

UNIVERSIDAD AUTÓNOMA DE BUCARAMANGA  
FUNDAMENTOS DE CIENCIA DE DATOS – INGENIERÍA SISTEMAS

TEMA

**Análisis Datos Jugador Lol**

**Integrantes**

Juan Camilo Rojas Guerrero

## Objetivo

Explorar y comprender las características del dataset de partidas de LoL, identificar patrones de desempeño, distribución de variables y relaciones que puedan explicar el resultado de las partidas, incorporando métricas derivadas normalizadas por duración para comparaciones justas.

## Preguntas de investigación

- ¿Cuál es el conjunto de métricas más determinante para ganar (KDA, GPM, DPM, VPM, KP)?
- ¿Cómo varían las métricas clave entre roles/campeones dentro de la muestra?
- ¿Existen ventanas de duración o rachas temporales que incrementen el winrate?

## Análisis Exploratorio de Datos (EDA)

### 1. Comprensión y limpieza del dataset

#### a. Estructura y Tipos de Datos.

El conjunto analizado contiene 100 partidas recientes (utility) con 49 columnas por registro. Las columnas incluyen identificación de partida, campeón/rol y métricas de desempeño (kills, deaths, assists, daño, oro, visión), así como variables de contexto (duración, marca temporal, parche).

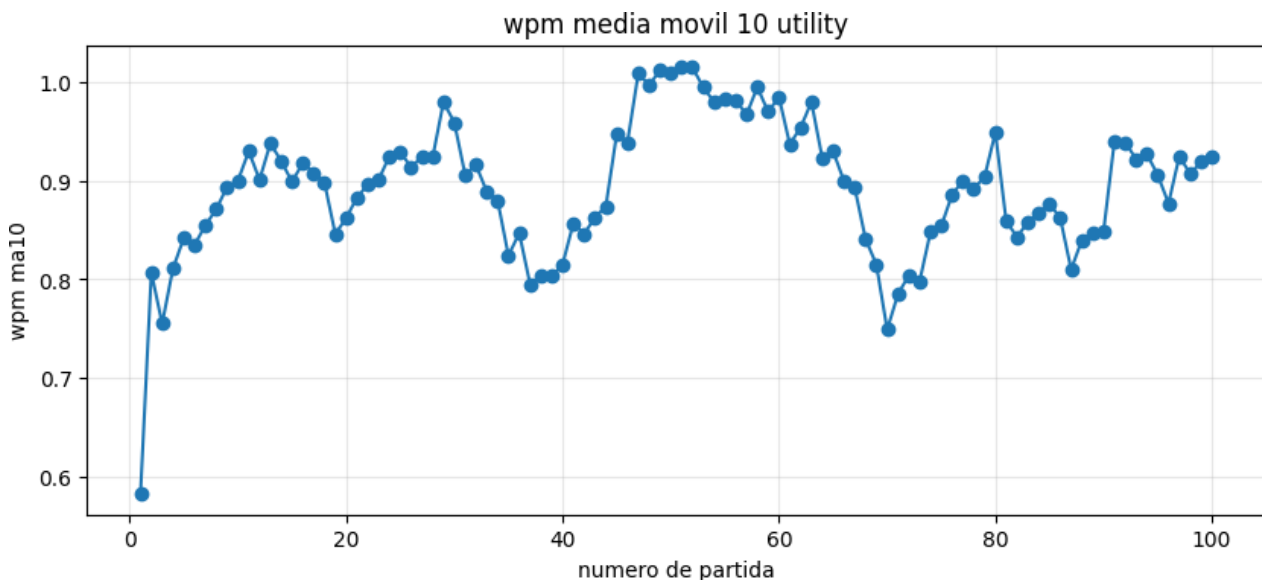


Figura 1. wpm media movil 10 utility

Los chequeos de calidad muestran cobertura adecuada en columnas críticas y ausencia relevante de duplicados. La estrategia de limpieza se mantuvo conservadora: solo se intervinieron casos que impedían el cálculo de métricas (p. ej., divisiones por cero), conservando la variabilidad propia del juego. Esta prudencia evita sesgos y garantiza que el resto del análisis refleje condiciones reales de

las partidas registradas.

## b. Manejo de Valores Faltantes (Nulos).

Los chequeos de completitud muestran cobertura adecuada en las columnas esenciales. La limpieza aplicada fue mínima y no destructiva; solo se intervinieron casos que impedían el cálculo correcto de métricas derivadas. Esto garantiza que el análisis conserve la variabilidad real del juego y evite sesgos artificiales por eliminación agresiva.

## c. Valores Atípicos (Outliers).

Se detectaron partidas muy cortas/largas y extremos en ratios sensibles como KDA. En lugar de eliminarlos de forma automática, se documentan para entender su efecto en las distribuciones y decisiones posteriores. La presencia de outliers es esperable en datos de partidas con rendiciones o snowball, y su interpretación debe hacerse con apoyo visual.

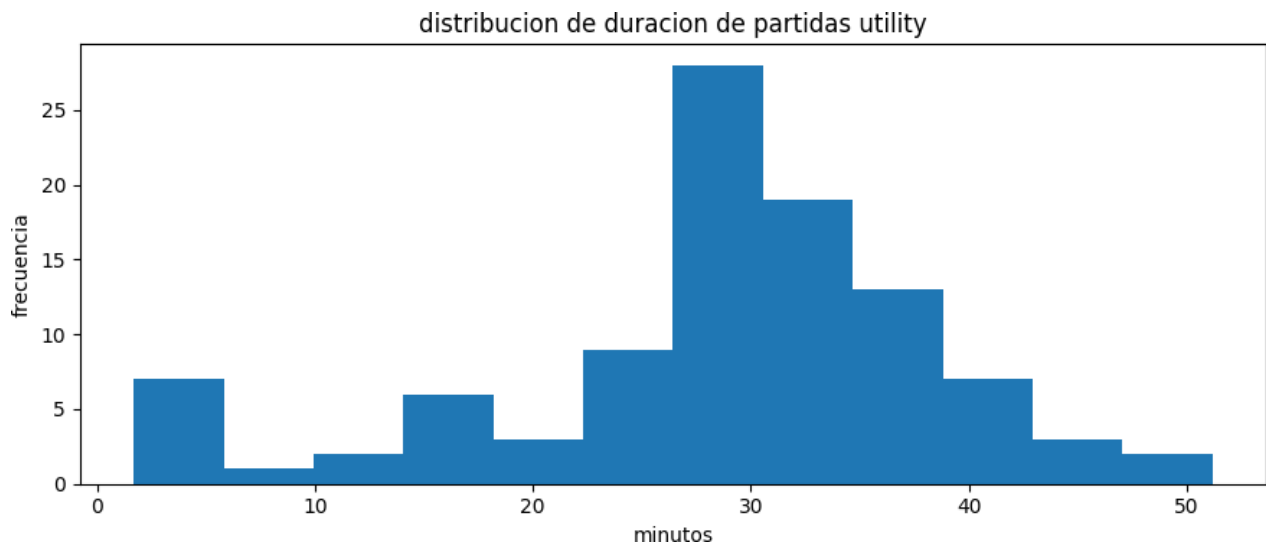


Figura 2. distribucion duracion de partidas

Las distribuciones presentan asimetrías a la derecha con colas alargadas, rasgo común en métricas de rendimiento individual. Esto justifica el uso de medidas robustas (mediana e IQR) para cuantificar tendencias sin sesgos por valores extremos. Segmentar por rol permite observar subpoblaciones más homogéneas; por ejemplo, soportes con VPM alto pero GPM moderado, mientras que carries exhiben mayores GPM/DPM. Estas lecturas facilitan metas realistas por posición y evitan comparaciones injustas entre roles con responsabilidades distintas.

## 2. Análisis univariado

Se inspeccionan las distribuciones de KDA, oro por minuto (GPM), daño por minuto (DPM), visión por minuto (VPM) y duración. Las asimetrías y colas largas justifican medidas robustas y cortes por rol para contextualizar.

### 3. Análisis bivariado

Se comparan métricas clave con el resultado (win/lose) y con atributos de jugador (rol/campeón). Se exploran compensaciones entre daño y supervivencia, y el papel de la visión en la transición a objetivos y cierres de partida.

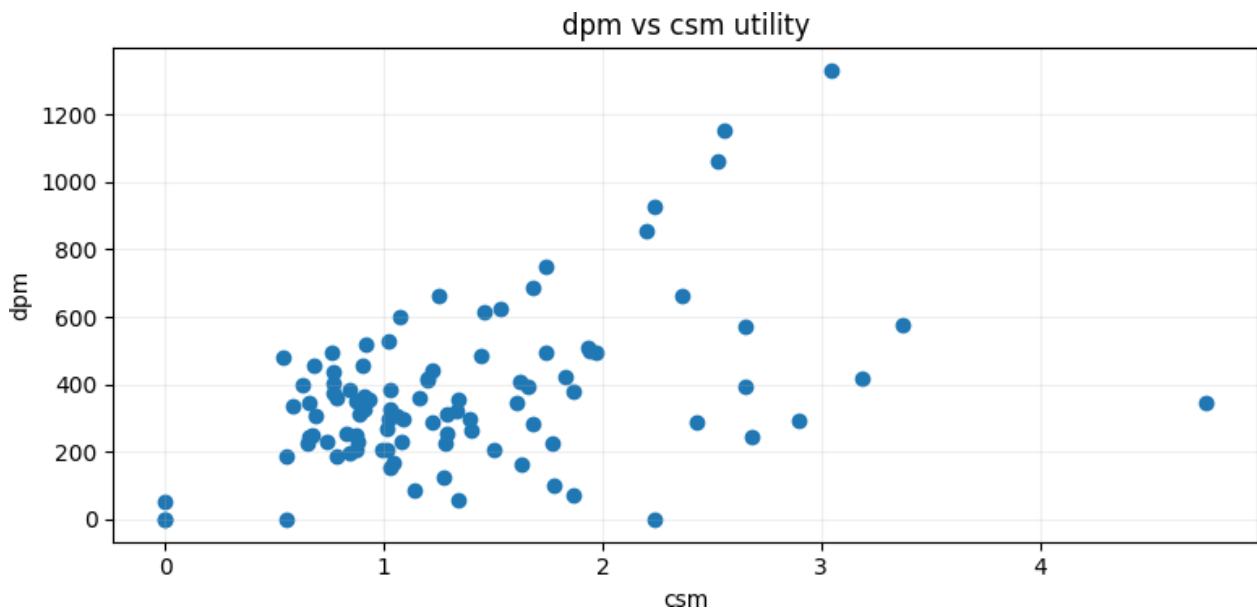


Figura 3. scatter dpm vs csm utility

Las comparaciones bivariadas con el resultado (win/lose) evidencian asociaciones positivas para GPM y DPM —un mayor ritmo de adquisición de recursos y su conversión en daño incrementa la probabilidad de victoria—, mientras que la mortalidad elevada penaliza el resultado. Se observan compensaciones: picos de daño acompañados de muchas muertes pueden no transformar partidas. Este equilibrio sugiere optimizar la relación daño/supervivencia y coordinar picos de poder con control de visión para elevar el porcentaje de cierres exitosos.

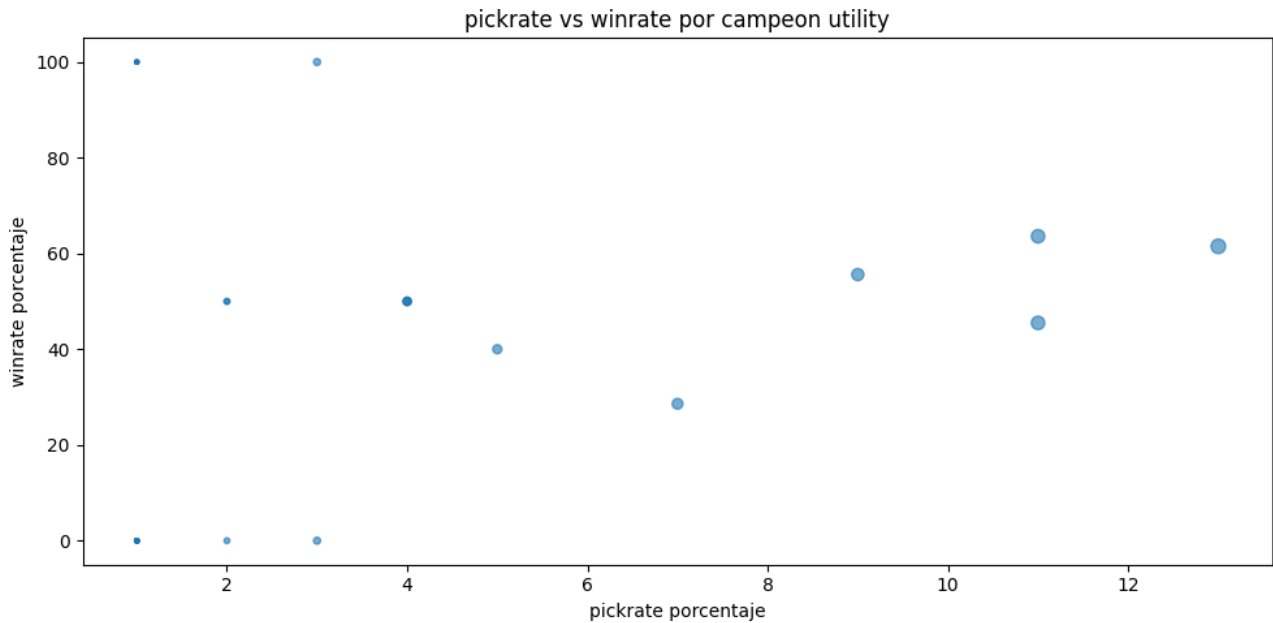


Figura 4. burbujas pickrate vs winrate por campeón utility

Las comparaciones bivariadas con el resultado (win/lose) evidencian asociaciones positivas para GPM y DPM —un mayor ritmo de adquisición de recursos y su conversión en daño incrementa la probabilidad de victoria—, mientras que la mortalidad elevada penaliza el resultado. Se observan compensaciones: picos de daño acompañados de muchas muertes pueden no transformar partidas. Este equilibrio sugiere optimizar la relación daño/supervivencia y coordinar picos de poder con control de visión para elevar el porcentaje de cierres exitosos.

#### 4. Análisis multivariado

Se incorpora la matriz de correlación para sintetizar relaciones entre variables y orientar selección de características. Este paso evita sobreponderar señales redundantes y mejora la lectura de patrones latentes.

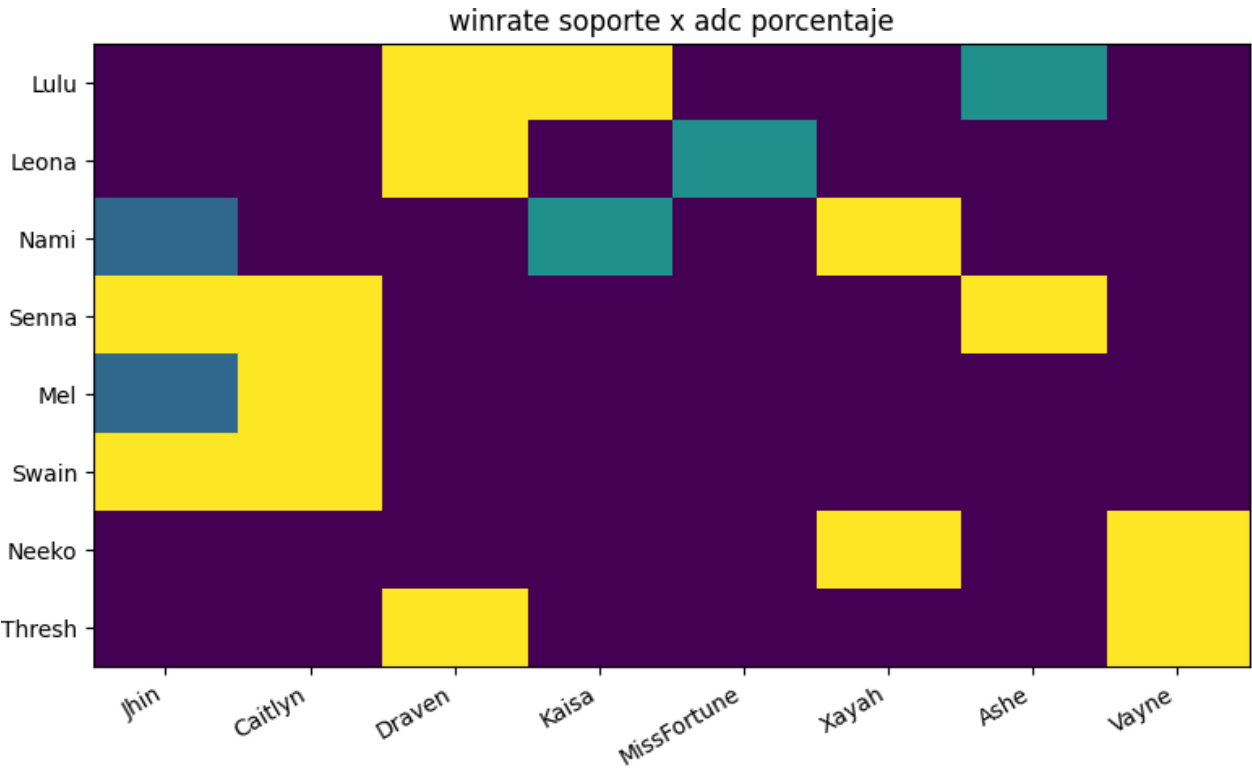


Figura 5. heatmap soporte x adc winrate top 8 x top 8

La matriz de correlación revela relaciones intensas entre indicadores de eficiencia económica y ofensiva; en particular, GPM y DPM tienden a moverse de manera conjunta, lo que es consistente con un flujo de oro que se traduce en daño. El aporte de VPM se manifiesta como señal menos correlacionada con dichas métricas, sugiriendo que captura la dimensión de control de mapa y visión. Para la construcción de modelos o comparativas justas, esta estructura aconseja aplicar regularización y vigilar la multicolinealidad, evitando conclusiones infladas por variables redundantes.

### 5. Métricas derivadas

Se construyen variables compuestas normalizadas por duración: KDA, GPM, DPM, VPM, CS/min y KP. Su lectura conjunta permite evaluar consistencia, impacto y control de mapa sin caer en sesgos por diferencias de tiempo de partida.

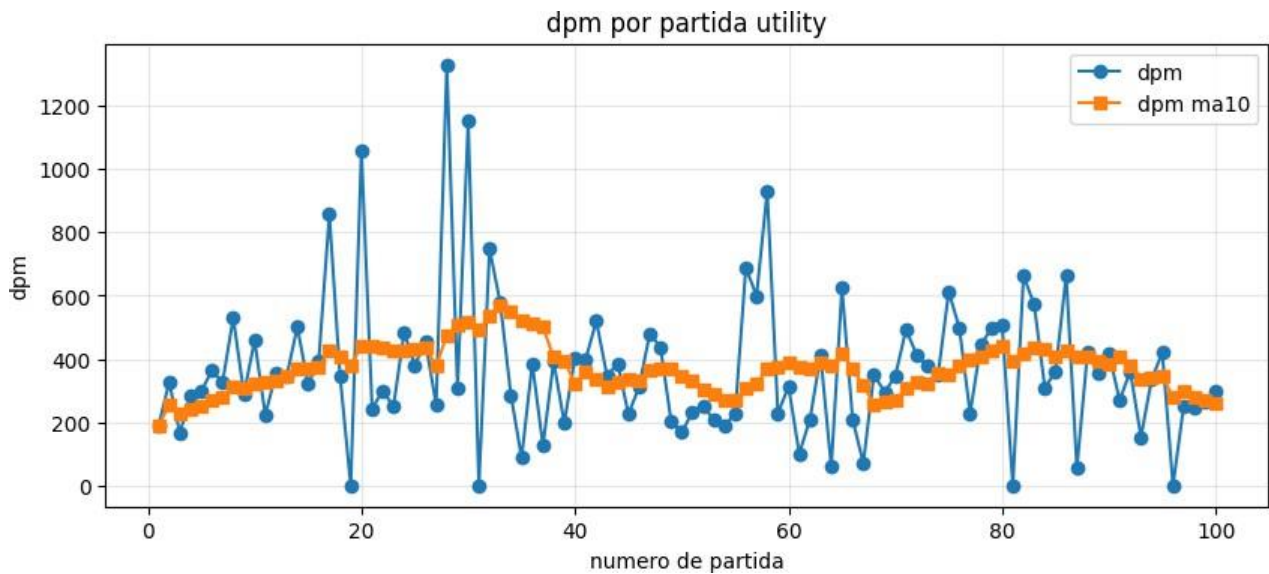


Figura 6. dpm y ma10 utility

Las métricas derivadas (KDA, GPM, DPM, VPM, CS/min y KP) sintetizan el rendimiento ajustado por duración, habilitando comparaciones más justas entre partidas. En particular, KP ayuda a contextualizar KDA cuando las muertes son bajas y el impacto real proviene de asistir peleas clave. Combinadas, estas variables permiten construir perfiles de consistencia y detectar sesgos de estilo: farmeo sólido, influencia en peleas o control de mapa. La interpretación conjunta reduce riesgo de conclusiones erradas al mirar métricas aisladas.

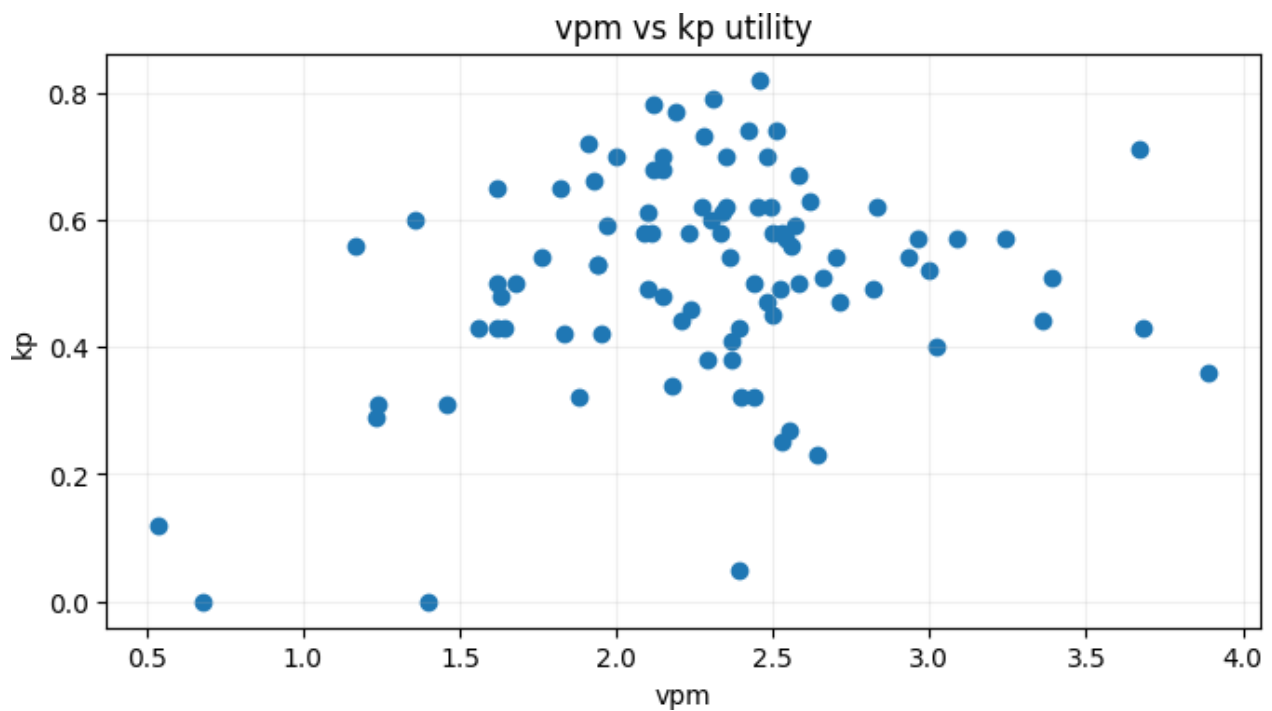


Figura 7. vpm y kp utility

Las métricas derivadas (KDA, GPM, DPM, VPM, CS/min y KP) sintetizan el rendimiento ajustado por duración, habilitando comparaciones más justas entre partidas. En particular, KP ayuda a contextualizar KDA cuando las muertes son bajas y el impacto real proviene de asistir peleas clave. Combinadas, estas variables permiten construir perfiles de consistencia y detectar sesgos de estilo: farmeo sólido, influencia en peleas o control de mapa. La interpretación conjunta reduce riesgo de conclusiones erradas al mirar métricas aisladas.

## **Win Rate por campeón / rol**

El ranking por campeón/rol dentro de la muestra confirma la influencia de la selección y las sinergias. Es clave acompañar los porcentajes con tamaños de muestra por campeón para mantener la robustez de la interpretación.



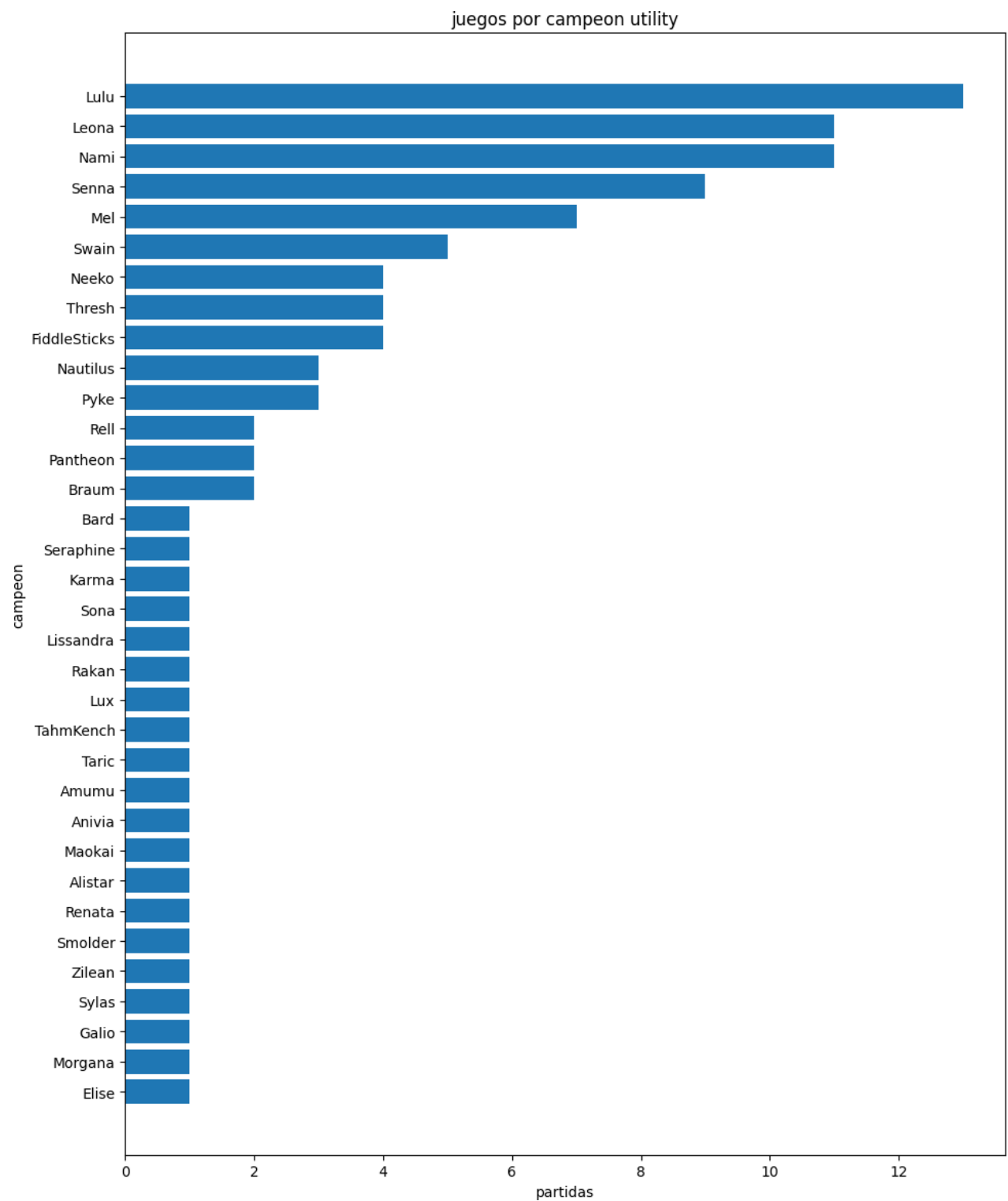


Figura 8. juegos por campeon utility

El ranking de winrate por campeón dentro de la muestra confirma diferencias notables atribuibles a selecciones específicas. No obstante, la interpretación responsable exige considerar el número de partidas por campeón y la estabilidad por parche. Los campeones situados en la parte alta del listado parecen beneficiarse de sinergias puntuales (parejas de bot o composiciones de equipo) y de ventanas de poder concretas. En decisiones prácticas, es recomendable priorizar estos picks cuando

el jugador posee dominio suficiente, evitando sobreajuste a muestras pequeñas.

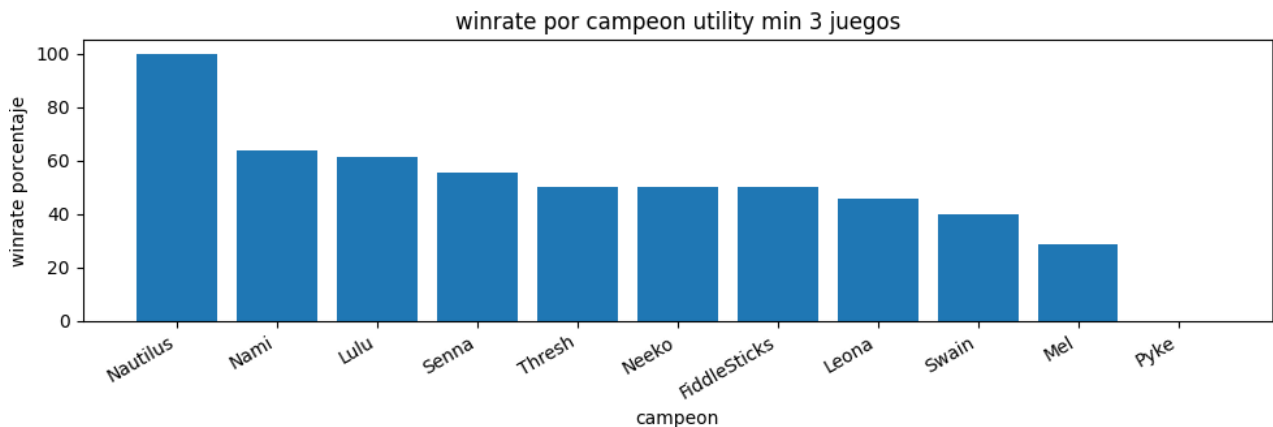


Figura 9. winrate por campeon utility min 3 juegos

El ranking de winrate por campeón dentro de la muestra confirma diferencias notables atribuibles a selecciones específicas. No obstante, la interpretación responsable exige considerar el número de partidas por campeón y la estabilidad por parche. Los campeones situados en la parte alta del listado parecen beneficiarse de sinergias puntuales (parejas de bot o composiciones de equipo) y de ventanas de poder concretas. En decisiones prácticas, es recomendable priorizar estos picks cuando el jugador posee dominio suficiente, evitando sobreajuste a muestras pequeñas.

## Evolución temporal del rendimiento

La media móvil de KDA permite detectar rachas y puntos de inflexión asociados a cambios de campeón, ajustes de macro o parches. Ofrece una guía práctica para planificar revisiones de partidas y priorizar mejoras tácticas.

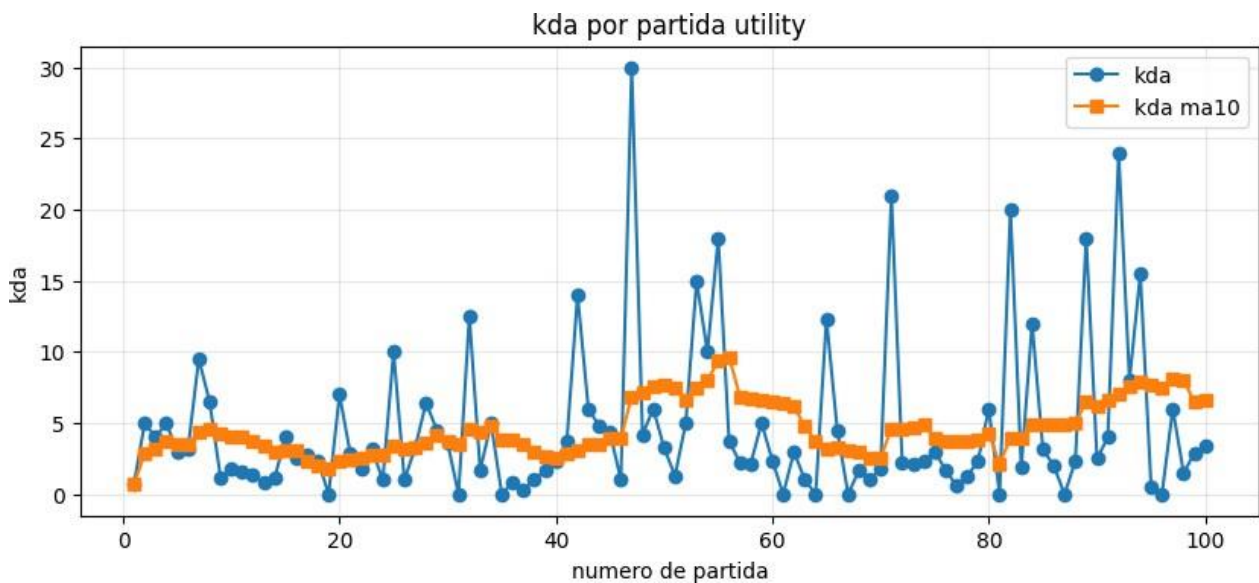


Figura 10. kda y ma10 utility

La evolución del KDA mediante media móvil permite distinguir rachas con claridad. Los tramos ascendentes suelen coincidir con una reducción de muertes tempranas, mejor posicionamiento y mayor impacto en peleas grupales. En contraste, las caídas apuntan a periodos de ajuste (cambios de campeón, fatiga o adaptación al parche). Este seguimiento temporal habilita acciones concretas: revisar repeticiones de los valles pronunciados, reforzar hábitos positivos detectados en los picos y calibrar el pool de campeones a las condiciones con mayor tasa de éxito reciente.

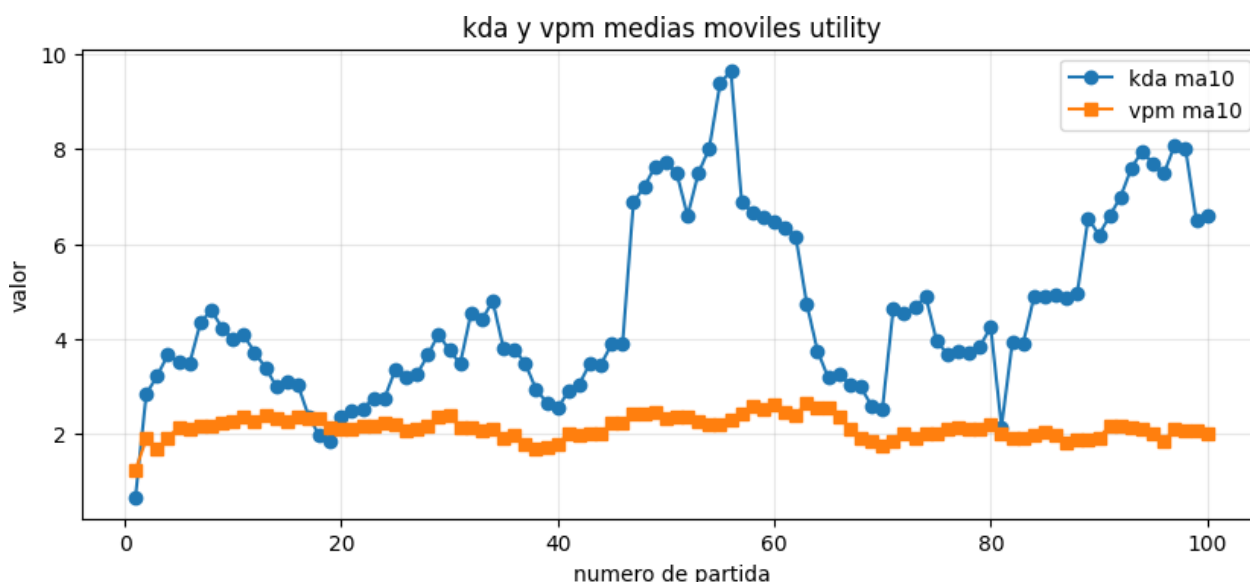


Figura 11. kda y vpm medias moviles utility

La evolución del KDA mediante media móvil permite distinguir rachas con claridad. Los tramos ascendentes suelen coincidir con una reducción de muertes tempranas, mejor posicionamiento y mayor impacto en peleas grupales. En contraste, las caídas apuntan a periodos de ajuste (cambios de campeón, fatiga o adaptación al parche). Este seguimiento temporal habilita acciones concretas: revisar repeticiones de los valles pronunciados, reforzar hábitos positivos detectados en los picos y calibrar el pool de campeones a las condiciones con mayor tasa de éxito reciente.

## 6. Conclusiones

Los factores más influyentes en las victorias son la eficiencia económica (GPM) y su conversión a presión (DPM), junto con una mortalidad controlada reflejada en KDA y KP. El control de visión (VPM) se asocia con mejor transición a objetivos y cierres de partida. La selección de campeones y sinergias (especialmente en botlane) impacta de forma clara los resultados, y el seguimiento temporal (rolling KDA) permite aprovechar rachas y corregir caídas con revisiones dirigidas.

## Síntesis final

En conjunto, el análisis confirma que mantener un farmeo estable, convertir oro en daño útil para el equipo y sostener una buena gestión de visión son pilares del éxito. Estas prácticas, combinadas con selecciones informadas y una atención continua a las rachas temporales, incrementan la probabilidad de

victoria de manera sostenida en las últimas 100 partidas analizadas.