHO

**Estadísticas Partidos:** En esta tabla busque expresar todo lo que tenga que ver con estadísticas, tanto individuales como colectivas. Esta tabla es sin lugar a duda la más importante del proyecto y siento que la he tratado como tal.

Esta creo yo que fue una de las tablas que mas transformaciones hubo que hacer, principalmente de cambiar el tipo de dato, los valores de las estadísticas venían en formato texto, por lo que hubo que cambiarlo a número entero, porcentaje o decimal según correspondiese. Además, tuve el trabajo de corregir todos los nombres tanto de



los equipos como de los jugadores ya que algunos estaban mal traducidos o con fallas ortográficas.

**Equipo:** Acá solamente se menciona el nombre de los equipos de la NBA. Fue simplemente un proceso de duplicación y de quitar otras columnas ya corregidas proveniente de la tabla “Estadísticas Partidos”.

**Amedida:** Aquí pude hacer medidas vistas durante la cursada que me han servido para llevar a la visualización en el dashbord. Estas son las medidas que utilice.

```
Cant_estadios = COUNT(Arena[ID_Arena])
Cant_jugadores = COUNT(Jugadores[ID_Jugador])
Jugador mas anotador = MAX('Estadisticas partido'[PTS])
Max_asistencia = MAX(Arena[Asistencia(miles)])
Prom_asistencia = AVERAGE(Arena[Asistencia(miles)])
```

**Partidos:** En esta tabla se expresa el equipo, resultado, asistencia, arena y fecha que tuvo lugar en encuentro. Esta tabla proviene de una combinación de las tablas (que ahora están ocultas) “Partidos Dic”, “Partidos Nov” y “Partidos Oct”. Hubo que hacer muchas transformaciones para poder arreglar las complicaciones en la columna fecha

Arena: Esta tabla fue un desafío ya que hubo que modificar todos los nombres de las arenas por problemas en la traducción. A su vez, tuve que agregar una columna en el que especifica la ciudad donde estaba el estadio. Por último, tuve que categorizar los datos de dos columnas, Ciudad (ciudad) y Arena (lugar).

**Calendario:** En esta tabla lleve a representación en otra tabla las fechas de la tabla partido. La tabla calendario se llevo a cabo gracias a las siguientes funciones:

```
Calendario = CALENDAR(MIN('Partidos Oct'[Fecha]), MAX('Partidos Dic'[FECHA]))
Año = Year(Calendario[Date])
Mes = MONTH(Calendario[Date])
Día = DAY(Calendario[Date])
Nombre del mes = FORMAT(Calendario[Date], "MMMM")
```

**FechaAct:** Esta tabla o mejor dicho función, decidí agregarla para mostrar cuando a sido la ultima fecha de actualización de los datos. La función es así:

```
DateTime.LocalNow()
```

## ***Visualización de los datos***

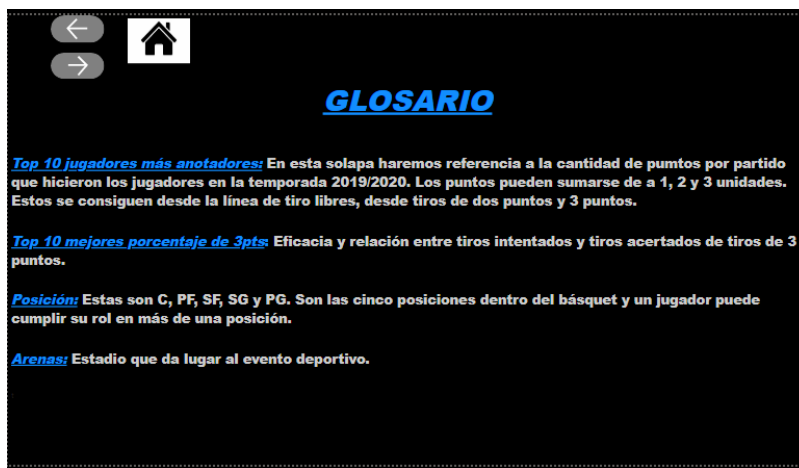
En el siguiente apartado de la documentación hablaremos de las solapas incluidas en la presentación.

## Solapa 1: Home



Lo que busca hacer en esta primera solapa es que el usuario pueda desde esta página llegar a las demás solapas. A su vez, de fondo quise poner una imagen representativa e histórica para el fanático de la NBA y de los Bulls de Michael Jordan.

## Solapa 2: Glosario



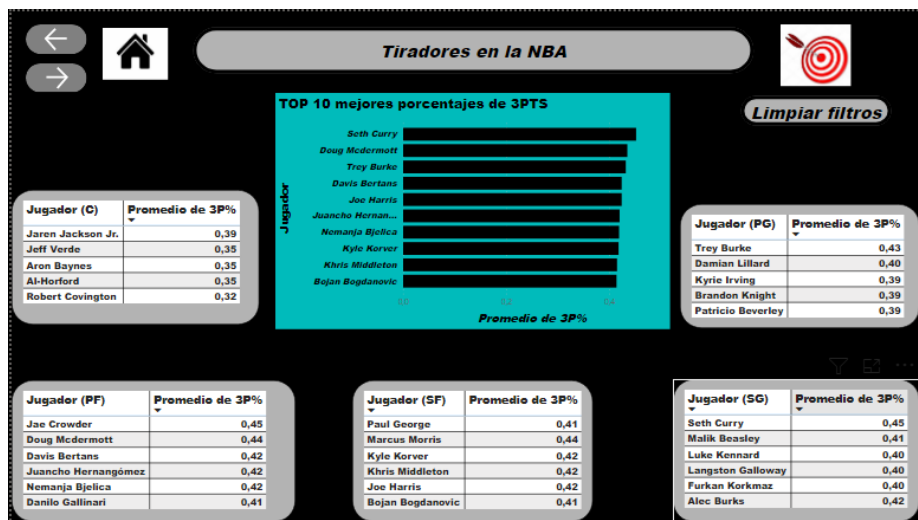
Aquí decidí poner las definiciones de los conceptos más utilizados a lo largo de la presentación. Además, incluí unos botones interactivos que se van a repetir por todas las solapas. La casita nos lleva a la solapa “Home” y las flechas nos hacen avanzar hacia atrás o hacia adelante respectivamente.

### Solapa 3: Top 10 jugadore más anotadores y Top 5 por posición



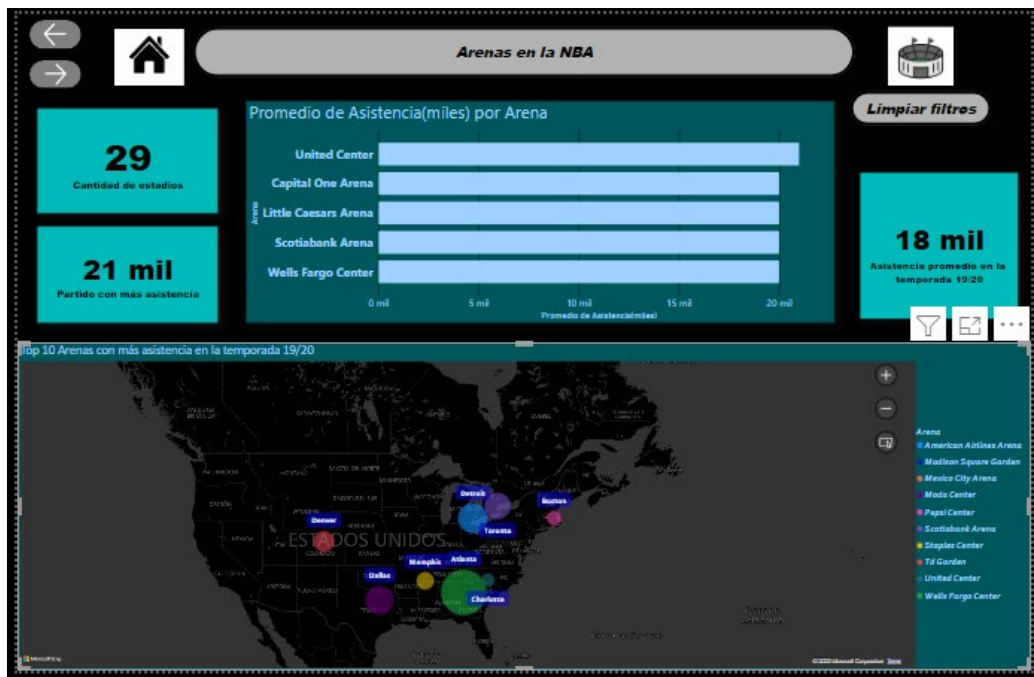
En esta solapa busque la visualización de, para mi gusto, la estadística mas representativa de los jugadores, los puntos por partido. Decidí agruparlo por los 10 máximos anotadores en general y por los 5 máximos anotadores por posición en la liga.

### Solapa 4: Top 10 mejores porcentajes de 3pts y mejores tiradores por posición



La idea de esta solapa es similar a la anterior, hacer un top de jugadores y por posición a partir de otra estadística, el porcentaje de tiro de 3 puntos. Por otro lado, agregue un botón interactivo que se va a repetir en las siguientes, la de limpiar filtros. Este lo hice a partir de la utilización de un marcador.

## Solapa 5: Arenas

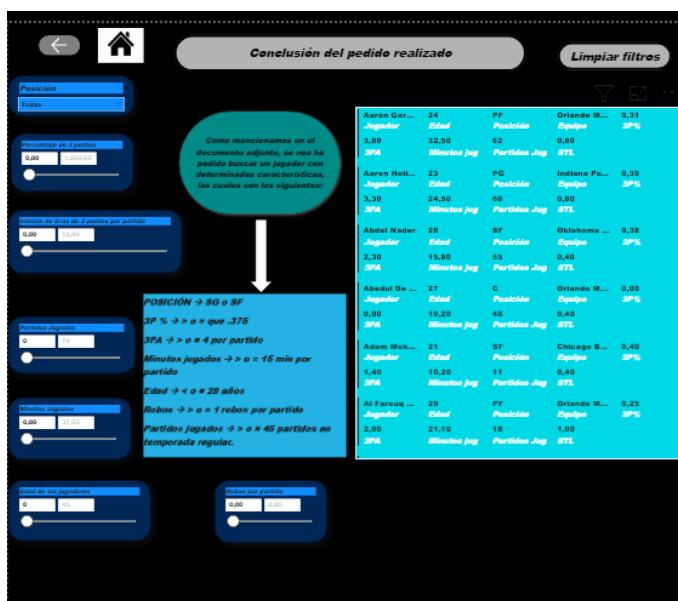


Aquí busque hacer un análisis de los estadios en la nba. Realice un top con los estadios con mas asistencia en promedio a lo largo de la temporada 2019/2020.

A su vez, implemente en esta solapa algunas medidas anteriormente mencionadas, como puede ser el partido con más asistencia, la asistencia promedio y la cantidad de estadios.

Por último, me pareció una gran idea incluir en el dashbord un mapa en el que se pueda visualizar la ciudad en la que contenga la arena correspondiente.

## Solapa 6: Conclusión



En la última solapa de la presentación, me guie a partir de dos conceptos, las segmentaciones utilizadas y requeridas y algo que escuche en la clase que considero que en esta solapa se lleva bien a cabo; el usuario puede meter mano e investigar el mismo a través de las segmentaciones lo que se nos pidió e incluso otras cosas. Si este por lo que sea necesita realizar otra filtración con características totalmente diferentes a partir de las mismas filtraciones, lo podría hacer tranquilamente y es algo que desde mi punto de vista suma mucho a la experiencia de usuario.

Una vez realizadas las filtraciones que se no han requerido, los jugadores encontrados son los expuestos en la siguiente tabla.

Alec Burks	28	SF	Golden Sta...	0,38
Jugador	Edad	Posición	Equipo	3P%
4,70	29,00	48	1,00	
3PA	Minutos jug	Partidos Jug	STL	
Justin Holli...	30	SF	Indiana Pa...	0,41
Jugador	Edad	Posición	Equipo	3P%
4,50	25,00	73	1,20	
3PA	Minutos jug	Partidos Jug	STL	
Kawhi Lee...	28	SF	Los Angel...	0,38
Jugador	Edad	Posición	Equipo	3P%
5,70	32,40	57	1,80	
3PA	Minutos jug	Partidos Jug	STL	
Paul George	29	SF	Los Angel...	0,41
Jugador	Edad	Posición	Equipo	3P%
7,90	29,60	48	1,40	
3PA	Minutos jug	Partidos Jug	STL	
Will Barton	29	SF	Denver Nu...	0,38
Jugador	Edad	Posición	Equipo	3P%
5,00	33,00	58	1,10	
3PA	Minutos jug	Partidos Jug	STL	

Por lo que, le llevaríamos la lista con estos jugadores a las altas jerarquías de la franquicia para que tome una decisión.

## **Futuras líneas**

A continuación, hablare de las posibles observaciones que me gustaría agregar al trabajo en un futuro.

1. Lo primero que se me ocurre es poder utilizar mas datos relacionados con la NBA, puede ser por ejemplo la visualización de los sueldos (estos son públicos de todos los jugadores), poder observar tendencias y a partir de esas tendencias poder encasillar las características de un jugador.
2. Otro aspecto que me gustaría comprender y poder llevar a cabo es como a partir de hechos que ya ocurrieron (estadísticas) poder

predecir hechos del futuro, una suerte de probabilidades.  
Igualmente creo que esto ya sobre pasa el área del analista de  
datos.