# A9. Práctica Final Visualización de Datos.

### 1. Presentación de la visualización.

La visualización realizada para la práctica final se titula: "DataViz: ¿Cómo juegan al League Of Legends los equipos profesionales?".

La visualización y el código utilizado para desarrollar los ficheros de la visualización y para generar los datasets que aprovecha la visualización se encuentran alojados en un repositorio de github: <a href="https://github.com/manurubo/manurubo a9\_vd">https://github.com/manurubo/manurubo a9\_vd</a>. Además, gracias a github pages, la visualización se encuentra disponible en <a href="https://manurubo.github.io/manurubo a9\_vd/manurubo final\_vd.html">https://manurubo.github.io/manurubo a9\_vd/manurubo final\_vd.html</a>.

La visualización es un dashboard que estudia diversos aspectos de como juegan diferentes equipos de LeagueOfLegends profesionales. Entre estos aspectos se incluyen el número de partidas por liga y equipo (circle packing), los campeones más jugados y su rendimiento con ellos (gráfico de barras con color), la evolución del ratio de victorias a lo largo de los meses para el lado azul y rojo (time-series multilínea), la relación entre la victoria y conseguir objetivos clave de la partida (parallel sets), y el perfil numérico en varias métricas de partida (parallel-coords). La visualización esta diseñada con la idea de que se puede seleccionar un equipo o liga en la primera visualización arriba a la izquierda, y esta actúa como filtro y modifica el resto de las gráficas en tiempo real.

Esto permite que la visualización nos permita estudiar tanto el perfil global de las partidas profesionales de League Of Legends como el perfil concreto de una liga o de un equipo en un torneo.

Este documento sirve como entrega de la práctica final de la asignatura de visualización de datos, en él se va a explicar las preguntas que responde la visualización, que gráficas se han escogido y por qué, cómo responden las preguntas, qué uso puede darle un usuario medio, y para qué y quién puede servir la visualización. Además, se hace una descripción técnica del proyecto llevado a cabo para realizar la visualización, donde se indican lenguajes, librerías, código, datos utilizados y transformaciones a estos datos.

### 2. Explicación de la visualización.

La visualización escogida de manera general ha sido un dashboard, ya que es útil realizar un dashboard para poder estudiar diversas gráficas que representan los datos a la vez y son interactivas. El dashboard nos permite extraer conclusiones en conjunto de todas las visualizaciones.

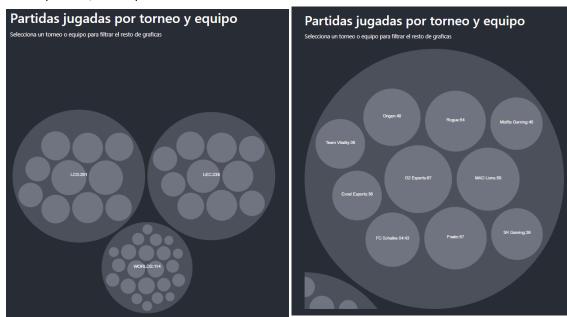
## 2.1. ¿Qué preguntas responde la visualización? ¿Qué visualización se ha escogido para responder estas preguntas?

Este dashboard está enfocado a responder la pregunta del título de la visualización "¿Cómo juegan al League Of Legends los equipos profesionales?". Ya que permite realizarse varias preguntas que responden a ideas o maneras de jugar al juego por parte de equipos profesionales y por tanto se engloban todas bajo esta temática.

Si recorremos las gráficas en sentido de izquierda a derecha y de arriba abajo, la primera que se encuentra nos permite responder a preguntas como: ¿Cuántas partidas se han jugado en cada torneo? ¿Cuántas partidas ha jugado un equipo en un torneo? ¿Ha jugado más partidas en un torneo un equipo Y que un equipo X?

Para responder a estas preguntas se ha escogido una visualización de Circle Packing. Se ha escogido esta visualización porque permite representar mediante un círculo las partidas que ha jugado un equipo, y fácilmente comparar entre equipos, puesto que la bola de un será más grande que la de otro equipo si ha jugado más partidas. Este tipo de visualización la he elegido también por que permite tener jerarquías respecto a los círculos, de manera que he podido agrupar todas las partidas por ligas, y dentro de cada liga los círculos de los equipos que participan en esta liga. También es interesante que esta visualización es muy fácil de interactuar con ella, ya que si el usuario quiere ver el círculo de un equipo, ha de pinchar en él, aprovechando esto se ha implementado que cuando un usuario selecciona un equipo o liga, se filtran el resto de gráficas acorde a este equipo o liga.

Mediante esta visualización, se podría responder ¿Cuántas partidas se han jugado en la LEC? A simple vista (izquierda) el usuario ya puede ver que han sido 239 y comparar su número con el de otras ligas. Para la pregunta ¿Cuántas partidas ha jugado G2 Esports en la LEC? El usuario tiene que seleccionar la bola de la LEC y visualiza las partidas de cada equipo en la LEC, en este caso (derecha) el usuario verá que son 67 partidas, 9 más que MAD Lions.



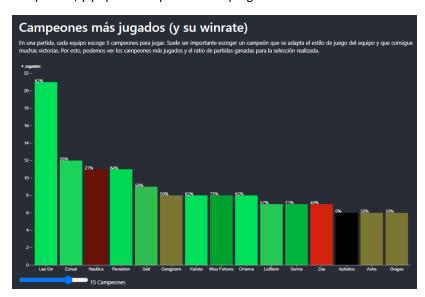
Actualmente, el usuario habrá seleccionado las partidas de LEC y el resto de gráficas se han filtrado acorde a esto, si quiere filtrar por G2 Esports en la LEC tendrá que seleccionar la bola de G2 Esports. En la parte de arriba del dashboard hay un header que indica la selección de filtrado.

### Mostrando datos de la liga: LEC

La segunda visualización es un gráfico de barras coloreadas acorde a una métrica. Este gráfico permite responder preguntas como ¿Cuáles han sido los 10 campeones más jugados en todo el año? ¿Y en una liga o por un equipo en concreto? ¿Los campeones más jugados por un equipo han sido buenas elecciones o malas elecciones para el equipo?

Se ha escogido un gráfico de barras porque simplemente queremos representar una variable numérica (veces jugado un campeón) respecto al nombre del campeón, y la manera más sencilla de realizarlo es esta. Además, tener las barras ordenadas permite que el usuario vea rápidamente los campeones más jugados, y se ha añadido una barra inferior para que el usuario pueda seleccionar entre solo el campeón más jugado hasta 20 campeones más jugados. Por último, cada barra es un gran rectángulo que ocupa espacio y esto permite añadir una tercera dimensión a la visualización, que en este caso es el ratio de victorias de ese campeón para la selección.

Por tanto, si nos preguntamos, ¿Cuáles son los 15 campeones más jugados de Mad Lions en la LEC? ¿Han sido buenas elecciones o escogen muchas veces campeones que no les funcionan? El usuario debería primeramente seleccionar a Mad Lions en LEC en el gráfico Circle Packing, y este se actualiza automáticamente para los datos de MAD Lions. A continuación, el usuario tiene que seleccionar abajo 15 campeones, y ya puede responder a las preguntas.



Los campeones más jugados son los 15 que aparecen en el eje X, destacando sobre todo Lee Sin con 21 partidas. Por otra parte, vemos que el equipo ha jugado generalmente campeones que les han dado buenos resultados y tienen un win ratio de más de 50 con la mayoría, sin embargo, han escogido a Nautilus 11 veces y solo tienen un win ratio de 27%, muy bajo para un personaje que escogen tanto.

La tercera gráfica es una time series multilínea que representa el win rate en el lado azul y el rojo a lo largo del tiempo. Esto nos permite preguntarnos ¿Qué lado tiene ventaja a lo largo del tiempo el azul o el rojo? ¿Es mejor un equipo en el lado azul o en el lado rojo?

Se ha escogido una time-series para esta gráfica porque es interesante ver el win rate en función del lado del mapa como algo que va cambiando, ya que el metajuego (campeones fuertes, ideas de jugar y plan de partida) evoluciona y a veces es beneficioso escoger el lado azul o el lado rojo. Generalmente se dice que hay ventaja para el lado azul, porque escoge primero y tiene más fácil el acceso a la zona del dragón.

Para responder a ¿Tiene ventaja el lado azul sobre el rojo? Simplemente se ha de ver la gráfica sin filtrar nada en el circle packing. Abajo vemos que generalmente el lado Azul ha tenido ventaja o ha estado casi igualado con el rojo, salvo en Octubre donde el lado rojo realmente parece tener ventaja.



Para responder a si por ejemplo, ¿El equipo Flyquest en la LCS gana más partidas en el lado azul que en el lado rojo? El usuaro tendría que seleccionar la bola de FlyQuest en la LCS y la visualización muestra como generalmente, FlyQuest es un equipo que salvo en Septiembre, siempre ha tenido mayor o igual winratio en azul que en rojo. Por tanto, si FlyQuest tiene que elegir un lado, debería escoger el lado azul.



La cuarta visualización nos permite responder preguntas como ¿Qué relación hay entre ganar una partida y los objetivos de una partida como conseguir el barón nashor, conseguir la primera torre o

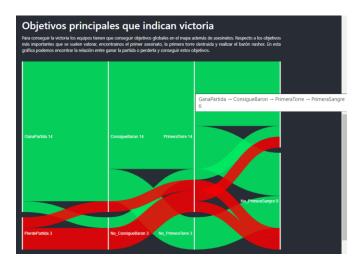
conseguir el primer asesinato? ¿Cuántas partidas que ganó el campeón mundial Damwon Gaming consiguió además de la victoria, el nashor, la primera torre y el primer asesinato?

Todas estas preguntas se pueden responder mediante una visualización de parallel set, se ha escogido porque no solo permite estudiar la cantidad de partidas que un equipo cae sobre una categoría (como sería un gráfico de barras para cada variable del parallel set) sino que también permite estudiar la relación entre estas variables, viendo el flujo de relaciones de partidas de los equipos entre una categoría y otra, por ejemplo, permite ver el flujo de partidas de equipos entre aquellos que ganan y aquellos que han conseguido el barón. Además, al tener varias variables podemos profundizar entre las relaciones y ver por ejemplo además de la victoria y el nashor, aquellos que equipos que no consiguieron la primera torre. Por último, se escoge porque los links entre categorías pueden tomar un color según una variable para distinguir a simple vista varios caminos, en nuestro caso para ver los links en función de si un equipo ganó o perdió.

Para responder a ¿Qué relación hay entre ganar una partida y los objetivos de una partida como conseguir el barón nashor, conseguir la primera torre o conseguir el primer asesinato? el usuario no tiene que seleccionar ningún círculo en el circle packing y visualizar esta gráfica. De esta manera y gracias a los links coloreados, puede ver como hay mucha relación entre conseguir la victoria y conseguir el barón, ya que el enlace verde entre ambas categorías es muy grande, sin embargo, al ver los links verdes que llegan a primera torre, se puede ver que hay bastantes menos que en consigue barón, por tanto el usuario puede concluir que hay mayor relación entre ganar partida y conseguir el barón que entre ganar y conseguir la primera sangre.



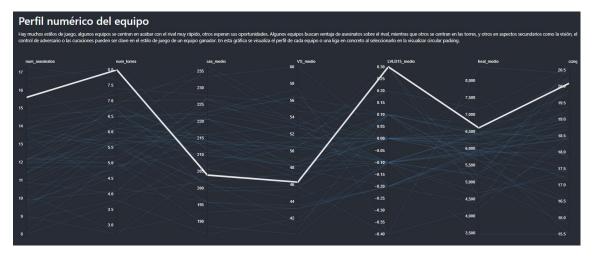
Para responder a ¿Cuántas partidas que ganó el campeón mundial Damwon Gaming consiguió también el nashor, la primera torre y el primer asesinato? El usuario tiene que seleccionar a Damwon Gaming en el circle packing bajo el torneo WORLDS y poner el ratón sobre el enlace de primera sangre a primera torre en verde, como se ve en la imagen. Se observa que Damwon Gaming consiguió la victoria y todos los objetivos en 6 partidas de las 14 que gano, lo que indica que fue muy dominador en sus partidas.



La última visualización nos permite responder a la pregunta ¿Cuál es el perfil numérico de juego de un equipo o de una liga respecto a los demás equipos y ligas? ¿Cuáles son sus fortalezas? ¿Y sus desventajas?

Para responder a estas preguntas se ha escogido un parallel coords, debido a que este gráfico nos permite representar varias métricas numéricas cercanas entre ellas en los mismos ejes y tener una línea para cada uno de los equipos / ligas del dataset. Así, podemos entender de un vistazo rápido como el perfil del equipo seleccionado y compararlo con el resto de los equipos. Las dimensiones que se observan son: el número de asesinatos por partida, el número de torres que destruyen, el número de súbditos que destruyen, la puntuación de visión media, la diferencia de nivel en el minuto 15 respecto al enemigo, la curación media por partida y el control de adversario medio por partida.

Para responder a las preguntas, un usuario tiene que seleccionar el equipo en el circle packing, y después visualizar la gráfica, el equipo o liga seleccionado se encuentra con su línea en blanco mientras que las otras disminuyen su visión. Si queremos responder las preguntas para el equipo Cloud 9 en la LCS, la visualización se ve:



El perfil numérico de Cloud 9 en la LCS es el que se observa con la línea blanca, podemos ver que sus fortalezas son el número de torres que consiguen, la diferencia de nivel medio al minuto 15 y el control del adversario medio. Lo que indica que son un equipo que consigue objetivos como las torres normalmente, que dominan los primeros 15 minutos de partida y que utilizan personajes que paran al enemigo. Sin embargo, sus desventajas respecto al resto de equipos recaen claramente en el número de súbditos asesinados por partida medio y en la puntuación de visión medio. Esto indica que los jugadores del equipo no consiguen mucho oro asesinando súbditos lo que les pone en desventaja de oro, y también que no controlan muy bien el mapa ni donde se encuentra el enemigo, porque no pueden verlos, debido a su poca visión.

### 2.2. ¿Para qué y quién puede servir la visualización?

Ya hemos visto que preguntas responde la visualización y como se puede interpretar cada gráfica para responder a las preguntas, pero no hemos indicado quién podría usar esta visualización, cual sería su uso medio ni para qué puede utilizarla.

Evidentemente, todo el mundo que esté interesado en el videojuego LeagueOfLegends y en su aspecto profesional puede utilizar la visualización, comprenderla y obtener conclusiones. Por una parte, un simple aficionado como el autor de esta visualización puede querer consultar la visualización como curiosidad para comprender el juego de los profesionales. Por otra parte, un periodista de deportes electrónicos también podría utilizar la visualización para estudiar los equipos y prepararse una entrevista con estos y preguntarles sobre sus personajes más escogidos o su perfil como equipo. También podría ser interesante la visualización para alguien interesado en las apuestas sobre partidos de League Of Legends profesional, ser capaz de comprender las diferencias entre equipos y sus fortalezas de un vistazo puede hacer que un usuario apueste a una apuesta que no lo haría normalmente, por ejemplo un equipo que suele dominar por diferencia de nivel en los primeros niveles, hacer una apuesta a que va a ir ganando en los primeros 15 minutos.

Independientemente de todos estos perfiles que pueden verse beneficiados por la visualización, hay que destacar que la visualización está desarrollada para que la utilicen usuarios que sean parte del cuerpo técnico de un equipo de League Of Legends profesional, ya sea para estudiar las capacidades de su equipo o para estudiar el perfil del equipo al que se tienen que enfrentar. Un ejemplo de uso medio, sería por ejemplo que al acudir al torneo de WORLDS, el mejor equipo de Europa (LEC) y el mejor torneo de NorteAmerica (LCS) se fueran a enfrentar al inicio. Por tanto, el cuerpo técnico del equipo de Europa, utilizaría la visualización para estudiar sus datos y los del rival. Haremos el ejemplo, por lo que supondremos que somos el equipo Esports de la LEC y queremos ver tanto nuestro rendimiento en LEC (antes de WORLDS) y el rendimiento de TSM en LCS (antes de WORLDS).

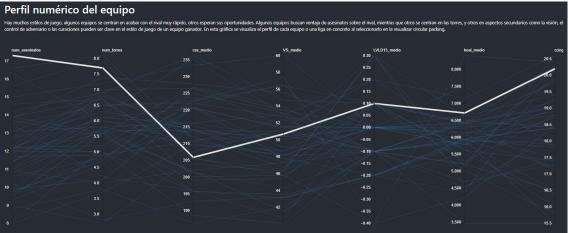
El usuario del cuerpo técnico tendría que seleccionar su equipo en el torneo de LEC.



A continuación, vemos como la mayoría de sus elecciones de campeones tienen un winratio positivo, con personajes con winratios incluso por encima del 70%, pero vemos también que un campeón de su top 10 más utilizados en la LEC es Aphelios con un winratio del 40%. El cuerpo técnico debería de plantearse escoger sus campeones más utilizados con mayor winratio, pero no seleccionar a Aphelios si quieren ganar la partida.

Las otras gráficas para G2 en LEC son:





En la de Winrate por lado, vemos que G2 ha tenido ventaja mayor a lo largo de los meses de verano en el lado rojo, por tanto, el cuerpo técnico debería de intentar seleccionar el lado rojo para la partida, y mejorar su estilo de juego desde el lado azul. En la gráfica de objetivos principales que indican la victoria, podemos ver que el equipo apenas consigue la primera torre, y es muy destacable, que de las 37 partidas que han conseguido el Barón, solo han perdido 2, por lo que es un objetivo que indica victoria para ellos, por tanto, deberían de centrarse en llegar vivos al minuto 20 que es cuando aparece este monstruo y tener posibilidades de conseguirlo.

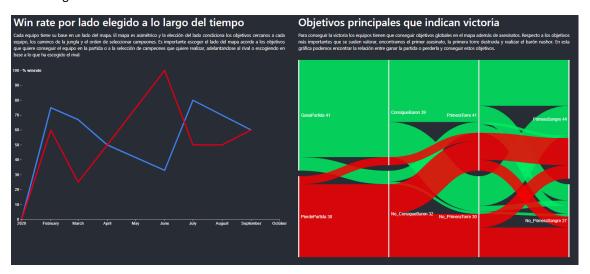
Por último, en su perfil numérico podemos ver que es un equipo que tiene como mayor fortaleza el num de asesinatos que consigue, siendo el mejor equipo en esto. Por tanto, el equipo debería centrarse en escoger campeones buenos al principio de la partida y pelear lo máximo al rival para ganarle a asesinatos.

Por otra parte, el mismo cuerpo técnico de G2 tendría que seleccionar a TSM en el círculo de la LCS, para estudiar a su rival en la partida.

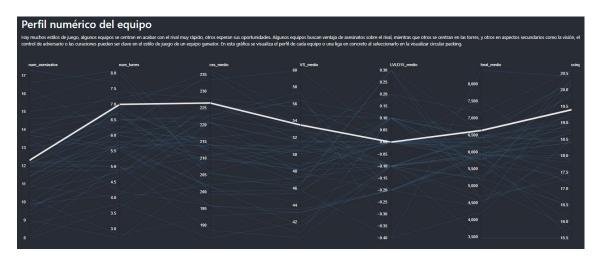


De primeras ya observamos respecto a los campeones, 3 aspectos que tendrá que tener en cuenta el cuerpo técnico de G2, lo primero esque ambos equipos utilizan a Sett como su mayor pick, y ambos tienen buenos winratios, por lo que será buena idea escoger al campeón antes que TSM. Por otra parte, el equipo de G2 deberá preparar algo contra Rakan o banear el campeón para la partida, por que es el segundo campeón más seleccionado de TSM y tienen casi un 90% de winrate con él. Por último tendrán que banear a Aphelios también, puesto que TSM lo juega bien con un buen winrate, y ya hemos visto que es el peor campeón de G2 en el top 10.

Las otras gráficas de TSM son:



Respecto al winrate a lo largo del tiempo, vemos que TSM ha ido mejor en ambos lados a lo largo del año, y por tanto lo lógico es que G2 Esports decida el lado en función de su propio rendimiento. Y para los objetivos principales, también vemos que TSM tiene mucha importancia sobre el barón nashor, porque igual que G2 casi no pierden partidas que consiguen el barón. Por tanto G2 tiene que centrarse en conseguir el objetivo por encima de su rival y tener una buena composición de equipo para el minuto 20.



Observando el perfil numérico de TSM, vemos que un equipo que destaca precisamente por el contrario que G2, TSM obtiene pocos asesinatos, pero son un buen equipo en los súbditos por minuto. Por tanto, el equipo de G2 tendrá que mantener su estilo de juego agresivo y conseguir muchos asesinatos para asegurarse que la partida se juega a su manera.

En resumen, la herramienta sirve como soporte a la decisión del cuerpo técnico de un equipo profesional de League Of Legends para tomar decisiones prepartido y conocer mejor su estilo de juego y el de sus rivales.

### 3. Descripción técnica del proyecto.

El proyecto se ha desarrollado a partir de los datos del dataset "Partidas profesionales de League of Legends correspondientes a de WORLDS, LEC y LCS 2020." que se encuentra en: <a href="https://zenodo.org/record/4265268#.YL5Dw\_kzZPZ">https://zenodo.org/record/4265268#.YL5Dw\_kzZPZ</a>. La descripción del dataset se ha realizado en la entrega anterior a esta, la A7 de la asignatura de Visualización de Datos.

Por otra parte, el código, el fichero de datos, los ficheros generados y los ficheros de la visualización se encuentran en: <a href="https://github.com/manurubo/manurubo">https://github.com/manurubo/manurubo a9 vd</a>.

A partir de estos datos, en Jupyter Notebook mediante Python 3.7.9, pandas y numpy se han procesado los datos hasta obtener un dataset para cada equipo para cada partida. De manera que el dataset ha pasado a tener casi la mitad de variables (algunas como el tiempo se duplican para los dos equipos) y el doble de filas, ya que para cada partida hay dos equipos. En definitiva, hay 1228 filas × 264 columnas.

A continuación, para cada visualización a realizar, se han tenido que generar los ficheros json que permitan generar las visualizaciones en d3 sin tener que procesar los datos en javascript mucho. Por tanto, se han generado los siguientes ficheros "campeones\_equipo.json", "objetivos\_gana.json", "parallel\_coords.csv", "partidas\_equipo.json" y "winrate.json". Estos ficheros se encuentran en el directorio DatasetsGenerados en el github de la entrega. Todos los datasets generados se obtienen de nuevo mediante pandas y numpy en Python. El código para realizar todo esto se llama "Lol\_ProfessionalGames\_DataPrep.ipynb". En todos estos ficheros se ha tenido que seguir el mismo procedimiento:

- 1. Determinar la visualización ha realizar.
- 2. Seleccionar las métricas que se quieren visualizar. Y juntarlas en un dataframe junto con las variables de torneo y equipo.
- 3. Agrupar para todas las partidas en general, para todas las partidas de cada torneo y para todas las partidas de cada equipo en cada torneo los datos.
- 4. Obtener el dato que se busca según la gráfica (el número de veces que ocurre un evento, la media de una métrica, etc) para cada agrupación.
- 5. Generar el fichero que se utilizará en D3 de acuerdo a la visualización ya generada en D3.

Una vez generados los datasets, se ha trabajado en el desarrollo de la visualización. La visualización se ha realizado en D3 v6, integrado en un fichero javascript que aporta el desarrollo e integración de las gráficas, junto con un fichero html que da la estructura de la visualización en formato web (con Bootstrap) y un fichero css para aportar estilo.

He escogido este desarrollo en D3 v6 porque esta librería de javascript permite desarrollar las visualizaciones casi al detalle del píxel, pudiendo personalizar las visualizaciones a desarrollar, los estilos de las visualizaciones y las interacciones entre ellas. Además, hay una gran cantidad de ejemplos de desarrollo de visualizaciones de D3 en la web observable.com que se pueden aprovechar para iniciar las visualizaciones y partir de una gráfica funcional, esto es lo que he hecho para las visualizaciones, primero las he encontrado en d3, posteriormente las he juntado en una función de javascript, he introducido los datasets generados a cada visualizaciones, he modificado su aspecto acorde a lo que me interesaba en la actividad y finalmente he añadido la interacción entre las gráficas.

Por último, he escogido esta librería D3 para hacer la práctica final frente a otras herramientas como Tableau, porque la libertad de desarrollo de visualizaciones no solo permite maximizar el estilo de la visualización, sino también la posibilidad de añadir complejidad a las visualizaciones casi sin límites, mientras que otras herramientas como Tableau, que son un software de una empresa se ven más limitas a visualizaciones simples para reportar datos. Tableau, podría haber permitido hacer la visualización de barras y la time-series. Mientras que el circle packing es muy difícil, puesto que está adaptado para visualizaciones de burbuja y para treemaps, pero no para algo que combine ambas. Por otra parte, hacer un parallel set o un parallel coords, necesitan de mucho toque de parámetros para realizar estas visualizaciones a partir de otras, lo que hace que realmente estas visualizaciones no sean tan efectivas ni reactivas, y que pierdan utilidad frente a una desarrollada en D3.

#### 4. Visualización de datos.

La visualización se encuentra disponible en: https://manurubo.github.io/manurubo a9 vd/manurubo final vd.html.

Ya se ha indicado los objetivos de esta visualización, que son identificar y visualizar las diferencias en como juegan las partidas de League Of Legends los equipos profesionales. Y así define el título de la visualización claramente. Además, hay una serie de anotaciones ante cada gráfica que ponen un poco en contexto lo que se visualiza en la gráfica y porqué es importante visualizar esto para un equipo de League Of Legends. También el layout de la página se ha diseñado para el usuario pueda interpretar las gráficas más complejas con más espacio y aquellas más simples ocupan menos porque el detalle se puede comprender igual.

Por otra parte, también se han explicado las preguntas, y respondido a ejemplos de estas.

También se ha indicado que la principal fuente de interacción en la visualización es que un usuario puede seleccionar un torneo o un equipo en la visualización circular packing, y el resto de gráficas se filtran automáticamente. Aunque hay otros elementos, como por ejemplo un slider para seleccionar el número campeones a mostrar en la visualización de barras, o en todas las gráficas es posible hacer un hover sobre los elementos y visualizar datos sobre la información que se presenta.

Gracias a haber desarrollado la visualización en D3, en una arquitectura HTML, CSS, JS, la actividad tiene una performance muy rápida, se puede ver incluso como se modifican las gráficas en tiempo real al hacer una selección.

Para el diseño del color es interesante mencionar, que se han utilizado el color rojo y azul en la gráfica de time series para el winratio por cada lado porque así es como se conocen a los lados en el juego (lado azul y lado rojo), de manera que el usuario puede interpretar rapidamente por el color a que lado pertenece la time-series. Además, para la visualización de campeones escogidos por el equipo, el color de cada barra representa el winratio del campeón en este equipo (que aparece también escrito encima de la barra), se ha escogido el rango de colores negro, rojo oscuro, rojo, verde, verde oscuro, dorado, para representar los win ratio de los campeones, de manera que no hace falta prestar atención directa al número de win

ratio y que el usuario puede interpretar rápidamente si un campeón muy escogido por un equipo tiene buen win ratio o no.

Por último, ya se ha indicado y visto que cada visualización tiene un título y texto explicativo de que se visualiza en ella para que un usuario no experto puede comprender lo que visualiza o porque es importante la visualización.

### 5. Entrega.

La entrega incluye un fichero zip con el dataset original, el fichero ipynb para procesar los datos y generar los datasets, los datasets generados, el fichero js con las visualizaciones en d3, el fichero css para dar estilo y el fichero html para dar la estructura a la entrega. Además se incluye este informe. De esta manera el usuario puede visualizar la práctica en su buscador copiando el path del fichero html. Además, la visualización se encuentra disponible en <a href="https://manurubo.github.io/manurubo">https://manurubo.github.io/manurubo</a> a9 vd/manurubo final vd.html.