**Control 5 Marketing Analytics SaaS**

Sebastian Gómez y Pedro García

**Parte 1: Análisis Python**

**Preprocesamiento de los datos**

En la primera etapa de nuestro análisis de datos en Python, realizamos una exploración exhaustiva del conjunto de datos. Comenzamos importando las librerías fundamentales para la manipulación de datos (Pandas), operaciones numéricas (NumPy) y visualizaciones (Seaborn, Matplotlib y Plotly). Luego, cargamos el archivo CSV y lo convertimos a un formato, en este caso pasamos la columna fecha a formato fecha para que python reconozca el tipo y haga el procesamiento adecuado para el análisis. A continuación, exploramos la estructura básica del dataset, identificando posibles inconsistencias o problemas en los datos, verificamos el número de filas y columnas y luego verificamos el tipo de datos que sigue cada columna. Para garantizar la calidad de los datos, realizamos una limpieza y preparación, manejando valores faltantes y creando nuevas métricas relevantes como el gross margin, marketing efficiency y clv cac ratio u otras métricas temporales para simplificar el análisis. Finalmente, llevamos a cabo un análisis exploratorio inicial, generando visualizaciones básicas para comprender la distribución de los datos, identificar correlaciones y descubrir patrones iniciales.

**Análisis Gráficos**

Los gráficos presentados ofrecen una visión general del desempeño de cuatro empresas que ofrecen servicios de Software as a Service (SaaS). El primer gráfico de líneas múltiples nos permite **visualizar la evolución de los ingresos** a lo largo del tiempo para cada una de estas compañías. Se observa una notable variabilidad en los ingresos entre las diferentes empresas, lo que sugiere que cada una enfrenta un contexto de mercado y desafíos únicos. Algunas empresas muestran un crecimiento sostenido, mientras que otras experimentan fluctuaciones más pronunciadas.

Los dos gráficos de caja inferiores profundizan en métricas clave relacionadas con la adquisición y el valor de los clientes. El primer gráfico de caja muestra la distribución del costo de adquisición del cliente (CAC) para cada empresa. Un CAC elevado indica que la empresa está invirtiendo una cantidad significativa de recursos para atraer nuevos clientes. El segundo gráfico de caja presenta la distribución del valor de vida del cliente (CLTV), que representa el valor económico total que un cliente genera a lo largo de su relación con la empresa. Un CLTV elevado sugiere que los clientes son más valiosos y rentables.

Al comparar los gráficos de CAC y CLTV, podemos obtener una idea de la eficiencia de las estrategias de adquisición de cada empresa. Idealmente, se busca un CLTV significativamente mayor que el CAC, lo que indica una estrategia de adquisición rentable. Sin embargo, los gráficos sugieren que existe una variabilidad considerable en este aspecto entre las diferentes empresas.

La primera gráfica muestra la distribución del churn rate (**tasa de cancelación**) para diferentes empresas SaaS. En general, las tasas de cancelación son similares entre las empresas, con medianas cercanas al 3%. Esto indica que ninguna empresa destaca significativamente en la retención de clientes. Además, los datos están concentrados en rangos estrechos, lo que sugiere estabilidad y predictibilidad en estas tasas. La ausencia de valores atípicos también confirma que las empresas mantienen un comportamiento consistente en términos de cancelación.

En la segunda gráfica, que analiza la distribución de customer satisfaction (**satisfacción del cliente**), se observa mayor variabilidad. Las medianas de satisfacción oscilan entre 8 y 8.5 en una escala, reflejando niveles moderados a altos de satisfacción. Sin embargo, SaaS Company 4 tiene una mediana inferior en comparación con las otras empresas, lo que podría indicar una oportunidad de mejora en la experiencia del cliente. Por otro lado, SaaS Company 2 y SaaS Company 3 destacan con niveles más altos de satisfacción y menor dispersión, lo que podría reflejar una gestión más efectiva de la experiencia del cliente.

Finalmente, la matriz de correlación identifica relaciones clave entre métricas empresariales. Se observa que el revenue(ingresos) está fuertemente correlacionado con indicadores como la tasa de conversión (conversion\_rate) y el valor de vida del cliente (customer lifetime value), mostrando que estas métricas son impulsores directos de los ingresos. También se destaca una relación negativa entre la satisfacción del cliente y la tasa de cancelación, lo que sugiere que mejorar la experiencia del cliente puede reducir significativamente la pérdida de clientes. Algunas métricas, como bounce rate y social posts, tienen correlaciones débiles con los resultados clave, indicando un impacto limitado en el desempeño general.

**Parte 2: Análisis R**

En R cargamos la base de datos, ya limpia que sacamos de python, con el fin de tener una constancia en los datos que analizaremos. Posterior a cargar los datos, lo que hicimos fue realizar ciertas transformaciones, dentro de las más relevantes, se incluyó la agrupación mensual de registros y la creación de métricas derivadas como el margen bruto, la eficiencia de marketing y el ratio CLV/CAC. Estas métricas no solo nos permitirán analizar el desempeño financiero, sino que también para la efectividad de las estrategias de adquisición de clientes.

### **Tendencias y Variabilidad Temporal**

El análisis temporal reveló patrones medianamente reconocibles de un comportamiento estacional. Algunas empresas mostraron picos y bajadas consistentes en los ingresos durante ciertos meses, especialmente a finales de año, lo que podría ser explicado por un efecto producido por navidad, lo que haría que empresas tuvieran mayores ingresos pero hay otras, como la compañía 1, que no tienen un pick de ingresos sino que todo lo contrario, tienen una bajada en los ingreso.

Con respecto al CAC de las 4 empresas no se puede notar patrones muy claros, salvo que se puede ver una leve tendencia a la alza en este índice. Sumado a lo anterior, la empresa que ha podido tener meses en los que su CAC haya sido menor es la empresa 3, que ha sobresalido de las otras compañías en este índice.

### **Análisis de Distribuciones**

Explorando la distribución de las métricas, se constató que tanto el retorno como el customer lifetime value no siguen una distribución normal en ninguna de las 4 compañías, sumado a eso no encontramos valores atípicos en el revenue ni en el CLV de ninguna compañía. Dentro la volatilidad en el revenue, podemos ver en el box plot de la empresa 1 que esta es levemente más grande que el de las otras empresas, pero en realidad la diferencia entre compañías no es muy significativa. Por otro lado con respecto a la volatilidad del CLV, podemos ver que la empresa que tiene un box plot más grande, y por ende una mayor volatilidad en su CLV, es la compañía 4, pero aun así tampoco la diferencia entre compañías es mucha. La poca diferencia en la volatilidad del CLV y del revenue puede deberse principalmente a que son compañías de la misma industria y por lo tanto están sujetas a los mismos efectos.

**Correlaciones y Relaciones Entre Variables**

Tal y como se puede ver en el gráfico de correlación, normalizada entre -1 y 1, hay varias relaciones entre diferentes variables. Dentro de las correlaciones que se puede destacar son la relación marketing spend y marketing roi, la cual es de -0.75, lo que nos quiere decir que este tipo de compañías SaaS, el que invirtamos más y más en marketing no significa que sea mejor y los más eficiente, otra correlación interesante para el mercado es que revenue tiene una correlación de 0.66 con el marketing Roi, lo que nos confirma que el hecho de tener un mejor ROI, hace que la compañía vea un beneficio real en los retornos de sus actividades.

### **Evaluación de la Eficiencia**

Según los gráficos que tenemos en R en la parte 6, podemos ver parte de las correlaciones que destacamos anteriormente, ya que como se puede ver, un mayor gasto en marketing no implica que el ROI sea mejor, de hecho mientras menos se gasta, se tiende a tener un mayor ROI. Por último, podemos decir que la compañía que tiene el mejor ROI en promedio es la compañía 2 con un ROI promedio de 4.677.

**Parte 3: Análisis Tableau**

**Dashboard 1:**

El dashboard analiza el desempeño de varias compañías SaaS en términos de revenue, desempeño general y distribución regional, permitiendo identificar fortalezas y debilidades específicas para cada compañía.

1. **Mapa de Calor de** **Revenue por Región:** Este gráfico resalta que el revenue no está distribuido equitativamente entre las regiones. La región Sur muestra consistentemente los tonos más oscuros, lo que indica que es la más rentable para todas las compañías menos la 1. Por otro lado, este y norte presentan tonos más claros, sugiriendo ingresos más bajos en estas áreas. Por compañía, parece que SaasCompany3 sobresale en regiones como sur y oeste, mientras que SaasCompany1 tiene un desempeño más débil en las regiones este y sur. Este gráfico sugiere que las estrategias regionales de marketing y ventas podrían estar más alineadas con el éxito en oeste y sur, mientras que otras regiones necesitan mayor atención.
2. **Gráfico de Líneas de Overall Performance Score por Compañía en el Tiempo:** El análisis de desempeño general a lo largo del tiempo muestra fluctuaciones más o menos constantes en el "Overall Performance Score" de las compañías. SaasCompany3 parece tener el desempeño más consistente y generalmente superior en comparación con las demás, mostrando menos caídas abruptas. En cambio, SaasCompany1 y SaasCompany4 tienen más variaciones en su desempeño, lo que podría indicar problemas de estabilidad en sus estrategias. Los picos significativos a mediados de 2022 y mediados de 2023 pueden corresponder a iniciativas específicas que impulsaron temporalmente el rendimiento, pero no fueron sostenidas en el tiempo.
3. **Gráfico de Líneas** **Revenue de Compañías en el Tiempo**: En este gráfico, se puede observar que SaasCompany3 y SaasCompany2 tienden a mantenerse por encima de las demás en términos de revenue. SaasCompany3 muestra una tendencia de crecimiento más sostenida y consistente, mientras que SaasCompany2 presenta un revenue más volátil con varios picos y caídas. Por otro lado, SaasCompany1 y SaasCompany4 muestran ingresos más bajos y menos consistentes, destacándose como las compañías con menor rendimiento. Este gráfico también sugiere que, a pesar de las fluctuaciones, hay oportunidades de crecimiento si las compañías pueden estabilizar sus ingresos, como lo ha hecho SaasCompany3.

**Dashboard 2:**

1. El dashboard analiza métricas avanzadas como la Tasa de Retención en el Tiempo y la Predicción de Revenue, ofreciendo una visión tanto histórica como proyectada para las empresas SaaS. En la tasa de retención, se observan fluctuaciones significativas, con picos en períodos específicos como junio de 2023 (155.142), reflejando esfuerzos exitosos, y caídas como en abril de 2023 (125.022), que podrían deberse a problemas operativos. Aunque muestra mejoras hacia finales de 2023, una línea de tendencia podría aportar mayor claridad.
2. Por otro lado, la predicción de revenue destaca que las compañías están consistentemente superando las expectativas, aunque una ligera disminución al final del período indica posibles retos futuros. En general, el dashboard ofrece información clave para evaluar la estabilidad y el crecimiento proyectado, pero se beneficiaría de un desglose más granular por compañía, anotaciones explicativas y mayor interactividad para identificar los factores detrás de los picos y caídas.

**Dashboard 3:**

El gráfico ofrece una visión útil para analizar y comparar el comportamiento de las métricas seleccionadas (actualmente el CAC, pero puede ser CLV y Revenue) entre las diferentes compañías SaaS. SaasCompany2 destaca por su gasto elevado y picos pronunciados, lo que podría implicar mayores costos de adquisición que deben analizarse en función de su retorno. Por otro lado, SaasCompany4 muestra un desempeño más moderado con los valores más bajos, indicando potencial eficiencia. Las fluctuaciones generales sugieren un mercado dinámico y estrategias de adquisición de clientes ajustadas periódicamente.

Para mejorar la utilidad del dashboard, sería valioso agregar líneas de tendencia, anotaciones en los eventos más relevantes y un desglose más granular. Esto ayudaría a identificar patrones más claros y los factores subyacentes que impulsan el comportamiento de las métricas a lo largo del tiempo. En su estado actual, el dashboard es una herramienta efectiva para comparaciones iniciales, pero puede fortalecerse para un análisis más profundo y estratégico.