

Informe de Análisis Predictivo de Ritmo de Carrera

Proyecto: Modelo Predictivo de Tiempo de Vuelta

Software Utilizado: JMP Standard 18.2.1

Analista: Javier Expósito

Fecha: 24/6/25

1. Resumen Ejecutivo

El objetivo de este análisis fue desarrollar un modelo predictivo robusto para cuantificar los factores que influyen en el tiempo de vuelta durante condiciones normales de carrera. Utilizando un riguroso proceso de validación manual, se construyó un modelo de Regresión Múltiple que logró explicar el **72.97%** de la variabilidad en los tiempos de vuelta en un conjunto de datos no visto previamente. El análisis reveló que, si bien el piloto y la velocidad instantánea son los predictores estadísticamente dominantes, el efecto de la degradación y el compuesto de los neumáticos es real y cuantificable, aunque su influencia se manifiesta a través de las variables de velocidad.

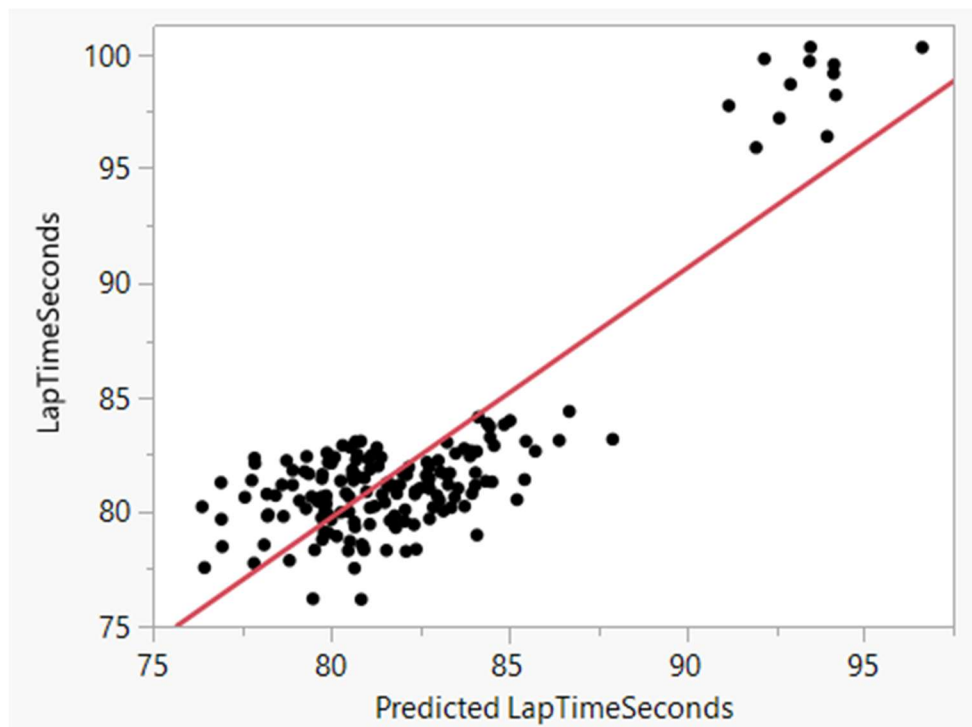
2. Metodología

- Preparación de Datos:** El conjunto de datos inicial se sometió a un proceso de limpieza para aislar el ritmo de carrera puro. Se excluyeron las vueltas de entrada y salida de pits, las vueltas bajo Safety Car, VSC o banderas amarillas, las vueltas eliminadas por límites de pista y la primera vuelta de la carrera.
- Ingeniería de Características:** Se creó una variable robusta, LapsOnStint, para medir de forma fiable el número de vueltas completadas en un stint, utilizando la función Col Rank de JMP.
- Validación del Modelo:** Para garantizar la robustez y evitar el sobreajuste, el conjunto de datos se dividió aleatoriamente en un conjunto de **entrenamiento (75%)** y un conjunto de **validación (25%)**.
- Modelado:** Se construyó un modelo de Regresión Múltiple utilizando la plataforma Fit Model de JMP, entrenado exclusivamente con el conjunto de datos de entrenamiento.
- Evaluación:** La fórmula predictiva del modelo se aplicó al conjunto de datos de validación para evaluar su rendimiento en datos nuevos, confirmando un RSquare de 0.7297.

3. Gráficos Clave a Incluir en el Informe

Gráfico 1: Validación del Modelo - Rendimiento Real vs. Predicho

- Gráfico de dispersión del conjunto de validación, mostrando el LapTimeSeconds real en el eje Y y el Predicted LapTimeSeconds en el eje X.
- La fuerte correlación lineal (puntos agrupados alrededor de la diagonal) y el RSquare de 0.7297 demuestran que el modelo es preciso y fiable.



"Figura 1: Correlación entre Tiempos de Vuelta Reales y Predichos en el Conjunto de Validación (RSq = 0.73)

Gráfico 2: Jerarquía de Influencia de los Factores (Effect Summary)

- Gráfico de barras que ordena los predictores del modelo según su significancia estadística (LogWorth).
- Comunica de forma instantánea cuáles son los factores más importantes. Permite ver de un vistazo que las velocidades y el piloto dominan, y abre la puerta a la discusión sobre la colinealidad.

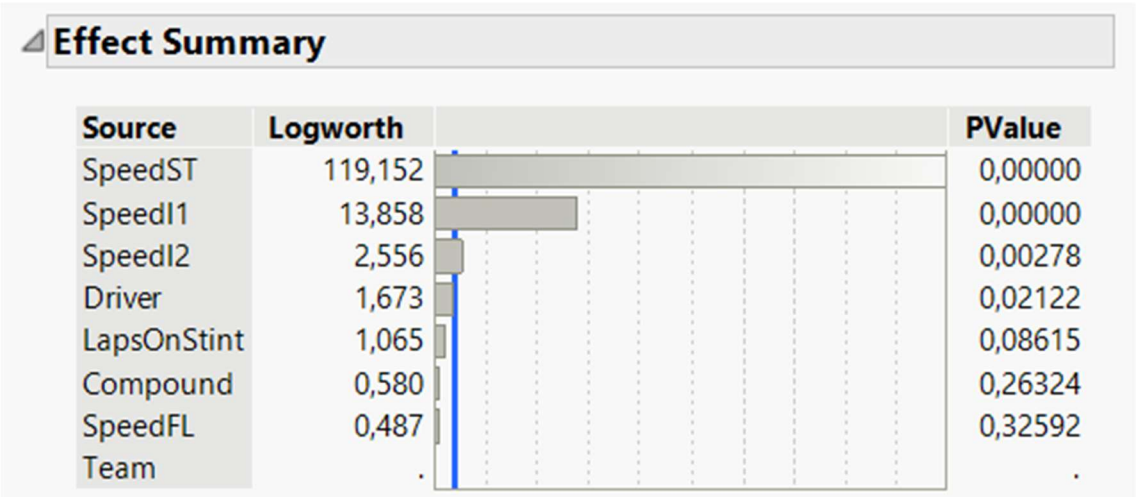


Figura 2: Resumen de Efectos del Modelo de Regresión, ordenado por Significancia Estadística (LogWorth)

Gráfico 3: Perfilador de Predicción - Aislamiento de Efectos Clave

- Una serie de capturas de pantalla del Prediction Profiler que demuestran el impacto individual de las variables clave.
- Es la herramienta más potente para comunicar los "insights". Aunque una variable no sea estadísticamente dominante en el modelo general, el profiler demuestra que su efecto es real y cuantificable.

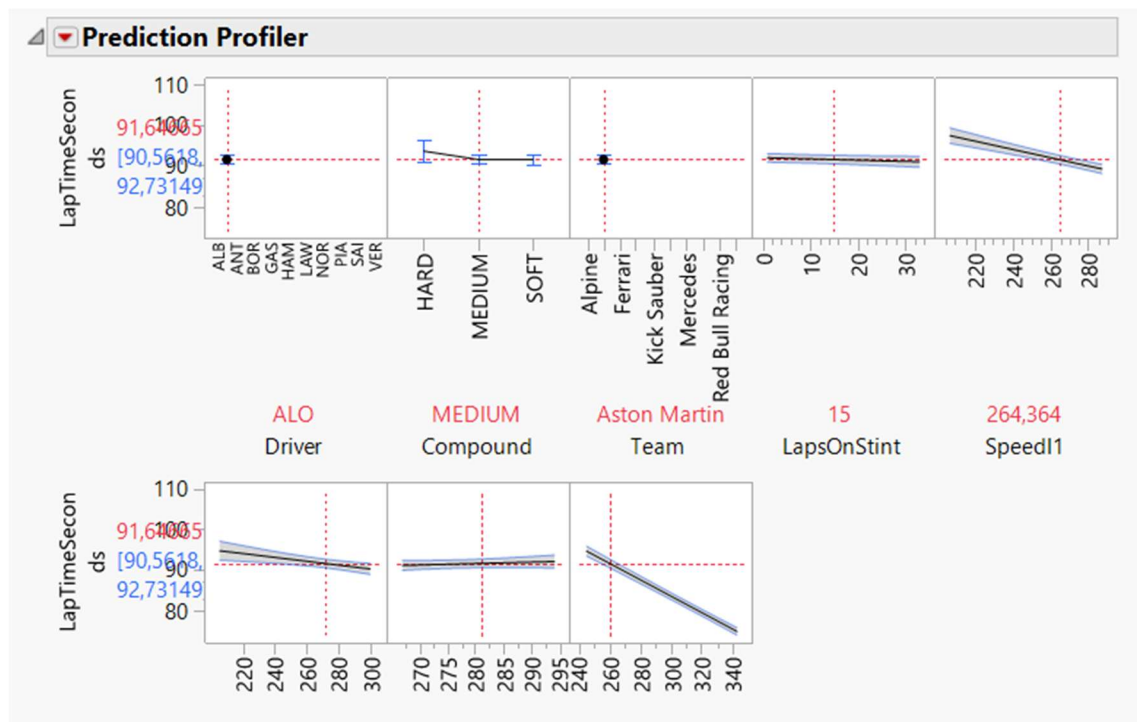


Figura 3: Perfiles de Predicción que Muestran el Impacto Aislado de la Degradación (a), el Compuesto (b) y el Piloto (c) en el Tiempo de Vuelta

4. Conclusiones Principales

1. El Talento del Piloto y la Velocidad en Pista son los Predictores Estadísticamente Dominantes.

El modelo de regresión múltiple, con un RSquare de validación del 73%, identifica que el piloto (Driver) y las velocidades instantáneas en puntos clave (SpeedST, SpeedI1, SpeedI2) son los factores con la mayor influencia estadística para predecir el tiempo de vuelta. Esto cuantifica que, incluso controlando por las capacidades del coche, el "factor piloto" sigue siendo una de las variables más decisivas en el rendimiento final, representando una diferencia de ritmo base cuantificable entre individuos.

2. El Efecto de la Degradación y el Compuesto del Neumático se Manifiesta a través de la Velocidad.

Aunque el compuesto (Compound) y la vida del neumático (LapsOnStint) no resultaron ser estadísticamente significativos en el modelo multivariado, el análisis a través del *Prediction Profiler* confirma su impacto real y cuantificable. El modelo revela que estas variables son causas raíz cuyo efecto es absorbido por las variables de velocidad, que actúan como un proxy del agarre disponible.

3. El Modelo Proporciona una Herramienta Cuantitativa para la Optimización Estratégica.

Este análisis trasciende la simple observación de promedios, proporcionando un modelo predictivo validado que puede ser utilizado como una herramienta estratégica. El *Prediction Profiler* permite simular escenarios "what-if" con una base de datos robusta. Por ejemplo, podemos cuantificar con precisión el delta de rendimiento esperado al cambiar de un compuesto a otro, de un MEDIUM a un SOFT en la primera vuelta o estimar el punto de inflexión en el que la degradación de un neumático hace que una parada en boxes sea la opción óptima. Esto transforma las decisiones estratégicas de una intuición a un cálculo basado en datos.