

Propuesta de Proyecto Final: Una mirada analítica al deporte ecuestre

Gabriela Jiménez Conde

10/05/2025

Descripción del caso de negocio

Los números cuentan las historias detrás del deporte. En las competiciones, el análisis de datos permite entender el desempeño de los atletas, anticipar resultados y mejorar la toma de decisiones. Sin embargo, en la hípica nos encontramos con un problema: no existe la medición precisa y significativa del rendimiento de los binomios (jinete y caballo) en las pruebas.

La hípica no es un deporte lineal, sino que en cada competición influyen múltiples factores, desde el nivel de los binomios, la dificultad del recorrido, las condiciones del terreno, el estado físico del caballo y del jinete, entre otras cosas. Todo ello convierte a la competición ecuestre en un deporte singular, en el que incluso los mejores binomios tienen todas las probabilidades en su contra.

Esta complejidad convierte a la hípica en un deporte difícil de comprender para quienes no están familiarizados con ella. Además, la escasa utilización y accesibilidad a los datos dificultan la toma de decisiones informadas tanto para los usuarios dentro del deporte, como jinetes, entrenadores u organizadores de eventos, como para los usuarios fuera del mismo, los fans. Este proyecto nace con el objetivo de modificar esta realidad, haciendo uso de los datos como herramienta estratégica e informativa, y para fomentar una mejor comprensión del deporte.

Objetivos e impacto esperado

El objetivo principal del proyecto es desvelar la historia detrás de los resultados de las competiciones. A través del análisis de datos, se busca aportar valor a la competición hípica, mejorando la comprensión, la toma de decisiones y la comunicación en el entorno del deporte ecuestre.

Es un proyecto enfocado a diferentes públicos. Para los fans, se busca mejorar la comprensión del deporte desde fuera. Para los jinetes, se realizará un análisis histórico de su rendimiento, enfocado en proporcionar feedback personalizado que permita optimizar los entrenamientos y realizar un establecimiento de objetivos en base a resultados pasados. Los organizadores de eventos y marcas tendrán a su disposición una forma de contar el deporte que permita mejorar el storytelling, llevando a mejorar el engagement de la audiencia. Los análisis estarán centrados en las tres disciplinas olímpicas: salto, doma y completo.

Para ello se va a desarrollar un dashboard interactivo en el que se representarán insights generales de la competición ecuestre, y se ofrecerán explicaciones sobre los resultados pasados. Por otro lado, se generará un sistema de predicción de resultados, que permita anticipar el rendimiento de los binomios en futuras competiciones.

Tecnologías y herramientas

El proyecto está dividido en varias fases, para las cuales se utilizarán diferentes herramientas:

- **Proceso de ETL:** se utilizará Web Scraping, para la obtención de datos, realizada a través de BeautifulSoup4 y Selenium, así como SQL con Python para la carga de los datos a la base, usando herramientas como Psycopg2.
- **Proceso de EDA:** se utilizará SQL y herramientas de Python, como Pandas, Matplotlib o Seaborn, con el fin de obtener insights relevantes.
- **Sistema predictivo:** se utilizarán herramientas de predicción a través de Python.
- **Dashboard e informe final:** se utilizará Power BI y Streamlit para la creación de insights de forma visual y entendible.

Fuentes de datos

Las fuentes de datos utilizadas en el proyecto se indican a continuación:

- **FEI Database** (<https://data.fei.org>): es la fuente principal del proyecto, de la cual se obtendrá información sobre las competiciones, los jinetes y los caballos de forma detallada.
- **Página web de la FEI** (<https://www.fei.org/>): de esta fuente se obtiene información sobre regulación y sobre el funcionamiento general y el reglamento de las pruebas y disciplinas analizadas.
- **Equiratings** (<https://www.equiratings.com/>): de esta fuente se obtiene información sobre el modelo predictivo y sobre el funcionamiento del análisis de datos aplicado a la competición hípica.

Next Steps

Una vez alcanzados los objetivos principales del proyecto, se plantean a continuación una serie de líneas de trabajo para desarrollar a futuro:

- Información sobre ingresos/costes a través de los FEI Annual Reports de cada competición hípica, de manera que se pueda realizar un análisis de ingresos/costes de las competiciones.
- Realizar los análisis a otras disciplinas hípicas fuera de las olímpicas.
- Crear los dashboard y predicciones de manera específica para las federaciones, inicialmente para la FHE (Federación Hípica Española) o la FDHM (Federación Hípica de Madrid). En este caso se obtendrían los datos de las webs oficiales de las federaciones (<https://rfhe.com> o <https://www.fhdm.es>) o de revistas del mundo hípico (Oxersport, Avanza Eventos, Ecuestre).