



UNIVERSIDAD DE VALLADOLID
ESCUELA DE INGENIERÍA INFORMÁTICA

DISEÑO Y EVALUACIÓN DE SISTEMAS INTERACTIVOS

PROYECTO DE VISUALIZACIÓN
INTERACTIVA

VISUALIZACIÓN NBA

Índice

1. Introducción	2
2. Análisis de los datos	3
2.1. Datos para el heatmap	3
2.1.1. Datos	3
2.1.2. Variables	3
2.1.3. Análisis de calidad de los datos	4
2.2. Datos para el donutchart	5
2.2.1. Datos	5
2.2.2. Variables	5
2.2.3. Análisis de calidad de los datos	5
3. Planificación de la visualización	7
3.1. Propósito:	7
3.1.1. Objetivo:	7
3.1.2. Función:	7
3.1.3. Tono:	7
3.1.4. Efecto:	7
3.2. Factores:	7
3.3. Enfoque:	7
4. Diseño	8
4.1. Diseño Heatmap	8
4.2. Diseño DonutChart	10
4.3. Interacción	10
4.4. Boceto	11
5. Fuentes documentales	12

1. Introducción

Para este proyecto de la asignatura de Diseño y Evaluación de Sistemas Interactivos he decidido realizar un mapa de calor (heatmap) o también llamado matrix chart junto con un donut chart.

He decidido realizar esta visualización porque quería realizar una visualización con datos relacionados con el deporte (un tema que me gusta mucho) y pensé que un gráfico relacionado con la NBA (Liga de Baloncesto de EEUU) podría ser una buena opción ya que en la NBA dan mucha importancia a las estadísticas.

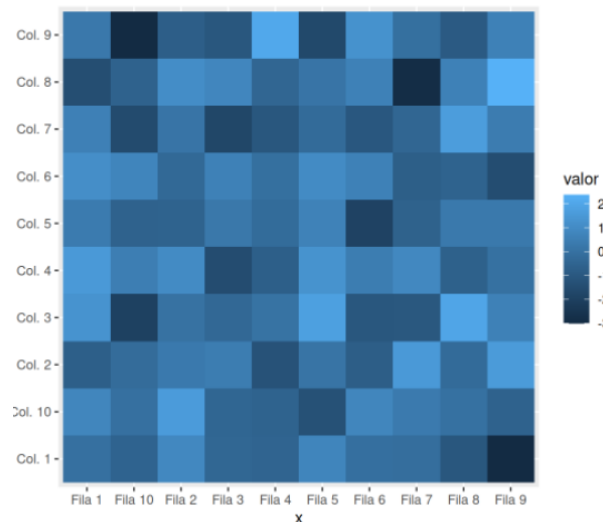


Figura 1: Ejemplo de Heatmap [1]

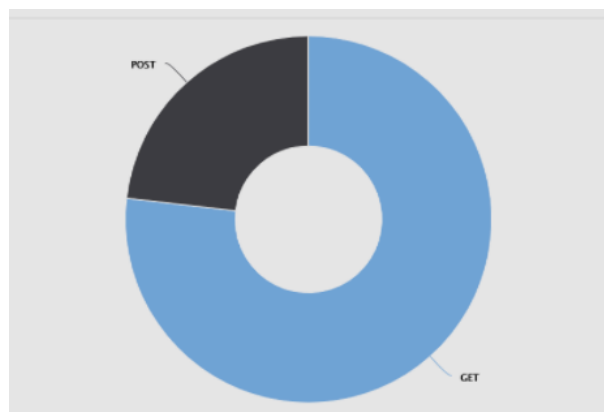


Figura 2: Ejemplo de DonutChart [2]

En el heatmap tenía pensado representar en las filas los jugadores de un equipo y como columnas variables las estadísticas más importantes en el baloncesto (tiros de campo anotados, asistencias, robos, tapones...).

En el donut chart mostrará la proporción de victorias y de derrotas de ese equipo seleccionado en la temporada que esté seleccionada.

Como interacción he pensado en que se pueda seleccionar el equipo y la temporada de la que quieres ver los datos (entre la temporada 2019-2020 o la 2020-2021).

2. Análisis de los datos

2.1. Datos para el heatmap

2.1.1. Datos

Los datos para el heatmap los he obtenido de la web kaggle (<https://www.kaggle.com/datasets>) ya que es una web que posee múltiples datasets abiertos de diversos temas en los que se encuentran varios sobre la NBA.

Más específicamente he escogido el dataset [3] el cual tiene los datos de todos los jugadores de la NBA en la temporada 2019-2020 y el dataset [4] el cual tiene los datos de todos los jugadores de la NBA pero de la temporada 2020-2021.

En los dos enlaces por cada temporada están varios datasets subidos en el que se muestran las mismas estadísticas pero uno es mostrando las estadísticas totales, otro dataset muestra las estadísticas por minuto, otro dataset estadísticas por partido...

Yo he escogido en ambas temporadas los datasets que muestra las estadísticas por partido ya que es como se suelen mostrar los datos de la NBA en la vida real y a los que están acostumbrados a ver los aficionados de la NBA.

En ambas temporadas muestran los datos referentes sólo a la temporada regular, no se muestran datos ni de los Play-Off ni de los Play-In. En la temporada 2019-2020 se incluyen los datos referentes a los partidos que se jugaron antes del parón del coronavirus junto a los partidos de temporada regular que se jugaron en la “Burbuja de Orlando”.

El formato de los dataset son csv ambos.

2.1.2. Variables

El dataset [3] referente a la temporada 2019-2020 contiene 651 observaciones y 29 variables y el dataset [4] referente a la temporada 2020-2021 contiene 497 observaciones y 29 variables, para ambos datasets escogeré estas 14 variables que son comunes en ambos datasets:

- Player: Nombre del jugador. Tipo categórico sin orden.
- Pos:(Position) Posicion en la que juega el jugador. Tipo categórico sin orden.
- Tm: (Team) Equipo al que pertenece el jugador. Tipo categórico sin orden.
- MP: (Minutes Played) Minutos jugados por partido. Tipo cuantitativo continuo.
- FG: (Field Goals Made) Tiros de campo anotados por partido. Tipo cuantitativo continuo.
- FG %: (Field Goal Percentage) Porcentaje de tiros de campo anotados. Tipo cuantitativo ratio.
- 3P: (3-Point Field Goals Made) Triples anotados por partido. Tipo cuantitativo continuo.
- 3P %: (3-Point Field Goal Percentage) Porcentaje de triples por partido. Tipo cuantitativo ratio.
- FT %: (Free Throw Percentage) Porcentaje de tiros libres. Tipo cuantitativo ratio.
- AST: (Assists) Asistencias por partido. Tipo cuantitativo continuo.
- ORB: (Offensive Rebounds) Rebotes ofensivos por partido. Tipo cuantitativo continuo.
- DRB: (Defensive Rebounds) Rebotes defensivos por partido. Tipo cuantitativo continuo.
- STL: (Steals) Robos por partido. Tipo cuantitativo continuo.
- BLK: (Blocks) Tapones por partido. Tipo cuantitativo continuo.

Las variables se pueden agrupar en varios grupos(una misma variable puede pertenecer a varios grupos(son conjuntos no disjuntos):

- Variables de tiros de campo = { FG, FG % }
- Variables de triples: {3P,3P % }
- Variables de tiro= { Variables tiro de campo, Variables de triple, FT % } = {FG, FG %, 3P,3P %, FT % }
- Variables de ataque = {Variables de tiro, AST,ORB } = { FG, FG %, 3P,3P %, FT %, AST,ORB }
- Variables de rebote = {ORB,DRB }
- Variables de defensa = {DRB, STL,BLK }

2.1.3. Análisis de calidad de los datos

Para realizar un heatmap necesitamos que los datos estén estandarizados ya que no se miden en las mismas unidades ni tienen la misma varianza las variables. Sino se realiza esto no se podría utilizar la misma paleta de colores para todo el heatmap.

Para estandarizar los datos utilizo esta fórmula:

$$z_i = \frac{x_i - \mu}{\sigma}$$

siendo x_i cada observación original y z_i cada observación ya normalizada.

	Player	Tm	MP	FG	FG.	X3P	X3P.	FT.	AST	ORB	DRB	STL	BLK
1	Steven Adams	OKC	26.7	4.5	0.592	0.0	0.333	0.582	2.3	3.3	6.0	0.8	1.1
2	Bam Adebayo	MIA	33.6	6.1	0.557	0.0	0.143	0.691	5.1	2.4	7.8	1.1	1.3
3	LaMarcus Aldridge	SAS	33.1	7.4	0.493	1.2	0.389	0.827	2.4	1.9	5.5	0.7	1.6
4	Kyle Alexander	MIA	6.5	0.5	0.500	0.0	NA	NA	0.0	1.0	0.5	0.0	0.0
5	Nickel Alexander-Walker	NOP	12.6	2.1	0.368	1.0	0.346	0.676	1.9	0.2	1.6	0.4	0.2
6	Grayson Allen	MEM	18.9	3.1	0.466	1.5	0.404	0.867	1.4	0.2	2.0	0.3	0.1
7	Jarrett Allen	BRK	26.5	4.3	0.649	0.0	0.000	0.633	1.6	3.1	6.5	0.6	1.3
8	Kadeem Allen	NYK	11.7	1.9	0.432	0.5	0.313	0.636	2.1	0.2	0.7	0.5	0.2
9	Al-Farouq Aminu	ORL	21.1	1.4	0.291	0.5	0.250	0.655	1.2	1.3	3.5	1.0	0.4
10	Justin Anderson	BRK	10.7	1.0	0.263	0.6	0.207	0.500	0.8	0.1	2.0	0.0	0.6
11	Kyle Anderson	MEM	19.9	2.3	0.474	0.4	0.282	0.667	2.4	0.9	3.4	0.8	0.6
12	Ryan Anderson	HOU	7.0	1.0	0.286	0.5	0.200	NA	1.0	0.0	3.5	0.5	0.0
13	Giannis Antetokounmpo	MIL	30.4	10.9	0.553	1.4	0.304	0.633	5.6	2.2	11.4	1.0	1.0
14	Kostas Antetokounmpo	LAL	4.0	0.6	1.000	0.0	NA	0.500	0.4	0.4	0.2	0.0	0.0
15	Thanasis Antetokounmpo	MIL	6.5	1.2	0.500	0.0	0.000	0.412	0.8	0.6	0.6	0.4	0.1
16	Carmelo Anthony	POR	32.8	5.8	0.430	1.5	0.385	0.845	1.5	1.2	5.1	0.8	0.5
17	OG Anunoby	TOR	29.9	4.1	0.505	1.3	0.390	0.706	1.6	1.2	4.1	1.4	0.7
18	Ryan Arcidiacono	CHI	16.0	1.6	0.409	0.9	0.391	0.711	1.7	0.3	1.6	0.5	0.1

Figura 3: Dataset temporada 2019-2020 con las columnas que he escogido

<pre>> numeroNA<-apply(datosSelected,2,function(x) sum(is.na(x)));numeroNA</pre>													
Player	Tm	MP	FG	FG.	X3P	X3P.	FT.	AST	ORB	DRB	STL	BLK	
0	0	0	0	2	0	35	33	0	0	0	0	0	0

Figura 4: Número de datos perdidos por columna(datos 2019-20)

Analizando los datos con R he visto como existen datos perdidos (NA) en las variables cuantitativas de tipo ratio (Porcentaje de tiros de campo, porcentaje de triples y porcentaje de tiros libres). Esto ocurre cuando no han tirado ningún tiro de campo, triple o tiro libre entonces para calcular el porcentaje dividen $\frac{0}{0}$ y ponen como porcentaje un NA, esto podría solucionarse de una

manera sencilla sustituyendo los datos perdidos por 0.

Otra transformación que voy a utilizar es la de agregación y la opción de agregación va a ser la del filtrado de los datos ya que tengo los datos de todos los equipos de la NBA y yo la visualización que voy a mostrar sólo se representarían las del equipo que se haya seleccionado.

```
> numplayersByTeam<- table(datosSelected$Tm);numplayersByTeam
ATL BOS BRK CHI CHO CLE DAL DEN DET GSW HOU IND LAC LAL MEM MIA MIL MIN NOP NYK OKC ORL PHI PHO POR SAC SAS TOR TOT UTA WAS
21  17  24  17  16  21  19  21  22  22  21  17  20  20  21  21  17  24  17  17  18  19  19  19  19  23  18  18  60  20  23
>
```

Figura 5: Número jugadores por equipo(datos 2019-20)

Como el número de jugadores por equipo no es el mismo y además me resulta excesivo mostrar un heatmap de tantos jugadores voy a realizar otro filtrado de escoger los 10 jugadores con más minutos por equipo, para que todos los heatmaps tengan el mismo número de filas y columnas.

Todas las transformaciones mencionadas se realizarán tanto en los datos de la temporada de 2019-2020 como en los datos de la temporada 2020-21.

2.2. Datos para el donutchart

2.2.1. Datos

Para los datos para la visualización del donutchart los he sacado de la Wikipedia, para la temporada 2019-2020 he seleccionado los datos de la tablas “Clasificación de los equipos participantes” y “Clasificaciones antes del parón” de la página [5] y los datos de la temporada 2020-2021 los he seleccionado de la página [6] de la tabla de clasificaciones.

En ambas temporadas muestran los datos referentes a la temporada regular, no se muestran datos ni de los Play-Off ni de los Play-In, igual que con los datos para el heatmap.

He utilizado una herramienta externa que transforma un HTML en csv para tener en csv las dos tablas de datos.

Una aclaración importante es que en la temporada 2019-2020 hubo un parón por el coronavirus y en la reanudación jugaron 8 partidos de temporada regular en la famosa “Burbuja de Orlando” en esta burbuja no participaron todos los equipos, sólo los que en ese momento todavía tenían opciones de entrar en Play-Off. Es por ello, que los datos de la temporada 2019-2020 los seleccioné de dos tablas, para los equipos que no fueron a la burbuja de Orlando escogí los datos de la tabla “Clasificaciones antes del parón” y para los equipos que fueron a la burbuja los obtuve “Clasificación de los equipos participantes” en los que sí cuentan esos 8 partidos que se jugaron allí.

2.2.2. Variables

Ambos datasets tendrán 30 filas(número de equipos en la NBA) y 3 columnas:

- Tm: (Team) Equipo de la nba. Tipo categórico sin orden.
- G:(Ganados) Número de partidos ganados en temporada regular. Tipo cuantitativo continuo.
- P:(Perdidos) Número de partidos perdidos en temporada regular. Tipo cuantitativo continuo.

2.2.3. Análisis de calidad de los datos

Tanto en este dataset como en el de la temporada 2020-2021 en el que se muestran los datos de las victorias y derrotas no se encuentra ningún NA.

	G	P	Tm
1	56	17	MIL
2	53	19	TOR
3	48	24	BOS
4	45	28	IND
5	44	29	MIA
6	43	30	PHI
7	35	37	BRK
8	33	40	ORL
9	25	47	WAS
10	21	45	NYK

Figura 6: Dataset partidos 2019-2020

```
> maxGames<- datos201920[(datos201920$Partidos)== max(datos201920$Partidos),c("Tm","Partidos")]
;maxGames
  Tm Partidos
22 DAL      75
> minGames<- datos201920[(datos201920$Partidos)== min(datos201920$Partidos),c("Tm","Partidos")]
;minGames
  Tm Partidos
29 MIN      64
> |
```

Figura 7: Equipo con más partidos y con menos partidos en la temporada 2019-2020

Analizando los datos, vemos como en la temporada 2020-21 todos los equipos jugaron el mismo número de partidos (72), ya que no hubo parón, pero en los datos de la temporada 2019-2020, cada equipo jugó un número diferente de partidos por culpa del parón, el que más partidos fue Dallas Mavericks con 75 partidos y el que menos Minnesota Timberwolves con 64.

Por este motivo, cuando comparemos equipos o cuando comparemos el mismo equipo en diferentes temporadas es más apropiado fijarse en si ha aumentado o disminuido el porcentaje de victorias ($\frac{Victorias}{Victorias+Derrotas} * 100 \%$) más que fijarse en si han aumentado o disminuido las victorias en total, porque puede que el motivo sea simplemente la diferencia en el número de partidos.

La transformación que voy a utilizar es la de agregación y la opción de agregación va a ser la del filtrado de los datos ya que tengo los datos de todos los equipos de la NBA y yo la visualización que voy a mostrar sólo se representaran las del equipo que se haya seleccionado.

3. Planificación de la visualización

3.1. Propósito:

3.1.1. Objetivo:

Comparar jugadores de un equipo de la NBA según las variables FG, FG %, 3P, 3P %,FT %, AST, ORB, DRB, STL y BLK e intentar buscar relaciones entre las variables. También comparar el porcentaje de victorias de cada equipo en la temporada antes del Covid-19 (temporada 2019-2020) frente a la temporada después del confinamiento(temporada 2020-2021)

3.1.2. Función:

Exploratoria. Ya que el objetivo es realizar un análisis visual de los datos y lo que queremos es que nos permita ver correlaciones , relaciones, conexiones...

3.1.3. Tono:

Analítico porque queremos que la visualización sea lo más legible posible.

3.1.4. Efecto:

- Buscar una posible relación entre las variables al tener tantas variables podemos buscar si hay relaciones entre las variables de tiro (FG, FG %, 3P, 3P %,FT %) o si hay relaciones entre las variables de ataque (FG, FG %, 3P, 3P %,FT %, AST, ORB) frente a las de defensa (DRB, STL,BLK)...
- Generar la pregunta de si los jugadores con más tiros anotados es por tener mejor porcentaje de tiro o si es porque tiran más.
- Comparar que jugadores siguen en el mismo equipo después de una temporada y ver como han mejorado o empeorado sus estadísticas y en que variables han sido.
- Generar la pregunta de a que equipos les ha afectado más la pandemia respecto al porcentaje de victorias.

3.2. Factores:

Habría que tener en cuenta las limitaciones y grado de conocimiento de la herramienta a usar y el tiempo (fecha de entrega). También tener en cuenta solo disponemos datos de dos temporadas por lo que no podemos hacer un análisis de varios años, sólo se pueden comparar esas dos temporadas entre sí.

3.3. Enfoque:

- Se mostrará relaciones entre los 10 jugadores con más minutos frente al resto de variables. Esta información se mostrará por equipos y por temporada.
- Se mostrará el porcentaje de victorias y de derrotas. Esta información se mostrará por equipo y por temporada.
- Se comparará las variables entre sí para ver cuales están relacionadas.
- **Interacción:** Se podrá seleccionar el equipo y la temporada (2019-2020 o 2020-2021) para que muestre los datos de ese equipo y de esa temporada en concreto.

4. Diseño

Arriba del todo se mostrará el nombre del equipo seleccionado, a la izquierda del nombre se mostrará su logo dentro del Donut Chart. Debajo del nombre se mostrará el heatmap del equipo seleccionado en la temporada seleccionada.

El menú de selección se mostrará a la derecha del heatmap y constará de los 30 equipos organizado en 6 filas de 5 equipos, cada fila será una de las 6 divisiones de la NBA División Atlántico, División Central y División Sureste (las 3 de la conferencia Este) y luego pondré las 3 divisiones de la conferencia Oeste (División Noroeste, División Pacífico, División Suroeste).

En este menú pondré los logos de los equipos seleccionables y cuando se clicke en uno de los logos cambiara el heatmap, el titulo del equipo seleccionado y el logo del equipo seleccionado. Dentro de cada división los equipos se mostrarán por orden alfabético. Encima habrá dos botones para elegir entre la temporada 2019-2020 o la temporada 2020-2021.

Por defecto saldrá seleccionado el primer equipo que en nuestro caso será Boston Celtics y la temporada seleccionada será 2019-2020.

4.1. Diseño Heatmap

Para el heatmap los datos que voy a utilizar van a ser los 10 jugadores que más minutos tienen del equipo que esté seleccionado en la temporada que este seleccionada. Cada fila representará a un jugador siendo el que más minutos el que más arriba se encuentre.

El heatmap básico al cuál le voy a aplicar cambios y en el que voy a basar mi visualización va a ser el siguiente:

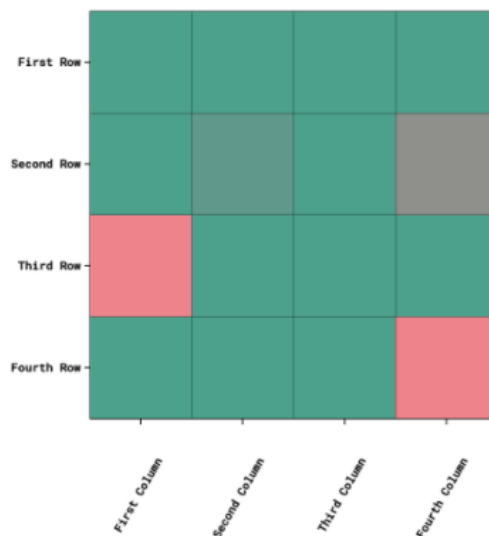


Figura 8: Heatmap basico en el que basaré mi visualización [7]

Las columnas serán las variables que he seleccionado (FG, FG %, 3P,3P % , FT %,AST, ORB, DRB,STL,BLK), El orden de las columnas no será aleatorio para que podamos mejor ver los grupos de variables y poder comparar mejor las variables.

Primero colocaré de izquierda a derecha las cinco variables de tiro FG, FG %, 3P,3P % y FT %, luego colocaré AST, ORB que junto a las variables de tiro forman las variables de ataque. A la derecha de ORB pondré DRB ya que son las dos variables relacionadas con los rebotes. Por último incluiré STL,BLK que junto a DRB son las variables de defensa así logro que todos los grupos de variables estén juntas pese a ser conjuntos no disjuntos.(Esto se verá gráficamente en el boceto).

Luego el valor de cada variable de cada jugador se mostrará estandarizada para que se pueda utilizar la misma paleta de colores en todas las variables.

Para la paleta de colores tenía dos opciones para escoger, la primera era utilizar la paleta Rojo-Amarillo-Verde siguiendo la metáfora del semáforo en la que el color verde indica un valor alto en la variable frente a rojo indica un valor bajo.

La otra opción es por cada equipo seleccionar su color principal de equipo como valor máximo y el blanco como valor mínimo, por tanto a más saturación de color implica mayor valor en la variable, esto puede ser más vistoso ya que por cada equipo el color del heatmap cambia pero es menos usable porque en cada heatmap el color del máximo cambia (lo que puede confundir al receptor) y no se aprecian tanto las diferencias entre valores como con la paleta de Rojo Amarillo y Verde.

Por este motivo he escogido la paleta Rojo-Amarillo-Verde porque he priorizado la usabilidad antes que el diseño.

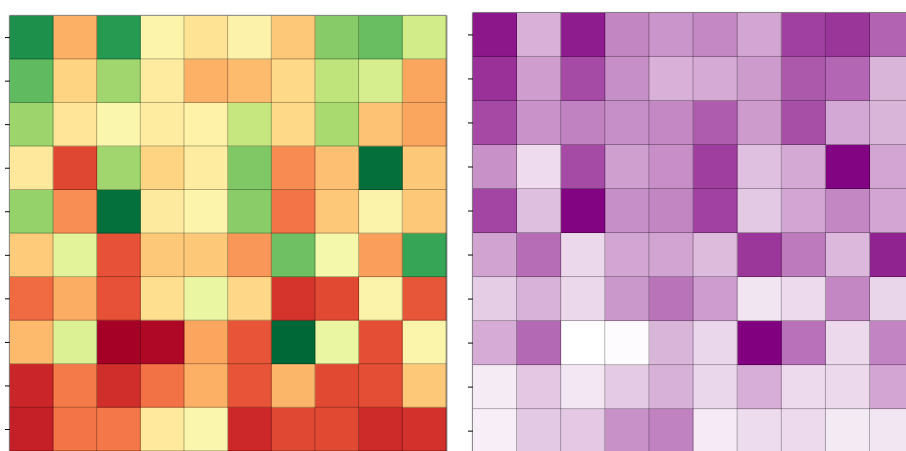


Figura 9: A la izquierda con la paleta(Rojo-Amarillo-Verde) y a la derecha la segunda versión variando la saturación del color principal del equipo. Son los mismos datos para los dos heatmaps

Todas las variables han sido escogidas para que cuanto más alto es el valor significa que es mejor, por ello he evitado variables como número de pérdidas o número de faltas que cuantas más tenga es peor y luego aunque utilice la paleta inversa de colores puede dar lugar a confusión al usuario.

He escogido un heatmap para representar estos datos porque pienso que es la mejor opción para realizar comparaciones entre varias variables para un mismo individuo y entre varios individuos para una misma variable. En un único gráfico puedo realizar las comparaciones deseadas que utilizando otros gráficos (como diagrama de barras) necesitaría varios gráficos y con el heatmap puedo tenerlo en un único gráfico.

Otra opción podría ser un diagrama de barras múltiples o un diagrama de barras apiladas pero la comparación entre jugadores y entre variables podría ser muy difícil de observar.

4.2. Diseño DonutChart

El DonutChart en el que voy a basar mi visualización va a ser el siguiente :

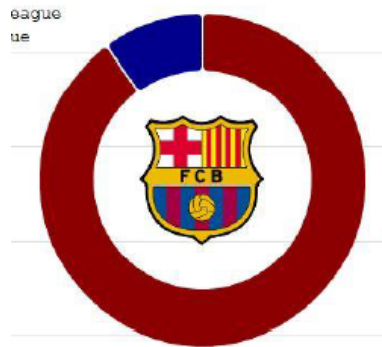


Figura 10: Donutchart en el que basaré mi visualización [8]

El cual es de un antiguo compañero del curso 18-19 llamado Alberto Calvo, he elegido este en concreto ya que me ha gustado como tiene el logo del equipo dentro del gráfico como voy a hacer pero con los logos de la NBA.

Los colores que voy a elegir van a ser:

- Para el porcentaje de victorias: El color principal del equipo (verde para Boston Celtics, azul para Orlando Magic...)
- Para el porcentaje de derrotas: Color blanco siempre, porque es un color que no voy a utilizar como color principal de ningún equipo y para que destaque el porcentaje de victorias frente al de derrotas y sea más fácil realizar la comparación entre equipos

La leyenda se encontrará a la izquierda del donut Chart ya que se dispone de más espacio que a la derecha donde se encontrará el nombre del equipo seleccionado.

Se ha elegido un un donut chart en vez de un gráfico de sectores porque gracias a ello puedo poner el logo del equipo dentro del gráfico lo cuál queda muy estético.

4.3. Interacción

La interacción que voy a utilizar es que cuando se seleccione el logo de algún equipo cambiará el color del fondo donde aparece el nombre del equipo seleccionado y se mostrará el logo del equipo seleccionado dentro del donut chart con el porcentaje de victorias y de derrotas de esa temporada. También cambiará el heatmap mostrando las estadísticas de los 10 jugadores con más minutos de ese equipo ordenado por minutos, el jugador con más minutos será el que esté más arriba.

Cuando se seleccione otro equipo se mostrará el heatmap de los 10 jugadores con más minutos de ese equipo seleccionado.

Otras interacciones serán que cuando pases por encima el ratón de un logo saldrá una etiqueta de tipo tooltip mostrando el nombre completo del equipo. Cuando pase por encima con el ratón de cada rectángulo del heatmap se mostrará un tooltip con el valor que toma en esa variable ese jugador en concreto (valor de la celda no estandarizada).

También cuando se pase el ratón por encima del nombre de los jugadores se mostrará información adicional como posición en la que juega o minutos jugados.

4.4. Boceto

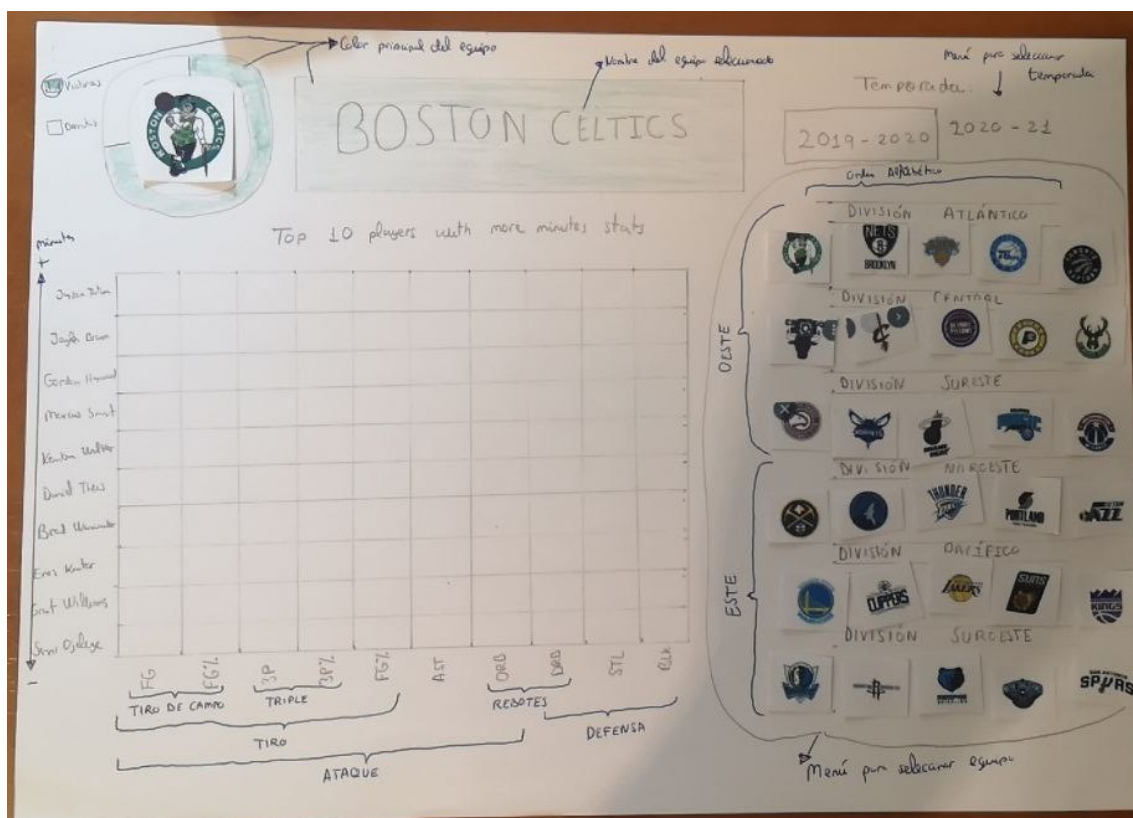


Figura 11: Boceto Inicial

Este es mi boceto inicial, en azul se muestran las anotaciones.

5. Fuentes documentales

Referencias

- [1] ggplot2 : Quick correlation matrix heatmap - r software and data visualization, stdh, URL <http://www.sthda.com/sthda/RDoc/figure/ggplot2/ggplot2-correlation-matrix-heatmap-geom-tile-data-visualization-1.png>.
- [2] Donut chart widgets, (10-05-2021), URL <https://docs.devo.com/confluence/ndt/latest/activeboards/activeboard-widgets/donut-chart-widgets>.
- [3] Kim N., Nba player stats (2019-20), kaggle, (15-10-2020), URL https://www.kaggle.com/nicklauskim/nba-per-game-stats-201920?select=nba_2020_per_game.csv.
- [4] Alpaydin U., Nba 2020-2021 season player stats, kaggle, (18-02-2021), URL https://www.kaggle.com/umutalpaydn/nba-20202021-season-player-stats?select=nba2021_per_game.csv.
- [5] Wikipedia, editada el 29-08-2021, URL https://es.wikipedia.org/wiki/Temporada_2019-20_de_la_NBA.
- [6] Wikipedia, editada el 13-10-2021, URL https://es.wikipedia.org/wiki/Temporada_2020-21_de_la_NBA.
- [7] O’Keeffe D., Build your own heat map with d3, gitconnected, (30-07-2020), URL <https://levelup.gitconnected.com/build-your-own-heat-map-with-d3-b5ce0284a256>.
- [8] Calvo A., Límite salarial - procedencia goles liga santander 17/18, (13-12-2018), URL https://gitlab.inf.uva.es/desi_18-19/futbol_presupuesto_clubes_vs_goles_marcados_por_jugadores_extranjeros.