

# **Diplomado en Analítica y Ciencia de Datos**

Universidad de Colima - TRANSICION

## **Proyecto Final**

### **Análisis de Precios de Laptops**

Equipo No. 11

Integrantes:

- Luis Alberto Pineda García
- Nazario Bustos Eusebio
- Julio Daniel Uribe Gómez
- Cesar Edwin Hernández Ochoa

# 1.- Selección y Justificación de la Base de Datos

- **Búsqueda**

La búsqueda de la fuente de datos se realizó en el sitio web Kaggle con dirección [www.kaggle.com](https://www.kaggle.com).

**Fuente de Datos elegida:** Dataset archivo CSV (laptop\_prices.csv)

**URL :** [https://www.kaggle.com/datasets/owm4096/laptop-prices?select=laptop\\_prices.csv](https://www.kaggle.com/datasets/owm4096/laptop-prices?select=laptop_prices.csv)

- **Justificación**

Revisando opciones entre distintos sitios de internet y diversos temas, se optó por elegir un dataset de interés general, ya que todas las profesiones y oficios estamos relacionados con el uso de equipos de cómputo. En este caso se trata de una lista de precios de laptops, tomando en cuenta diversos datos como la compañía a la que pertenecen, marca de CPU, modelo de CPU, Sistema Operativo, características gráficas, tamaño de monitor entre otras.

Además de esto el conocimiento adquirido puede ayudar de alguna manera a la toma de decisiones en cuestión de compra.

- **Objetivo**

Realizar un análisis exploratorio de los datos a través de lenguaje R, y visualización de los resultados en software Tableau para observar de manera intuitiva y clara el comportamiento de los precios de laptops según sus características y también cuales de éstas mencionadas son las que los fabricantes toman más en cuenta para el diseño de sus equipos de cómputo.

## 2.- Preparación y Limpieza de Datos

- **Descripción de la Base de Datos**

La Base de datos elegida (archivo CSV) cuenta con 1275 registros y 23 columnas descriptivas de los equipos.

Tabla de Columnas y Tipos de Datos

No.	Columnas	Descripción	Tipo Dato
1	Company	Nombre de Compañía o Fabricante	chr
2	Product	Nombre del Producto (Laptop)	chr
3	TypeName	Tipo de Laptop	chr
4	Inches	Pulgadas tamaño del display monitor	num
5	Ram	Memoria Ram	Int
6	OS	Sistema Operativo	chr
7	Weight	Peso	<u>num</u>
8	Price_euros	Precio en euros	num
9	Screen	Tecnología de Display o Pantalla	Chr
10	ScreenW	Ancho de Display o Pantalla	Int
11	ScreenH	Alto de Display o Pantalla	int
12	Touchscreen	Pantalla tipo táctil	chr
13	IPspanel	Si la computadora portátil tiene o no un IPspanel.	chr
14	RetinaDisplay	Si la computadora portátil tiene o no un Retina Display	chr
15	CPU_company	Marca o compañía de CPU	chr
16	CPU_freq	Frecuencia en Hz de CPU	num
17	CPU_model	Modelo de CPU	chr
18	PrimaryStorage	Almacenamiento Primario en Gigabytes	Int
19	SecondaryStorage	Almacenamiento Secundario en Gigabytes	Int
20	PrimaryStorageType	Tecnología o Tipo Almacenamiento Primario	chr
21	SecondaryStorageType	Tecnología o Tipo Almacenamiento Secundario	chr
22	GPU_company	Compañía o marca de Tarjeta Gráfica	chr
23	GPU_model	Modelo de Tarjeta Gráfica	chr

- **Limpieza de los datos**

1. Identificación y manejo registros faltantes

La Base de datos elegida no contaba con ningún dato erróneo ni descripciones repetitivas o mal redactadas. Por lo tanto nuestra base de datos a usar la renombramos como:

✓ laptop\_prices\_depurada.csv

2. Corrección de errores en los datos

Se realizaron agrupamientos y no se encontraron descripciones e inconsistencias en los datos como descripciones con caracteres especiales.

3. Transformaciones necesarias

El dataset elegido cuenta con los suficientes campos, para realizar las estadísticas deseadas que no fue necesario generar nuevas columnas calculadas al dataset base. Sin embargo, en la creación de resúmenes estadísticos se agregaron las columnas necesarias para generar los mismos.

### 3.- Análisis Exploratorio de Datos (EDA)

Apoyados del lenguaje de programación "R" se realizaron los siguientes análisis.

Estadísticas Generales de Precios de Laptops (Euros)

Mínimo	1er Qu.	Mediana	Promedio	Moda	3er Qu.	Máximo
174	609	989	1135	1099	1496	6099

Durante este proceso se realizaron archivos csv adicionales con agrupamientos necesarios para realizar estadísticas específicas.

Estadísticas de Precios de laptops INTEL por Compañía  
(Resumen\_precios\_Company\_INTEL.csv)

Company	Precio.mediana	Precio.promedio	Precio.maximo	Precio.minimo	Recuento
Dell	999	1199.2251	3659.4	274.9	291
Lenovo	949	1127.3429	4899	229	273
HP	1049.26	1136.997	4389	209	245
Asus	1094	1142.6109	3975	191.9	141
Acer	613	652.2627	2599	174	91
MSI	1599	1728.9081	2799	839	54
Toshiba	1211.5	1267.8125	2799	447	48
Apple	1339.69	1564.1986	2858	898.94	21
Samsung	1674	1507.75	1849	269	8
Mediacom	265	295	389	239	7
Razer	2899	3346.1429	6099	1029	7
Microsoft	1569.5	1612.3083	2589	989	6
Vero	206.85	217.425	260	196	4
Xiaomi	1099.45	1133.4625	1399.95	935	4
Chuwi	248.9	314.2967	449	244.99	3
Fujitsu	739	729	799	649	3
Google	1559	1677.6667	2199	1275	3
LG	2099	2099	2299	1899	3
Huawei	1424	1424	1499	1349	2

Estadísticas Precios de laptops AMD por Compañía  
(Resumen\_precios\_Company\_AMD.csv)

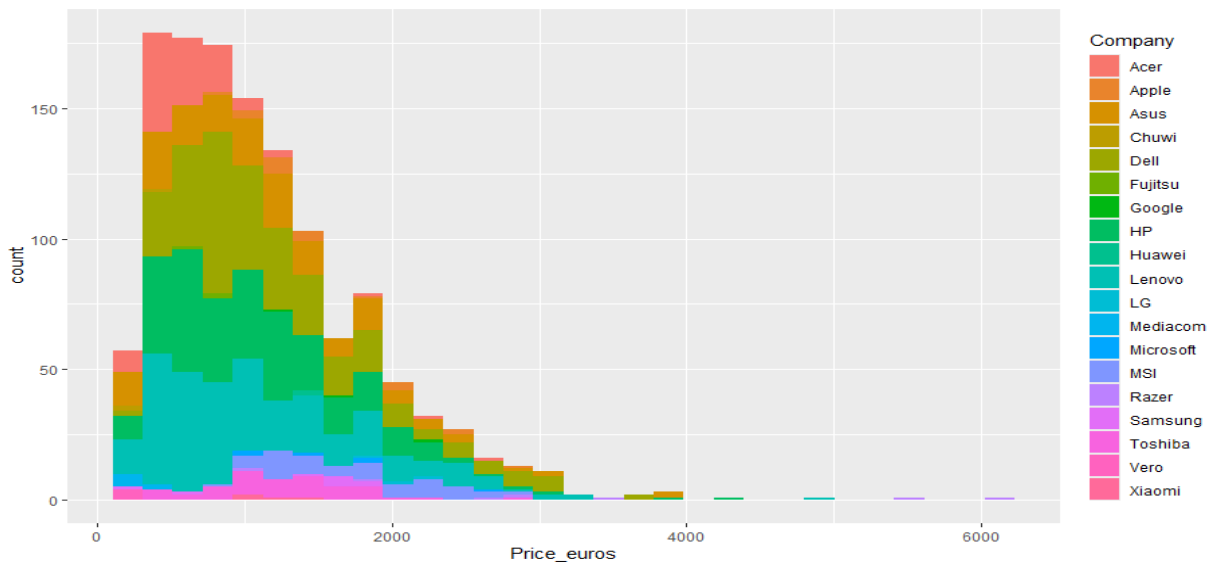
Company	Precio.mediana	Precio.promedio	Precio.máximo	Precio.mínimo	Recuento
Acer	437.5	462.4	659	387	10
Asus	575	883.09	2199	199	11
HP	488.99	476.5247826	655.01	258	23
Lenovo	459	522.598125	999	298	16

Estadísticas de Precios de laptops por CPU Model  
(Resumen\_precios\_cpu\_model.csv)

CPU_model	Precio.mediana	Precio.promedio	Precio.maximo	Precio.minimo	Recuento
Core i5 7200U	846.5	919.3181	2330	393.9	193
Core i7 7700HQ	1679	1762.3512	3659.4	779	147
Core i7 7500U	1224	1348.9742	2824	579	133
Core i3 6006U	469	492.6089	849	339	81
Core i7 8550U	1103	1206.54	2499	609	73
Core i5 8250U	857	944.0381	1869	521.47	72
Core i5 6200U	1083.5	1044.2382	1690	398	68
Core i7 6500U	1399	1380.7098	2339	629	43
Core i7 6700HQ	1561	1700.989	3949.4	799	41
Core i3 7100U	597	599.6163	949	384	35
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.

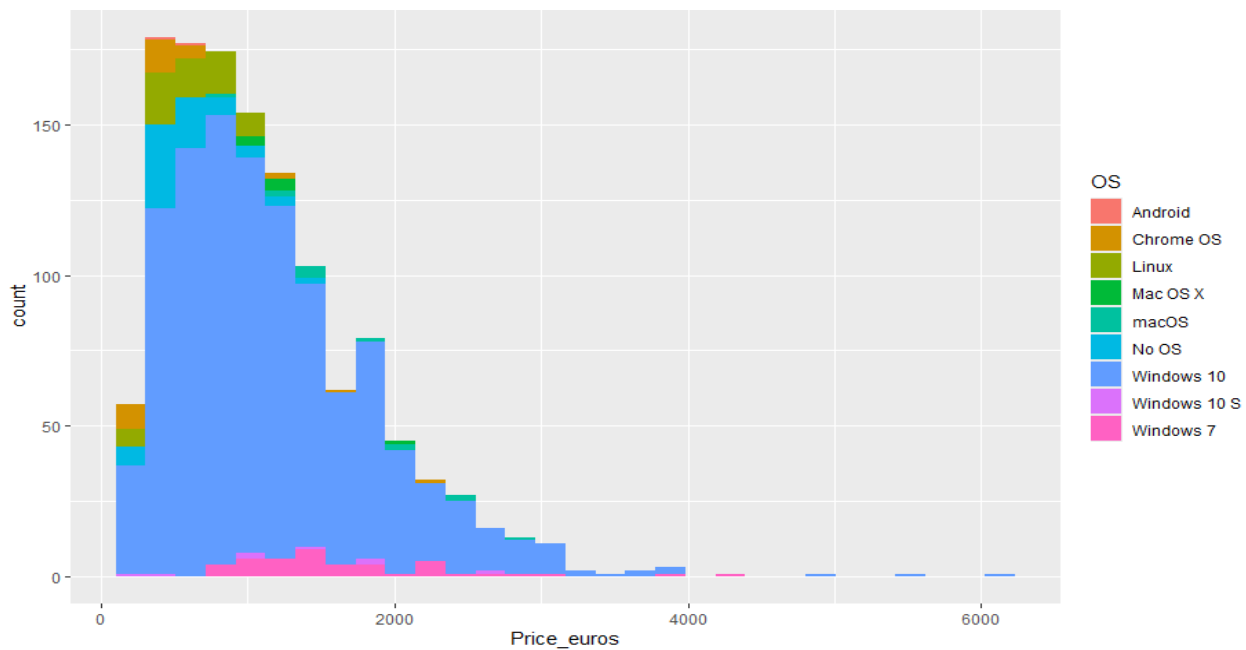
## Histogramas, gráfico de barras y Distribución de los datos

### *Histograma de Precios por Compañías*



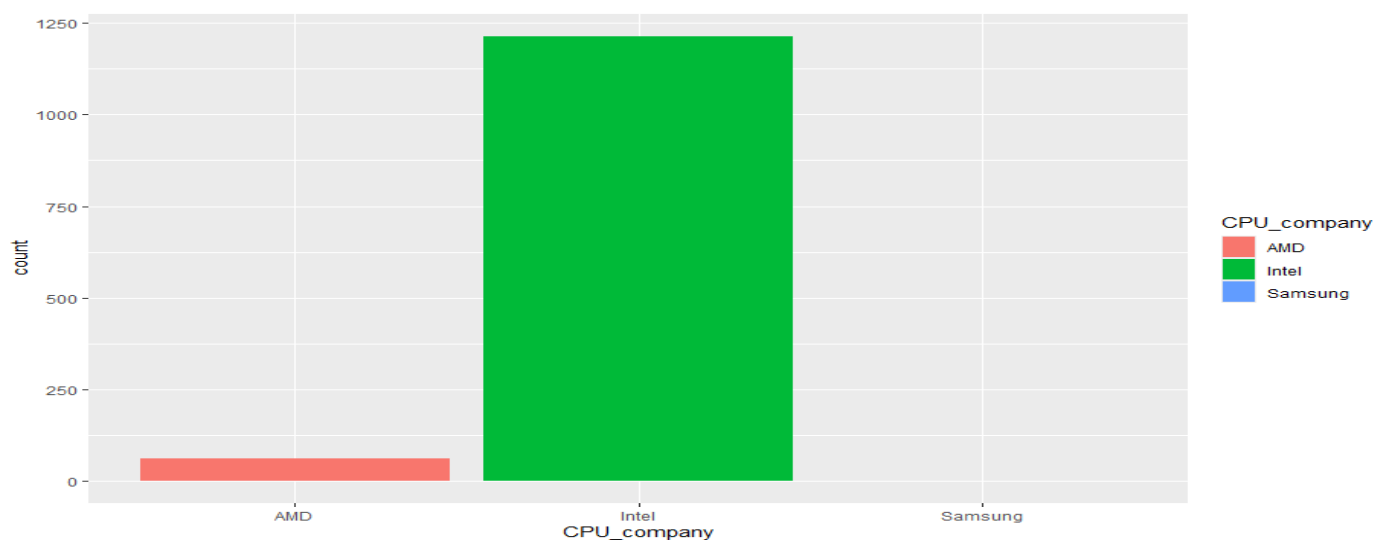
Observaciones: Existe un grupo grande de laptops con rango de precios entre el 1er Cuartil que equivale a 609 Euros a 2,000 Euros aproximadamente.

### *Histograma de Precios por Sistema Operativo*



**Observaciones:** Existe un grupo grande de laptops con sistema operativo Windows 10

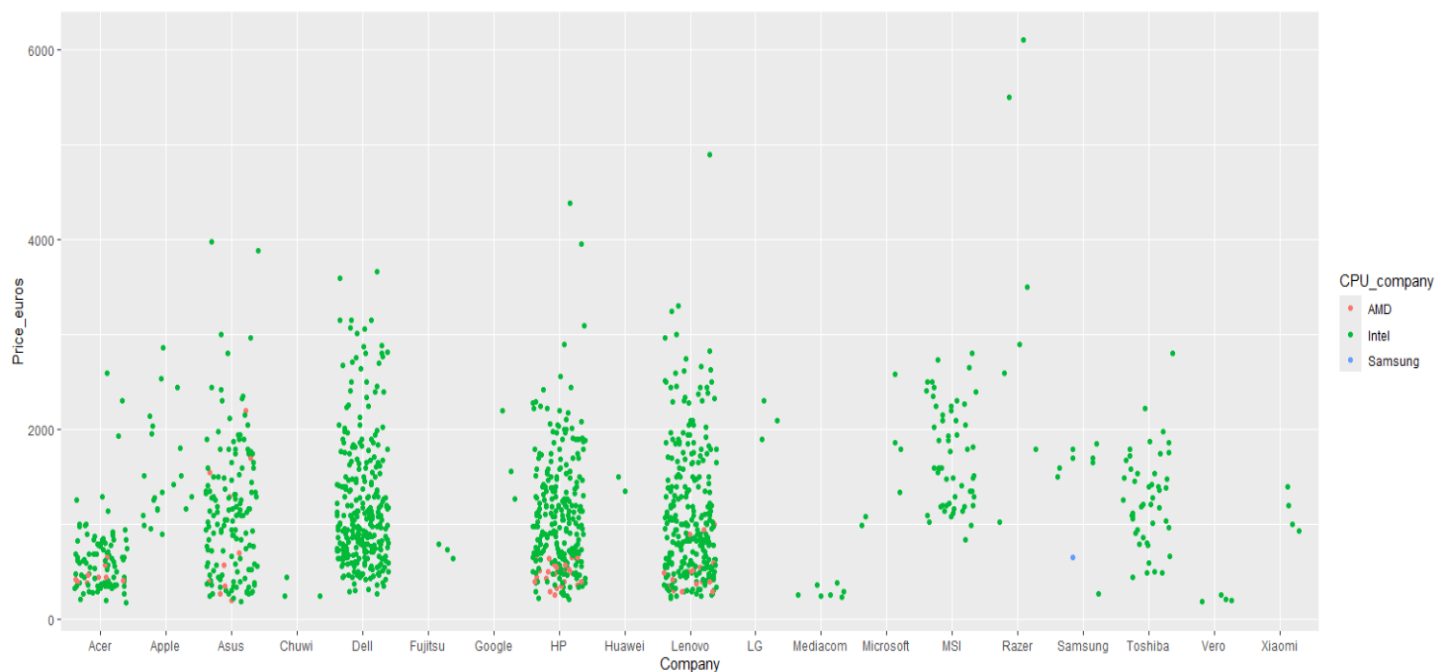
*Gráfico de Barras - Recuento de Laptops por CPU Company*



**Observaciones:**

Es evidente que la mayoría de los equipos laptops de esta base de datos, la mayoría cuentan con un CPU de la compañía INTEL.

*Diagrama de dispersión de los datos por Precio, Compañía y CPU Compañía*

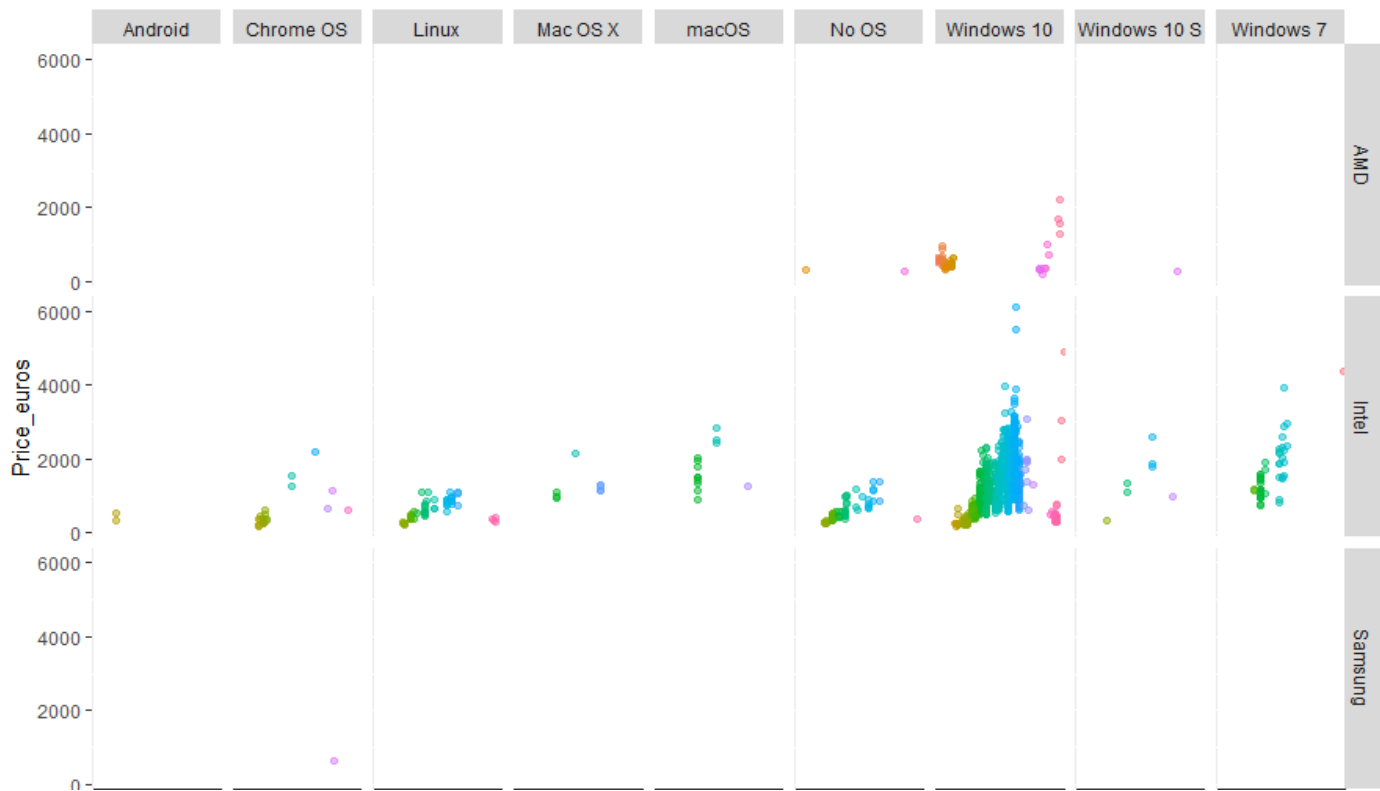


**Observaciones:**

Existe una gran distribución de Laptops con CPU Intel, de las Compañías Dell, Lenovo, HP y Asus



### *Distribución de datos en Matriz Grid de Sistema Operativo x CPU Company*



#### **Observaciones:**

Existe una gran distribución y concentración en equipos con Windows 10 y la mayoría cuentan con CPU INTEL que es el de mayor demanda o preferencia en el mercado hoy en día por el desempeño y durabilidad.

## 4.- Visualización de Datos

- **Herramienta de visualización de datos**

Software	<u>Tableau</u>
URL Público:	<a href="https://public.tableau.com/app/profile/.....">https://public.tableau.com/app/profile/.....</a>
Archivo fuente	ProyectoFinal.twb

Software: Tableau

URL Tableau público:

FOTOS DE DASHBOARD

FOTOS DE DASHBOARD

FOTOS DE DASHBOARD

## 5.- Interpretación y Conclusiones

- **Resultados del Análisis**

1. Existe una mayor preferencia por los fabricantes de laptop la integración de CPU de la compañía INTEL.
2. La Mayoría de los fabricantes de laptops, incluyen Sistema Operativo Windows 10 como instalación base.
3. Existe una gran distribución de Laptops con CPU Intel, de las Compañías Dell, Lenovo, HP y Asus.
4. El precio promedio de las laptops independientemente de sus características está en 1,135 Euros.
5. Existe un grupo grande de laptops con rango de precios entre el 1er Quartil que equivale a 609 Euros a 2,000 Euros aproximadamente.

- **Conclusiones**

Sin duda alguna, la realización del análisis de los datos nos conduce a la mejor toma de decisiones. En esta ocasión, logramos comprender el objetivo que se pretendía, sobre conocer cuales equipos de cómputo de los que se encuentran en el mercado están más accesibles en cuanto a precio, entender que el sistema operativo más confiable según fabricantes es Windows 10. Cabe resaltar que solo se está tomando en cuenta la estadística en cuanto a existencia de modelos y su precio del universo de datos analizado, sabiendo de antemano que las mejores características que un equipo de cómputo pueda tener, el conjunto de ellas son las que en realidad son las definirán su costo.

Sugerencias para trabajos futuros: Sería interesante contrastar este análisis de precios de laptops con un dataset de ventas y observar su preferencia de acuerdo a zonas geográficas como Países, Estado, etc.