

Módulo 2. Aprendizaje basado en un problema real (ABP) en el fenómeno de *Sports Analytics* (I)

2.1. Introducción

En la actualidad, aprender todo lo que abarca el análisis del deporte y la aplicación de la estadística en cursos introductorios como este no es una tarea fácil. Por ello, el objetivo de este módulo será conocer e introducir conceptos de *Sports analytics*, mediante un ejercicio de aprendizaje basado en problemas (ABP). El ABP es una metodología de aprendizaje que proviene de una universidad de Canadá, donde el alumno aprende conceptos de una forma más autónoma, trabajando un caso o un problema en concreto (Servant-Miklos, 2018). El ABP es concretamente un enfoque pedagógico que utiliza casos y problemas como punto de partida para adquirir los objetivos de aprendizaje deseados. El papel de un tutor en el aprendizaje basado en problemas es diferente del papel de un instructor de curso en un curso didáctico o basado en conferencias, que es más tradicional.

Como se ha comentado anteriormente en el módulo 1, muchos interesados en *Sports Analytics* tienen ganas de analizar datos, visualizar e interpretarlos, pero una de sus claves es tratar, primero, de conocer cuál es realmente el problema y, posteriormente, intentar realizar buenas preguntas. Para ello, en este módulo se presentará un caso ficticio donde, a partir de este, se incorporarán conceptos de *Sports Analytics*.

2.2. Caso: Pedro, candidato a analista del Barça B

Este caso es un texto que explica la prueba de selección a la que está expuesto Pedro, un chico que quiere trabajar como analista cuantitativo de fútbol para el equipo del F.C. Barcelona B. Ustedes no tienen que pensar que son Pedro, hay que leer, como primera medida, todo el caso y, posteriormente en cada párrafo, anotar todas las palabras, conceptos o frases que creen que no tienen del todo asimiladas, que estén relacionadas con el deporte y análisis de datos. De hecho, el objetivo inicial será intentar sacar las mejores preguntas de conceptos que estén relacionados con el deporte y la estadística en cada uno de los párrafos de este caso. A partir de estas “preguntas” se intentará, de forma autónoma, dar pistas para intentar contestar adecuadamente a ellas en todo lo relacionado con el fenómeno de *Sports Analytics*.

Texto del Caso: Pedro, candidato a analista del Barça B





Figura 1: Pedro

Me llamo Pedro, soy exfutbolista aficionado, exjugador de ajedrez y póker profesional, además de haber realizado mi carrera en Ciencias de Actividad Física y Deporte. Durante el último año he querido ampliar mis conocimientos de video análisis con un curso en línea introductorio de *data science*. Después de actualizar mi perfil en *LinkedIn*, *Researchgate*, y *Twitter*, el jefe de analistas de fútbol del FCB me llamó directamente para una entrevista laboral.

Es ya mi tercera entrevista y ahora mismo estoy en la Planta 3 de la Ciudad Deportiva del F.C. Barcelona, delante de 7 personas responsables (jefe del área, científico de datos, analista táctico, académico *sports scientist*, jefe de *scouting*, jefe área médica *Innovation Hub* y un periodista de datos) del departamento de *Sports Analytics* del área de ciencias del deporte.

Al comienzo de la entrevista, el responsable de esta área me pregunta mi opinión sobre los clubes de fútbol que empiezan a tener departamentos de *Data Science* y cómo creo yo que puede ayudar la estadística en ellos.

Minutos más tarde, el científico de datos me habla de la importancia de estar familiarizado con la investigación, y cómo esta se adapta a los fenómenos de *Big Data* y *Data Science*. También me explica que su grupo se reúne cada martes para comentar algunos artículos científicos de interés relacionados con el fútbol. Me comenta que, en el inicio de la liga, le ha sorprendido el número de tarjetas amarillas y, por este motivo, le gustaría saber cómo podrían influir en el fútbol. Rápidamente le comento que me suena un estudio que se publicó hace unos años en la revista *Journal of Quantitative Analysis in Sports* (Anders, A. y Rotthoff, K. W. (2011). "Yellow Cards: Do They Matter?". *Journal of Quantitative Analysis in Sports*, 7(1)).

El científico de datos tiene un poco de prisa y me presenta a uno de los analistas tácticos o de juego del FCB del primer equipo. Este me facilita una base de datos (validada por el equipo de IT) de los últimos 15 partidos de liga de todos los equipos (no han podido obtener los 10 primeros partidos), donde hay diferentes características de los equipos (ejemplo: Estadio, Home/Away), formación inicial (ejemplo: 4-3-3, 4-4-2), porcentaje de posesión, número de remates a la izquierda, número de remates al medio, número de remates a la derecha, tasa de conversión, número de corners, número de faltas, resultado final, para que pueda ver algunos patrones y describir algunos comportamientos. El analista agrega que no tiene una pregunta concreta para hacerme, pero hace tiempo que le gustaría saber si existe una asociación entre el hecho de marcar antes que el contrario y el resultado final del partido.

Más tarde me presentan a un analista con perfil más académico y, como prueba, me da una hoja con el siguiente *abstract* de un estudio científico (Mitrotasios, M. y Armatas, V. (2014). *Analysis of goal scoring patterns in the 2012 European Football Championship*. United States Sports Academy) . A partir de su lectura, me invita a que lo charlemos tomando un café.

ABSTRACT Purpose: The aim of the present study was to investigate the factors associated with goal scoring in top class football matches. **Methods:** 76 goals analyzed from 31 matches from the European football Championship in Poland-Ukraine in 2012. The Sport Scout performance analysis tool was used to time code when goal scored and the associated behaviors relating to goal scoring. **Results:** It was presented that 57.9 % of the goals were scored in the second half and most of them in the last 15-min period, while the first goal seems to play an important role in the final result of the match. Most goals were scored during open play (72.4%) but 27.6% scored after a set play, mainly following corner kick and free kick. Concerning initiation zone of the attack, it was presented that 56.6% began from the attacking half. Most goals were scored after a cross (43.7%) while regarding goal scoring area, over 90% of goals were scored inside the penalty area and specifically from the area between penalty spot and goal zone (42.1%). Finally, almost ½ of goals were scored without any pressure from defenders. **Conclusions:** The present study presented guidelines to coaches in order to design trainings similar to the actual competition and adapt the game style with regard to match status. Moreover, coaches can use this information to establish objectives for players and teams with a special reference to the offensive or defensive play.

Llevo ya unas 3 horas en la Planta 3 y me presentan al analista con perfil *scouting*, que parece preocupado porque, últimamente, los resultados del equipo no son los esperados, y no sabe cuál es el efecto del cambio de entrenador en equipos profesionales de fútbol. Además, también tiene interés en conocer el rendimiento de los porteros. Me comenta que quiere saber, por ejemplo, si el número de pases que realizan los guardametas fuera de casa puede diferir si juegan como locales. Y a partir de ese momento me muestra las siguientes tablas de unos estudios que se realizaron anteriormente para que los analice con ellos.

A hablar de rendimiento, el jefe médico de F.C. Barcelona *Innovation Hub* y el periodista de datos del club se suman a la entrevista y me muestran la siguiente foto sobre lesiones de la liga inglesa para que haga, brevemente, mi radiografía analítica final.

Figura 2: Lesiones de la liga inglesa



Fuente: [Imagen sin título sobre lesiones de la liga inglesa]. 2016. Recuperado de <http://sportsdiscovery.net/journal/2016/08/05/interview-head-of-sports-science-at-swansea-city-afc-jonny-northeast/>

Ya son las 19:00, y espero que les haya causado una buena impresión y me pueda integrar en este equipo de analistas.

2.3. Análisis del 1º párrafo del CASO

Después de leer detenidamente el primer párrafo, voy a poner en negrita las palabras que me puedan ayudar a hacer preguntas para consolidar algunos conceptos relacionados indirectamente con *Sports Analytics*. Como primer ejemplo, una de las primeras preguntas para formular sobre el primer párrafo sería:

“Me llamo Pedro, soy **exfutbolista** amateur, exjugador de **ajedrez y póker** profesional, además de haber realizado mi carrera de **INEFC**. Durante el último año he querido **ampliar mis conocimientos** de **video análisis** con un **curso en línea** introductorio de **data science**. Después de actualizar mi perfil en **Linkedin, Researchgate y Twitter**, el jefe de analistas de fútbol del FCB me llamó para una entrevista laboral”.

1. ¿Qué habilidades, conocimientos y formación pueden ayudar para ser mejor analista en el futuro?

¿Se había formulado una pregunta similar? Puede ser que no, pero probablemente ha pensado en preguntas similares que lo hacen cuestionarse cómo responder y adquirir más información sobre ello. Durante unos minutos puede pensar en formular las

siguientes posibles preguntas. Después de estos minutos, le presentamos algunas que se han pensado leyendo el primer párrafo. Esperemos que sean similares y esté interesado en contestarlas para poder profundizar más sobre el fenómeno de *Sports Analytics*.

1. **¿Qué habilidades, conocimientos y formación pueden ayudar para ser mejor analista en el futuro?**
2. **¿Qué es *Data Science*?**
3. **¿Qué diferencia hay entre *Data Science* y *Science*? ¿Y entre *Data Science* y *Statistics*? Aprenda nuevas vías de conocimiento de análisis.**
4. **¿Tener conocimientos y hacer un buen uso de las redes sociales puede ayudarme?**

Vamos a intentar responder a las preguntas.

1. **¿Qué habilidades, conocimientos y formación pueden ayudar para ser mejor analista en el futuro?**

No existe una habilidad, conocimiento, formación o receta concreta, pero sí varias experiencias y consejos que pueden ayudar. Antes que nada, tenemos que preguntarnos qué idea general hay de un profesional que se dedica a ser analista del deporte. La siguiente figura nos puede dar alguna pista.

Figura 3: La visión que tienen algunas personas de lo que representa un *performance analyst*



Fuente: [Imagen sin título sobre la visión de algunos de lo que representa *performance analyst*]. 2012. Recuperado de <http://thevolleyballanalyst.blogspot.com/2012/02/performance-analyst-perception.html>

La figura anterior nos hace pensar que, probablemente, falte más conocimiento sobre las actuales habilidades del analista de hoy. La evolución de este perfil ha sido importante, y hay que tener en cuenta varias de estas habilidades. Algunas de las características del analista de fútbol (también aplicable a otros deportes) serían:

1. Conocer el fútbol en profundidad (no solo aspectos tácticos o técnicos, sino también físicos, psicológicos, sociales).
2. Tener conocimientos de la plantilla y cuerpo técnico al que pertenece y competición en la que está involucrado el club.
3. Ser emprendedor y tecnológico. Es fundamental que sepa adaptarse a todas las opciones de software, plataformas y herramientas que hay en la actualidad.
4. Ser experto cuantitativo y cualitativo, capaz de dar sentido a la cantidad de estadísticas y números relacionados con el fútbol, para transmitir correctamente a la plantilla y al cuerpo técnico.
5. Estar alineado con el entrenador y el club en cuanto al modelo de juego.

Y estos profesionales ¿qué pueden analizar? Realmente pueden focalizar sus objetivos en el **rendimiento físico**, el rendimiento técnico (ejemplo: pases y zonas donde se realizan, tiros a puerta, cualquier evento técnico), el **rendimiento táctico** (ejemplo: posesiones realizadas, zonas y características; acciones a balón parado, a favor, en contra; relaciones entre jugadores), rendimiento psicológico, prevención de lesiones y mucho más.

El alumno que está realizando este curso probablemente se esté preguntando: "¿tengo que saber de todo?". Al principio el analista solamente hacía fotos o **capturas de pantalla de videos**. Esto implicaba muchas horas de trabajo pero, gracias a nuevos *softwares* disponibles de video análisis, se facilitó este trabajo (ej: Nacsport, Longo match, ERIC Sport, Sportcode). Las **nuevas tecnologías**, el boom de la información y la **profesionalización** del deporte en general han creado la necesidad de preguntarse cómo formar mejores analistas. Muchos no saben si soy un ¿video analista? o ¿foto analista? ¿Analista táctico o de rendimiento? ¿Analista de datos? Esto nos da una idea de la evolución de perfiles que puede haber en un departamento de análisis del deporte en un mismo club de fútbol.

A continuación, realizaremos una actividad para ver si tiene una idea de los diferentes perfiles de un analista. Si nos fijamos en las siguientes fotos, ¿qué perfil de entrenador sería afín, por ejemplo, al *Football Analytics*?

Figura 4. Uno de ellos es un entrenador afín al *Sports Analytics*

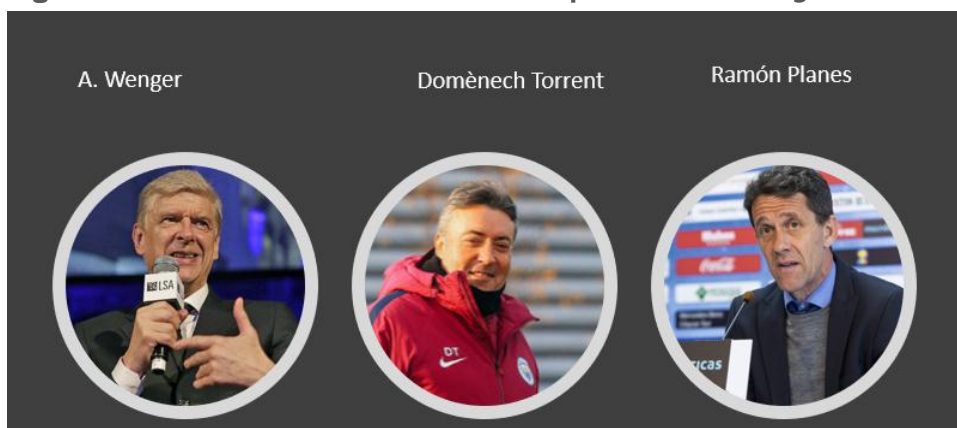


Fuente: elaboración propia.

La respuesta correcta sería Jürgen Klopp.

¿Y el analista con perfil de *scouting*?

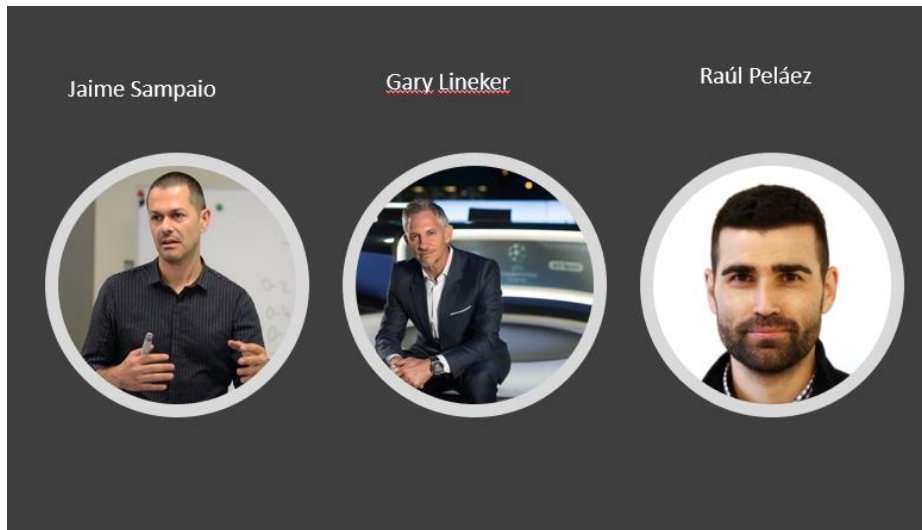
Figura 5: Uno de ellos es un analista con perfil de *scouting*



Fuente: elaboración propia.

La respuesta correcta sería Ramón Planes si nos hacemos esta pregunta en la temporada 2019/20. Si estamos interesados en conocer un perfil analista más académico y nos fijamos en la siguiente figura, podríamos decir que el perfil de ejemplo podría ser Jaime Sampaio.

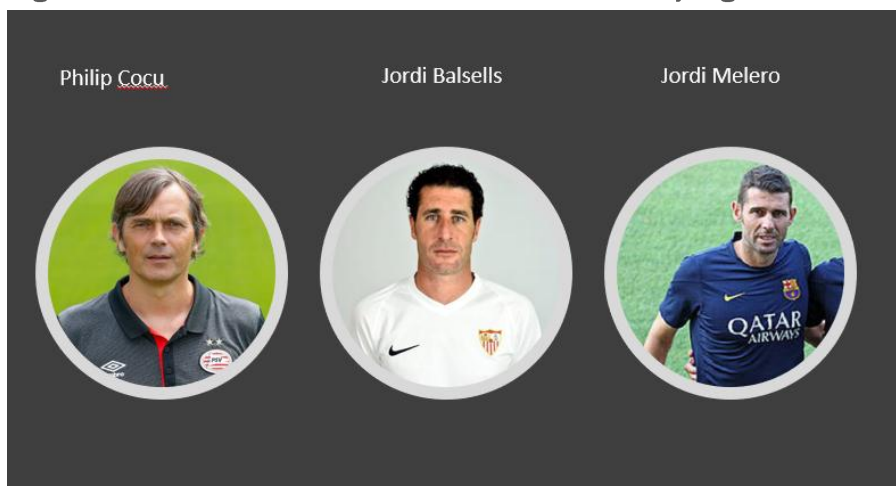
Figura 6: Uno de ellos es un analista académico



Fuente: elaboración propia.

Otra posibilidad sería tener en un club un perfil de analista más táctico o de juego.

Figura 7: Uno de ellos es un analista táctico o de juego

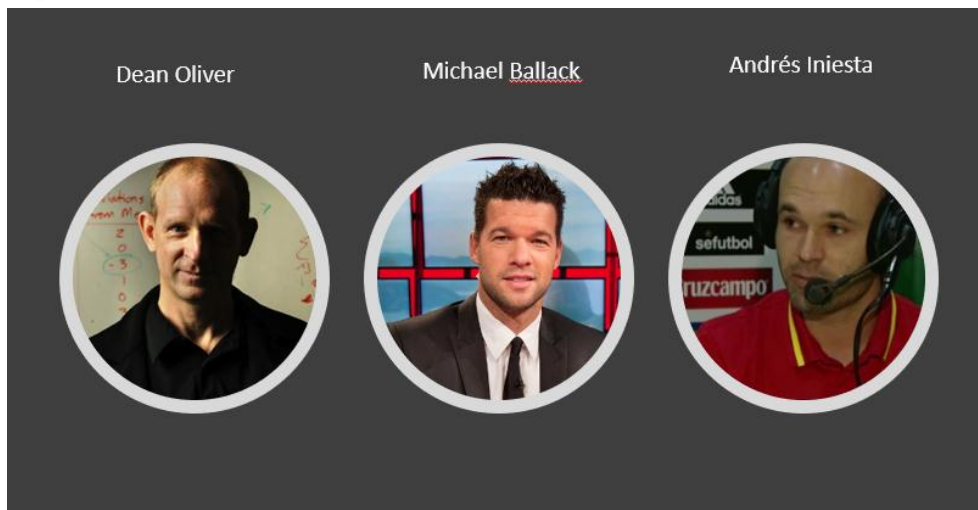


Fuente: elaboración propia.

La respuesta correcta sería Jordi Melero si nos hacemos esta pregunta en la temporada 2019/20.

¿Y podemos encontrar analistas comentaristas? Sí, uno de ellos, mundialmente conocido, es un exjugador de fútbol y actual comentarista especializado, que debutó en la ESPN, Michael Ballack.

Figura 8: Uno de ellos es comentarista analista de fútbol



Fuente: elaboración propia.

Por último, y no menos importante, nos podríamos preguntar si existen analistas de datos o estadístico con perfil científico. En este perfil no hemos puesto ninguna figura, ya que es un perfil con una gran evolución y recorrido, y que seguro aprenderemos pronto de ellos. Es relevante recordar que estos expertos con habilidades más concretas (ejemplo: estadísticos, matemáticos) tienen que estar siempre involucrados desde el inicio del estudio y que se puede acudir a ellos como expertos siempre que sea necesario.

En esta primera pregunta, hemos visto qué posibles habilidades puede tener un analista deportivo y qué especialidades puede haber en el mundo del fútbol, como ejemplo. Otro de los puntos que queremos resaltar para contestar nuestra pregunta inicial: “**¿Qué habilidades, conocimientos y formación pueden ayudar para ser mejor analista en el futuro?**” es respecto a la formación. En este sentido, muchas veces, cuando asistimos a cursos de formación como este, hay que tener en cuenta cuánta información aprendida se podrá retener y utilizar a partir de este momento. Por ello es interesante pensar en la pirámide de Edgar Dale, de 1969, la que básicamente informa cuál es la manera de recordar un poco pasiva y activamente.

Figura 9: La pirámide del aprendizaje
The Cone of Learning

sparkinsight.com



Fuente: [Imagen sin título sobre la pirámide del aprendizaje] s.f. Recuperado de <https://eingleses.com/la-piramide-del-aprendizaje/>

Esto es muy interesante porque, básicamente, hace pensar, por ejemplo, qué idea uno se hace sobre el fenómeno *Moneyball* después de ver la película o, por el contrario, leer su libro. También vemos la importancia de perfiles como los ponentes, profesores o expertos que tienen que explicar, en sus presentaciones, conceptos de interés, ya que sus posibilidades de aprendizaje pueden ser elevadas.

En resumen, y contestando a la primera pregunta, si quiere dedicarse al análisis deportivo se tiene que:

1. Conocer muy bien el **deporte**.
2. Tener un **mentor**.
3. **Formarse constantemente**.
4. Contar con grandes habilidades en **estadística, analytics y programación computacional**.
5. Conocer adecuadamente algún **software de análisis estadístico** (R, SAS, Stata, Python, Matlab), de **gestión de datos** (Python, SQL, GitHub), de **visualización** (Tableau, Power BI, Qlikview, Plotly).
6. **Tener las aptitudes de comunicación según Bucheit (2017)**.

Posteriormente abarcaremos las siguientes preguntas que están muy relacionadas.

2. ¿Qué es *Data Science*?

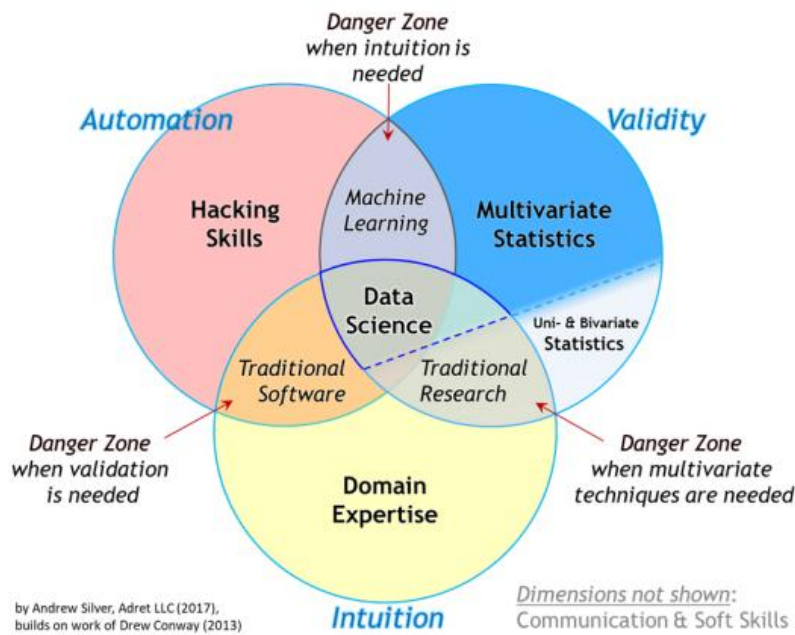
3. ¿Qué diferencia hay entre *Data Science* y *Science*? ¿Y entre *Data Science* y *Statistics*?
Aprende nuevas vías de conocimiento de análisis.

En el año 2012, en la revista *Harvard Business Review* ya se comentaba cuál sería el trabajo más sexy del siglo XXI: *Data Scientist*. Desde un comienzo, se ha pensado que se podía definir fácilmente *Data Science* (Varian, 2009; Davenport y Patil, 2012; Paarlberg, 2015) como:

- Data science is a **sexy job. The salaries are high**, the work is interesting, and there's significant prestige that comes with the title. As a result, many people want to be data scientists.
- A data scientist is someone who is **better at statistics than any software engineer** and **better at software engineering than any statistician**.
- The term data science has caused **excitement, confusion and controversy**. Some of the confusion is from the lack of a consistent definition. There is an **ecosystem of related terms (e.g. analytics, business intelligence, big data, data mining)**.
- The problem is that the **field has grown too fast**: there are now far too **many data scientists with very little experience** heading into a job market that has very few experts. Times have changed. When the term was still relatively new, companies were accepting candidates with only a basic knowledge of data, and getting them to learn on the job.

Las aproximaciones de *Data Science* descritas anteriormente no son correctas y es justamente el siguiente diagrama de Venn descrito por Drew Conway (2013) y Andrew Silver *et al.* (2017) que nos muestra qué engloba *Data Science* y también sus "peligros".

Figura 10: Diagrama de Venn de *Data Science*



Fuente: Silver, A. 2017.

Vemos que definir *Data Science* es más complejo de lo que parece, y las siguientes aproximaciones serían más reales:

- A data scientist is someone who **uses data to solve problems**.
- **A multidisciplinary field.** Data scientists come from a variety of academic backgrounds: computer science, physics, statistics, and many others. What matters most is **having a creative mind coupled with first rate critical thinking skills**.
- Intersection of **three areas: maths/statistics, computation, particular domain** (e.g: Sports) (Conway, 2010; Blei and Smyth, 2017).
- Intersection of **six areas called "Greater Data Science"**: 1- Data gathering, preparation and exploration; 2- Data representation and transformation; 3- Computing with data; 4- Data modeling; 5- Data visualization and presentation (Donoho, 2017).
- **I don't know.** Data science is an awesome profession, but there are definitely some serious frustrations that come with it.
- **"Data Science" could disappear, to be eclipsed by the next buzzword.**

Una de las principales diferencias entre *Science* y *Data Science* es que “**Science** starts with a **question**; **data science** starts most of times with the data” (Carmichael y Morrison, 2019; Donoho, 2017). *Data Science* es un proceso global y su profesión se está subespecializando en ingenieros *data scientist*, estadísticos o analistas *data scientist*, etc. Más información sobre este tema puede ser encontrado en la bibliografía de Carmichael, I. y Marron, J. S. (2018), Donoho, D. (2017), o incluso en Healy si nos remontamos al año 1978. Actualmente existe buena información sobre la visión de *statistical thinking* en *Data Science* (Ollé, Casals y Fernandez, s.f).

4. ¿Tener conocimientos y hacer un buen uso de las redes sociales puede ayudarme?

Según la bibliografía actual, es importante saber cómo hacer un buen uso de las redes sociales y compartir conocimiento, ya que se puede compartir conocimiento con diferentes grupos de personas afines a tu interés y crear sinergias e, incluso, obtener oportunidades de un nuevo trabajo.

Esta presentación vía *slideshare* (ver: <https://es.slideshare.net/xavierlasauca/science-dissemination-20-social-media-for-researchers-126297683>), de Xavier Lasauca, muestra sus verdaderas posibilidades como ejemplo.

2.4. Análisis del 2º párrafo del CASO

Vamos a leer detenidamente el siguiente párrafo del CASO. Ahora que ya sabemos cómo funciona esta metodología ABP, intentaremos formular algunas preguntas. Al cabo de unos minutos, puede ver, a continuación, algunas de ellas (no quiere decir que tengan que ser exactamente iguales), pero probablemente puedan ayudar.

“Es ya mi tercera entrevista y ahora mismo estoy en la Planta 3 de la Ciudad Deportiva del FCB delante de 7 personas responsables (jefe del área, científico de datos, analista táctico, académico *sports scientist*, jefe de *scouting*, jefe del área médica *FCB Innovation Hub* y periodista de datos) del departamento de *Sports Analytics* del área de **ciencias del deporte**. Al comienzo de la entrevista, el responsable de esta área me pregunta mi opinión sobre los clubes de fútbol que empiezan a tener **departamentos de *Data Science*** y cómo creo que puede ayudar la **estadística** en ellos”.

A partir del párrafo surgen las siguientes preguntas:

5. ¿Qué es *FCB Innovation Hub*?

6. ¿Qué es *Sports Analytics*? ¿*Soccer Analytics*?

7. ¿Qué es *Sports Sciences*? ¿Ciencia e investigación? ¿Método científico?

8. ¿Qué supone tener un departamento de *Data Science*?

9. ¿Qué es la estadística? ¿Y cómo puede ayudar?

5. ¿Qué es FCB Innovation Hub?

FCB *Innovation Hub* es un gran laboratorio deportivo para el futuro del deporte. Si te has inscrito en este curso, probablemente has visitado la plataforma web (<https://barcainnovationhub.com/es/>) y su blog y noticias que se van actualizando.

6. ¿Qué es Sports Analytics? ¿Soccer Analytics?

Creemos que, en el módulo anterior, ha quedado claro qué representa *Sports Analytics*. Si nos centramos en soccer o *football analytics*, nos damos cuenta de que cada deporte tiene su propia complejidad. En el año 2016, Masoud Nikravesch mencionó *Sports Analytics* como:

The processes that identify and acquire the knowledge and insight about potential players' performances based on the use of a variety of data sources such as game data and individual player performance data. These advanced and sophisticated type of analytics should be able to extract valuable actionable insights for the coaches and managers to utilize.

Y respecto a *Soccer Analytics* como:

It is the art of creating insights and actionable decisions using soccer related data. While predictive analytics uses big data to determine the probability or the likelihood of a certain outcome, intelligent descriptive analytics looks at big data and analyzes it using machine learning and artificial intelligence methods to come off with suggestions that will improve the likelihood of a desired outcome.

Libros como el de Ben Alamar (2013) nos acercan probablemente a estos dos fenómenos. De todas formas, si nos centramos en *football* o *soccer analytics*, en las últimas dos décadas ha habido un cambio en cómo se observa, graba y analiza el rendimiento en el deporte y, evidentemente, también en el fútbol. El proceso del video analista al analista cuantitativo está muy bien detallado por Rob Carroll en su página web: (<http://thevideoanalyst.com/contact/>).

Ahora estamos menos acostumbrados a oír hablar de análisis notacional, se habla más de *analytics*. En el año 2011, Ben Alamar comenta que *Sports Analyst* tiene tres componentes: *Data management, Predictive models and Information systems*. Jeremy Abramson, ya en 2014, define *Sports Analytics* como "the discovery and communication of meaningful patterns in data". Bill Gerard, en el año 2016, entiende *Sports Analytics* como "analysis of tactical data to support tactics-related sporting decisions". Y más adelante, Ben Alamar, en el año 2018, lo define como "Data and analysis are reducing the subjectivity of player valuation, strategic decisions, and player development". Otro

ilustre del *basketball analytics*, Dean Oliver contesta a Ben Alamar diciendo: "Reducing the subjectivity is a good phrase. You don't eliminate subjectivity (and don't really want to), but there are issues where analytics definitely limits errors subjectivity can make". Estas primeras definiciones y mensajes de estos autores nos acercan a pensar que, probablemente, es un fenómeno en continuo desarrollo. Bill Gerard nos da algunos consejos para tener en cuenta: (<https://winningwithanalytics.com/2016/12/05/bridging-the-gap-improving-the-coach-analyst-relationship-part-2/>).

- **Start simple** when first introducing data analytics as a **coaching** tool.
- Analytical results are usually presented most effectively to coaches by using **data visualization and story-telling**.
- Data analytics is **only one input into decision making by coaches**, albeit a potentially very important one if **used effectively**.
- Data analytics is **suffering from a fixation with big-data analytics**. Big-data, context-generic statistical analysis **must be translated into practical solutions to small-data**.
- **Sports analytics is most effective when the analyst understands the specific operational context of the coach**, produces relevant data analysis and translates that analysis into practical recommendations.

En resumen, las reflexiones actuales para el *soccer analyst* y según Brian Prestidge serían:

The analysts are getting smarter. Instead of relying on generic data — such as shots on target, possession statistics or heat maps— from companies such as Prozone and Opta, **they are creating their own metrics and building statistical models/algorithms to fit the club's philosophy and tactics.** It is this approach that could be described as a brains race.

7. ¿Qué es *Sports Sciences*? ¿Ciencia e investigación? ¿Método científico?

La **ciencia** puede definirse como "the process of observing, asking questions, and finding things, out about the world around you".

Para investigar, es importante tener la capacidad de hacerse preguntas y utilizar el **método científico**. No servirá, pues, cualquier conocimiento intuitivo o suposición no fundamentada. El método científico es la sucesión de pasos (observación, inducción, hipótesis; prueba o test de la hipótesis por experimentación; demostración o refutación

de la hipótesis; y tesis o teoría científica) que debemos dar para descubrir nuevos conocimientos y para comprobar hipótesis desconocidas hasta el momento. La investigación y las etapas del método científico están relacionadas con las principales secciones que encontramos a menudo en un artículo científico (ver Tabla) (Mabrouki, y Bosch, 2007).

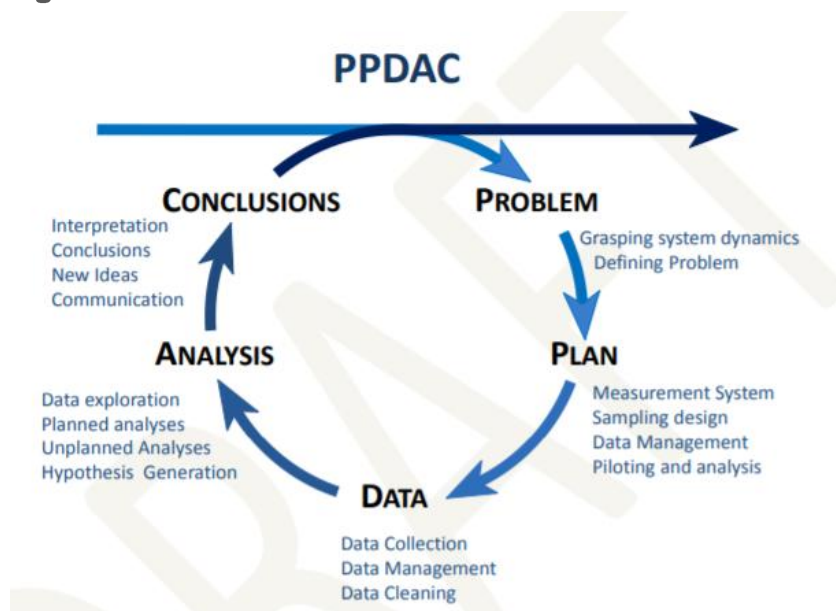
Tabla: Etapas del Método científico

ETAPAS EN EL MÉTODO CIENTÍFICO	PRINCIPALES SECCIONES DE UN ARTÍCULO CIENTÍFICO
Comprender el problema que hay que estudiar	Introducción
Establecer una hipótesis	
Recoger datos	Material y Métodos
Analizar datos	Resultados
Interpretar resultados	Discusión
Conclusiones	

Fuente: elaboración propia.

El método científico está ligado al PPDAC (*Problem-Plan-Data-Analysis-Conclusion*) que se mencionó en el módulo 1 y que hace referencia a la siguiente figura de David Spiegelhalter.

Figura 11: El ciclo PPDAC



Fuente: <https://www.pinterest.es/pin/391320655096691083/?lp=true>

Un aspecto clave para investigar es saber hacer buenas preguntas y, a partir de ahí, aplicar el método científico, conociendo bien los diferentes diseños, utilizando una correcta metodología y, sobre todo, siguiendo las guías de consenso (por ejemplo, en el

campo de la medicina: Strengthening the Reporting of Observational Studies in Epidemiology (STROBE), Consolidated Standards of Reporting Trials (CONSORT), Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), WHO Injury Surveillance Guidelines).

De todas formas, en la ciencia hay todavía margen de mejora, como ya nos indicaban John Ioannidis (2005) y Ben Goldacre (2010) en sus respectivos trabajos.

Definir *Sports Sciences*, y posteriormente *Sports Scientist*, no es fácil. Según Xavi Schelling, Director de Ciencias aplicadas al deporte (San Antonio Spurs; NBA):

An applied sports scientist is a multidiscipline integrator, who uses the scientific method to extract the signal from the noise. He/she will ask the right questions to the right people (experts), and will explain the new knowledge to the decision-makers in understandable terms.

Los diferentes roles de un staff y, en concreto, de un *sports scientist*, han ido evolucionando en los últimos años a raíz del propio desarrollo tecnológico. En uno de los artículos de Casals (2019), donde se menciona una analogía entre el baloncesto y un equipo multidisciplinar de investigación, se describen varios roles de un posible nuevo staff, entre ellos, el *sports scientist*.

8. ¿Qué supone tener un departamento de *Data Science*?

Hace unos años cuando visitábamos el post de la página web de *nbastuffer* titulado: "NBA Teams That Have Analytics Department", era difícil encontrar más de una o dos personas en cada franquicia de la NBA. Los departamentos de *Sports Analytics* o *Data Science* han ido creciendo recientemente, han incorporado una academia especializada (*statisticians, mathematicians, computer scientists*) en la industria del deporte. Franquicias de la NBA, como Raptors, Sixers, Spurs, entre otros, han ayudado a ello. Parece ser que en el fútbol, diferentes clubes europeos están, poco a poco, apostando fuerte. Una buena noticia es la incorporación de Michael López en la NFL como director de *Sports Analytics*, dado que es, posiblemente, uno de los mejores *sports analyst* académico hasta el momento que puede ayudar a mejorar la brecha entre la industria y la academia del análisis del deporte.

9. ¿Qué es la estadística? ¿Y cómo puede ayudar?

Ya le hacían esta pregunta a la Royal Statistical Society en el año 1838. ¿Tenemos una respuesta única? ¿Podemos hablar de una única disciplina o profesión?

Guadalupe Gómez, Professor Statistics and Operations Research (Universitat Politècnica de Catalunya-BarcelonaTECH), ya contestaba sobre algunos puntos de interés que gran parte de la sociedad desconoce; la profesión del estadístico existe hace como tal.

¿Qué hacen los estadísticos? Recoger datos, analizarlos, colaborar con científicos en todas las etapas de la investigación y desarrollar teorías y métodos.

¿Qué formas toma la estadística? Puede aceptar un modelo científico derivado de una teoría matemática o puede construir un modelo y tomar este por bueno y correcto.

¿Qué rol juegan los estadísticos? Rol activo/pasivo en la planificación, rol activo en el análisis o como personal de soporte.

¿Dónde trabajan los estadísticos? En el mundo del deporte, economía, salud pública, política, psicología, etc. No es extraño que sea difícil encontrar la estadística como una disciplina única y unificada.

Definir la estadística, una ciencia joven, no es tan sencillo. Si Karl Pearson, en los años 1857-1936 definía la estadística como la gramática de la ciencia, algunos, actualmente, ya la definen de forma más simple como la ciencia que utiliza datos para entender la incertidumbre sobre un proceso y tomar decisiones informadas haciendo uso de la teoría de la probabilidad. De todas formas, la American Statistical Association (ASA) la define como: "the science of learning from data, and of measuring, controlling and communicating uncertainty". La mayoría de las definiciones mencionan los "datos" y la "incertidumbre", aunque es difícil tener un único consenso al definir estadística.

Según la ASA, **los estadísticos**:

- Help companies make sense of the world around us by **analyzing** data.
- Use data to **solve** complex **problems**, in fields like **business**,
- **medicine, public service, sports...**
- and more.

¿Estás preparado para leer el próximo párrafo e intentar hacerte buenas preguntas para ir aprendiendo más contenido de este curso?

Tercer párrafo:

Minutos más tarde, el científico de datos me habla de la importancia de estar familiarizado con la investigación y cómo esta se adapta a los fenómenos de *Big Data* y *Data Science*. También me explica que su grupo se reúne cada martes para comentar algunos artículos científicos de interés relacionados con el fútbol. Me comenta que, en el inicio de la liga, le ha sorprendido el número de tarjetas amarillas y por este motivo le gustaría saber cómo podrían influir estas en el fútbol. Rápidamente le comento que me

suenan un estudio que se publicó hace unos años en la revista *Journal of Quantitative Analysis in Sports* (Anders, A. y Rotthoff, K. W. (2011). Yellow Cards: Do They Matter? *Journal of Quantitative Analysis in Sports*, 7(1)).

Referencias

Alamar, B. (2013). *Sports analytics: A guide for coaches, managers, and other decision makers*. Columbia University Press.

Anders, A. y Rotthoff, K. (2011). Yellow Cards: Do They Matter?. *Journal of Quantitative Analysis in Sports*, 7(1).

Bik, H. y Goldstein, M. (2013). An introduction to social media for scientists. *PLoS biology*, 11(4), e1001535.

Blei, D. y Smyth, P. (2017). Science and data science. *Proceedings of the National Academy of Sciences*, 114(33), 8689-8692.

Buchheit, M. (2017). Want to see my report, coach? Sport science reporting in the real world. *Aspetar Sports Medicine Journal*.

Carmichael, I. y Marron, J. (2018). Data science vs. statistics: two cultures?. *Japanese Journal of Statistics and Data Science*, 1(1), 117-138.

Carroll R. (<http://thevideoanalyst.com/contact/>).

Casals, M. y Nielsen, R. (2019). Who and what can contribute to improve the statistical thinking in sports injury research? A humorous analogy between basketball and members of the multidisciplinary research team. *Apunts Medicina de l' Esport (English Edition)*, 54(203), 81-84.

Conway, D. (2010). The data science Venn diagram. (2010). <http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>.

Conway, D. (2013). The data science Venn diagram.

<http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>

Davenport, T. y Patil, D. (2012). Harvard Business Review. Data Scientist: the Sexiest Job of the 21st Century. October 2012. <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/>

Donoho, D. (2017). 50 years of data science. *Journal of Computational and Graphical Statistics*, 26(4), 745-766.

Felt, M. (2016). Social media and the social sciences: How researchers employ Big Data analytics. *Big Data & Society*, 3(1), 2053951716645828.

Goldacre, B. (2010). *Bad science: Quacks, hacks, and big pharma flacks*. McClelland & Stewart.

Healy, M. (1978). Is statistics a science? *Journal of the Royal Statistical Society: Series A (General)*, 141(3), 385-393.

[Imagen sin título sobre la pirámide del aprendizaje]. s.f. Recuperado de <https://eingleses.com/la-piramide-del-aprendizaje/>

[Imagen sin título sobre la visión de algunos de lo que representa *performance analyst*]. 2012. Recuperado de

<http://thevolleyballanalyst.blogspot.com/2012/02/performance-analyst-perception.html>

[Imagen sin título sobre lesiones de la liga inglesa]. (2016). Recuperado de <http://sportsdiscovery.net/journal/2016/08/05/interview-head-of-sports-science-at-swansea-city-afc-jonny-northeast/>

Mabrouki, K. y Bosch, F. (2007). *Redacción científica en biomedicina. Lo que hay que saber*. Barcelona, ES: Prous Science.

Mitrotasios, M. y Armatas, V. (2014). *Analysis of goal scoring patterns in the 2012 European Football Championship*. United States Sports Academy.

Ollé, J., Casals, M. y Fernández, D. (s.f). La figura del científico de datos - mitos y revolución profesional. Recuperado de <https://investigacondatos.com/2019-dia2-charla05-cientifico-datos-marti-casals-daniel-fernandez/>

Paarlberg, B. (2015). The measurement standard. Why “Data Scientist” Is Being Called The Sexiest Job of The 21st Century. July 2015. Recuperado de: <http://www.themeasurementstandard.com/2015/07/why-data-scientist-is-being-called-the-sexiest-job-of-the-21st-century/>

Servant-Miklos, V. (2018). Fifty Years on: A Retrospective on the World's First Problem-based Learning Programme at McMaster University Medical School. *Health Professions Education*.

Silver, A. et al. (2017). The Essential Data Science Venn Diagram. Recuperado de: <https://towardsdatascience.com/the-essential-data-science-venn-diagram-35800c3bef40>

Van Noorden, R. (2014). Online collaboration: Scientists and the social network. *Nature news*, 512(7513), 126.

Varian, H. (2009). Flowing Data. Google's Chief Economist Hal Varian on Statistics and Data. Recuperado de <https://flowingdata.com/2009/02/25/googles-chief-economist-hal-varian-on-statistics-and-data/>

Voytek, B. (2017). Social Media, Open Science, and Data Science Are Inextricably Linked. *Neuron*, 96(6), 1219-1222.