



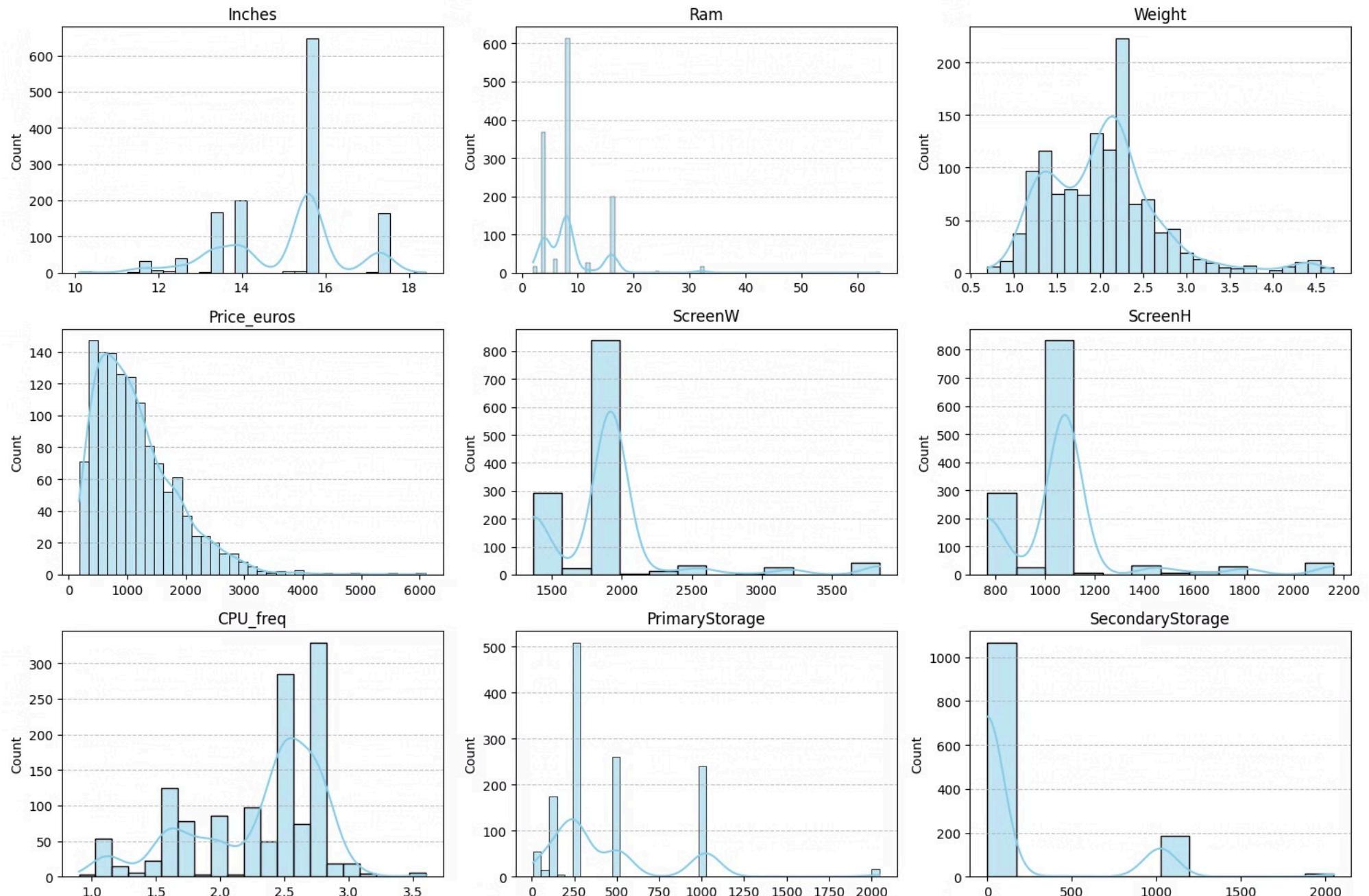
Proyecto final de Estadística: Precio de Laptops

Análisis del precio de laptops. Datos de [Kaggle](#).

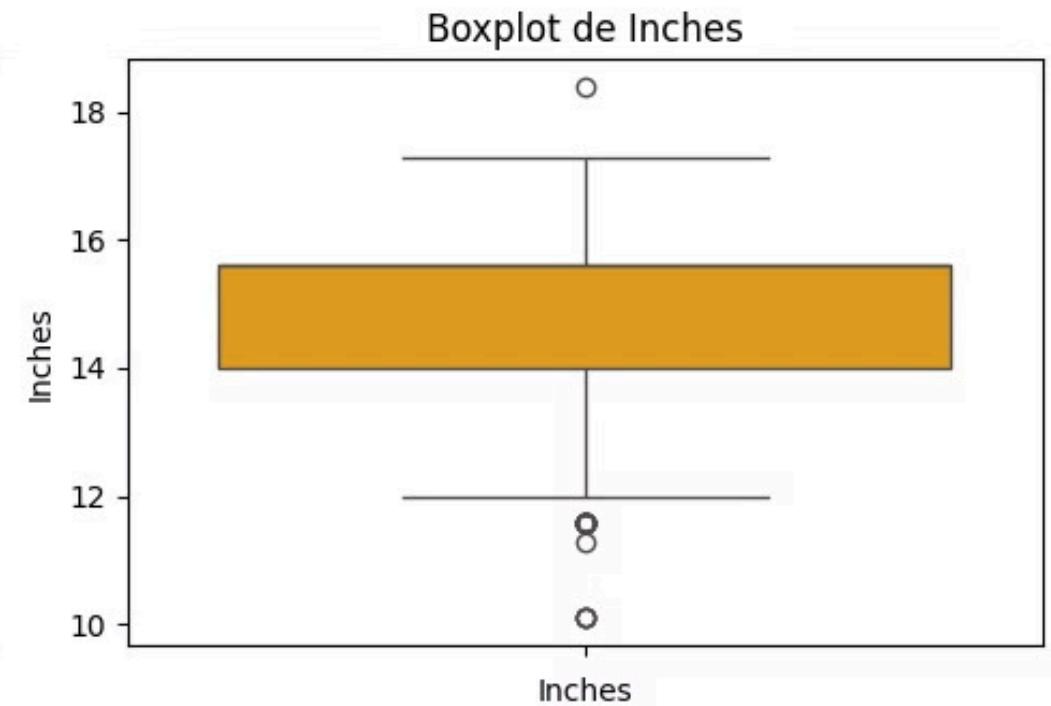
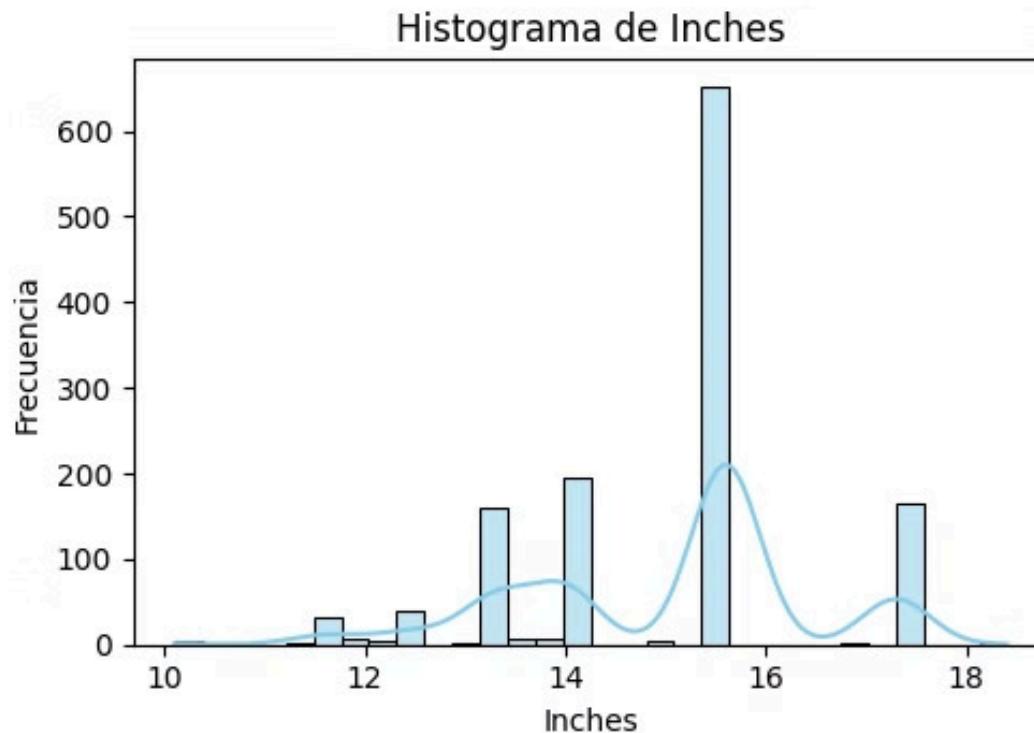
Integrantes:

- Carlos Daniel Largacha Leal C312
- Gabriel Andrés Pla Lasa C311

Análisis Descriptivo: Datos Numéricos



Inches (Tamaño de Pantalla)



1 Rango Típico

La mayoría de las laptops tienen tamaños de pantalla entre 13 y 15 pulgadas, lo que refleja una preferencia por tamaños estándar.

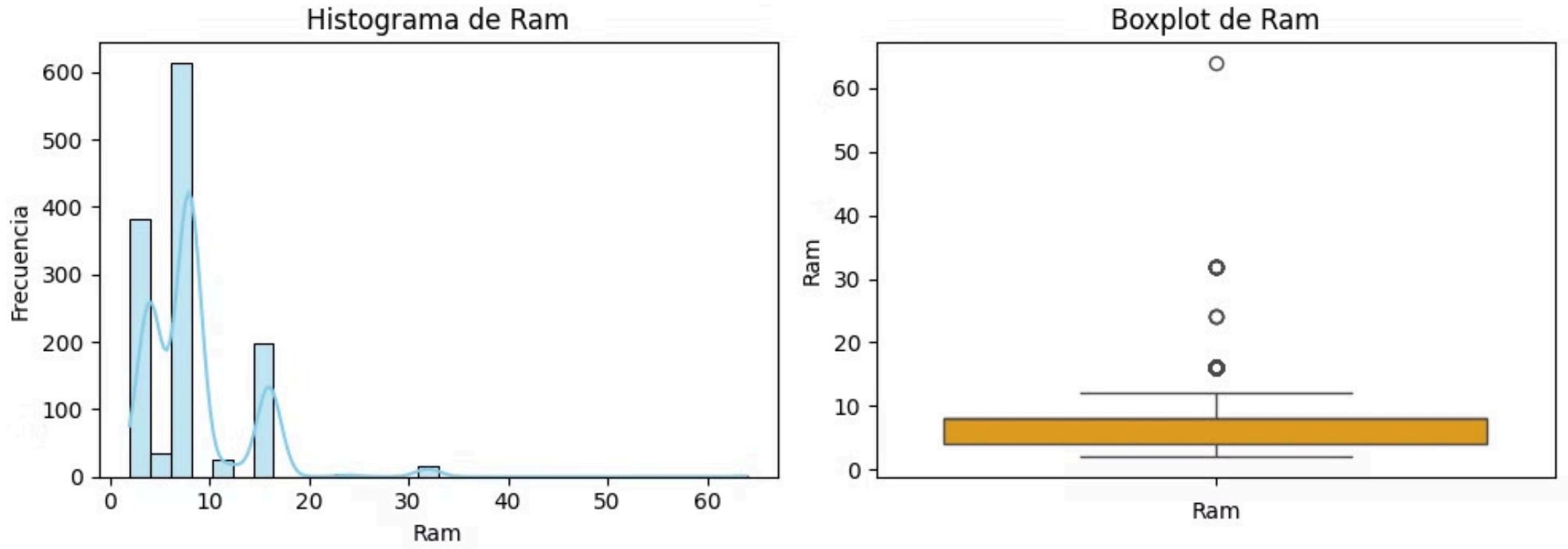
2 Valores Atípicos

Se observan laptops pequeñas (11 pulgadas) y algunas significativamente grandes (17 pulgadas o más). Estos extremos suelen estar orientados a la portabilidad extrema (pequeñas) o al mercado de gaming/diseño (grandes).

3 Distribución

Leve sesgo hacia tamaños más pequeños, indicando que laptops más compactas son ligeramente más frecuentes.

RAM



1 Rango Típico

La mayoría de las laptops tienen entre 8 GB y 16 GB de RAM, que es la configuración estándar en laptops modernas.

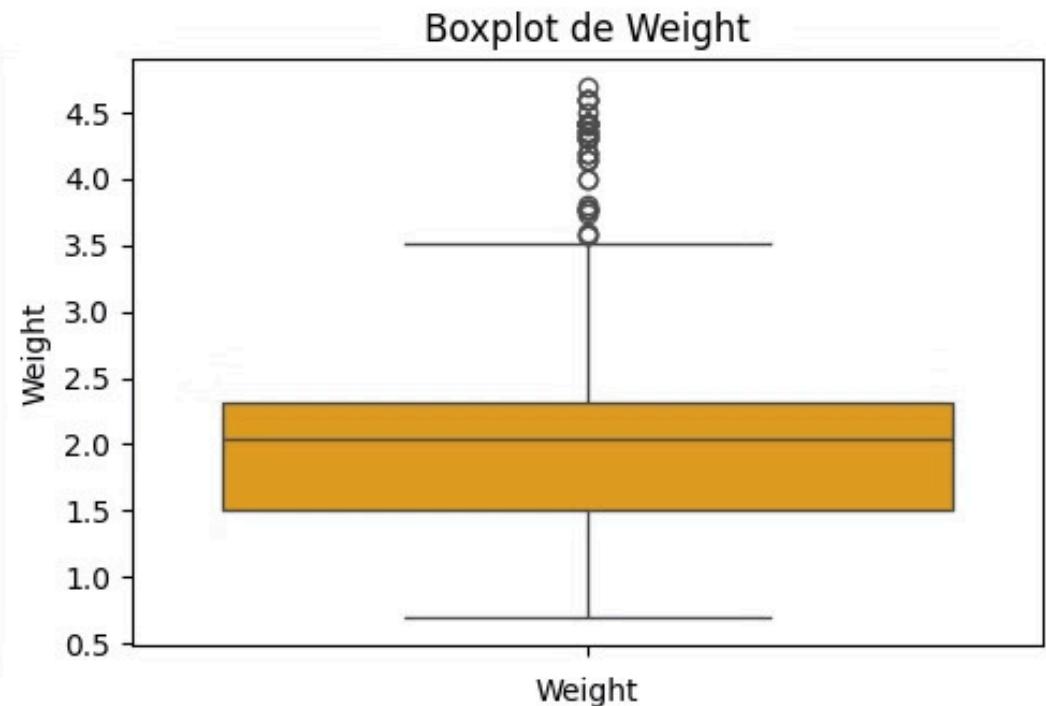
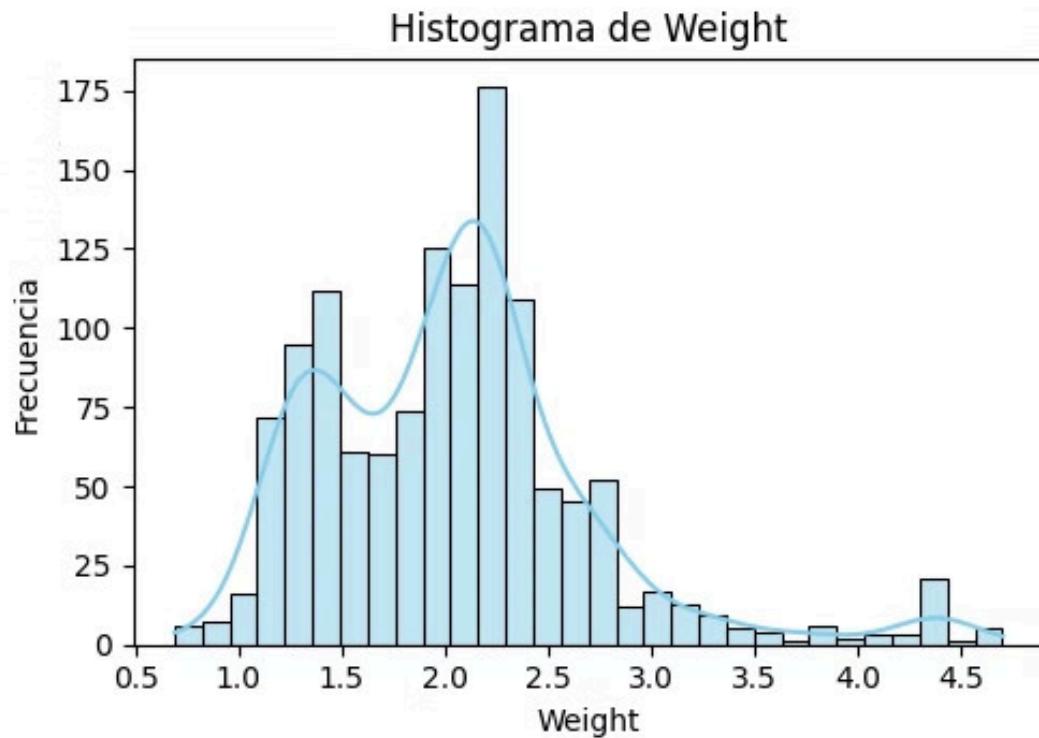
2 Valores Atípicos

Se observan laptops con RAM alta (32 GB o 64 GB), típicas de estaciones de trabajo o gaming de alta gama.

3 Distribución

No hay un sesgo claro, pero los valores extremos resaltan el incremento en la demanda de RAM en modelos premium.

Weight (Peso)



1 Rango Típico

La mayoría de las laptops pesan entre 1 y 2 kg, lo que muestra una tendencia hacia la portabilidad.

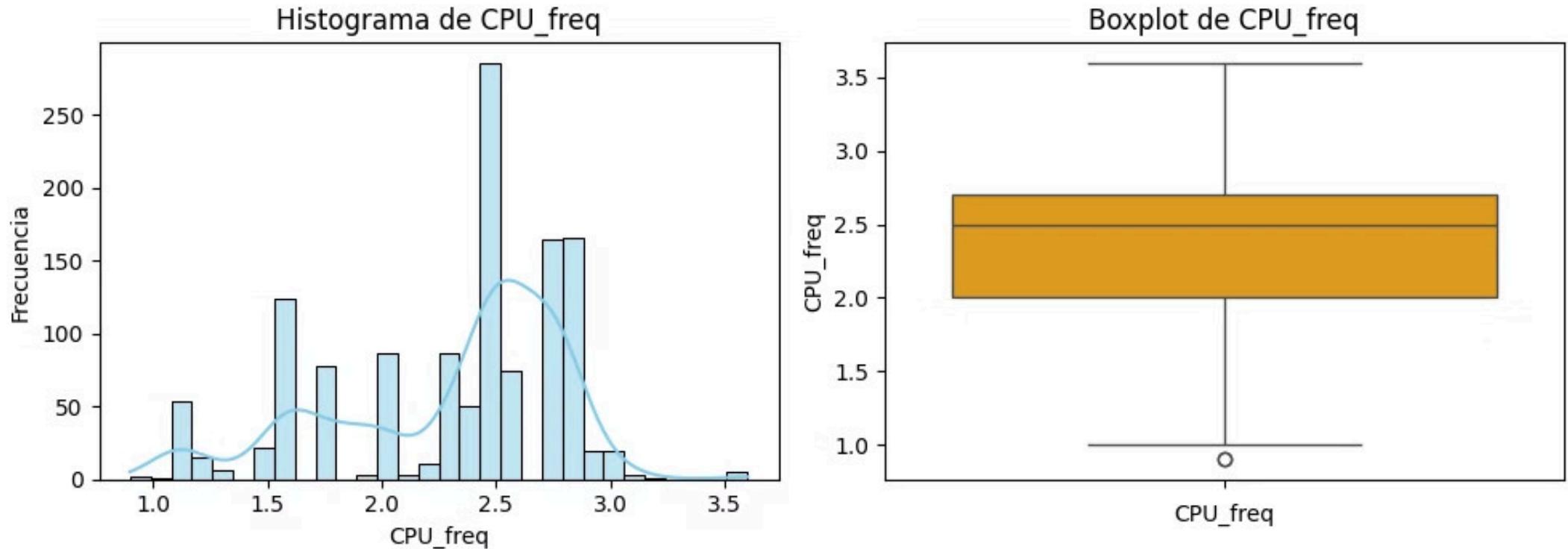
2 Valores Atípicos

Laptops ultraligeras (< 1 kg) o muy pesadas (> 2.5 kg), estas últimas suelen estar relacionadas con gaming o estaciones de trabajo.

3 Distribución

Sesgo leve hacia pesos más bajos, con más frecuencia en laptops ligeras.

Cpu_Freq (Frecuencia del Procesador)



1 Rango Típico

La mayoría de los procesadores tienen frecuencias entre 2.0 y 3.0 GHz.

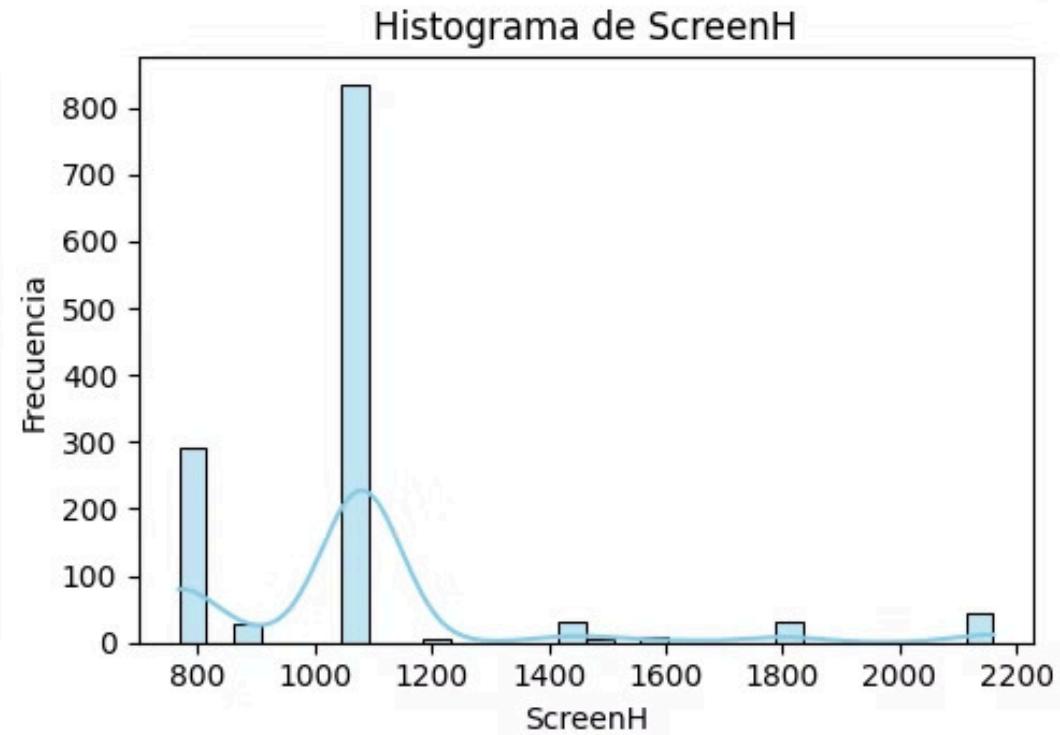
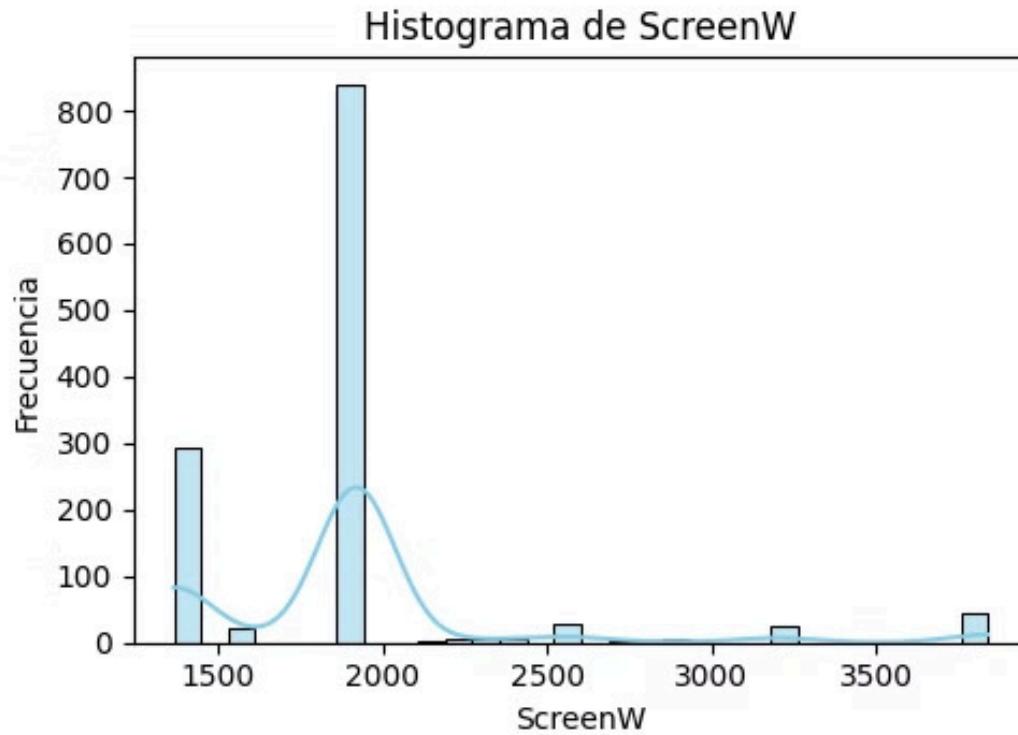
2 Valores Atípicos

Frecuencias superiores (> 4.0 GHz) son raras y corresponden a modelos de alto rendimiento.

3 Distribución

Los datos se concentran en valores intermedios, reflejando un equilibrio entre potencia y eficiencia energética.

Screen Height y Width (Resolución de Pantalla)



1 Rango Típico

Resoluciones estándar Full HD (1920x1080) son las más comunes, con variaciones hacia resoluciones más altas en laptops premium.

2 Valores Atípicos

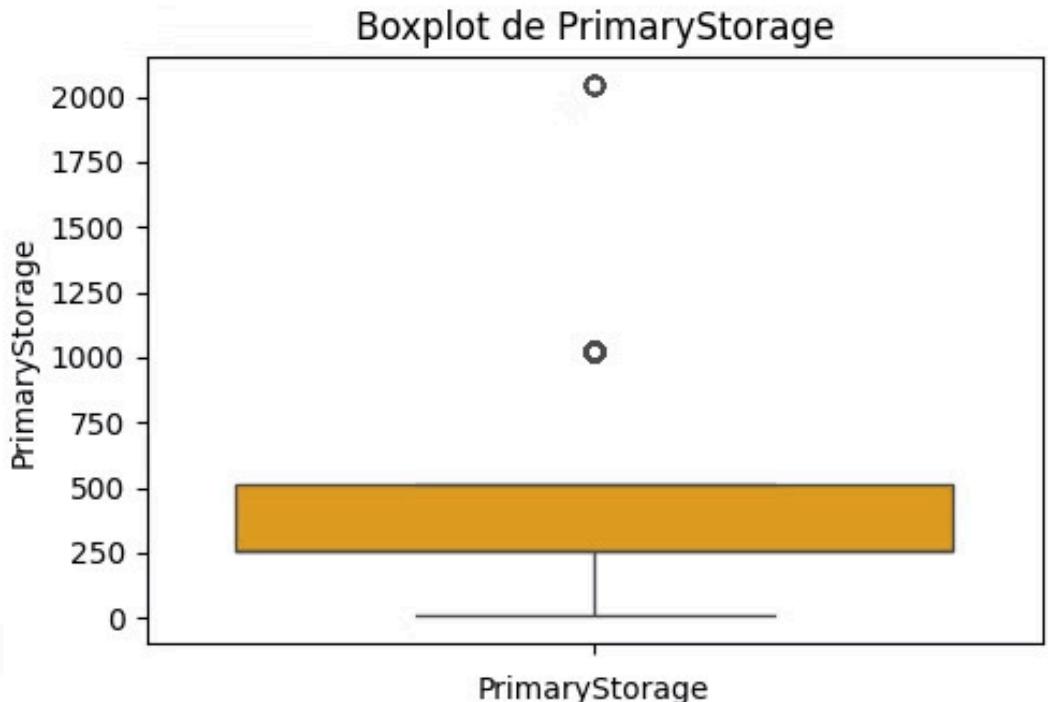
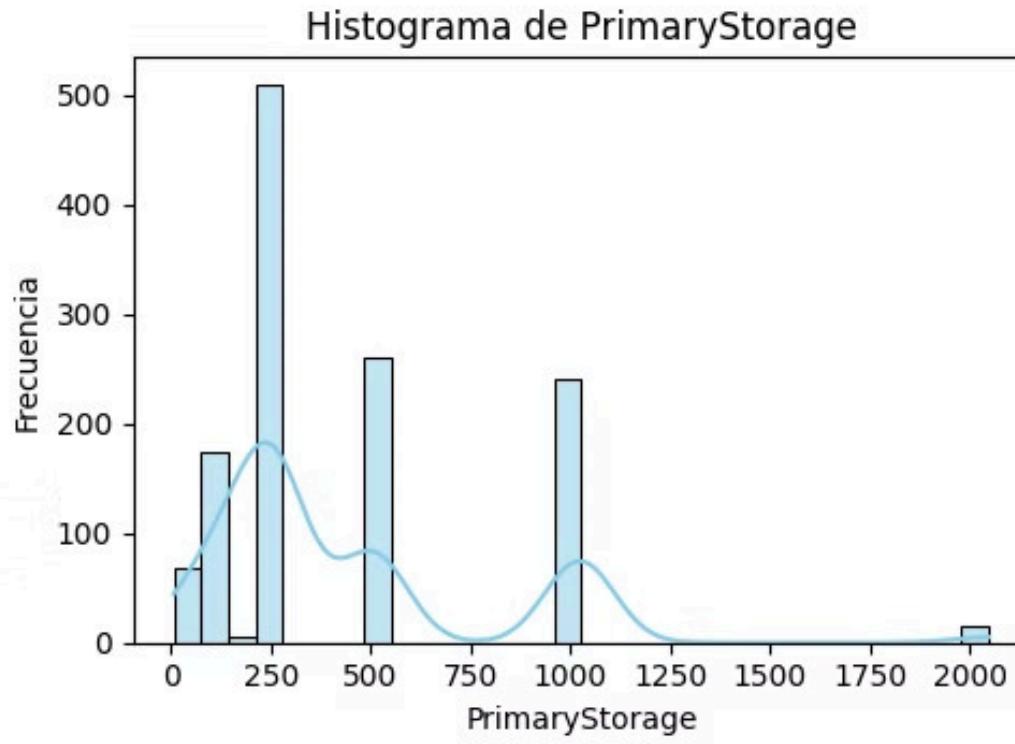
Algunas laptops tienen resoluciones superiores (4K y más), destinadas a tareas profesionales o gaming.

3 Distribución

Las alturas y anchos tienden a alinearse proporcionalmente, con frecuencias más altas en relaciones estándar como 16:9.



Primary Storage (Almacenamiento principal)



1 Rango Típico

Entre 256 GB y 512 GB (SSD) son los más frecuentes.

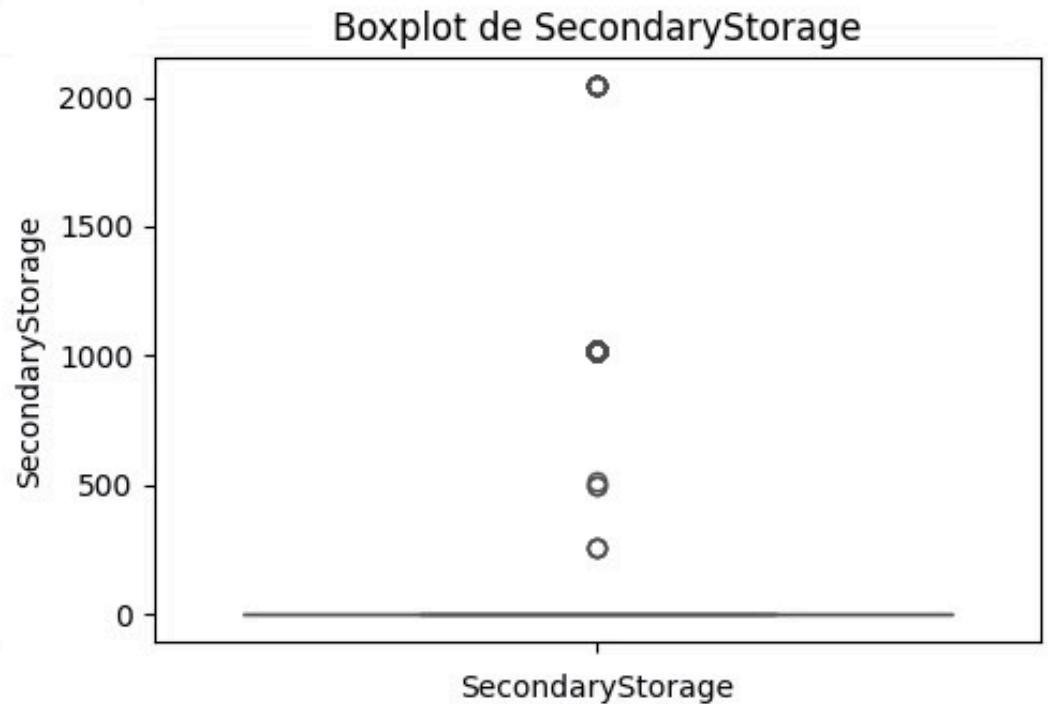
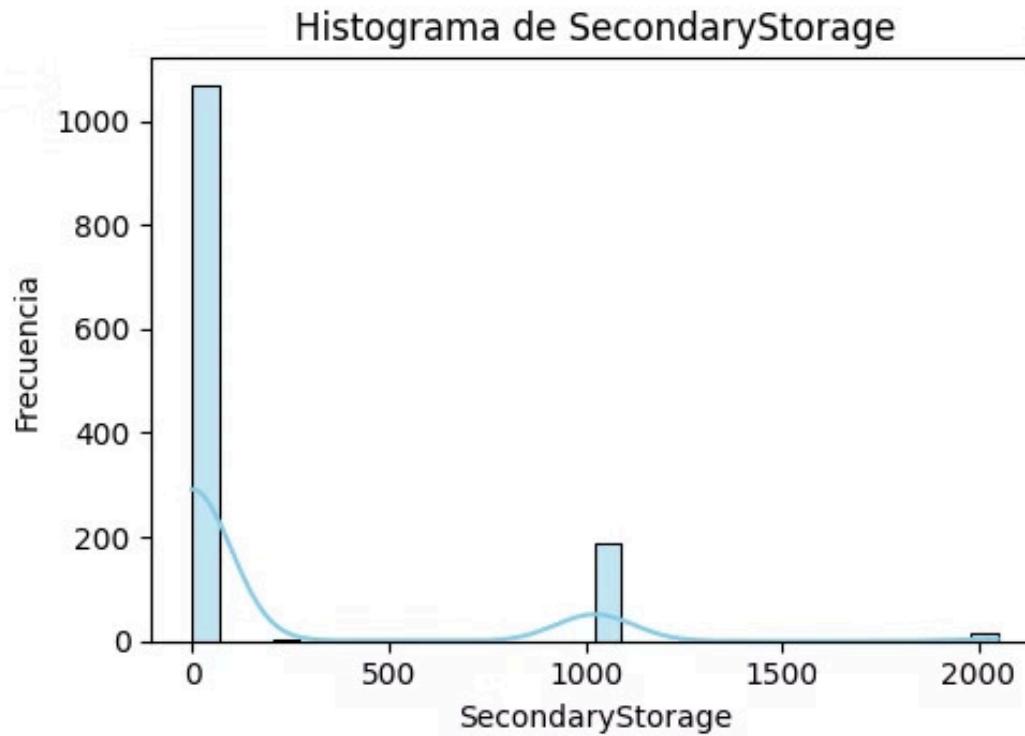
2 Valores Atípicos

Modelos con almacenamiento de 1 TB o más, diseñados para usuarios que manejan grandes volúmenes de datos.

3 Distribución

Se observa una amplia variabilidad, indicando opciones diversas para necesidades específicas.

Secondary Storage (Almacenamiento secundario)



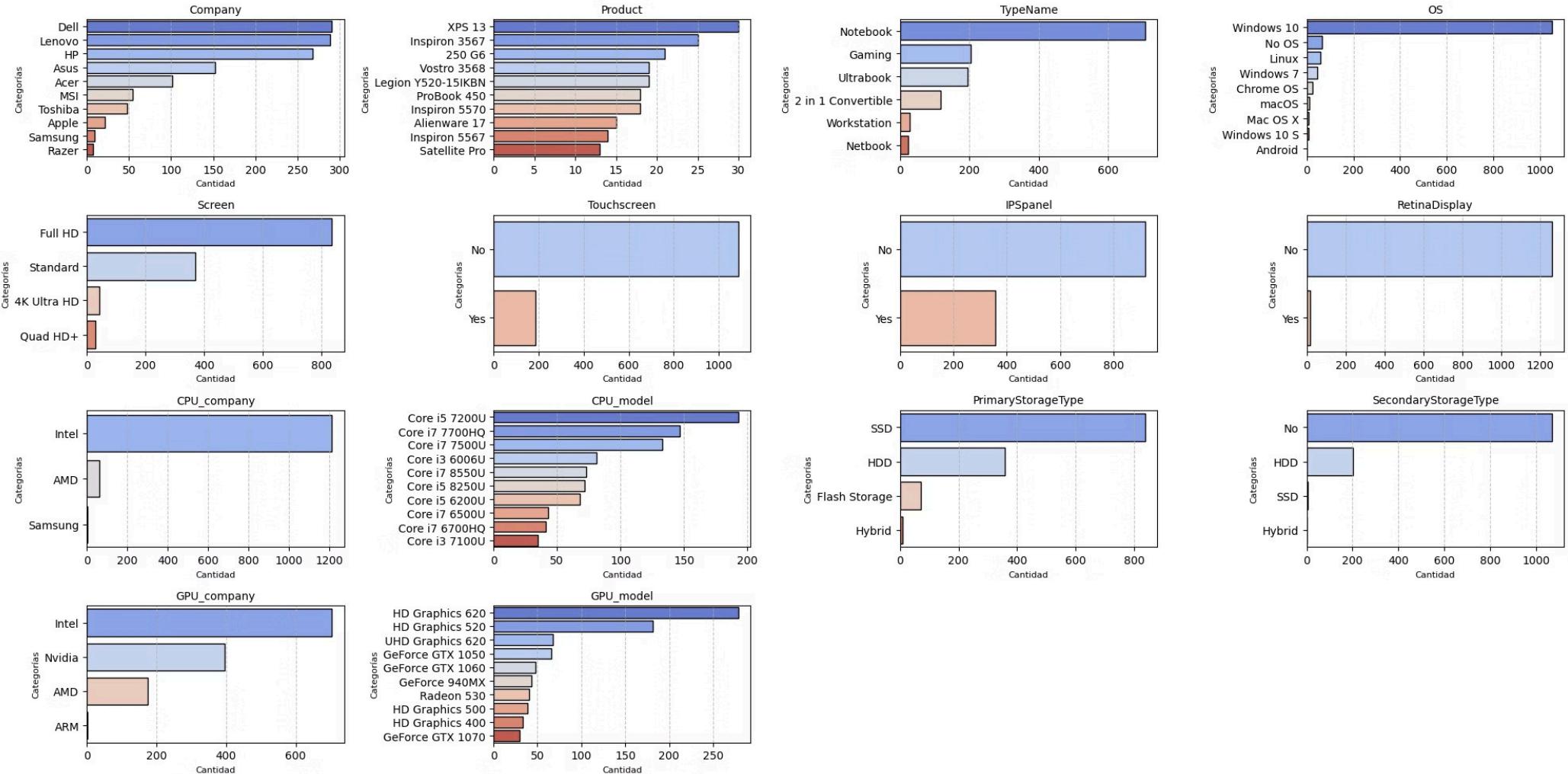
1 Rango Típico

Muchas laptops no tienen almacenamiento secundario. Aquellas que sí lo tienen suelen usar discos HDD de alta capacidad (> 1 TB).

2 Distribución

La presencia de almacenamiento secundario es menos común y depende del enfoque del modelo (gaming o estaciones de trabajo).

Análisis Descriptivo: Datos Categóricos



Análisis Descriptivo: Columnas Categóricas

1 Marcas y Fabricantes

- **Marcas dominantes:** HP, Dell y Lenovo son las marcas más representadas, lo que indica su liderazgo en el mercado.
- **Marcas premium:** Apple tiene una menor frecuencia, probablemente debido a su enfoque en un segmento premium o de nicho.
- **Diversidad de fabricantes:** Hay varias marcas menos conocidas, pero su presencia es minoritaria.

2 Modelos y Productos

- **Alta variedad:** La distribución de modelos es muy dispersa, con muchos modelos específicos que aparecen pocas veces.
- **Modelos populares:** Algunos modelos como **Inspiron** (Dell) y **Pavilion** (HP) destacan en frecuencia debido a su popularidad.
- **Mercado diverso:** La alta variedad de modelos refleja un mercado con muchas opciones, dificultando identificar líderes claros más allá de las marcas principales.

3 Tipos de Laptops

- **Portátiles ligeros:** Las categorías más comunes son **Notebook** y **Ultrabook**, lo que indica una preferencia por dispositivos portátiles y versátiles.
- **Segmentos especializados:** Las categorías **Gaming** y **Workstation** son menos frecuentes, dirigidas a usuarios con necesidades específicas.
- **Tendencia híbrida:** La categoría **2 in 1 Convertible** destaca, reflejando la popularidad de dispositivos híbridos.

4 Sistema Operativo (OS)

- **Dominio de Windows:** Windows es el sistema operativo más común, seguido de macOS, lo que confirma su popularidad en laptops comerciales y de uso general.
- **Nichos específicos:** Sistemas como Linux y Chrome OS tienen una representación marginal, probablemente dirigidos a usuarios especializados.

5 Tarjetas Gráficas (GPU_company y GPU_model)

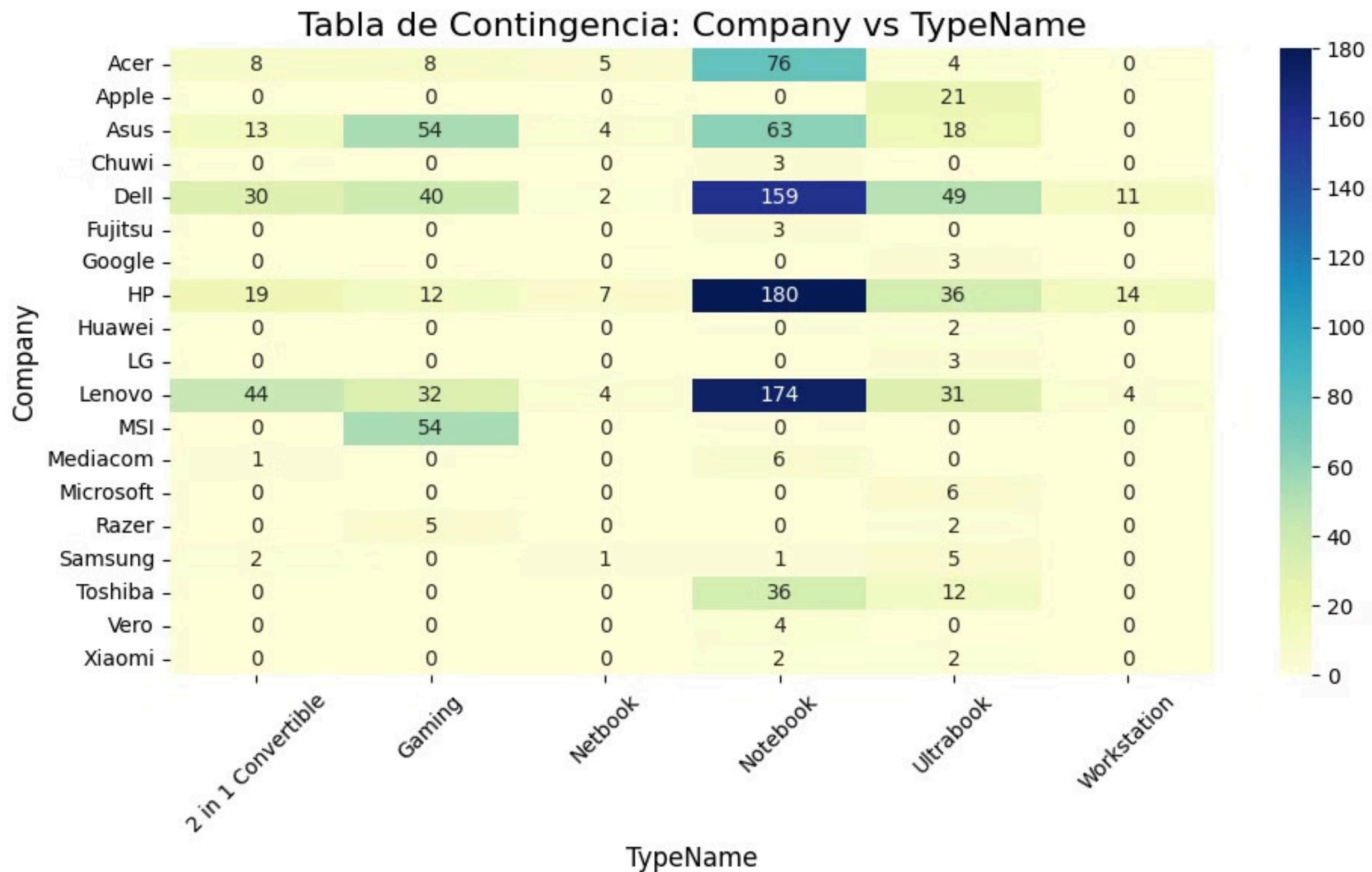
- **NVIDIA líder en GPUs dedicadas:** NVIDIA domina en GPUs dedicadas, especialmente en el segmento de gaming y computación gráfica.
- **Intel en GPUs integradas:** Intel lidera en GPUs integradas, siendo una opción estándar para tareas no intensivas.
- **AMD minoritaria:** AMD tiene una presencia menor, pero aún relevante en el mercado de GPUs.

6 Procesadores (CPU_company y CPU_model)

- **Dominio de Intel:** Intel lidera ampliamente el mercado de procesadores, seguido por AMD en una proporción significativamente menor.
- **Diversidad de modelos:** Existe una gran variedad de modelos de CPU, con **Core i5** e **i7** de Intel probablemente liderando por su equilibrio entre costo y rendimiento.

Análisis de los tipos de laptop según la compañía

Primeramente, vamos a realizar un análisis de como se encuentra distribuido el mercado de las laptops según su tipo y compañía



Conclusiones de la tabla de contingencia

Dominancia de Notebooks

La categoría "Notebook" es la más representada, especialmente en HP (180), Lenovo (174), Dell (159) y Acer (76). Refleja el enfoque de los fabricantes en esta línea de productos.

Especialización en Gaming

MSI es un fabricante especializado en laptops gaming, con 54 productos. Asus y Dell también tienen una fuerte presencia en esta categoría con 54 y 40 productos, respectivamente.

Ultrabooks

Apple lidera en la categoría de Ultrabook con 21 productos, seguido por Dell (49) y HP (36). Refleja el enfoque de estos fabricantes en laptops premium.

2-in-1 Convertibles

Lenovo destaca con 44 productos en esta categoría, seguido por Dell (30) y HP (19). Refleja la innovación en laptops versátiles.

Netbooks

Esta categoría tiene poca representación, con solo algunos productos de Acer (5), Asus (4) y HP (7). Sugiere que el mercado de netbooks está en declive.

Workstations

Las Workstations tienen una representación limitada, siendo HP (14) y Dell (11) los principales fabricantes. Refleja que es una categoría de nicho para usuarios profesionales.

Analisis de los test sobre la tabla de contingencia

Test de Homogeneidad

Evalúa si las distribuciones de los tipos de laptops son similares entre los fabricantes.

Hipótesis:

- H_0 : La distribución de los tipos de laptops es la misma entre los fabricantes.
- H_1 : La distribución de los tipos de laptops difiere entre los fabricantes.

Resultado: p-value muy bajo (7.66×10^{-95}), rechazamos H_0 . Implica que las distribuciones difieren significativamente entre fabricantes.

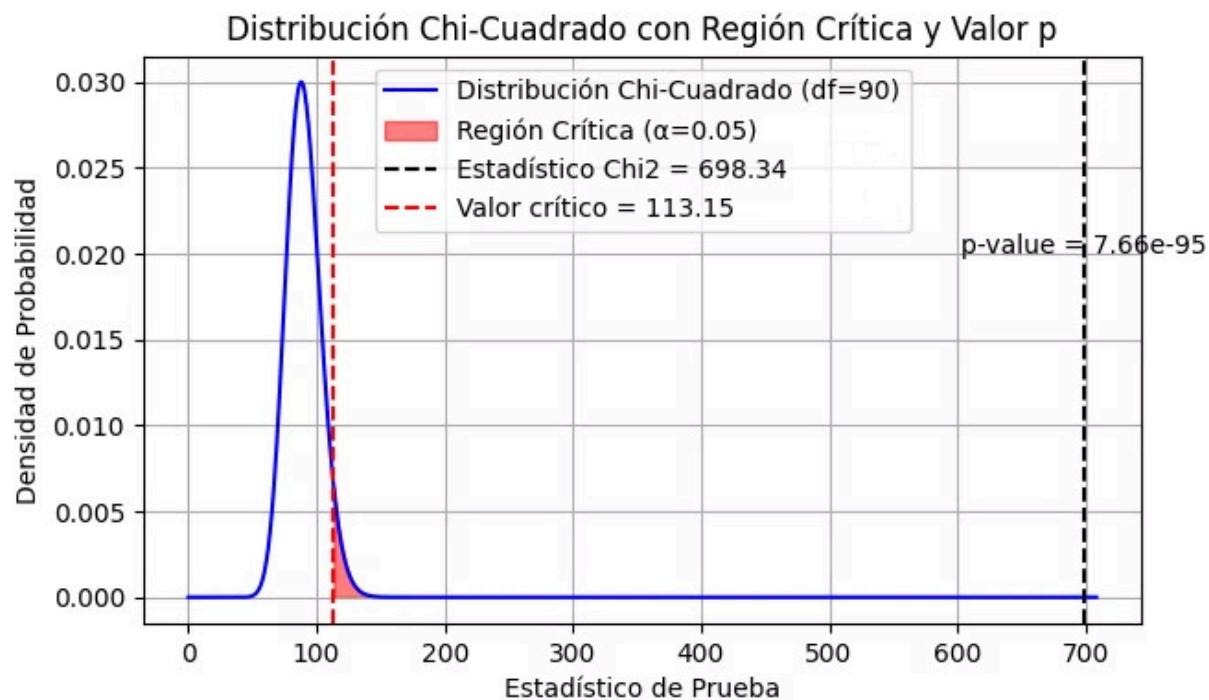
Test de Independencia

Evalúa si hay relación entre el fabricante (Company) y el tipo de laptop (TypeName).

Hipótesis:

- H_0 : Las variables Company y TypeName son independientes.
- H_1 : Las variables Company y TypeName están relacionadas.

Resultado: p-value muy bajo (7.66×10^{-95}), rechazamos H_0 . Sugiere que existe una relación significativa entre fabricante y tipo de laptop.

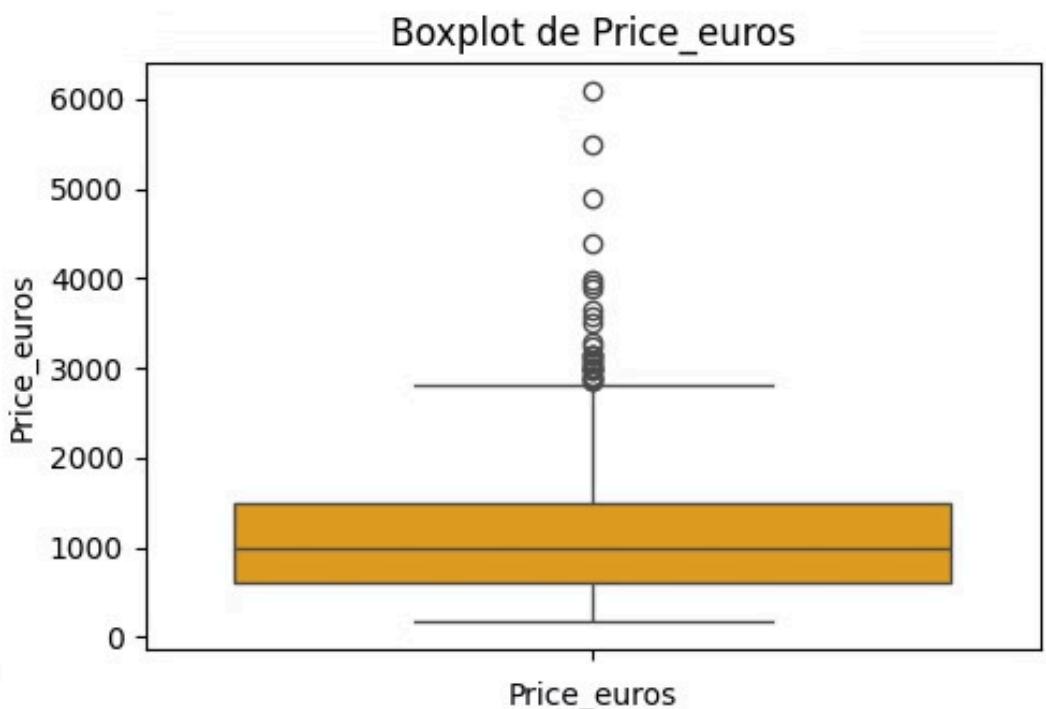
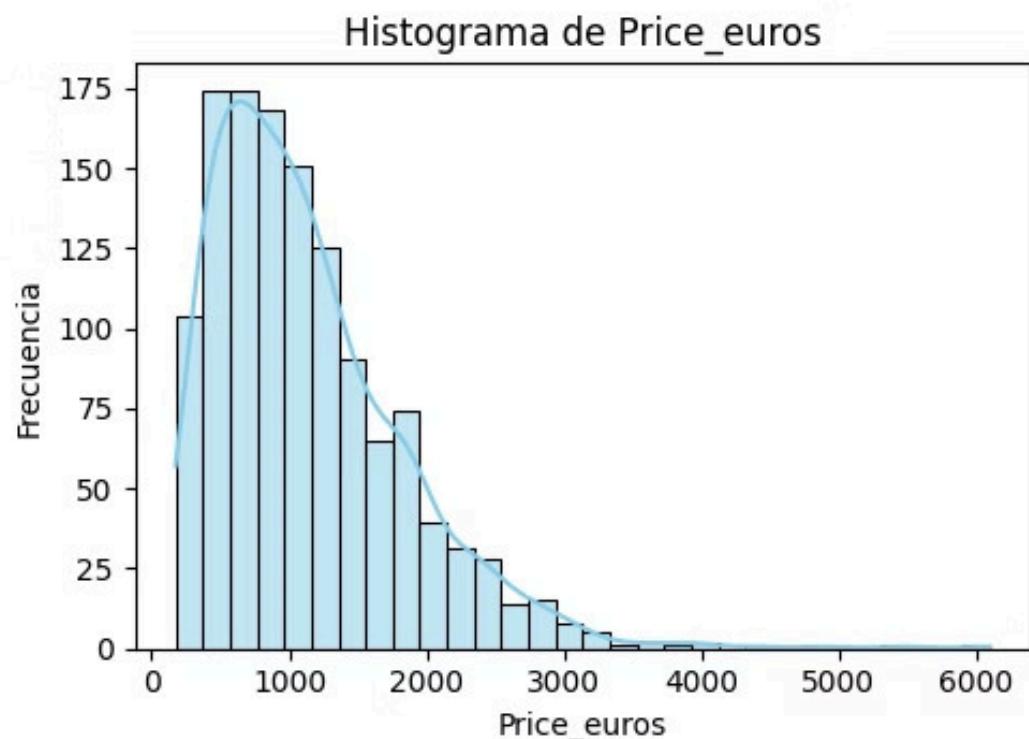


Conclusiones

Distribuciones no homogéneas: Los tipos de laptops no se distribuyen de manera homogénea entre los fabricantes. Algunos se especializan como lo muestran el análisis de la tabla de contingencia.

Relación significativa: Existe una relación significativa entre el fabricante y el tipo de laptop. Esto refleja estrategias de mercado y especializaciones claras de las empresas.

Analisis de los precios



Histograma de Price_euros

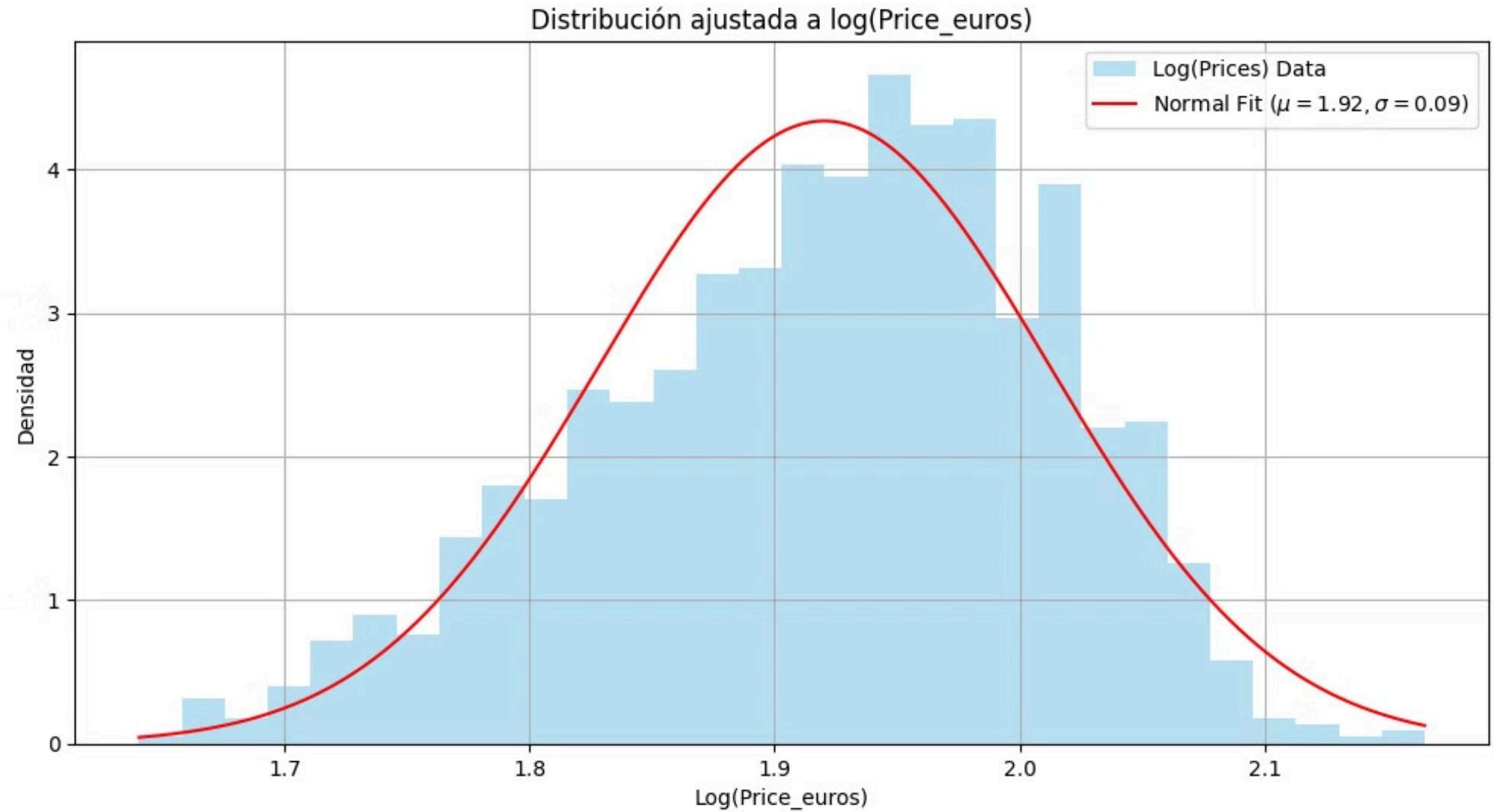
- Distribución asimétrica (sesgo positivo):**
 - El histograma muestra que la mayoría de los precios están concentrados en el rango entre 500 y 1500 euros.
 - La cola derecha indica la presencia de algunos laptops con precios significativamente más altos (modelos premium).
- Densidad baja en precios altos:**
 - Hay una caída drástica en la frecuencia conforme los precios aumentan más allá de los 2000 euros, sugiriendo que estos modelos no son comunes en el mercado.
- Pico en el rango intercuartílico (IQR):**
 - Los valores más densos parecen estar entre 800 y 1200 euros, que es el rango típico para laptops de gama media.

Boxplot de Price_euros

- Presencia de valores atípicos (outliers):**
 - Hay múltiples puntos fuera del límite superior del bigote, indicando modelos con precios excepcionalmente altos, es decir equipos diseñados para propósitos específicos como gaming extremo o edición profesional.
- Rango intercuartílico (IQR):**
 - El rango típico de precios (IQR) está entre 500 y 1500 euros, lo que incluye la mayoría de las laptops estándar.
- Mediana de precios:**
 - La línea dentro del box indica que la mediana de precios está por debajo del centro del IQR, lo que confirma un sesgo positivo.
- Simetría del rango inferior:**
 - El bigote inferior está más cerca del rango intercuartílico, lo que refleja que hay pocos modelos con precios extremadamente bajos.

Distribución de prices_euros

Según el histograma de la columna *prices_euros*, parece que esta variable no sigue una distribución normal debido a su asimetría positiva (sesgo hacia la derecha). Sin embargo, con base en su forma, podría seguir una distribución Log-Normal.



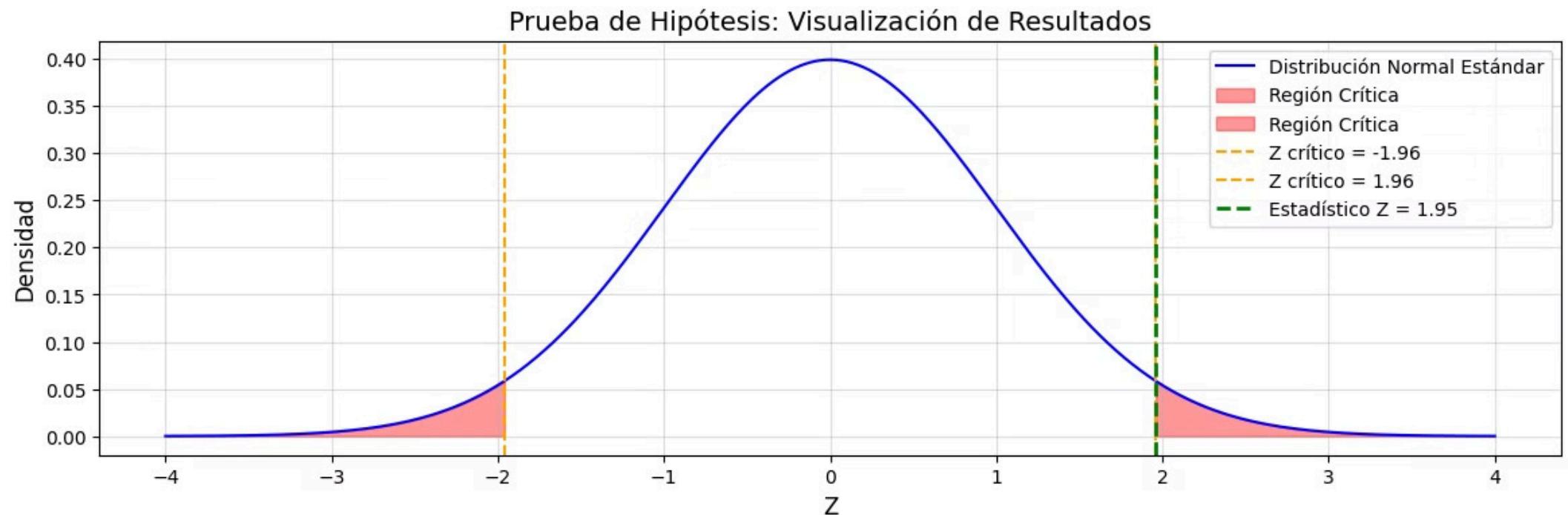
Con base en el ajuste realizado, los datos transformados mediante el logaritmo natural parecen seguir una distribución normal razonablemente bien. Esto implica que los precios originales se ajustan a una distribución log-normal.

Una distribución log-normal significa que los precios bajos son más frecuentes, mientras que los precios altos ocurren con menor frecuencia pero tienen un impacto significativo. Este comportamiento es común en mercados donde los valores extremos (como laptops premium) tienen gran importancia en comparación con productos regulares.

**Nota: a partir de aquí se tratará a Price_euros como el logaritmo y el original se denominará Price_euros_ori

Prueba de Bondad de Ajuste de Chi-Cuadrado

Entonces con el objetivo de corroborar lo concluido, se va realizar una prueba de bondad de ajuste contra una distribución log-normal.

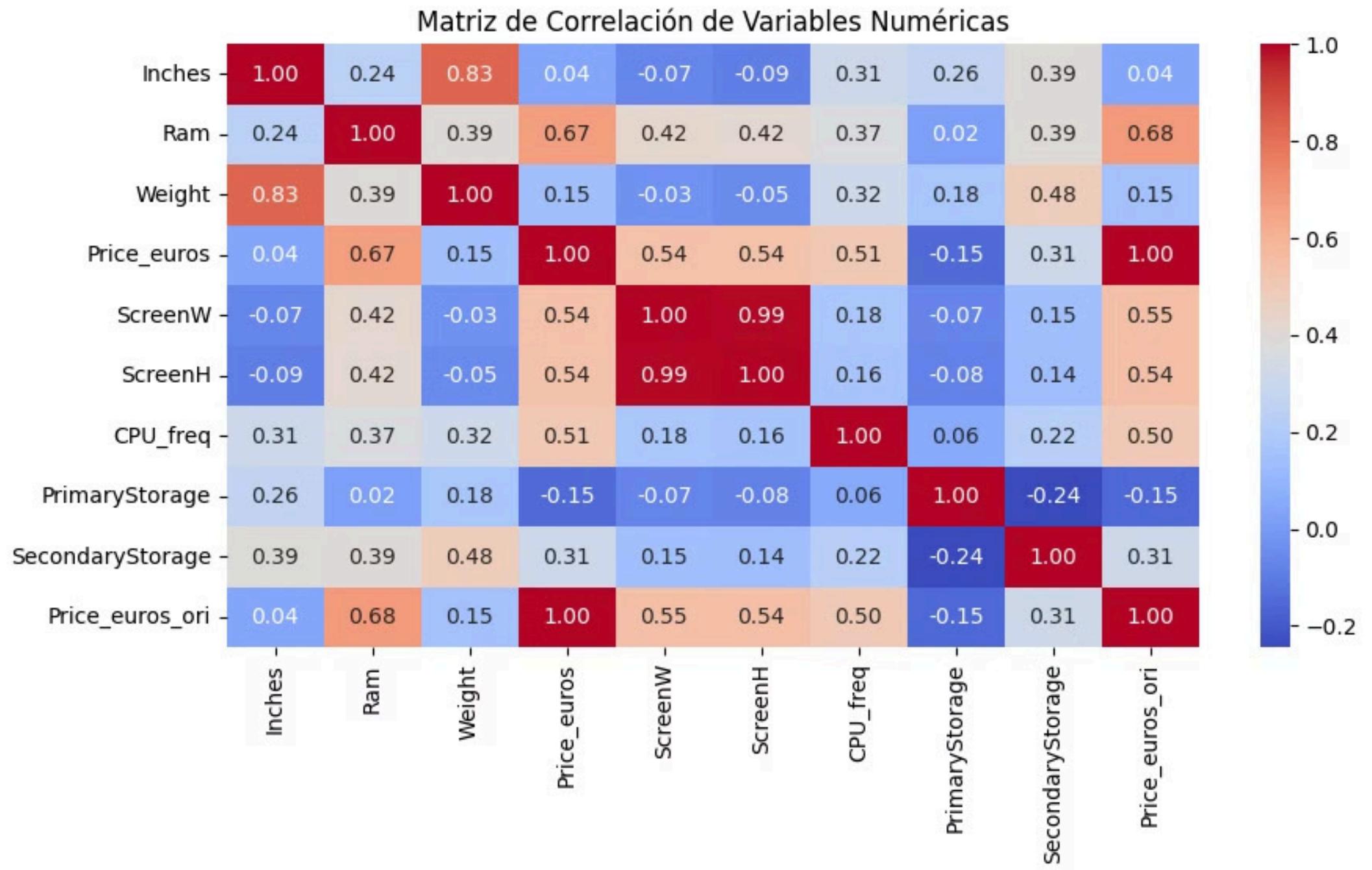


A pesar que los datos no siguen una distribución log-normal, se puede notar que el estadístico está cercano al umbral, lo cual indica que los datos casi pasan la prueba de bondad de ajuste para una distribución log-normal, pero no lo hacen de manera concluyente.

Es decir, este valor sugiere que hay cierta evidencia de que la distribución log-normal podría ser un modelo razonable, pero no es suficientemente fuerte como para rechazar la hipótesis nula de que los datos no siguen una distribución log-normal.

Matriz de correlación de variables numéricas

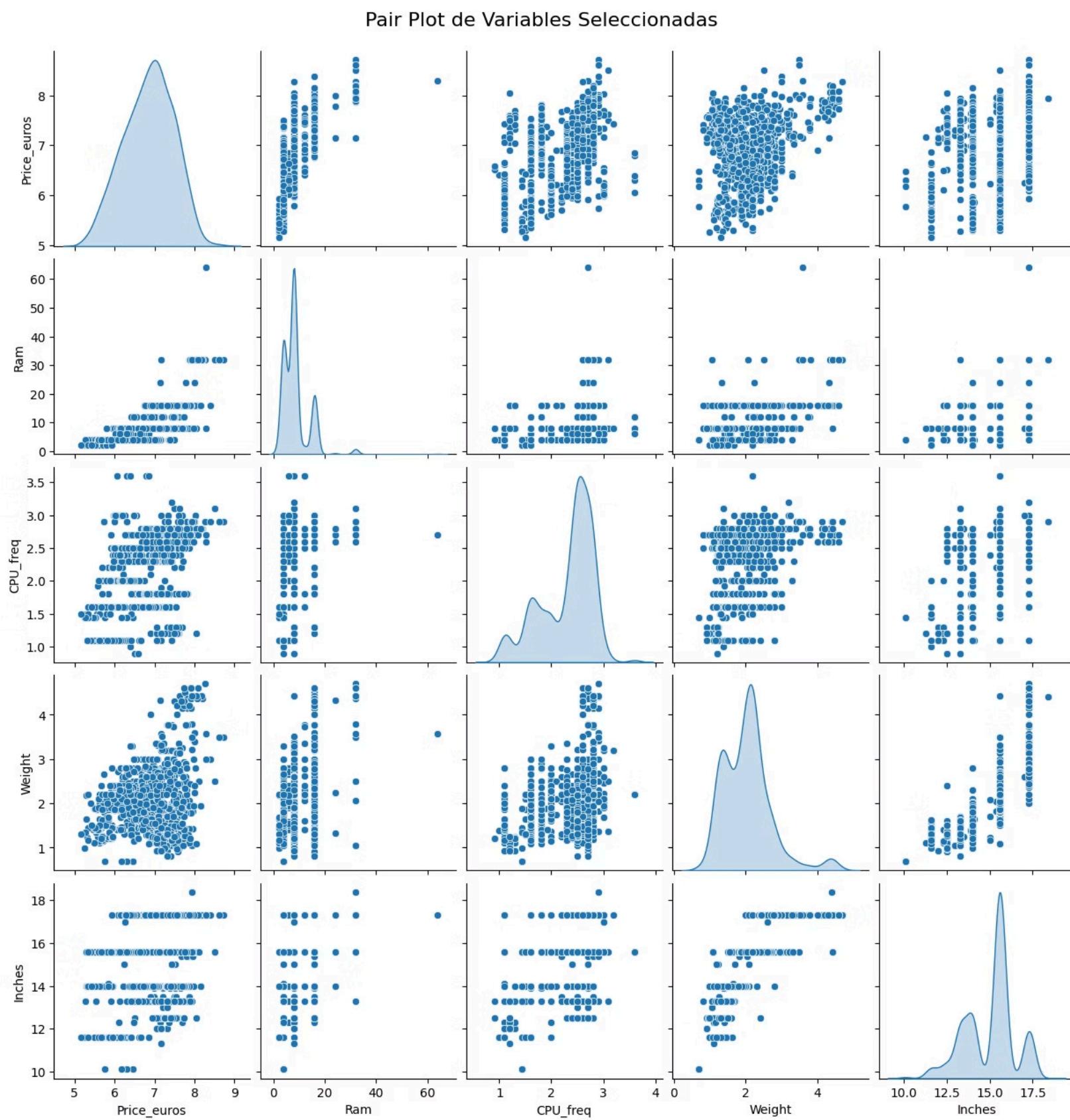
Luego de analizar la distribución que posee Price_euros, queremos ver cuales son las variables que mas influyen en Price_euros. Y para ello grafiquemos la matriz de correlación correspondiente.



Basandonos en las conclusiones del análisis de la matriz de correlación, las variables más prometedoras e interesantes para analizar en el pair plot son aquellas que presentan correlaciones fuertes o moderadas y tienen una relación significativa con el precio o las características físicas y de rendimiento. Estas variables son:

1. **Ram**: Tiene una correlación moderada-alta con el precio.
2. **CPU_freq**: Relacionada moderadamente con el precio.
3. **Weight**: Fuerte relación con características físicas como el tamaño de la pantalla.
4. **Inches**: Aunque su correlación con el precio es baja, tiene una fuerte correlación con el peso.

Pair Plot de las variables relacionadas con los precios



Observaciones sobre la relación entre variables y precios

1 RAM:

Los valores de RAM están concentrados en ciertos puntos específicos (4GB, 8GB, 16GB, 32GB), lo que refleja las configuraciones más comunes en laptops.

3 Peso (Weight):

La distribución es sesgada a la derecha, con la mayoría de las laptops pesando entre 1 y 2.5 kg, pero con algunas ultraligeras y otras más pesadas.

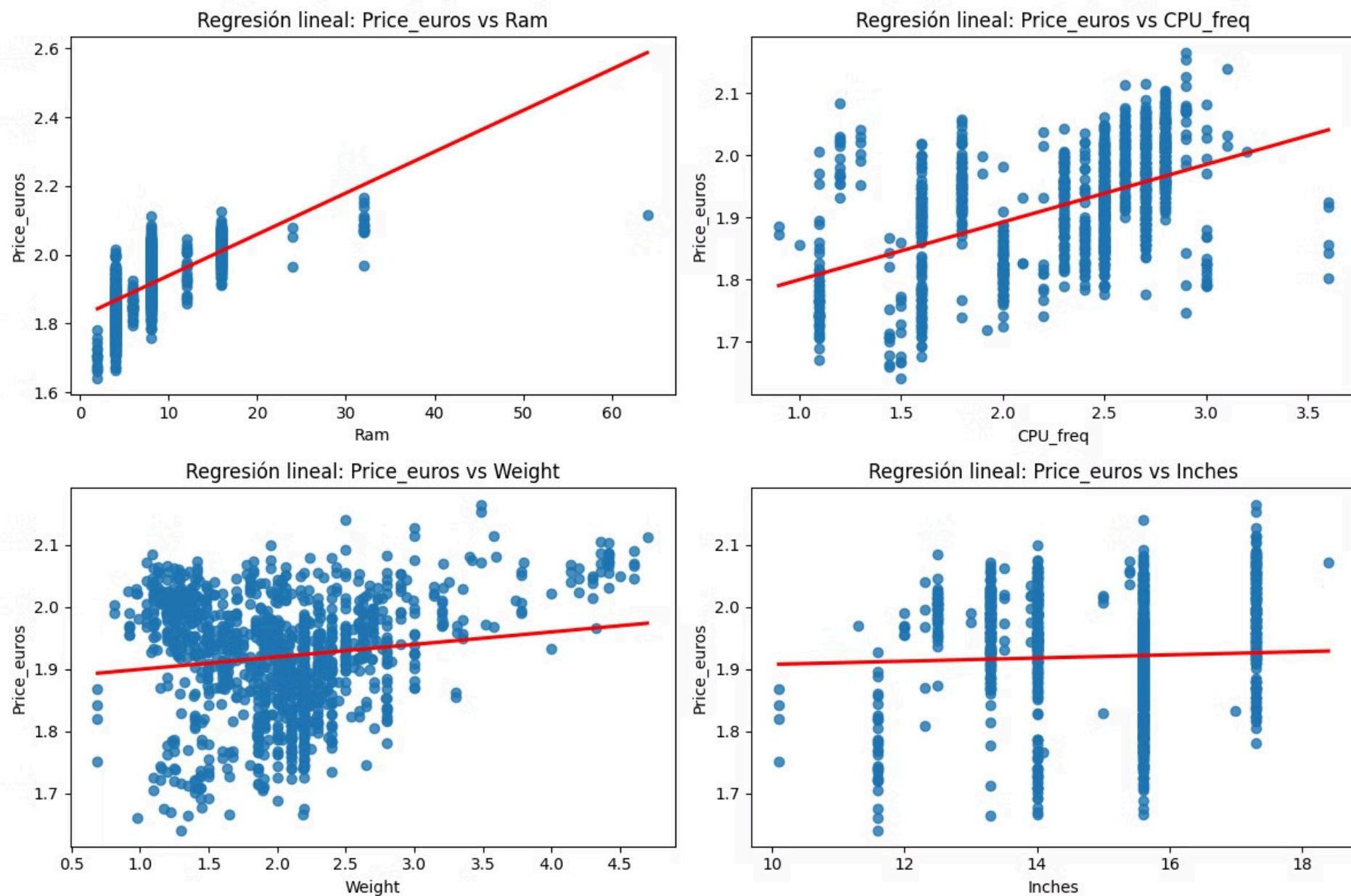
2 Frecuencia del CPU (CPU_freq):

Tiene una distribución unimodal, con una mayoría de procesadores operando entre 1.5 GHz y 3.0 GHz, y pocos valores extremos más altos.

4 Tamaño de Pantalla (Inches):

Los tamaños de pantalla muestran una distribución multimodal, con concentraciones en 13, 15 y 17 pulgadas, que son los tamaños más comunes.

Linealidad variables relacionadas con los precios



1 RAM (`Ram`):

Existe una **fuerte relación positiva** entre la cantidad de RAM y el precio. A medida que aumenta la RAM, el precio también tiende a incrementarse significativamente. Esto indica que la RAM es un factor importante que influye en el costo de los laptops.

3 Peso (`Weight`):

Existe una **relación positiva muy débil** entre el peso del laptop y su precio. Esto podría deberse a que el peso no es un factor determinante para el costo, o que los laptops más pesados pertenecen a gamas más económicas o intermedias.

2 Frecuencia del CPU (`CPU_freq`):

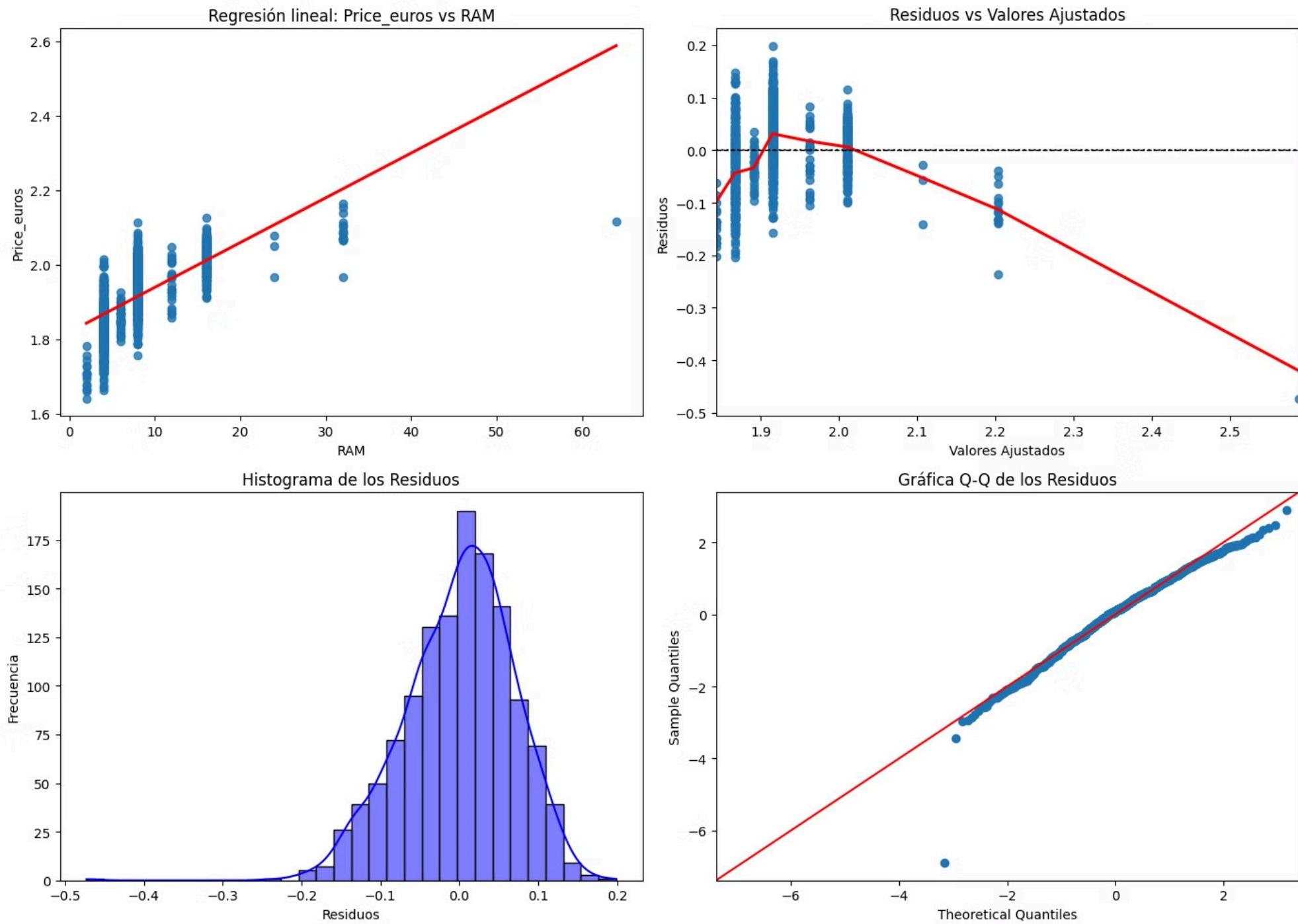
La relación es **positiva pero débil**, lo que sugiere que una mayor frecuencia de CPU tiende a asociarse con precios más altos, aunque no de manera tan pronunciada como la RAM. Esto podría deberse a que otros factores del CPU (como el número de núcleos o la arquitectura) tienen más peso en el precio que la frecuencia por sí sola.

4 Tamaño de la pantalla (`Inches`):

La relación es **muy débil y prácticamente nula**. El tamaño de la pantalla no parece influir significativamente en el precio de los laptops. Esto podría indicar que los laptops con pantallas más grandes no necesariamente son más caros, ya que el precio puede depender más de otros factores internos (hardware).

Analisis de la linealidad entre la RAM y los precios

Dado que la RAM y Price_euros paracen tener una fuerte relacion, comprobemos cuan lineal es realmente esa relacion.



Analizando las gráficas generadas podemos sacar las siguientes conclusiones:

Residuos vs Valores Ajustados

Se observa un patrón curvo en los residuos, lo que indica que la relación podría no ser estrictamente lineal. Además, la dispersión de los puntos no es uniforme, estrechándose hacia valores altos de precios, lo que sugiere heterocedasticidad.

Histograma de Residuos

Si bien el histograma muestra una distribución aproximadamente normal de los residuos, con una curva suave y simétrica, existen pequeños picos y ligeras desviaciones que podrían indicar que la suposición de normalidad no se cumple del todo.

Por lo tanto, aunque la RAM parece tener una fuerte relación con los precios, estos análisis preliminares sugieren que la relación podría no ser completamente lineal.

Análisis de los supuestos en el modelo de regresión

Y para corroborar los resultados de la regresión hasta ahora, vamos a comprobar los supuestos.

▼ 1. Independencia de los Errores

Se aplica el **test Durbin-Watson**:

- H_0 : No hay correlación entre los residuos.
- H_1 : Existe autocorrelación en los residuos.

Resultado: Durbin-Watson = 1.97 → **Los residuos son independientes.**

▼ 2. Esperanza de los Errores Igual a Cero

Se utiliza un **test de media poblacional**:

- $H_0: \mu_0 = 0$.
- $H_1: \mu_0 \neq 0$.

Resultado: P-value = 1.00 → **No hay evidencia para rechazar H_0 , por lo que el valor esperado de los residuos es 0.**

▼ 3. Homocedasticidad

Se emplea el **test de Breusch-Pagan**:

- H_0 : Existe homocedasticidad (varianza constante).
- H_1 : No hay homocedasticidad (presencia de heterocedasticidad).

Resultado: P-value = 1.18×10^{-7} → **Se rechaza H_0 , indicando heterocedasticidad en los residuos.**

▼ 4. Normalidad de los Errores

Se aplica el **Normality Test** (para $n \geq 30$):

- H_0 : Los residuos siguen una distribución normal.
- H_1 : Los residuos no siguen una distribución normal.

Resultado: P-value = 3.29×10^{-22} → **Se rechaza H_0 , indicando que los datos no siguen una distribución normal.**

Análisis de los resultados del modelo de regresión

OLS Regression Results						
Dep. Variable:	Price_euros	R-squared:	0.566			
Model:	OLS	Adj. R-squared:	0.564			
Method:	Least Squares	F-statistic:	413.3			
Date:	Fri, 14 Feb 2025	Prob (F-statistic):	4.64e-228			
Time:	19:41:07	Log-Likelihood:	1765.3			
No. Observations:	1275	AIC:	-3521.			
Df Residuals:	1270	BIC:	-3495.			
Df Model:	4					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	1.8445	0.025	72.556	0.000	1.795	1.894
Ram	0.0108	0.000	28.252	0.000	0.010	0.012
CPU_freq	0.0656	0.004	17.553	0.000	0.058	0.073
Weight	-0.0108	0.005	-2.247	0.025	-0.020	-0.001
Inches	-0.0096	0.002	-4.462	0.000	-0.014	-0.005
Omnibus:	60.516	Durbin-Watson:	2.012			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	110.743			
Skew:	-0.347	Prob(JB):	8.96e-25			
Kurtosis:	4.266	Cond. No.	270.			

Análisis de Modelo General

1 R-cuadrado (R^2) = 0.566

Significa que aproximadamente el 57.8% de la variabilidad en el precio de las laptops se explica por las variables predictoras.

2 Estadístico F = 434.9

Con un valor p extremadamente bajo, lo que indica que el modelo es significativo en conjunto.

Interpretación de Coeficientes

1 Constante (6.4084)

Representa el valor promedio del logaritmo del precio cuando todas las variables predictoras son cero.

2 RAM (0.0753)

Cada GB adicional de RAM aumenta el precio promedio en aproximadamente 7.53%.

3 CPU_freq (0.4239)

Cada 1 GHz adicional en la frecuencia del procesador aumenta el precio promedio en 42.39%.

4 Weight (-0.0624)

Cada 1 kg adicional en el peso de la laptop se asocia con una disminución del precio en 6.24%.

5 Inches (-0.0693)

Cada pulgada adicional en el tamaño de pantalla se asocia con una disminución del precio en 6.93%.

Diagnóstico de Residuos

1 Durbin-Watson (2.018)

Sugiere que no hay autocorrelación significativa entre los residuos.

2 Omnibus y Jarque-Bera

Los residuos no siguen una distribución normal, lo que puede afectar las inferencias.

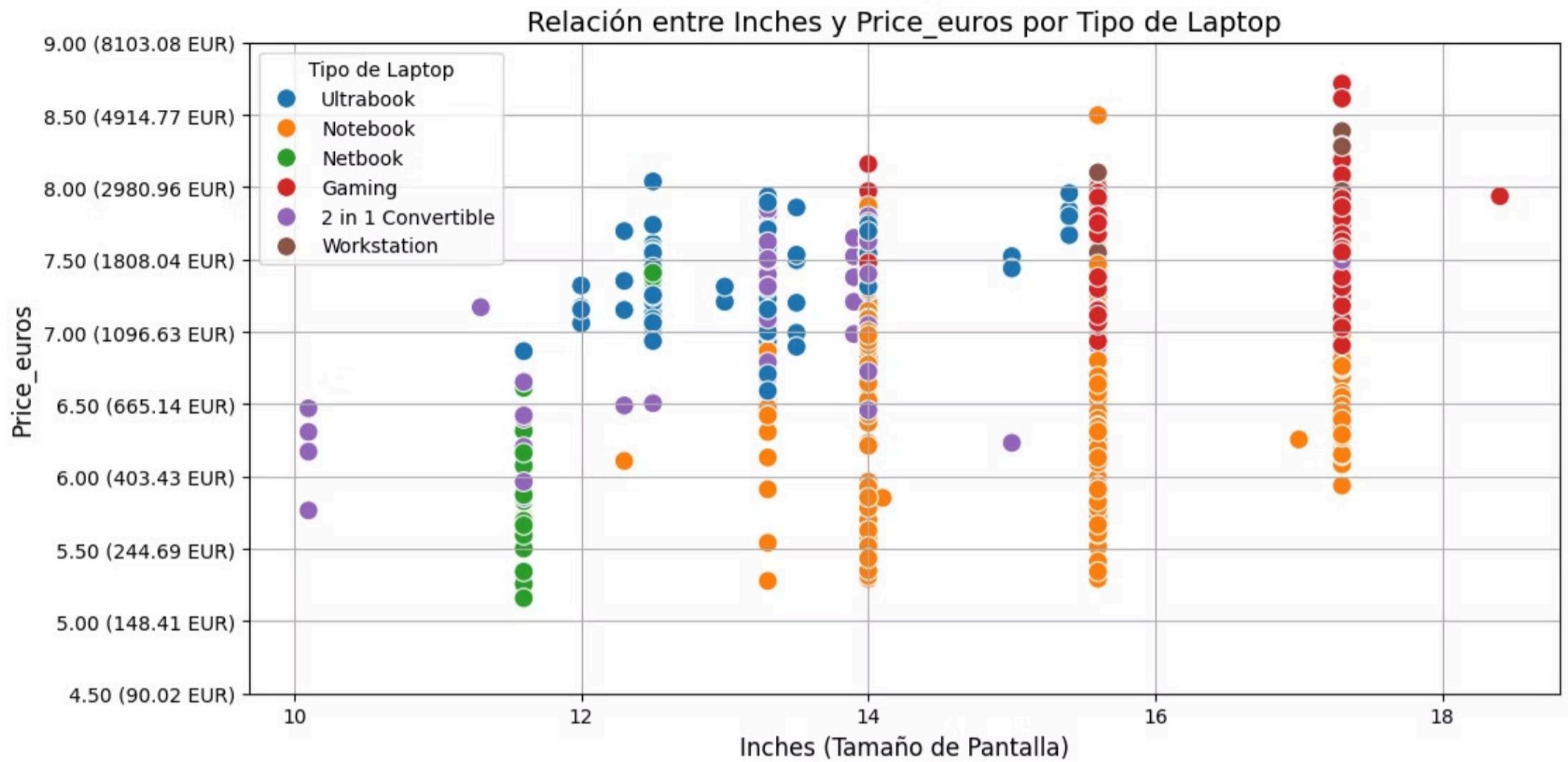
3 Condición del Modelo (270)

El número de condición es alto, lo que indica cierta multicolinealidad que podría afectar la estabilidad del modelo.

En resumen, el modelo ajustado explica bien la variabilidad del precio de las laptops, siendo las variables RAM y CPU_freq las más importantes. Aunque el peso y el tamaño de pantalla también tienen efectos, son menos significativos. La multicolinealidad no es un problema severo, pero los residuos no son normales, lo que puede afectar algunas inferencias.

Analisis de la relación entre los inches y los precios

Según análisis de los coeficientes del modelo de regresión, cada pulgada adicional en el tamaño de pantalla se asocia con una disminución del precio en 6.93%. Por lo tanto sería interesante analizarlo mas a fondo contrastando ahora por el tipo de laptops.



Análisis del Gráfico de dispersión

1 Menores a 13"

Las Netbooks y las 2 in 1 Convertible parecen tener precios más bajos, aunque existen excepciones, como ciertos dispositivos Ultrabook.

2 Entre 13" y 15"

Los precios muestran una mayor dispersión, reflejando una combinación de laptops más costosas (principalmente Gaming y Ultrabook) junto con dispositivos más accesibles ('Notebook')

3 Mayores 15"

Tienen precios variados, aunque hay un grupo de precios más bajos dominado por Notebook. Y las laptops Gaming y Workstation y Ultrabook predominan con precios muy elevados

Conclusiones

1 Relación inversa entre los precios y los inches

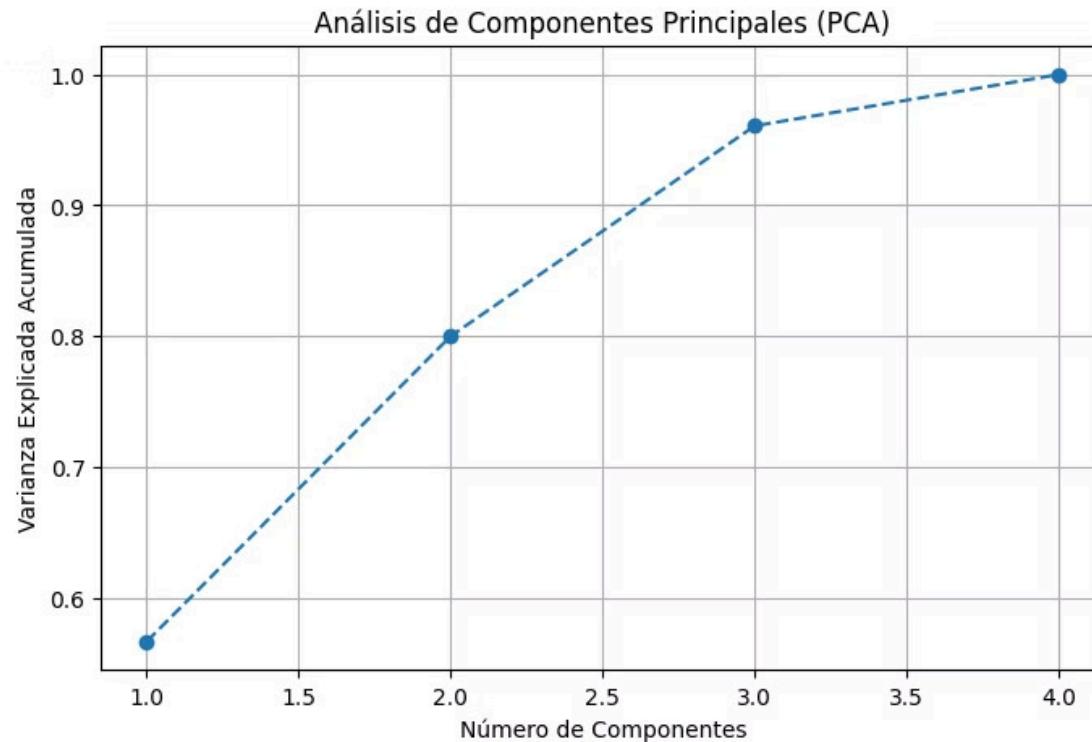
La tendencia general muestra que los precios no siempre aumentan con el tamaño de la pantalla. De hecho, las laptops grandes tienen una media de precios similar a las pequeñas, lo que sugiere que otros factores (como el tipo de laptop o el hardware interno) son más influyentes.

2 Tipo de Laptop

Los dispositivos Gaming y Workstation parecen ser los más costosos en cualquier tamaño de pantalla, mientras que Notebook y Ultrabook dominan los rangos de precios más accesibles.

Análisis de las componentes principales

Entonces para identificar las características que más influyen en el precio de las laptops y revelar patrones subyacentes de manera clara y eficiente, vamos a realizar un PCA sobre las variables: **Ram, CPU_freq, Weight, Inches**.

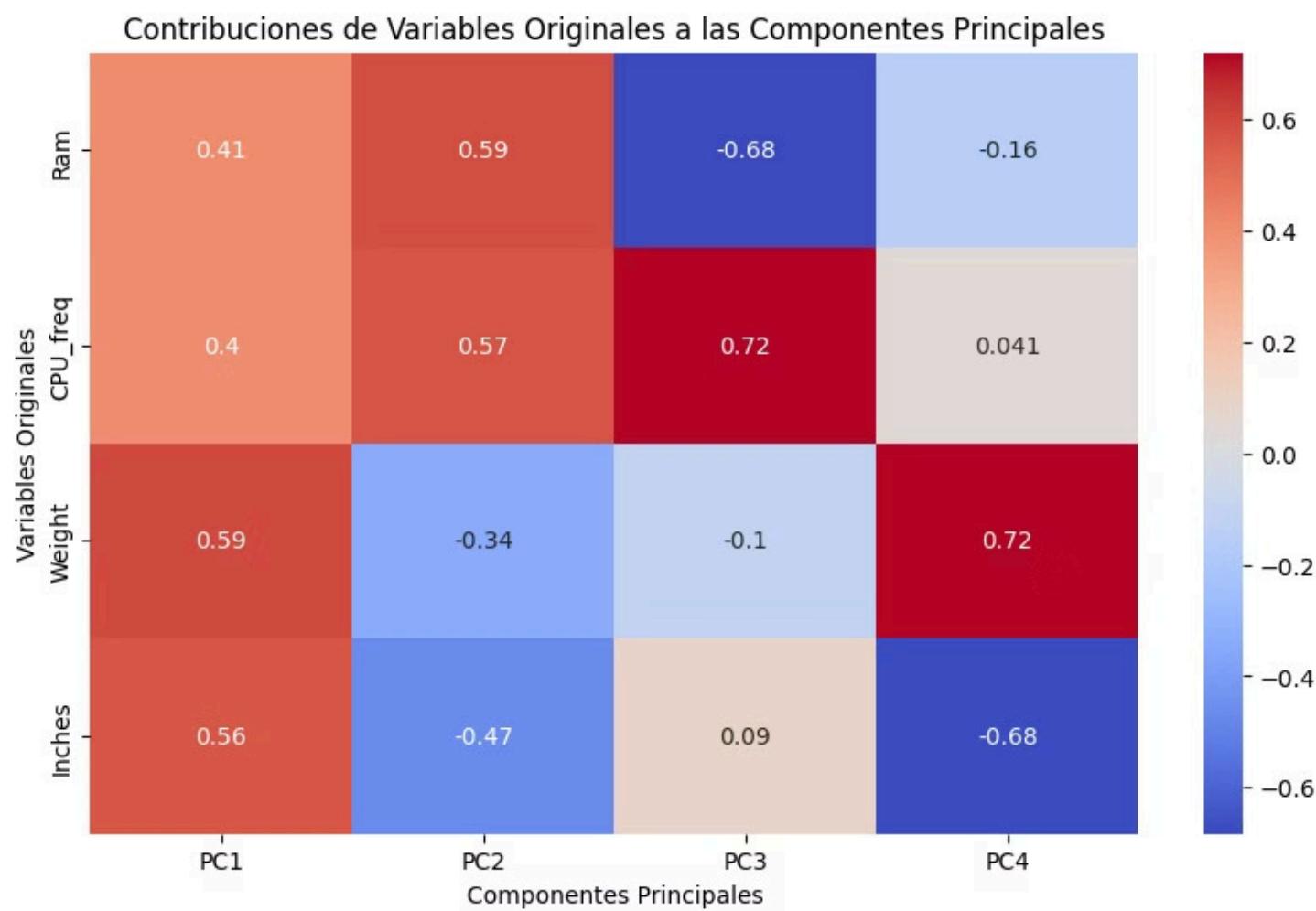


Conclusiones

Los resultados muestran que los dos primeros componentes principales explican aproximadamente el 80% de la varianza (56.64% + 23.36%), lo que indica que estos dos componentes capturan la mayor parte de la información contenida en las variables originales (Ram, CPU_freq, Weight, Inches). Esto permite reducir la dimensionalidad del conjunto de datos sin perder demasiada información.

Análisis de las contribuciones de las variables al PCA

Entonces, después de conocer cuáles son las componentes que más influyen en la varianza, es necesario determinar las contribuciones de cada variable a cada componente.



1 PC1 (Componente Principal 1)

Variables con mayores contribuciones positivas:

- **Weight (0.59)**
- **Inches (0.56)**

Estas variables sugieren que PC1 está relacionada principalmente con el tamaño físico del dispositivo. Pero a su vez sugiere que no descarta el rendimiento de la laptop.

2 PC2 (Componente Principal 2)

- Variables con mayores contribuciones positivas:

- **Ram (0.59)**
- **CPU_freq (0.57)**

PC2 parece estar asociada al rendimiento del dispositivo, en contraposición con sus características físicas.

3 PC3 (Componente Principal 3)

Variables con mayores contribuciones positivas y negativas:

- **CPU_freq (0.72, positiva)**
- **Ram (-0.68, negativa)**

PC3 refleja una relación inversa entre la capacidad del procesador y la memoria RAM.

4 PC4 (Componente Principal 4)

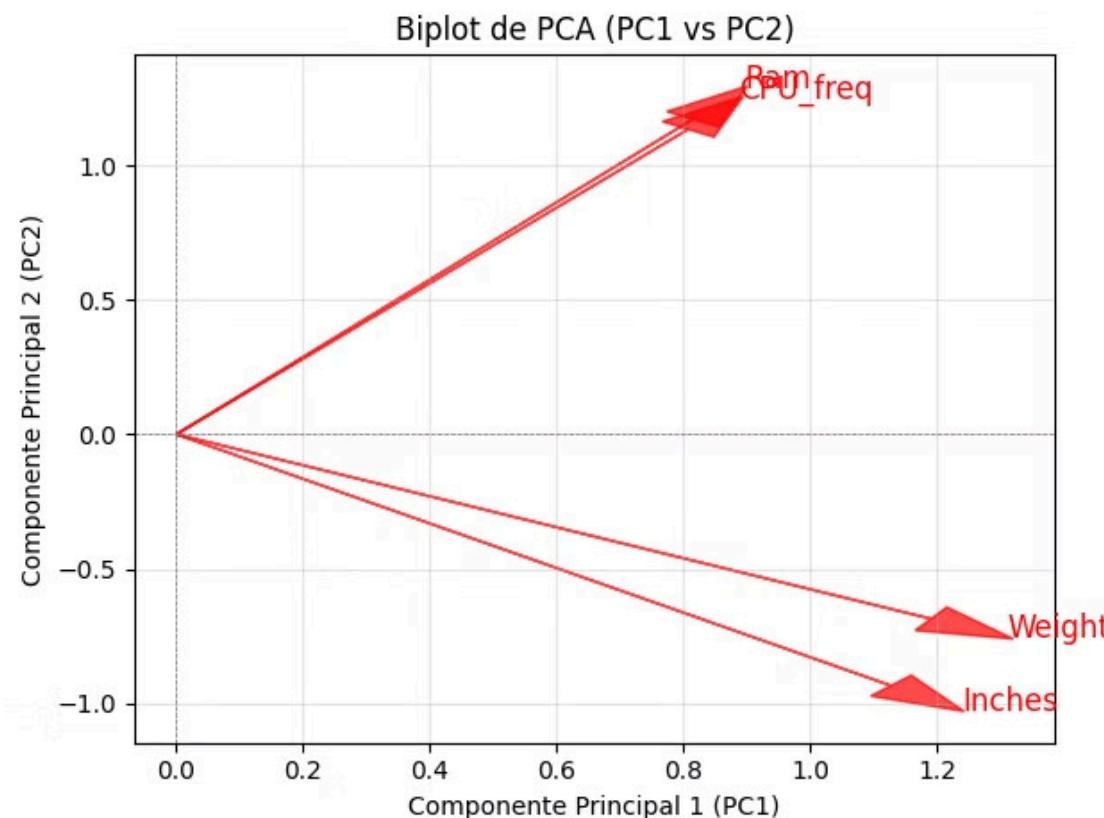
Variables con mayores contribuciones positivas y negativas:

- **Weight (0.72, positiva)**
- **Inches (-0.68, negativa)**

PC4 parece capturar variaciones entre peso y tamaño de la pantalla.

Análisis del Biplot de la PC1 vs PC2

Luego analicemos el biplot de las componentes principales 1 y 2, el cual permite visualizar simultáneamente la distribución de las laptops y la influencia de las variables originales, facilitando la interpretación de las relaciones entre las características y su impacto en las componentes principales.



Relación entre Ram y CPU_freq

- Los vectores de **Ram** y **CPU_freq** están muy cercanos y apuntan en direcciones similares, indicando una correlación positiva fuerte entre estas variables.
- Esto significa que laptops con mayor RAM suelen tener también procesadores con mayor frecuencia, ambas características clave para el rendimiento.

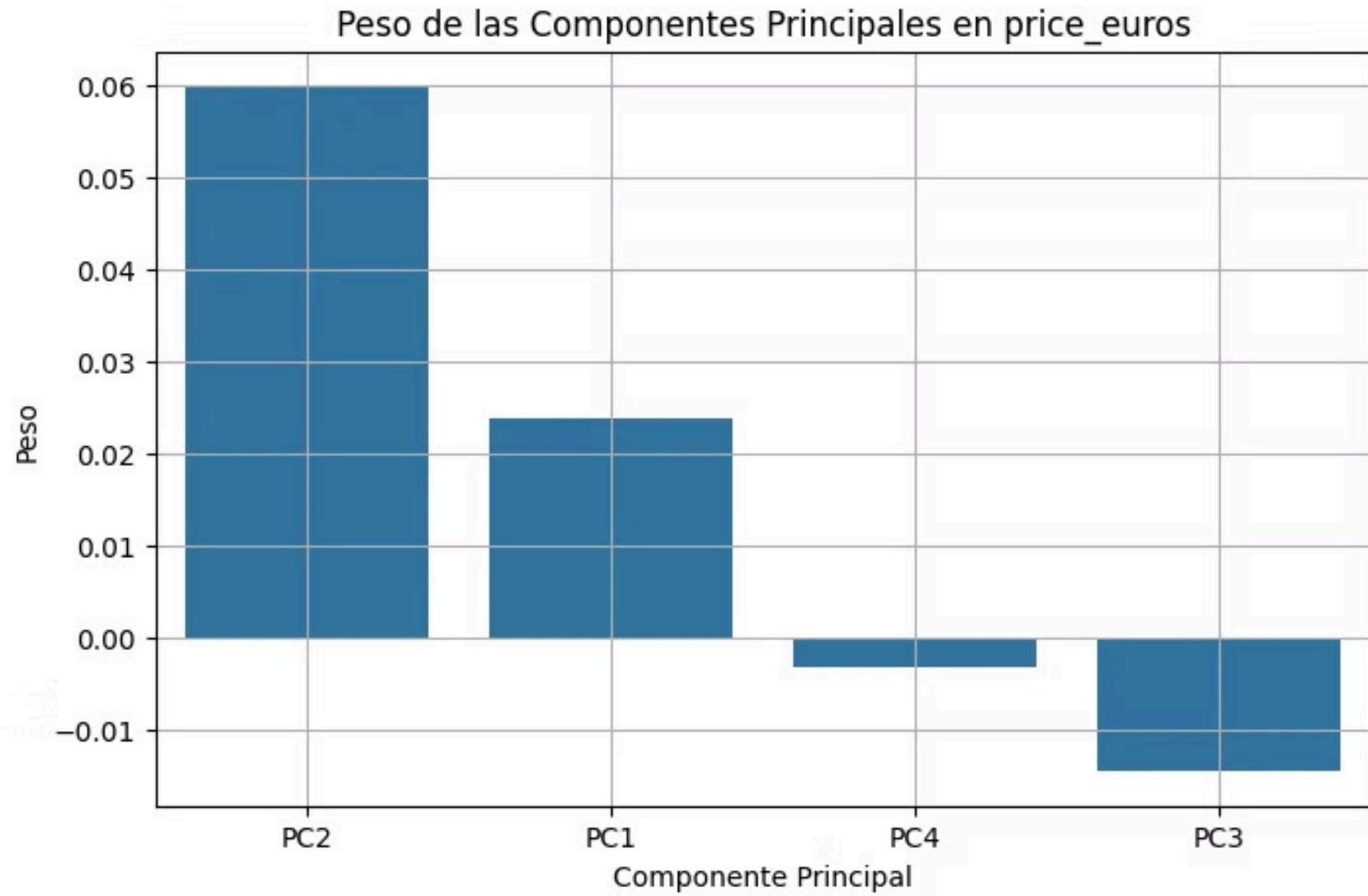
Relación entre Inches y Weight

- Los vectores de **Inches** y **Weight** también están muy cercanos entre sí, lo que refleja que estas variables están fuertemente relacionadas.
- Laptops más grandes en tamaño de pantalla tienden a ser más pesadas.

Separación entre grupos de variables:

- Existe una separación evidente entre los vectores de rendimiento (**Ram** y **CPU_freq**) y los de características físicas (**Weight** e **Inches**).
- Esto indica que las variables físicas no están correlacionadas directamente con las variables de rendimiento, lo que implica que diseñar una laptop liviana y compacta con alto rendimiento puede requerir compromisos tecnológicos y de diseño.

Análisis del peso de cada componente sobre los precios



1. Influencia de las Componentes Principales en el Precio:

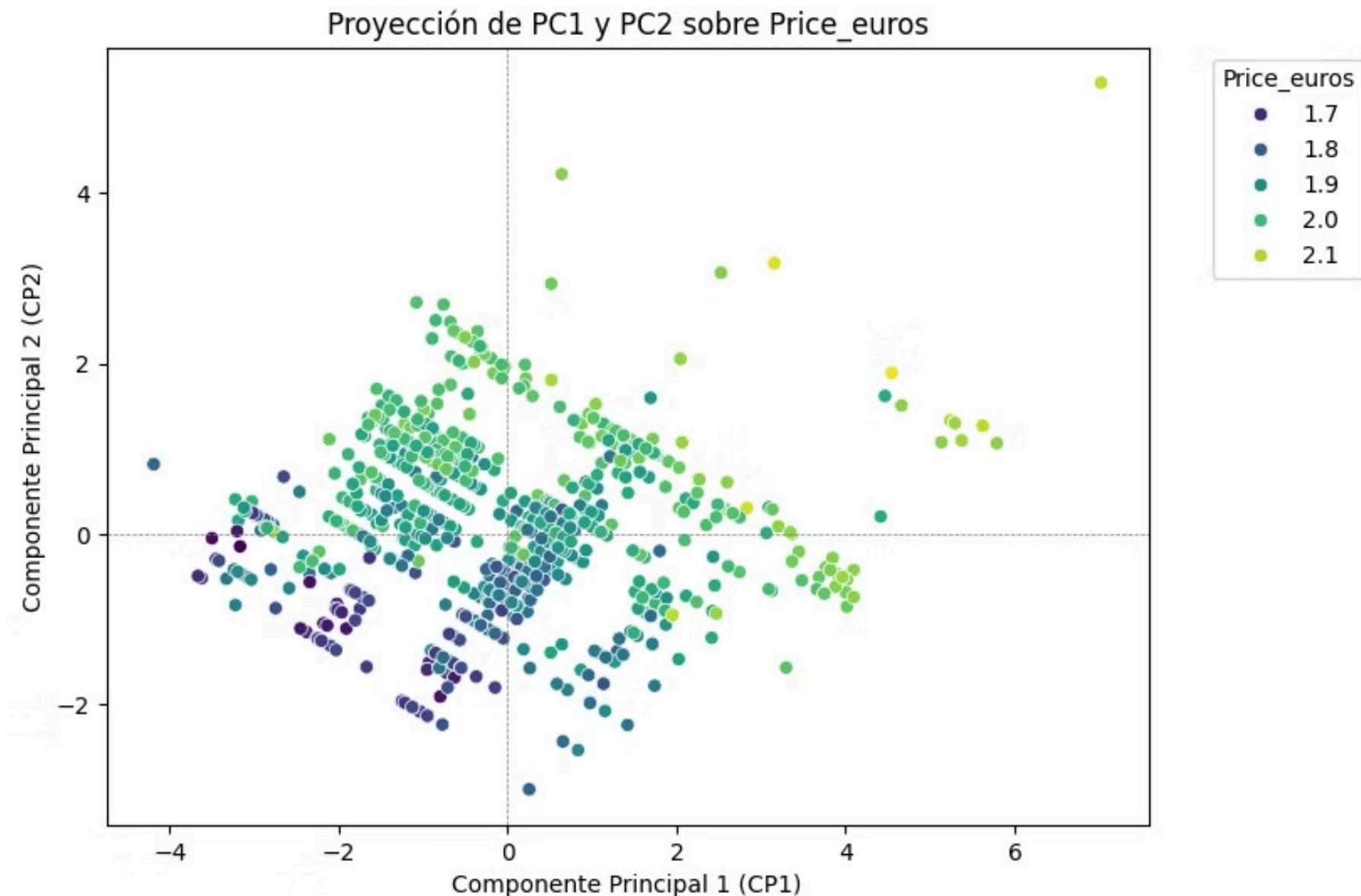
- **PC2 (Componente Principal 2)** tiene la mayor relación positiva con el precio (0.406297). Las variables que contribuyen fuertemente a PC2 son clave para explicar los precios.
- **PC1 (Componente Principal 1)** también tiene una influencia positiva significativa (0.161732), aunque menor que PC2.
- **PC3 (Componente Principal 3)** tiene una relación negativa considerable (-0.014660) con el precio. Las variables que contribuyen a PC3 están inversamente relacionadas con los precios.
- **PC4 (Componente Principal 4)** tiene una influencia positiva marginal (-0.113632) sobre el precio.

2. Importancia Relativa de las Componentes Principales:

- Aunque **PC1 explica la mayor varianza** en los datos, su influencia directa sobre el precio es menor que la de **PC2**.
- Esto demuestra que las componentes principales no siempre están en el mismo orden de importancia para una variable dependiente específica (como el precio). La varianza explicada no implica automáticamente mayor relevancia para la regresión.

Análisis de la proyección de la PC1 y 2 sobre los precios

Finalmente, veamos la proyección de las PC1 y PC2 sobre los precios y poder analizar sus resultados.



PC1 vs Precio

Los puntos con valores altos de PC1 (hacia la derecha) tienden a asociarse con laptops de precios más altos (colores amarillos y verdes). Esto indica que las laptops con mayor peso y tamaño físico o características generales de alto rendimiento tienden a tener un precio más elevado.

PC2 vs Precio

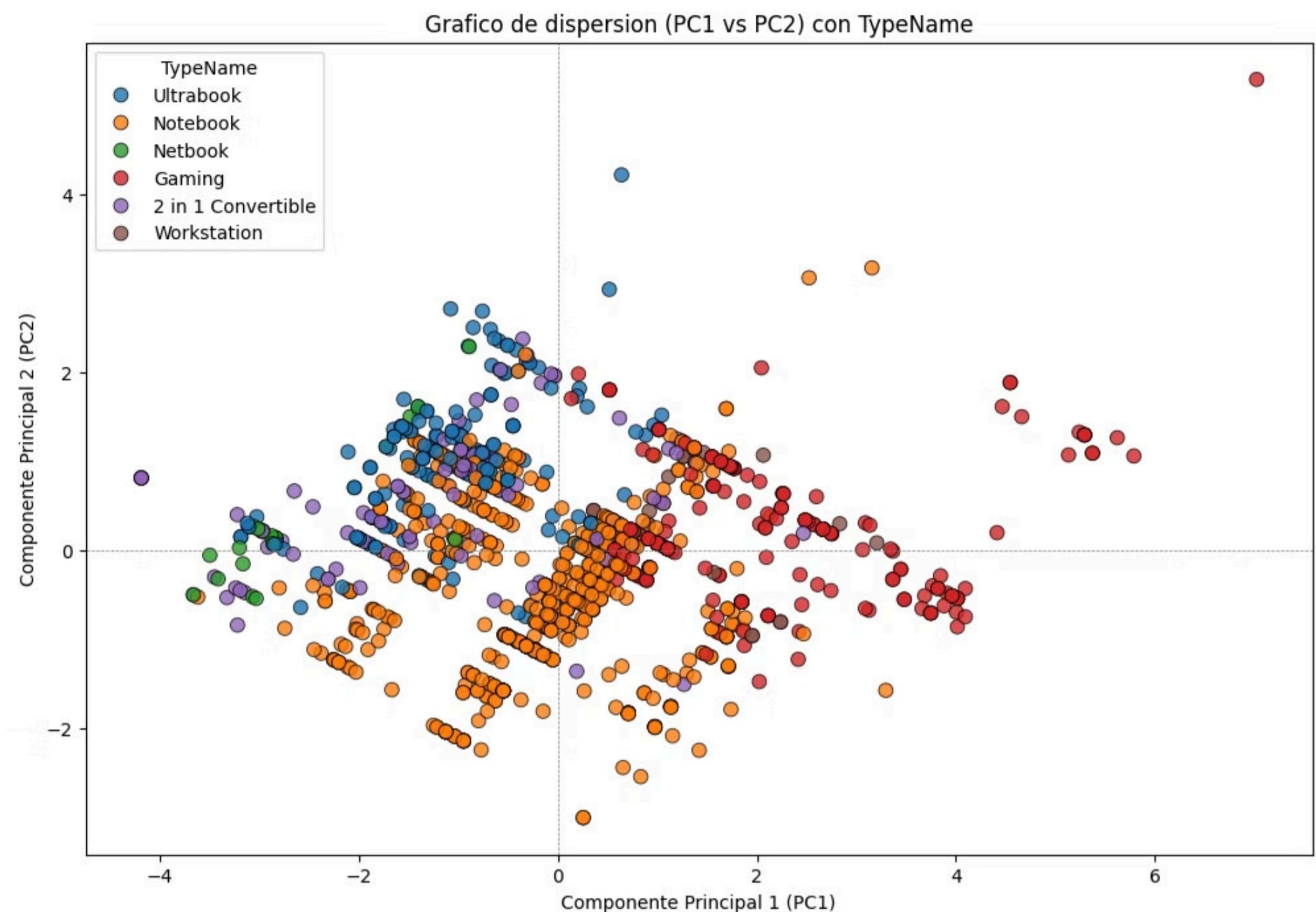
Los valores de PC2 parecen separar ligeramente las laptops de alto precio (colores amarillos) de las de precio más bajo (colores morados), aunque esta relación no es tan clara como en PC1. Esto sugiere que los precios intermedios están asociados con laptops que equilibran alto rendimiento y peso reducido.

Distribución de Precios

Los puntos con precios más bajos (morados) están más concentrados hacia el centro y los extremos negativos de PC1 y PC2, donde están las laptops más ligeras, pequeñas y con menor rendimiento. Los puntos con precios más altos (verdes claros y amarillos) se ubican en la parte positiva de PC2 y PC1, donde están las laptops de mayor rendimiento y características físicas.

Análisis de la proyección de la PC1 y PC2 según el tipo

Pero también, sería interesante analizar como sería la proyección de las PC1 y PC2 según el tipo de laptop y poder contrastar su distribución entre los precios.



Computadoras Gaming

Las laptops clasificadas como Gaming tienden a concentrarse en los extremos positivos de la Componente Principal 1 (PC1), lo que indica que están altamente relacionadas con características de rendimiento y diseño físico.

Ultrabook

Las Ultrabooks se agrupan en las regiones intermedias, con valores cercanos al centro de las componentes principales. Esto sugiere que combinan características balanceadas de rendimiento y diseño físico, siendo ligeras y delgadas.

Notebook

Las Notebooks están dispersas a lo largo del eje central, con una ligera inclinación hacia valores negativos en la PC1. Esto indica que priorizan características físicas como menor peso y mayor portabilidad en lugar de alto rendimiento.

Netbook

Las Netbooks aparecen principalmente en las regiones de valores negativos de PC1 y PC2, asociándose con menor rendimiento y diseño físico reducido.

Workstation

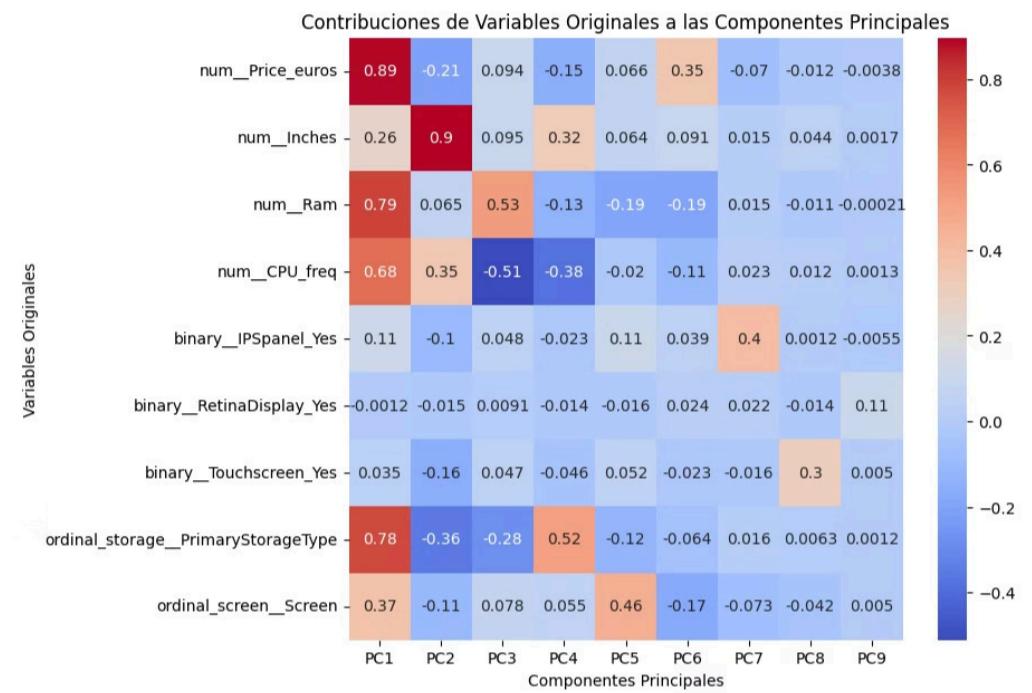
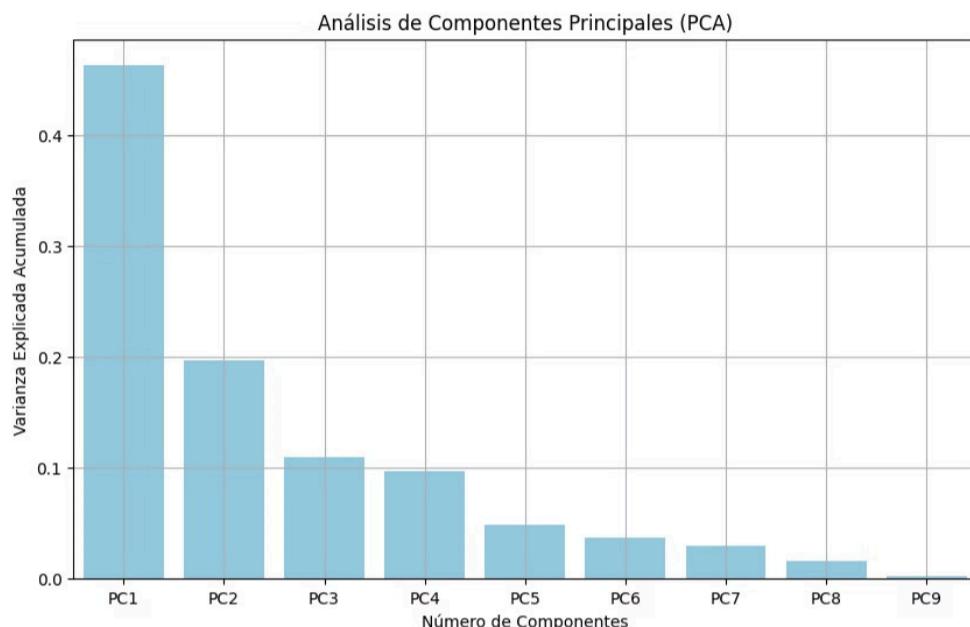
Las Workstations se encuentran en la parte positiva de la PC1 y tienen una dispersión moderada en PC2. Esto refleja su alta relación con rendimiento y menos importancia en las características físicas.

2 in 1 Convertible

Este tipo se agrupa hacia valores bajos de PC2 y cercanos al centro en PC1, estando más relacionado con características físicas como menor peso y flexibilidad.

Analisis expandido de las componentes para las agrupacion de las laptops por gamas

A partir de los resultados que obtenidos en las componentes principales, seria también interesante hacer una clusterización de la laptop con la intención de poder clasificar las laptop según su gama. Para ello es necesario añadir mas variables al análisis de las componentes principales, puesto que para la clasificación de la gama de una laptop influyen tanto características físicas, ya sea los materiales que están compuestos como del rendimiento y potencia. Es decir, que sera necesario añadir variables categóricas y hacerle una correspondiente codificación en binario u ordinal.



Y luego llegamos a estas conclusiones

PC1 (46.29%)

Representa una dimensión de rendimiento y costo, influenciada por precio, RAM, CPU y almacenamiento SSD. Valores altos indican laptops de gama alta, valores bajos laptops económicas.

PC2 (19.65%)

Refleja principalmente el tamaño de pantalla, con laptops grandes en un extremo y compactas en el otro, también influenciado por frecuencia de CPU.

PC3 (11.02%)

Podría indicar una compensación entre RAM y frecuencia de CPU, sugiriendo que algunas laptops optimizan uno de estos aspectos sobre el otro.

PC4 (9.66%)

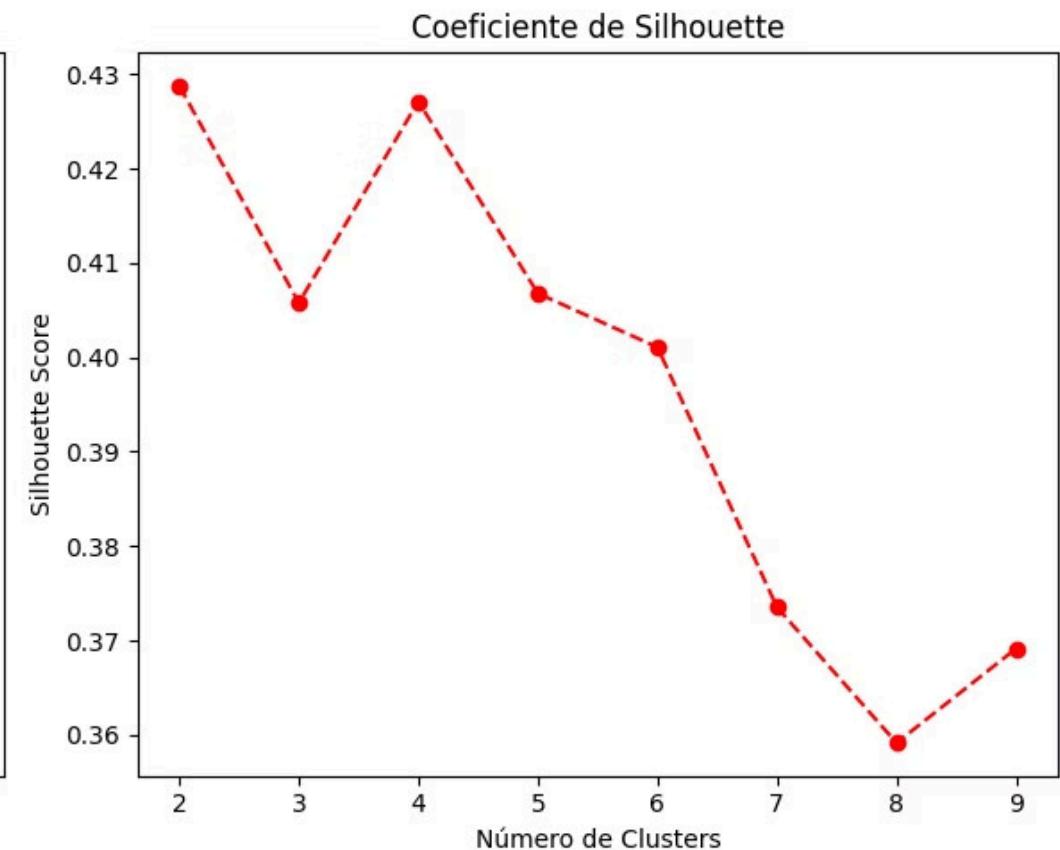
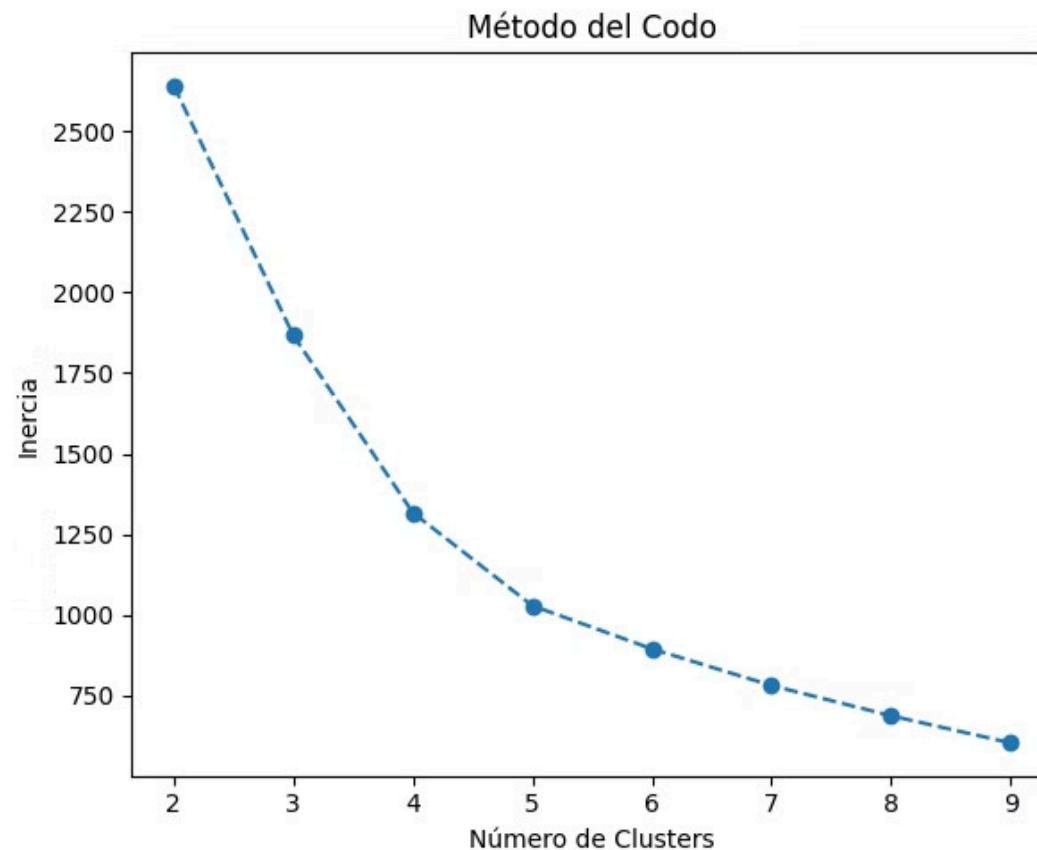
Refleja una dimensión relacionada con el tipo de almacenamiento (SSD vs HDD) y el tamaño de la laptop.

Otros Componentes

Los componentes restantes se asocian con características como pantalla IPS, Retina y táctil, aunque con menor importancia.

Análisis para obtener el numero de clúster adecuado

El método del codo y el coeficiente de Silhouette son técnicas clave para determinar el número óptimo de clústeres en un análisis de agrupamiento. El método del codo evalúa la variabilidad explicada por los clústeres en función de su cantidad, buscando el punto donde la mejora se estabiliza. Por su parte, el coeficiente de Silhouette mide qué tan bien diferenciados están los clústeres, indicando una mejor calidad de agrupamiento cuando los valores son cercanos a 1.



Y a raíz del análisis de las graficas de ambos métodos podemos concluir que el mejor numero de clúster es: 4

Análisis de los resultados de la clusterización

Entonces, estos son los resultados y las conclusiones que podemos sacar de la clusterización

Cluster	Inches	Ram	Weight	Price_euros	ScreenW	\
3	13.776761	4.211268	1.629577	1.777603	1537.492958	
1	15.739952	5.946731	2.208426	1.864020	1681.263923	
2	13.668539	8.308989	1.477298	1.968752	2094.449438	
0	16.020055	13.049451	2.561187	1.992478	2099.576923	

Cluster	ScreenH	CPU_freq	PrimaryStorage	SecondaryStorage	Price_euros_ori	\
3	879.154930	1.421831	331.718310		0.000000	
1	945.995157	2.315496	708.842615		39.670702	
2	1188.943820	2.257865	302.280899		42.426966	
0	1182.494505	2.676648	327.725275		530.219780	

Cluster	Price_euros_ori
3	5.931789
1	6.458372
2	7.169567
0	7.345732

Cluster 3 (Gama baja)

Pantalla pequeña (13.7"), resolución baja. RAM baja (4GB), CPU lento (1.42 GHz). Almacenamiento principal HDD/Flash (335 GB), sin secundario. Peso ligero (1.62 kg), precio medio €429.

Conclusión: Laptops económicas, básicas, para tareas simples.

Cluster 0 (Gama alta)

Pantalla pequeña (13.6") pero alta resolución. RAM alta (8GB), CPU potente (2.26 GHz). Almacenamiento principal modesto (302 GB), algo secundario (42 GB). Peso muy ligero (1.47 kg), precio medio €1378.

Conclusión: Laptops premium, portátiles y con pantalla de calidad, para profesionales.

Cluster 1 (Gama media)

Pantalla grande (15.7"), resolución media. RAM moderada (6GB), CPU decente (2.31 GHz). Almacenamiento principal amplio (700 GB), algo secundario (41 GB). Peso moderado (2.21 kg), precio medio €679.

Conclusión: Laptops versátiles, equilibradas en rendimiento y precio, para estudiantes u oficina.

Cluster 2 (Gama premium)

Pantalla grande (16") y alta resolución. RAM muy alta (13GB), CPU rápido (2.68 GHz). Gran almacenamiento principal (328 GB) y secundario (533 GB). Peso elevado (2.56 kg), precio medio €1707.

Conclusión: Laptops de alto rendimiento, para gaming, edición o desarrollo intensivo.

Graficas de las gamas

Y finalmente el análisis de las gamas se puede ver reflejado en estas graficas y corroborar los resultados.

