



Analisis de Datos

FORMULA 1

Trayectoria de pilotos en la historia de la Fórmula 1

CONTENIDO

Dataset Seleccionado

Inconsistencias

Descripción del dataset

Ajustes del Dataset

Lectura de datos (SQL)

Analisis grafico

Nulos y duplicados

Conclusiones & insights

Verificacion de outliers

Dashboard Power BI

DATASET SELECCIONADO

- **Formula 1 Drivers Dataset**
- **Historia de pilotos:** Nacionalidad, años activos y logros.
- **Logros clave:** Campeonatos, poles, victorias y podios.
- **Rendimiento:** Ratios y puntos promedio por carrera.
- **Análisis:** Comparación del éxito en la F1.



DESCRIPCIÓN DEL DATASET

Tamaño

```
print(f'Cantidad de registros: {len(f1_df)}')
print(f'Cantidad de columnas: {len(f1_df.columns)}')

Cantidad de registros: 868
Cantidad de columnas: 22
```

Principales características

- Driver: Nombre del piloto.
- Nationality: Nacionalidad.
- Seasons: Años activos en la F1.
- Championships: Campeonatos ganados.
- Race_Wins: Carreras ganadas.
- Podiums: Podios alcanzados.
- Points: Puntos totales acumulados.
- Win_Rate: Proporción de victorias.
- Points_Per_Entry: Puntos promedio por carrera.



LECTURA DE DATOS (SQL)

- 10 pilotos con mas carreras ganadas

Driver	Carreras ganadas
Lewis Hamilton	103
Michael Schumacher	91
Sebastian Vettel	53
Alain Prost	51
Ayrton Senna	41
Max Verstappen	38
Fernando Alonso	32
Nigel Mansell	31
Jackie Stewart	27
Jim Clark	25

Nationality	Total_Nationality
United Kingdom	164
United States	180
Italy	99
France	72
West Germany	39

LECTURA DE DATOS (SQL)

- 5 pilotos con mas campeonatos mundiales

Driver	Total_Championships
Michael Schumacher	7
Lewis Hamilton	7
Juan Manuel Fangio	5
Sebastian Vettel	4
Alain Prost	4

NULOS Y DUPLICADOS

● Verificación de nulos

Championship Years = 834 valores nulos

EXPLICACION

Si el piloto nunca gano un campeonato se toma como nulo

SOLUCION

Reemplazar nulos por = "No Championship"

● Verificación de registros duplicados

NO SE ENCONTRARON REGISTROS DUPLICADOS



VERIFICACIÓN DE OUTLIERS

Analisis estadistico del dataset

Llama la atencion ver que el 75% de los pilotos ganan 0.38 pts por carreara. Y el que mas gana tiene un promedio de 14.19. Veamos quien es ese piloto...

```
# Muestra el piloto con mas puntos promedio por carrera
highest_points_per_entry = f1_df.loc[f1_df['Points_Per_Entry'].idxmax()]
highest_points_per_entry
```

Lewis Hamilton

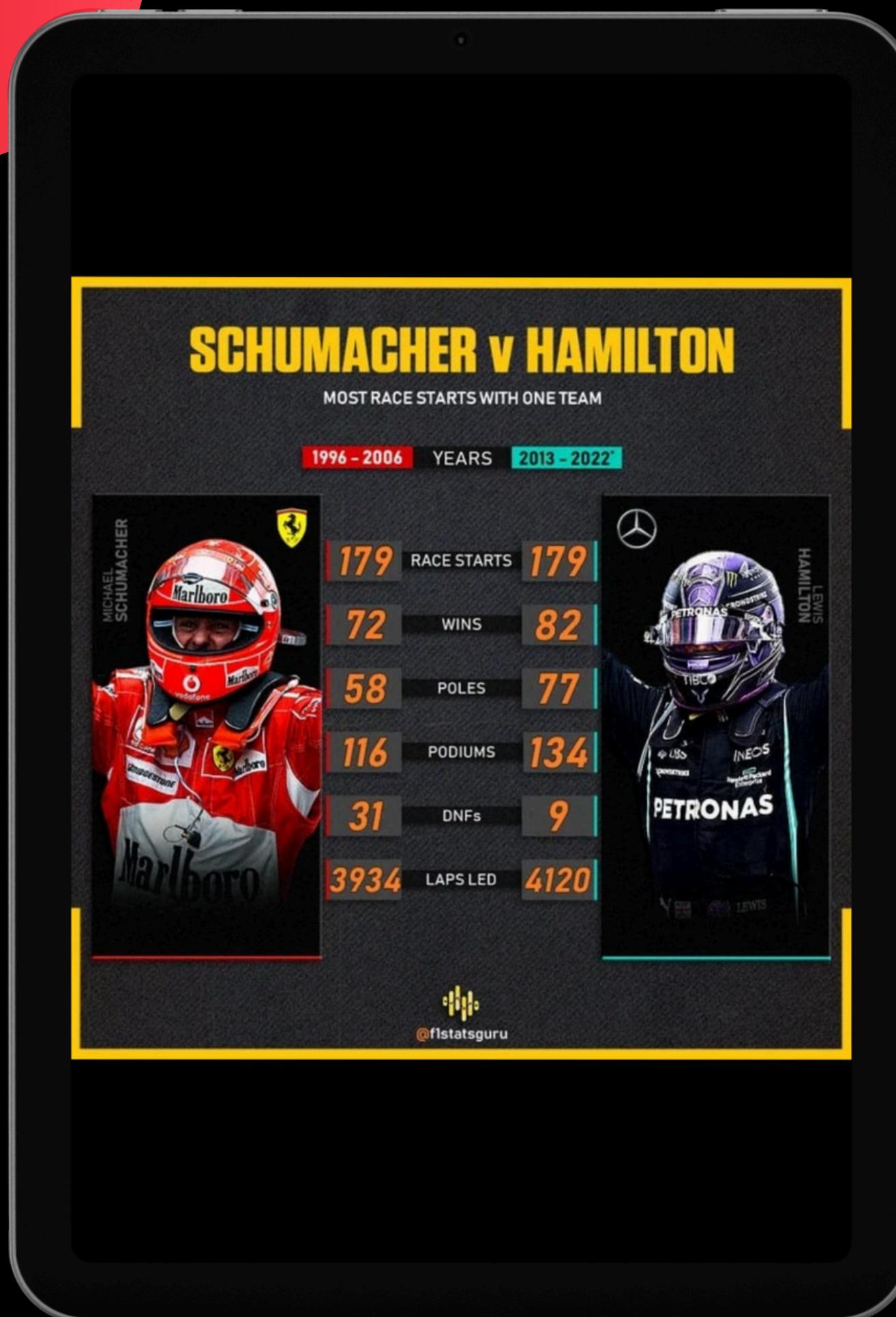
El piloto con mas ptos promedio por carreara es Lewis Hamilton. Se podria determinar que este valor es un outlier con respecto al resto del dataset. Esto es debido a que Lewis Hamilton = 



VERIFICACIÓN DE OUTLIERS

Analisis estadistico del dataset

- Hay 72 registros que tienen una cantidad de años activos mayor al percentil 90
- Hay 9 registros que tienen una cantidad de carreras ganadas mayor al percentil 99
- Hay 44 registros que tienen una cantidad de inscripciones en carreras mayor al percentil 95





INCONSISTENCIAS

Se validarán valores negativos de:

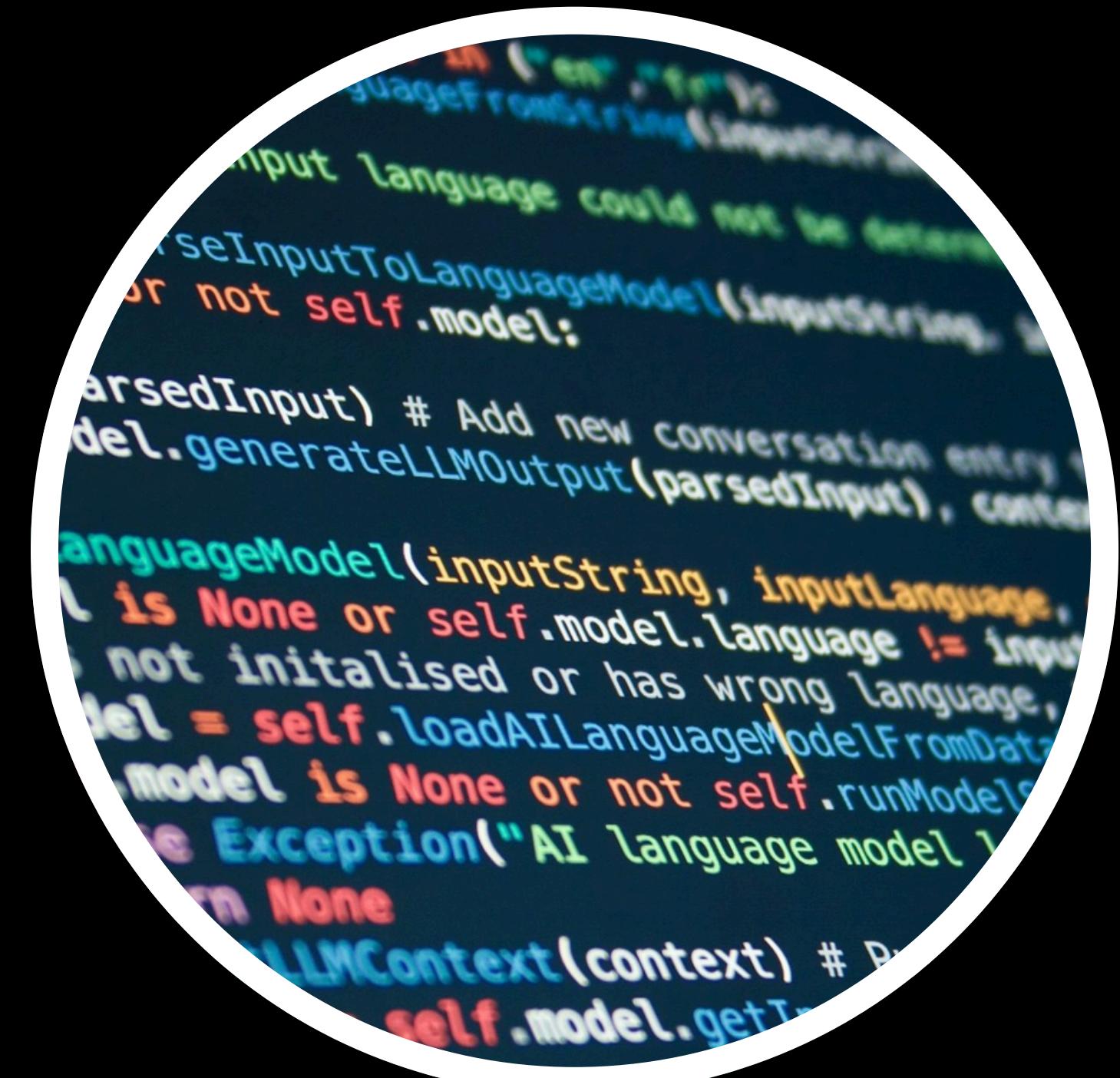
1. Race_Wins
2. Race_Entries
3. Podiums
4. Years_Active

No se encontró ninguna inconsistencia en los valores analizados dentro del Dataset



AJUSTES DEL DATASET

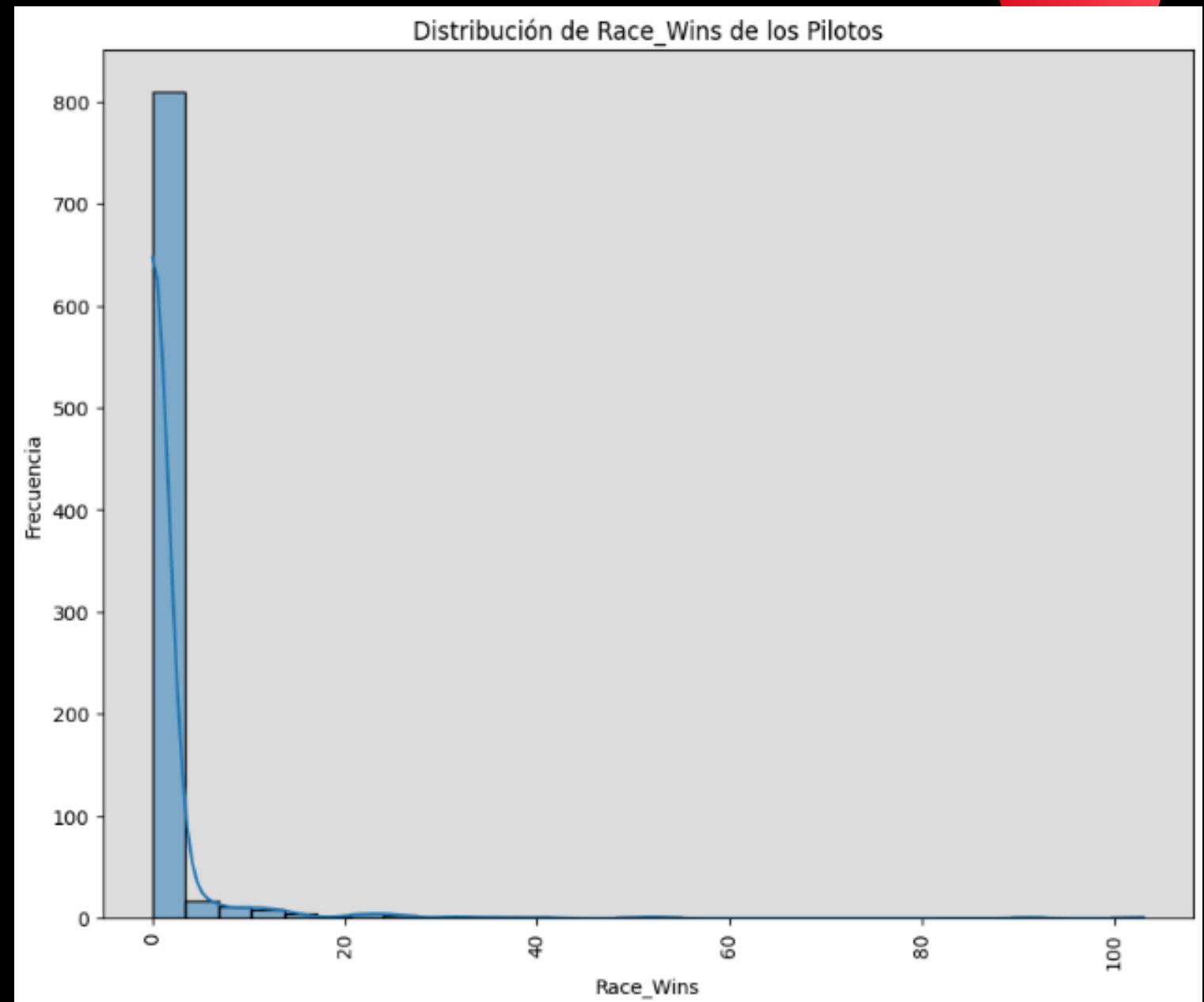
- Se completan los valores nulos de **Championship Years** con "No Championship"
- Se elimina la columna "Champion" ya que tiene la misma informacion que la columna "Championship Years"
- Se elimina la columna "Race_Entries" ya que tiene casi la informacion que la columna "Race_Starts"



ANALISIS GRAFICO

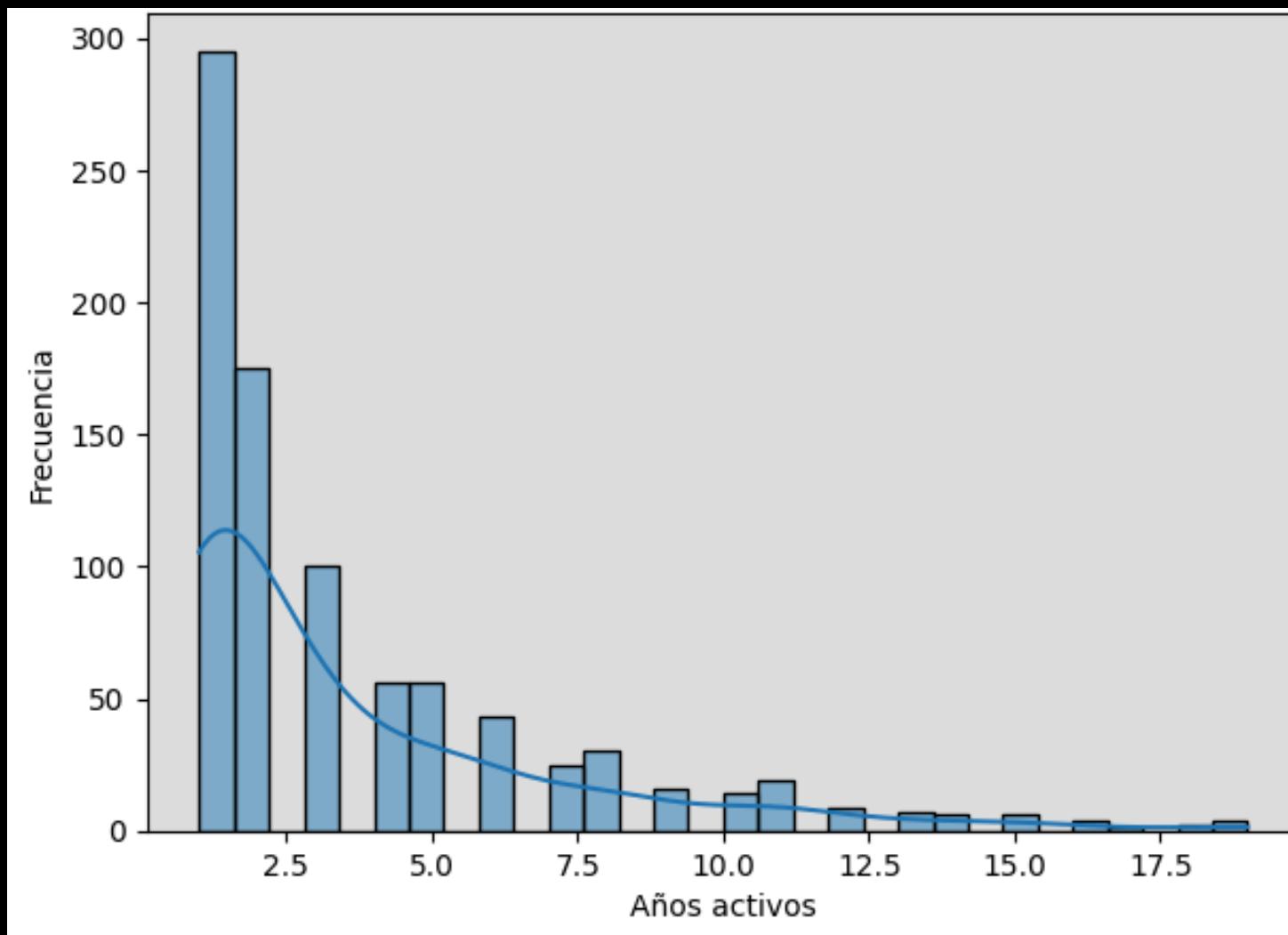


● Matplotlib/Seaborn

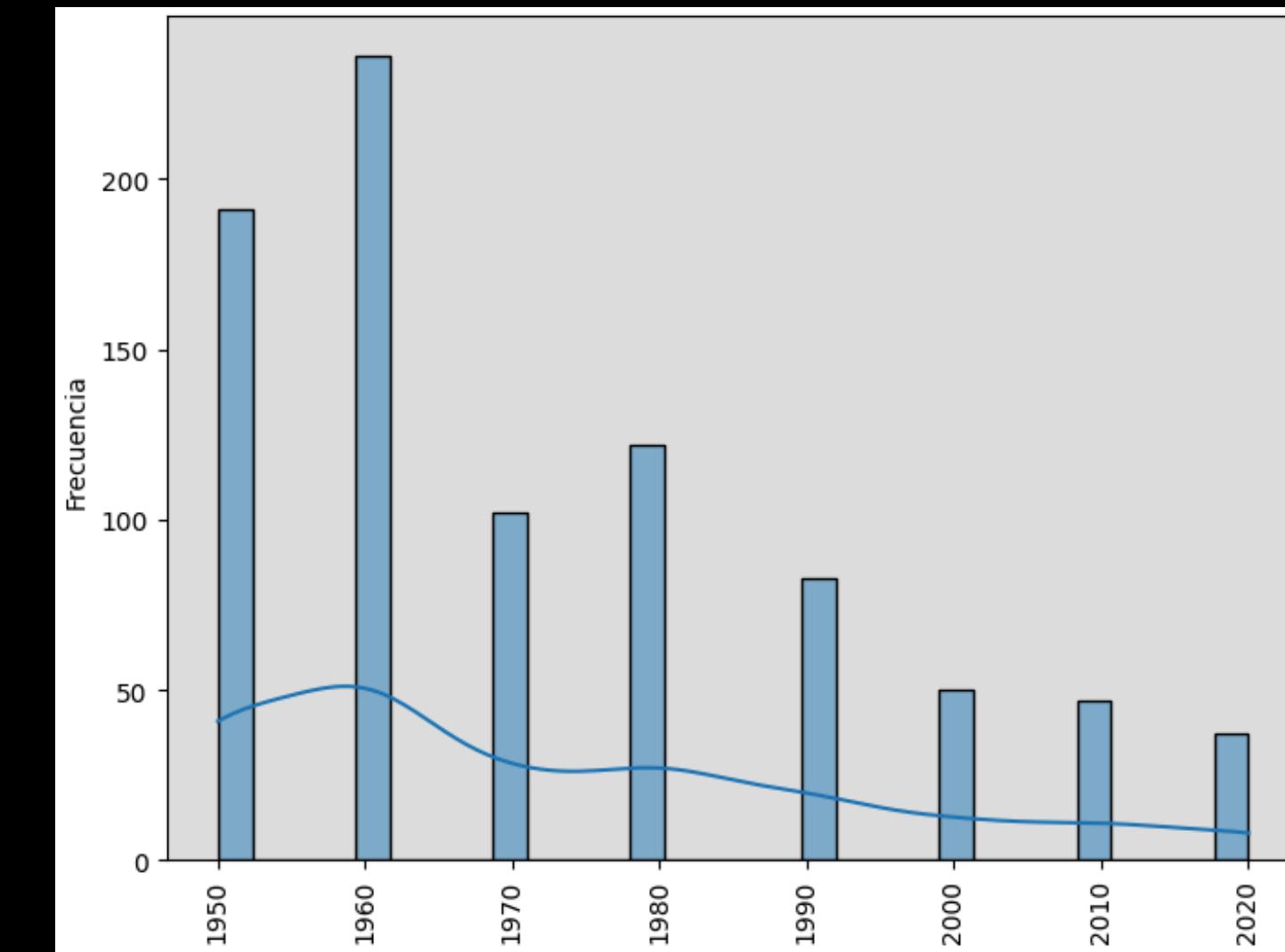


La GRAN parte de los pilotos que corren un carrera de Formula 1 NUNCA ganan una carreara en toda su vida.

Años activos de pilotos



Décadas de los pilotos



La carrera de un piloto de F1 suele ser muy corta, dura por lo general 1 / 2 años.

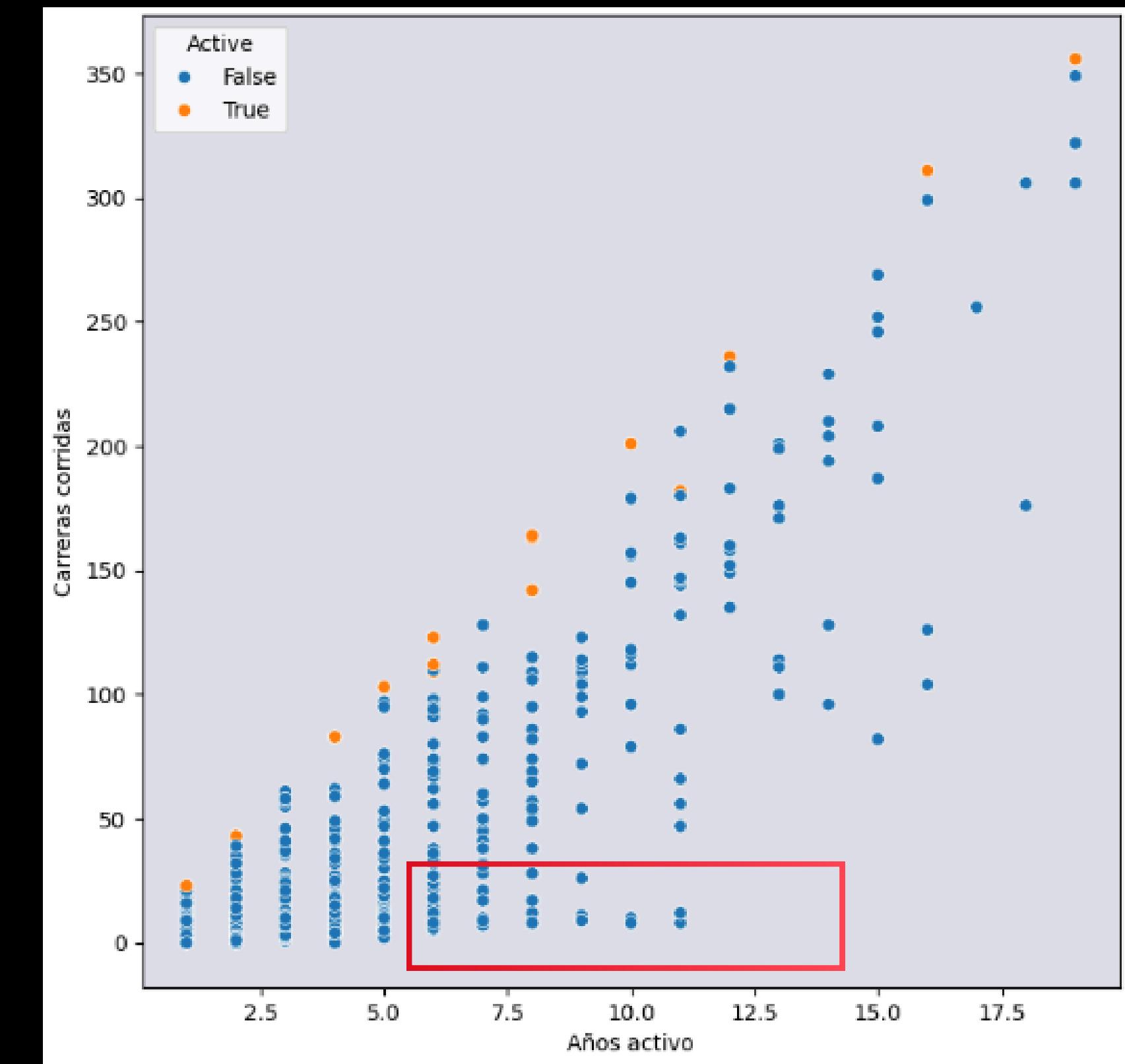
En 1960 la rotación de pilotos era cerca de 200 pilotos al año, un número muy elevado comparado con el año 2020 que solo es un 10% del mismo.

Años activos de pilotos VS Carreiras corridas

Pilotos activos en Naranja (True)

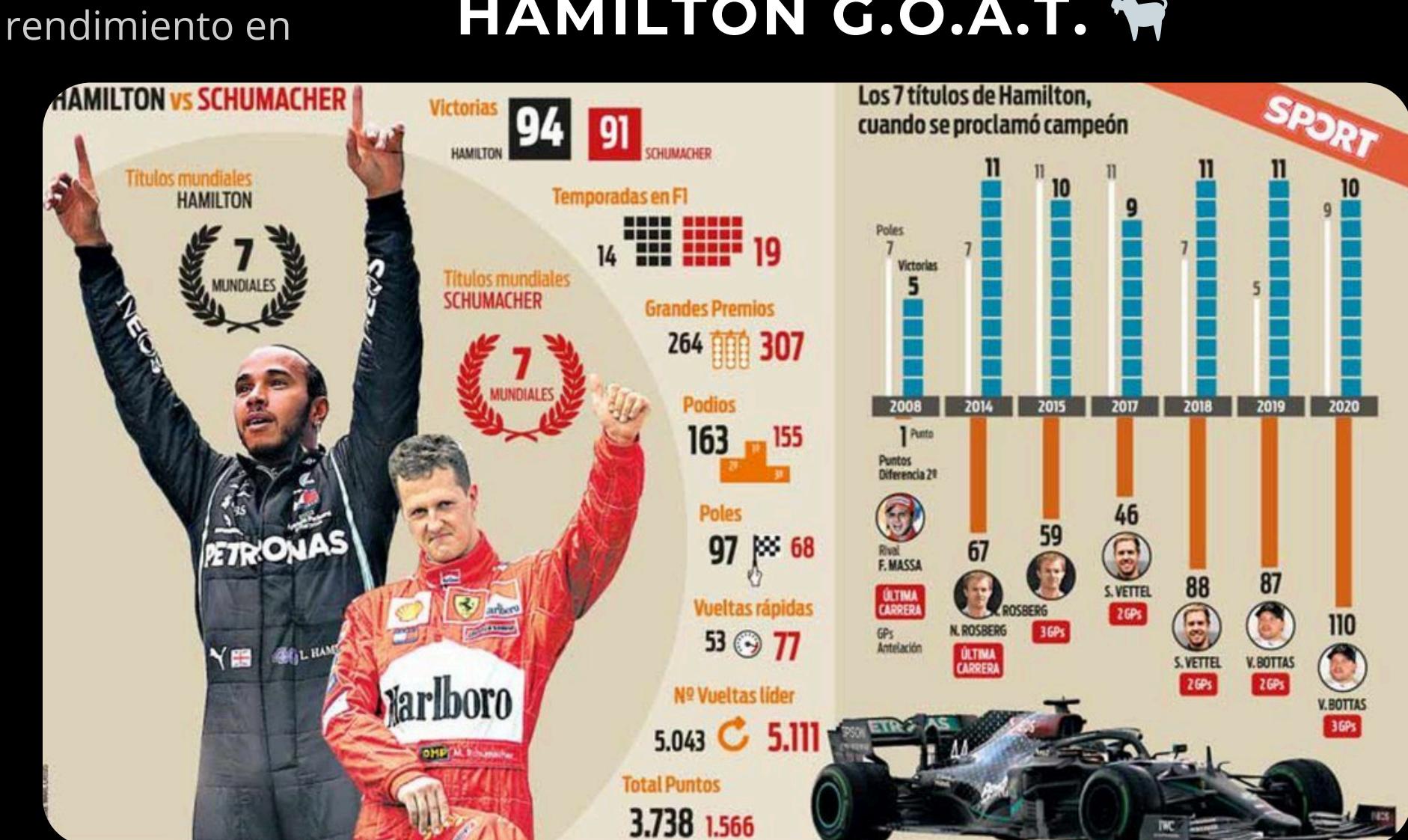
Se puede ver como hay una tendencia lineal ascendente entre los años activos y las carreiras corridas

De todos modos, hay valores fuera de lo normal que no siguen esta tendencia. Tienen mas de 10 años corriendo y muy pocas carreiras



CONCLUSIÓN & INSIGHTS

- **El mejor piloto:** Lewis Hamilton lidera como el mejor piloto, seguido por Michael Schumacher. El tercer lugar es debatido entre Fangio, Prost, Senna y Vettel.
- **Relaciones clave:** Pole positions, victorias y podios están altamente relacionados, destacando la importancia de un buen rendimiento en clasificación.
- **Exclusividad de la F1:** Muy pocos pilotos logran competir en F1, y menos aún destacan con victorias o títulos. Mantenerse relevante es extremadamente difícil.
- **Conclusión:** La Fórmula 1 es el pináculo del automovilismo, exigiendo talento, consistencia y resiliencia excepcionales.



DASHBOARD POWER-BI



<http://surl.li/dqlmkw>

CONSULTAS?

INTEGRANTES

Bianco, Matias

E-MAIL

mbianco922@gmail.com

Cerdan, Joaquín

[joaquincedan232@gmail.com](mailto:joaquincerdan232@gmail.com)

Ferrer, Joaquín

joacoferre00@gmail.com



MUCHAS GRACIAS

2024 Data Analysis Presentation