

Módulo 3. Análisis estadístico en Google Sheets y visualización con Looker Studio

Introducción

En el módulo 2, aprendiste a aplicar conceptos estadísticos básicos y crear visualizaciones usando **Microsoft Excel** y **Microsoft Power BI**. Ahora veremos cómo replicar esas mismas ideas en el entorno de Google, concretamente con **Google Sheets** y **Looker Studio** (anteriormente conocido como Google Data Studio).

Nuestro objetivo es que puedas **analizar datos** (por ejemplo, de entrenamientos deportivos) usando Google Sheets, aplicando medidas de tendencia central, dispersión, intervalos de confianza, cambio mínimo detectable (CMD) y promedios móviles (*rolling average*). Después, en Looker Studio, aprenderás a crear paneles (*dashboards*) dinámicos que permitan visualizar de manera efectiva dichos análisis y extraer conclusiones sólidas.

Al igual que en el módulo 2, trabajaremos con una base de datos de entrenamiento deportivo que contiene duración de sesión, distancia recorrida, esfuerzos de alta velocidad, frecuencia cardíaca, etc. Esto es muy relevante para quienes planifican entrenamientos, quieren prevenir lesiones y buscan optimizar la carga de trabajo en deportistas.

El módulo se dividirá en tres secciones:

1. Sección 1: Estadística aplicada en Google Sheets

- Conceptos básicos y funciones de Google Sheets
- Cálculo de medidas de centralidad y dispersión
- Intervalos de confianza (IC) y cambio mínimo detectable (CMD)
- Uso de *rolling average* y ACWR (agudo/crónico)
- Ejercicios prácticos con datos reales de entrenamiento

2. Sección 2: Visualización interactiva con Looker Studio

- Introducción y diferencias con Power BI
- Conexión a Google Sheets y configuración de fuentes de datos
- Creación de paneles interactivos con filtros



- Ejercicios prácticos para mostrar tendencias y comparaciones

3. Sección 3: Reflexión y conclusiones

- Comparación de Microsoft vs. Google
- Ventajas de la nube y la colaboración en tiempo real
- Recomendaciones finales y siguientes pasos

Unidad 3.1 Estadística aplicada en Google Sheets

1. Introducción a Google Sheets

Google Sheets es la hoja de cálculo de Google, que funciona 100 % en la nube. Algunas ventajas son:

- **Colaboración en tiempo real:** varios usuarios pueden editar al mismo tiempo, con actualizaciones instantáneas.
- **Accesibilidad:** solo necesitas tu cuenta de Google y conexión a internet para trabajar desde cualquier dispositivo.
- **Integración con otros servicios de Google:** Drive, Forms, Looker Studio, entre otros.
- **Compatibilidad:** muchas de las funciones de Excel también existen en Google Sheets (aunque el nombre puede variar).

Si vienes de Excel, la lógica de **celdas**, **filas**, **columnas**, **fórmulas** y **referencias** se mantiene. Solo cambia la interfaz y algunos detalles.

2. Conceptos básicos de Google Sheets

- **Celdas:** identificadas por letra de columna y número de fila (A1, B2, C15, etc.).
- **Referencias absolutas y relativas:**
 - Relativas: A1 (al copiar/pegar, se ajusta).
 - Absolutas: \$A\$1 (permanece fija al mover la fórmula).
- **Fórmulas:** empiezan con "=" y se pueden arrastrar o copiar.
- **Formato de datos:** es importante asegurarte de que los números estén bien reconocidos como números y las fechas como fechas, para evitar errores en los cálculos.
- **Historial de versiones:** Google Sheets guarda automáticamente las versiones, lo que esto permite retroceder cambios en caso de error.

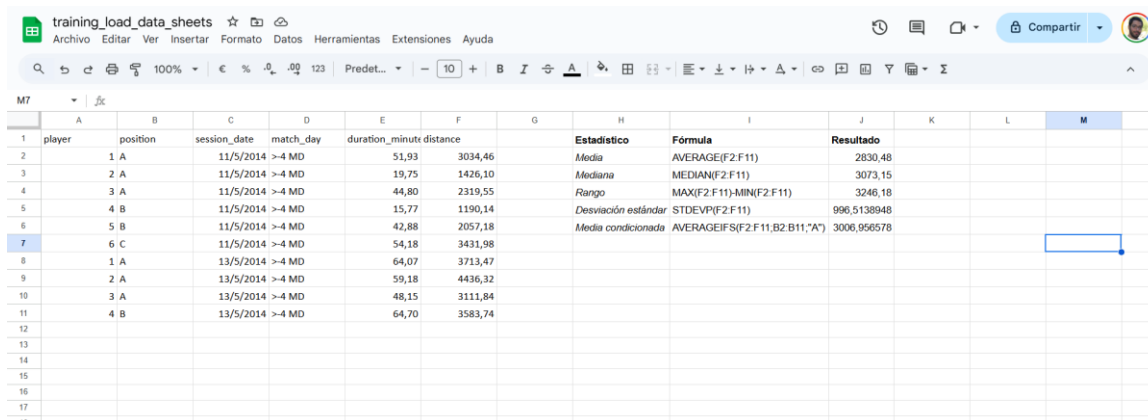


3. Medidas de centralidad y dispersión

Para empezar a analizar datos, solemos calcular:

- **Media** =AVERAGE(rango_datos)
- **Mediana** =MEDIAN(rango_datos)
- **Rango** (máximo - mínimo) =MAX(rango_datos)-MIN(rango_datos)
- **Desviación estándar** =STDEVP(rango_datos)
- **Promedio condicional** =AVERAGEIFS(rango_promedio; rango_criterio1; criterio1; [rango_criterio2]; [criterio2];...)

Figura 1: Medidas de centralidad y dispersión



1	player	position	session_date	match_day	duration_minut	distance	Estadístico	Fórmula	Resultado
2	1 A		11/5/2014	>-4 MD	51,93	3034,46	Media	AVERAGE(F2:F11)	2830,48
3	2 A		11/5/2014	>-4 MD	19,75	1426,10	Mediana	MEDIAN(F2:F11)	3073,15
4	3 A		11/5/2014	>-4 MD	44,80	2319,55	Rango	MAX(F2:F11)-MIN(F2:F11)	3246,18
5	4 B		11/5/2014	>-4 MD	15,77	1190,14	Desviación estándar	STDEVP(F2:F11)	996,5138948
6	5 B		11/5/2014	>-4 MD	42,88	2057,18	Media condicionada	AVERAGEIFS(F2:F11,B2:B11,"A")	3006,956578
7	6 C		11/5/2014	>-4 MD	54,18	3431,98			
8	1 A		13/5/2014	>-4 MD	64,07	3713,47			
9	2 A		13/5/2014	>-4 MD	59,18	4436,32			
10	3 A		13/5/2014	>-4 MD	48,15	3111,84			
11	4 B		13/5/2014	>-4 MD	64,70	3583,74			
12									
13									
14									
15									
16									
17									
18									

Fuente: elaboración propia.

4. Intervalo de confianza y cambio mínimo detectable

El **intervalo de confianza (IC)** nos dice el rango en el que se encuentra la media poblacional con un nivel de confianza dado (normalmente 95 %). La fórmula se expresa del siguiente modo:

$$IC = X^- \pm Z * (\sigma / \sqrt{n})$$

- X^- : media muestral.
- Z: valor que corresponde al 95 % (aprox., 1,96).
- σ : desviación estándar.
- n: tamaño de la muestra.

En Google Sheets:

1. Calcula la media con =AVERAGE(...).



2. Calcula n con =CONTAR(...) o =COUNTA(...).
3. Calcula la desviación estándar con =STDEVP(...).
4. Error estándar = (Desviación Estándar) / SQRT(n).
5. Límite superior = Media + (1.96 * Error estándar).
6. Límite inferior = Media - (1.96 * Error estándar).

Esto te indica, con un 95 % de confianza, entre qué valores se ubicaría la media real de la población.

El **cambio mínimo detectable** ayuda a discernir cuándo un cambio entre dos mediciones (por ejemplo, la distancia recorrida en dos entrenamientos) es significativo y no solo ruido de medición. Una fórmula típica es:

$$CMD = 0.2 * \text{Desviación Estándar}$$

En Google Sheets, podrías poner lo siguiente en una celda:

$$CMD = 0.2 * STDEVP(\text{rango_datos})$$

Si el valor del CMD es, por ejemplo, 0,2 m/s en velocidad, y la diferencia que ves entre dos sesiones es de 0,15 m/s, podríamos decir que no supera el cambio mínimo detectable.

Figura 2: Intervalo de confianza y cambio mínimo detectable

	A	B	C	D	E	F	G	H	I	J	K
1	player	position	session_date	match_day	duration_minute	distance		Estadístico	Fórmula	Resultado	
2	1	A	11/5/2014	>4 MD	51,93	3034,46		Media	AVERAGE(F2:F51)	2760,6	
3	2	A	11/5/2014	>4 MD	19,75	1426,10		Desviación estándar	STDEVP(F2:F51)	996,5	
4	3	A	11/5/2014	>4 MD	44,80	2319,55		Tamaño muestra	COUNTA(F2:F51)	50	
5	4	B	11/5/2014	>4 MD	15,77	1190,14		Error estándar	STDEVP(F2:F51)/COUNTA(F2:F51)	28,9	
6	5	B	11/5/2014	>4 MD	42,88	2057,18		Límite + IC	AVERAGE(F2:F51)+1,96*STDEVP(F2:F51)/COUNTA(F2:F51)	2817,2	
7	6	C	11/5/2014	>4 MD	54,18	3431,98		Límite - IC	AVERAGE(F2:F51)-1,96*STDEVP(F2:F51)/COUNTA(F2:F51)	2704,1	
8	1	A	13/5/2014	>4 MD	64,07	3713,47		Cambio Mínima Detectable	0,2*STDEVP(F2:F51)	288,5	
9	2	A	13/5/2014	>4 MD	59,18	4436,32					
10	3	A	13/5/2014	>4 MD	48,15	3111,84					
11	4	B	13/5/2014	>4 MD	64,70	3583,74					
12	5	B	13/5/2014	>4 MD	66,02	3451,24					
13	6	C	13/5/2014	>4 MD	82,03	4511,29					

Fuente: elaboración propia.

6. Rolling average y ACWR

Rolling average (promedio móvil): permite suavizar fluctuaciones diarias y ver tendencias. Se usa mucho en entrenamiento para calcular promedios móviles a 7 días (carga aguda) o 28 días (carga crónica):

- **7 días (agudo):** sumas (o promedias) los valores de los últimos 7 días.
- **28 días (crónico):** lo mismo, pero con ventana de 28 días.

ACWR (acute:chronic workload ratio):



ACWR = promedio 7 días / promedio 28 días

Un valor >1,5 a menudo indica riesgo de sobrecarga.

En Google Sheets, si tus datos están ordenados por fecha y cada fila corresponde a un día, puedes hacer en la fila 8 algo como lo siguiente:

=AVERAGE(B2:B8)

(asumiendo que B tiene la distancia o carga) y luego arrastrar esa fórmula hacia abajo. Para 28 días, lo mismo pero con B2:B29, etc. Si hay discontinuidades en los días o varios jugadores, se complica un poco, pero a nivel básico, la técnica principal es la misma.

Figura 3: Rolling average y ACWR

	A	B	C	D	E	F	G	H	I	J	K	L
1	session_date	duration_minut	distance	rolling_7	rolling_28	ratio A:C	fórmula 7	fórmula 28	fórmula ratio			
2	11/5/2014	51,93	3034,46	3034,46	3034,46	1,00	AVERAGE(FS2:F2)	AVERAGE(FS2:F2)	G2/H2			
3	12/5/2014	19,75	1426,10	2230,28	2230,28	1,00	AVERAGE(FS2:F3)	AVERAGE(FS2:F3)	G3/H3			
4	13/5/2014	44,80	2319,55	2260,04	2260,04	1,00	AVERAGE(FS2:F4)	AVERAGE(FS2:F4)	...			
5	14/5/2014	15,77	1190,14	1992,56	1992,56	1,00	AVERAGE(FS2:F5)			
6	15/5/2014	42,88	2057,18	2005,49	2005,49	1,00	AVERAGE(FS2:F6)			
7	16/5/2014	54,18	3431,98	2243,24	2243,24	1,00	AVERAGE(FS2:F7)			
8	17/5/2014	64,07	3713,47	2453,27	2453,27	1,00	AVERAGE(FS2:F8)			
9	18/5/2014	59,18	4436,32	2653,54	2701,15	0,98	AVERAGE(F3:F9)			
10	19/5/2014	48,15	3111,84	2894,35	2746,78	1,05	AVERAGE(F4:F10)			
11	20/5/2014	64,70	3583,74	3074,95	2830,48	1,09	AVERAGE(F5:F11)			
12	21/5/2014	66,02	3451,24	3397,97	2886,91	1,18			
13	22/5/2014	82,03	4511,29	3748,55	3022,28	1,24			
14	23/5/2014	43,87	2763,10	3653,00	3002,34	1,22			
15	24/5/2014	54,85	4305,01	3737,50	3095,39	1,21			
16	25/5/2014	44,07	2988,12	3530,62	3088,24	1,14			
17	26/5/2014	63,07	4087,50	3670,00	3150,69	1,16			
18	27/5/2014	54,23	3315,00	3631,61	3160,36	1,15			
19	28/5/2014	69,45	3858,16	3689,74	3199,12	1,15			
20	29/5/2014	52,92	3500,10	3545,28	3214,96	1,10			
21	30/5/2014	55,77	4208,02	3751,70	3264,62	1,15			
22	31/5/2014	63,02	3504,06	3637,28	3276,02	1,11			
23	1/6/2014	53,75	3453,36	3703,74	3284,08	1,13			
24	2/6/2014	49,28	3164,86	3571,94	3278,90	1,09			
25	3/6/2014	51,45	2750,20	3491,25	3256,87	1,07			
26	4/6/2014	0,67	68,28	2949,84	3129,32	0,94			
27	5/6/2014	0,63	66,16	2459,28	3011,51	0,82			
28	6/6/2014	9,60	1013,41	2002,90	2937,51	0,68			
29	7/6/2014	9,50	1003,80	1645,72	2868,44	0,57	...	AVERAGE(FS2:F29)	...			
30	8/6/2014	9,47	970,34	1291,01	2794,73	0,46	...	AVERAGE(F3:F30)	...			
31	9/6/2014	9,38	1009,56	983,11	2779,85	0,35	...	AVERAGE(F4:F31)	...			
32	10/6/2014	45,18	2380,17	930,25	2782,01	0,33			
33	11/6/2014	45,13	1991,32	1204,97	2810,63	0,43			

Fuente: elaboración propia.

7. Ejercicios prácticos con Google Sheets

Ejercicio 1:

- Copia la base de datos de entrenamiento en Google Sheets.
- Calcula la media, la mediana, el rango y la desviación estándar de la variable “distancia recorrida”.
- Interpreta en qué rango se sitúa la mayoría de los valores.



Ejercicio 2:

- Calcula el IC al 95 % para la distancia recorrida.
- ¿El intervalo es amplio o estrecho? ¿Qué implicaciones podría tener?

Ejercicio 3:

- Calcula el CMD para la misma variable.
- Compara la diferencia entre dos sesiones consecutivas. ¿Supera o no el CMD?

Ejercicio 4:

- Calcula el *rolling average* a 7 y 28 días (columnas nuevas).
- Calcula ACWR como (Promedio_7días / Promedio_28días).
- Identifica valores de ACWR >1,5 y comenta qué recomendaciones podrías hacer al cuerpo técnico.

Unidad 3.2 Visualización interactiva con Looker Studio

1. Introducción a Looker Studio

Looker Studio, antes Google Data Studio, es una aplicación web para crear informes y paneles (*dashboards*) que se pueden compartir fácilmente mediante un enlace. Como Power BI, permite combinar datos de diferentes fuentes (Google Sheets, BigQuery, etc.) y crear gráficos dinámicos.

Ventajas:

- Todo es web, no hay que instalar nada.
- Conexión sencilla a cuenta de Google Sheets.
- Facilidad para compartir y colaborar.
- Gran variedad de visualizaciones básicas (barras, líneas, dispersión, tablas, mapas...).

2. Diferencias principales con Power BI

- **Power BI** normalmente se instala en tu PC (aunque existe versión en línea) y está muy integrado con el ecosistema Microsoft. Tiene más funcionalidades corporativas y posibilidades de modelado de datos.
- **Looker Studio** funciona en el navegador y se integra con herramientas de Google, y ofrece un acceso rápido y una curva de aprendizaje más suave para usuarios de Google.

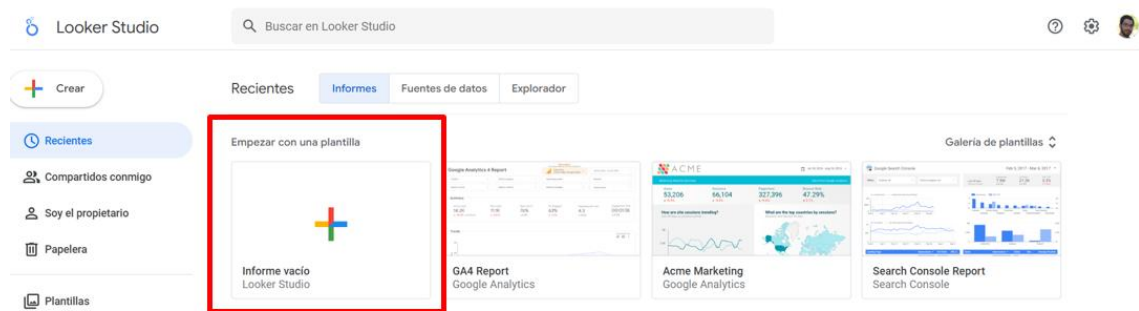


- Ambas permiten hacer paneles con filtros y segmentaciones, y ambas tienen versiones gratuitas (con limitaciones).

3. Conexión de Looker Studio a Google Sheets

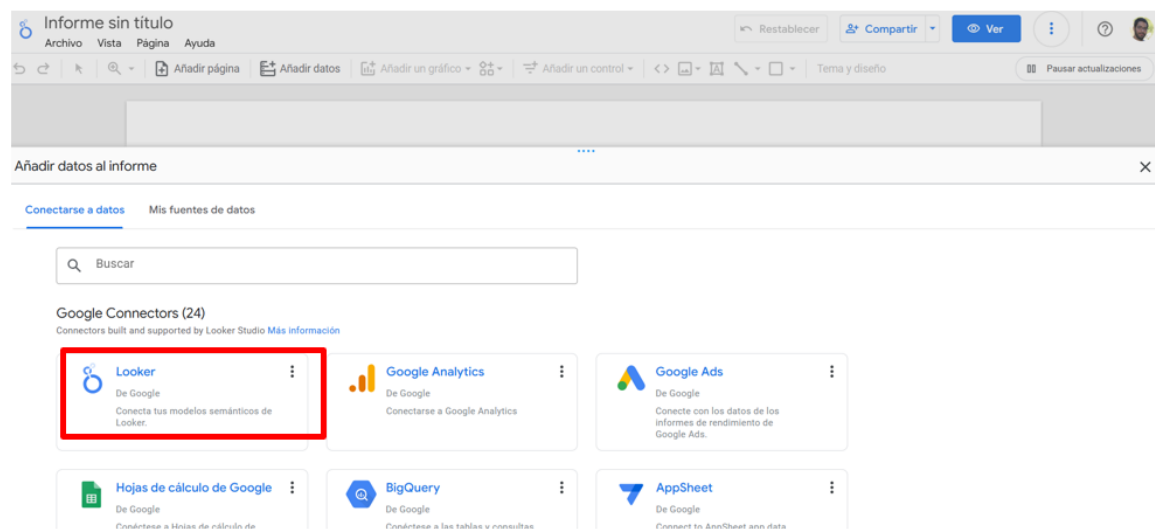
1. Ve a Looker Studio en tu navegador.
2. Inicia sesión con tu cuenta de Google.
3. Crea un “Informe en blanco”.
4. Elige “Google Sheets” como fuente de datos.
5. Selecciona la hoja de cálculo que contiene tu base de datos.
6. Haz clic en “Conectar”.
7. Looker Studio mostrará una vista previa de campos. Revisa que las columnas tengan el tipo correcto (numérico, fecha, texto...).

Figura 4: Conexión de Looker Studio a Google Sheets



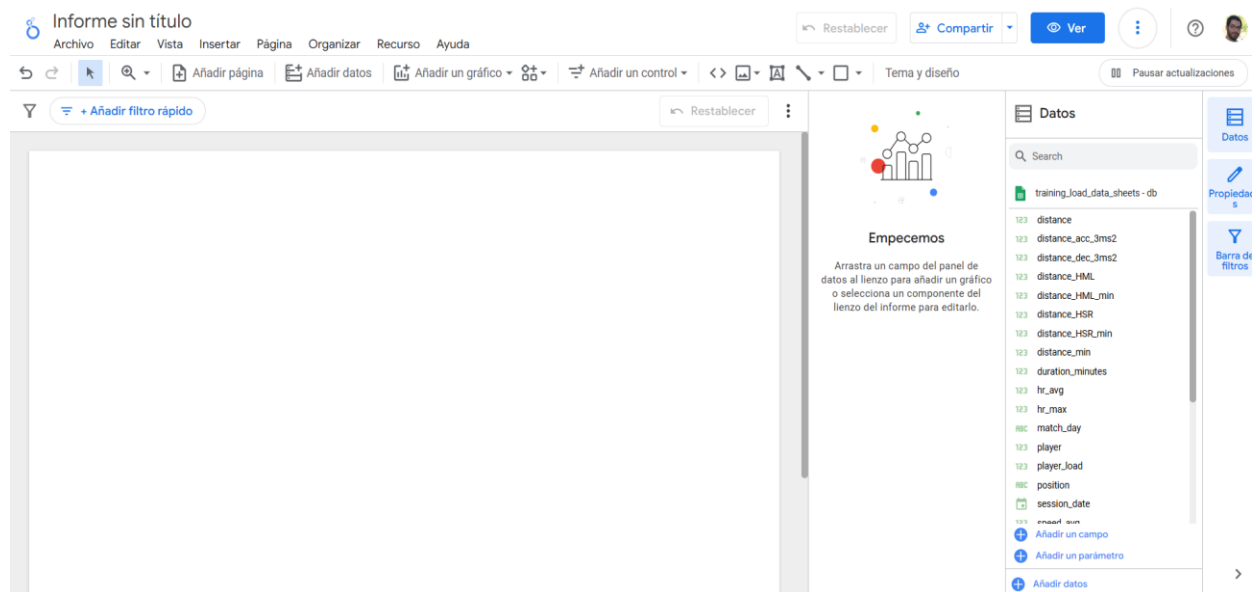
Fuente: captura de pantalla de Looker Studio (<https://lookerstudio.google.com/u/0/>).

Figura 5: Conexión de Looker Studio a Google Sheets II



Fuente: captura de pantalla de Looker Studio (<https://lookerstudio.google.com/u/0/>).

Figura 6: Conexión de Looker Studio a Google Sheets III

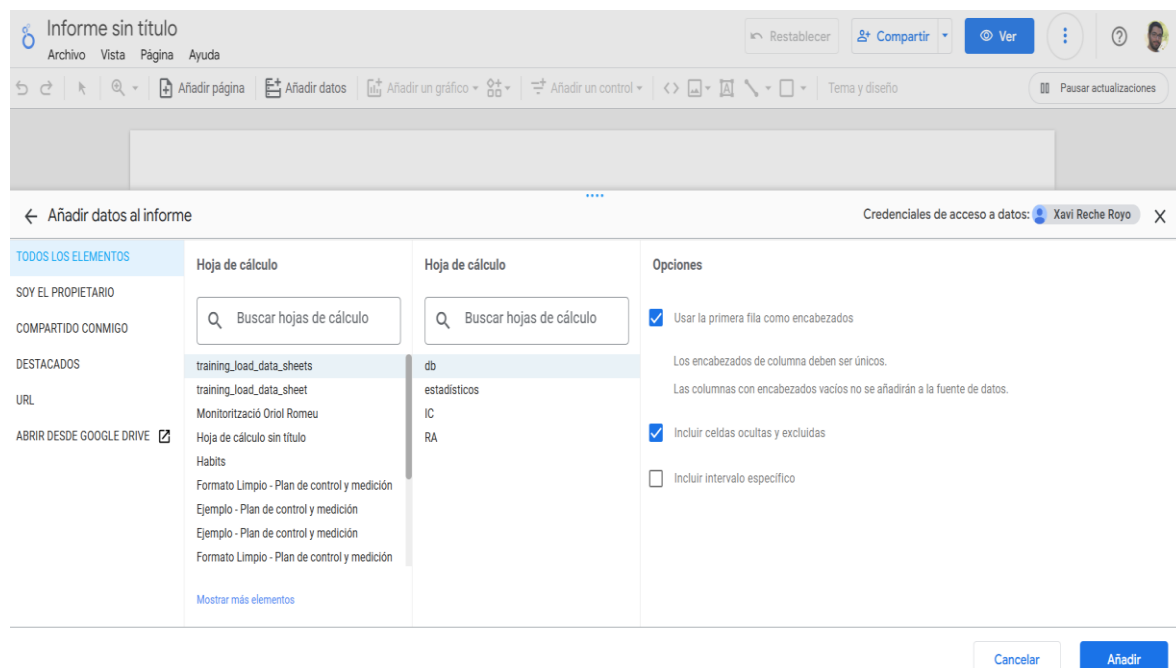


Fuente: captura de pantalla de Looker Studio (<https://lookerstudio.google.com/u/0/>)

4. Creación de visualizaciones

Una vez seleccionada la fuente de datos, podrás **insertar** distintos gráficos:

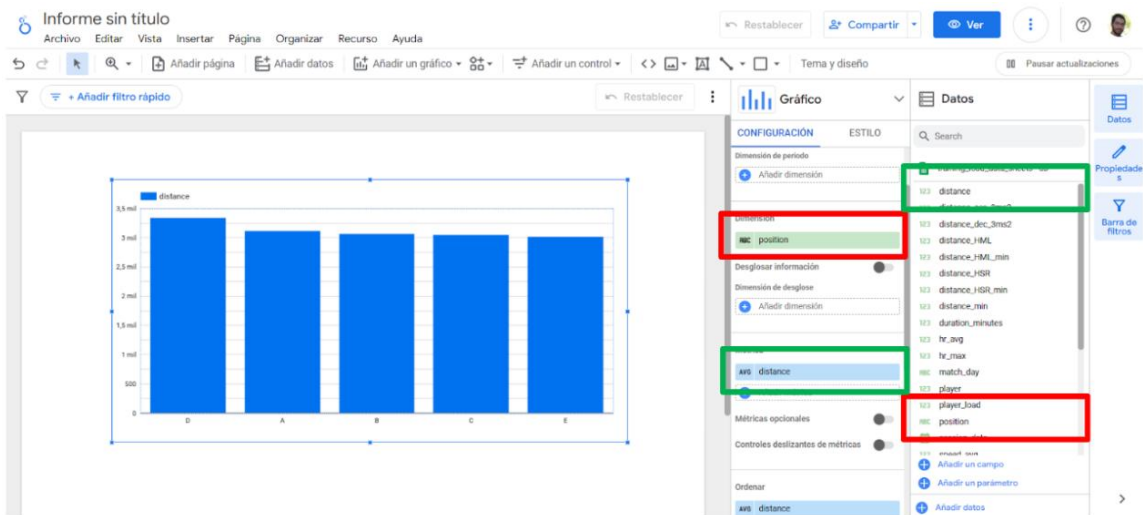
Figura 7: Creación de visualizaciones



Fuente: captura de pantalla de Looker Studio (<https://lookerstudio.google.com/u/0/>).

- **Gráfico de barras:** útil para comparar grupos (p. ej., distancia por posición).

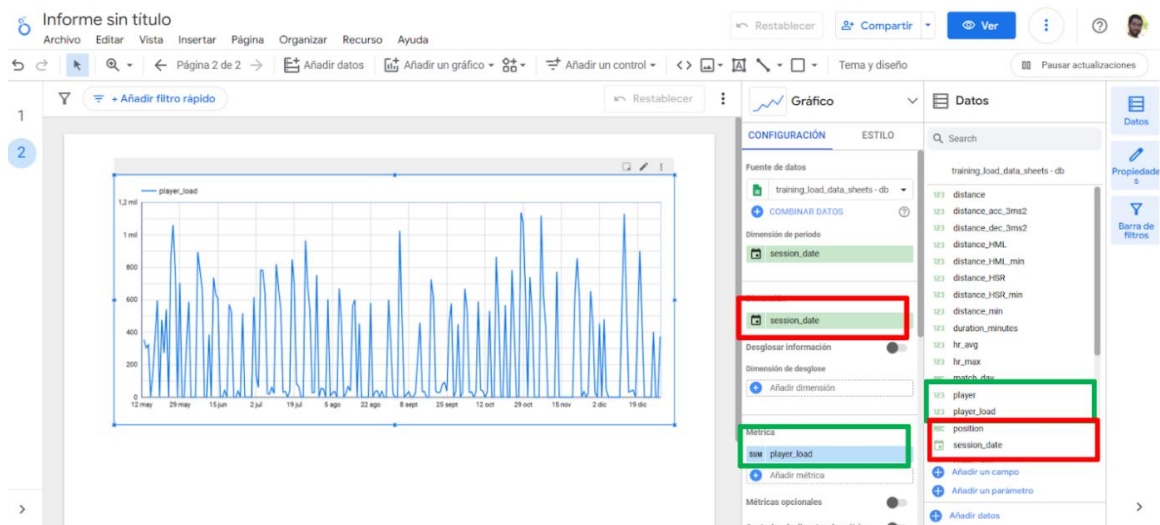
Figura 8: Gráfico de barras



Fuente: captura de pantalla de Looker Studio (<https://lookerstudio.google.com/u/0/>).

- **Gráfico de líneas:** útil para tendencias en el tiempo (p. ej., evolución del player_load).

Figura 9: Gráfico de líneas

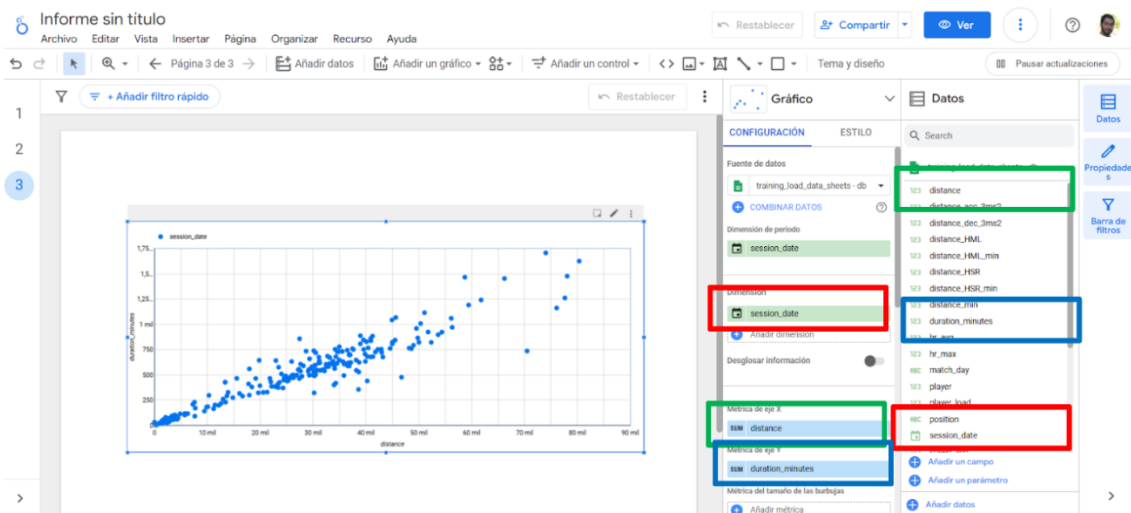


Fuente: captura de pantalla de Looker Studio (<https://lookerstudio.google.com/u/0/>).

- **Gráfico de dispersión:** para ver correlaciones (duración vs. distancia).

Figura 10: Gráfico de dispersión

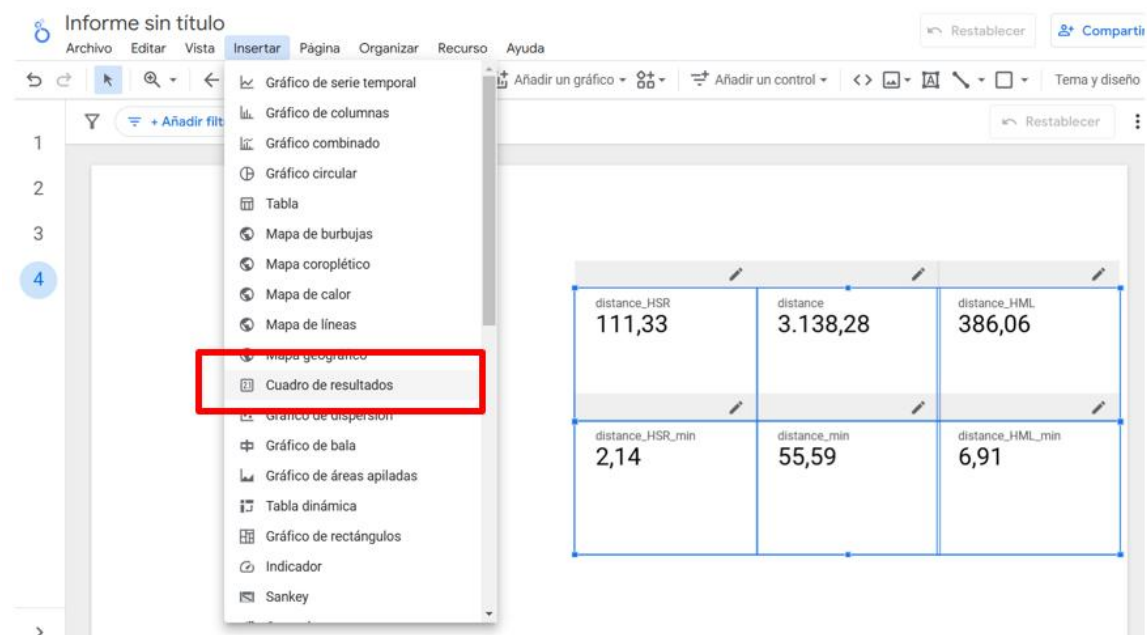




Fuente: captura de pantalla de Looker Studio (<https://lookerstudio.google.com/u/0/>).

- **Tablas o *scorecards***: para métricas clave (p. ej., promedio diario de un indicador).

Figura 11: Gráfico de dispersión



Fuente: captura de pantalla de Looker Studio (<https://lookerstudio.google.com/u/0/>).

En la derecha de Looker Studio, tendrás un panel donde ajustar:

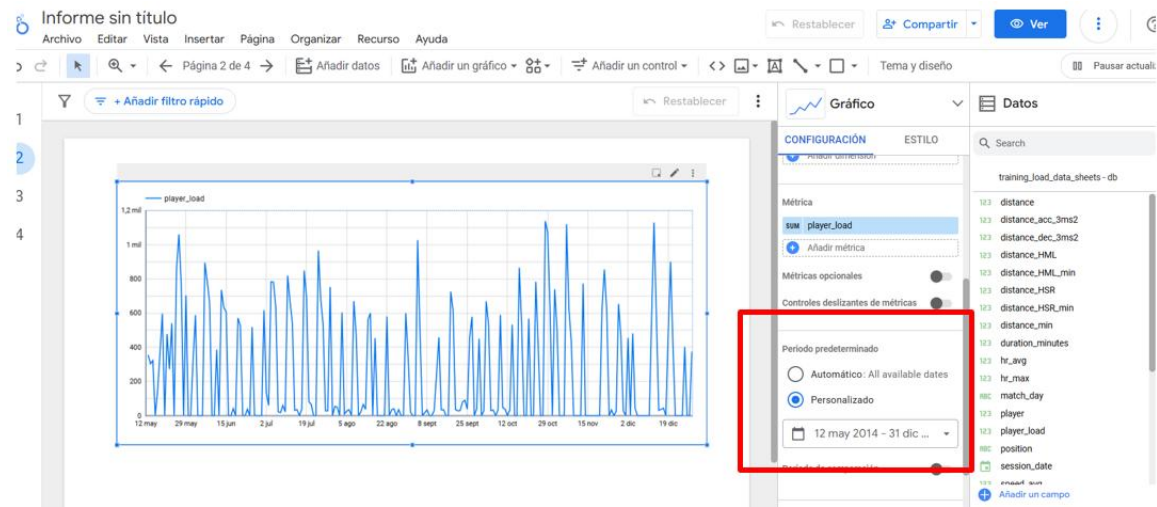
- **Dimensiones** (categorías como posición, jugador, fecha, etc.).
- **Métricas** (valores numéricos con sumas, cuentas o promedio).
- **Estilo** (colores, formas, etc.).

5. Paneles interactivos y filtros

La gran ventaja de Looker Studio es la **interactividad**:

- **Control de fecha:** permite al usuario seleccionar un rango de fechas.

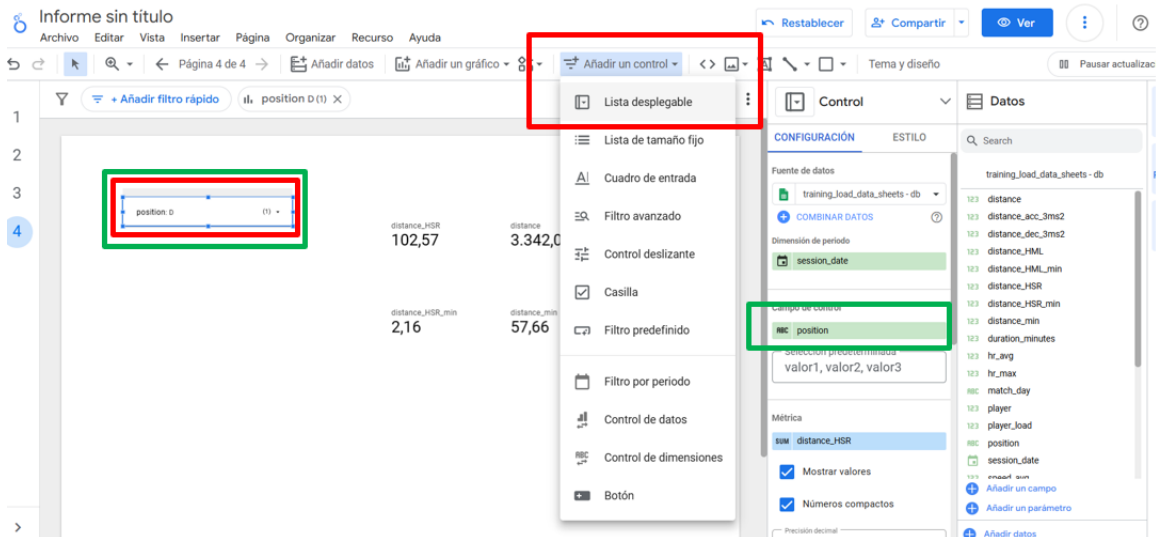
Figura 12: Control de fecha



Fuente: captura de pantalla de Looker Studio (<https://lookerstudio.google.com/u/0/>).

- **Control de filtros** (p. ej., filtrar por posición, por jugador, etc.).

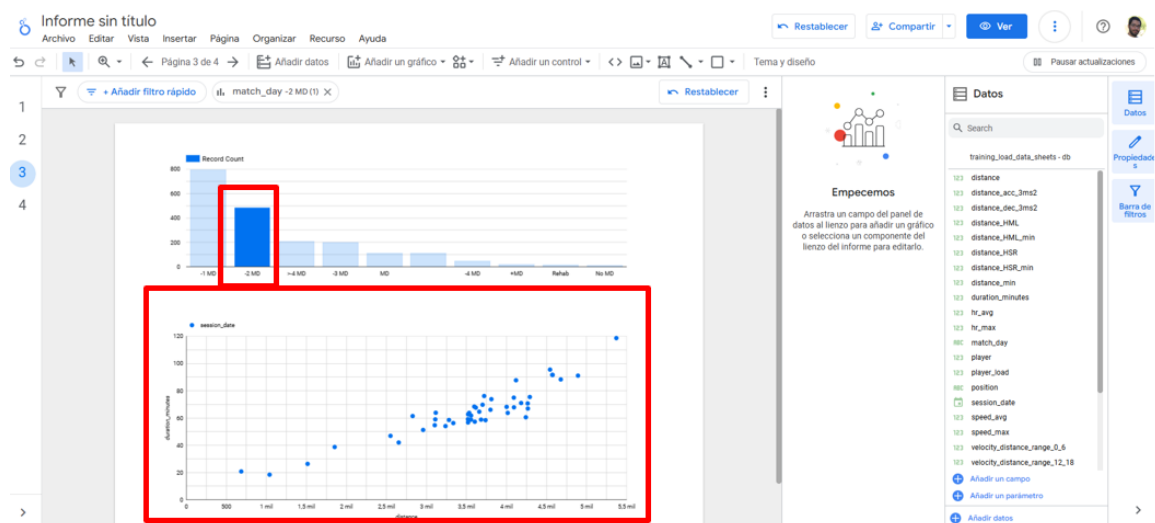
Figura 13: Control de filtros



Fuente: captura de pantalla de Looker Studio (<https://lookerstudio.google.com/u/0/>).

- **Filtros en gráfico:** si clicas en una barra que representa “Delanteros”, el resto del panel se ajusta a ese filtro (por ejemplo, seleccionar MD-2 desde el gráfico de barras).

Figura 14: Filtros en gráfico



Fuente: captura de pantalla de Looker Studio (<https://lookerstudio.google.com/u/0/>).

Esto hace que el informe sea un *dashboard* vivo, donde el analista o el entrenador pueda cambiar parámetros y ver resultados al instante.

6. Ejercicios prácticos en Looker Studio

Ejercicio 1

- Crea un informe en Looker Studio.
- Conecta tu hoja de Google Sheets.
- Verifica los tipos de datos (fechas, números, texto).

Ejercicio 2

- Inserta un **gráfico de líneas** mostrando cómo evoluciona la Distancia_total a lo largo del tiempo (dimensión: Fecha; métrica: Distancia_total).
- Agrega un **control de fecha** para filtrar periodos específicos.

Ejercicio 3

- Crea un **gráfico de barras** que compare la Distancia_alta_velocidad promedio según la posición (dimensión: Posición; métrica: Promedio de Distancia_alta_velocidad).
- Añade un **filtro** para poder elegir un solo jugador o un grupo de jugadores.

Ejercicio 4

- Crea un **gráfico de dispersión** que relacione Duración_sesión (eje X) y Distancia_total (eje Y).

- Segmenta por colores según la posición, para ver si hay patrones distintos entre defensas, mediocampistas y delanteros.

De este modo, tendrás un panel que te permita analizar rápidamente el rendimiento de los deportistas y aislar la información que más te interese en cada momento.

Unidad 3.3 Reflexión y conclusiones

A lo largo de este módulo 3, hemos aprendido a **replicar los análisis estadísticos** y a **crear visualizaciones interactivas** en el ecosistema de Google (Sheets y Looker Studio), comparando con lo que hicimos en el módulo 2 (Excel y Power BI).

1. Similitudes y diferencias

- **Estadística descriptiva:** sumas, promedios, desviaciones estándar, intervalos de confianza y CMD se calculan igual en Google Sheets que en Excel, solo cambian algunos nombres de funciones y la forma de filtrar.
- **Visualización:** Looker Studio y Power BI ofrecen paneles dinámicos con filtros. La diferencia principal es que Looker Studio está 100% basado en la nube y se integra fácilmente con Google Sheets, mientras que Power BI suele usarse más en entornos de escritorio integrados a Microsoft 365.

2. Ventajas de la nube

- **Colaboración en tiempo real:** varios usuarios pueden editar el archivo de Google Sheets al mismo tiempo.
- **Posibilidad de compartir paneles:** con Looker Studio bastan unos clics para compartir tu tablero.
- **Historial automático:** hay menos riesgo de perder datos.
- **Acceso en cualquier lugar:** con tu cuenta de Google, puedes trabajar desde distintos dispositivos.

3. Importancia para el ámbito deportivo

Ya sea con Excel o en Google Sheets (y, por extensión, con Power BI o Looker Studio), la posibilidad de calcular **rolling average**, **ACWR**, **intervalos de confianza** y **CMD** es fundamental para lo siguiente:

- **Identificar patrones de fatiga y riesgo de lesión.**
- **Personalizar la carga de entrenamiento** según resultados reales.
- **Tomar decisiones basadas en datos**, no en impresiones subjetivas.



La **visualización** facilita la interpretación de estos indicadores, lo que permite al cuerpo técnico (entrenadores, preparadores físicos, etc.) ver rápidamente quién está en riesgo de sobrecarga, cómo evoluciona la distancia recorrida en el tiempo, o si hay correlación entre duración de la sesión y alta velocidad.

4. Recomendaciones finales

- **Practica:** la clave es dedicar tiempo a poner en práctica todo lo visto, ya sea con datos reales o ejemplos.
- **Cuida la calidad de los datos:** asegúrate de que tus columnas estén bien definidas (números como números, fechas como fechas), porque errores de formato llevan a resultados estadísticos engañosos.
- **Explora integraciones:** si tu base de datos crece, Google Sheets podría quedarse corto, y podrías enlazar con Google BigQuery o bases de datos más grandes para que Looker Studio trabaje sin problemas de rendimiento.
- **Comparte:** involucra a tus compañeros o entrenadores, creando *dashboards* que sean claros y centrados en las métricas más relevantes.

5. Conclusión

Este módulo demuestra que el **conocimiento estadístico** y de **visualización de datos** es **independiente** de la herramienta empleada. Dominar Excel/Power BI o Google Sheets/Looker Studio te brinda flexibilidad y te permite adaptarte a diferentes entornos laborales.

- **Excel y Power BI:** tienen gran robustez e integración con Microsoft 365; son muy usados en entornos corporativos.
- **Google Sheets y Looker Studio:** ofrecen sencillez, colaboración en tiempo real y fácil acceso a cualquier persona con cuenta de Google.

En todos los casos, el “corazón” del análisis (la estadística descriptiva, la interpretación de la información y las métricas como IC, CMD, *rolling average* y ACWR) **sigue siendo el mismo.**



Referencias

Google. (s.f.). *Looker Studio*. Google. <https://lookerstudio.google.com>

