

TAREA EN CLASE SOBRE REGRESIÓN LINEAL MÚLTIPLE

Integrantes:

- Oliver Roberto Saraguro Remache

DataSet: [laptop_data_cleaned.csv](#)

Para el dataset provisto que contiene información sobre características y precios de laptops, realizar:

- Construir un modelo de regresión lineal múltiple para predecir el **precio** de las laptops.
- Explorar el **mejor modelo** usando **máximo 5 variables** independientes.
- Procurar que la variabilidad explicada por el modelo (**Ajusted R-squared**) sea de al menos **80%**
- Aplicar lo aprendido hasta ahora para lograr el mejor modelo

Solución

Código R

```
getwd()
setwd("/Users/oliversaraguro/Desktop/CicloVI/analisisDeDatos/bimestre02/semana
03")

datos <- read.csv("laptop_data_cleaned.csv", sep = ";", dec = ",")  
  
str(datos)
unique(datos$Company) # Categorica
unique(datos$TypeName) # Categorica
unique(datos$Ram) # Bimodal
unique(datos$Weight) # Numerica
unique(datos$Price) # Numerica
unique(datos$TouchScreen) # Categorica
unique(datos$Ips) # Categorica
unique(datos$Ppi) # Numerica
unique(datos$Cpu_brand) # Categorica
unique(datos$HDD) # Bimodal
unique(datos$SSD) # Bimodal
unique(datos$Gpu_brand) # Categorica
unique(datos$Os) # Categorica  
  
datos$TypeName <- as.factor(datos$TypeName)
datos$Ram <- as.factor(datos$Ram)
datos$Cpu_brand <- as.factor(datos$Cpu_brand)
datos$SSD <- as.factor(datos$SSD)  
  
model <- lm(Price ~ TypeName + Ram + Ppi + Cpu_brand + SSD, data = datos)  
  
summary(model)
```

Resultados, interpretación, y conclusiones

```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools
RStudio
Breakpoints cannot be set until the file is saved.
10 unique(datos$Modelo) # Numerico
11 unique(datos$Peso) # Numerico
12 unique(datos$TouchScreen) # Categorico
13 unique(datos$Tecnologia) # Categorico
14 unique(datos$SistemaOperativo) # Categorico
15 unique(datos$Cpu_brand) # Categorico
16 unique(datos$Ram) # Numerico
17 unique(datos$SSD) # Numerico
18 unique(datos$Cpu_brand) # Categorico
19 unique(datos$Os) # Categorico
20
21 datos$TypeName <- as.factor(datos$TypeName)
22 datos$Ram <- as.factor(datos$Ram)
23 datos$Cpu_brand <- as.factor(datos$Cpu_brand)
24 datos$SSD <- as.factor(datos$SSD)
25
26
27 model <- lm(Price ~ TypeName + Ram + Ppi + Cpu_brand + SSD, data = datos)
28
29
30
31 summary(model)
32
33
34 ## F-test 1

```

R > 4.4.1.1 -> (Desktop/Ciclos/AnalisisDeDatos/semanas2)

SD06	-0.374954	0.272996	-1.374	0.169810
SD0128	0.160476	0.3220518	6.480	0.00000000 ***
SD0400	0.232094	0.3220518	6.899236 ***	
SD0240	1.236995	0.2761229	4.478	8.21e-06 ***
SD046	0.230851	0.3220518	10.710	<2e-16 ***
SD0512	0.230851	0.3220518	10.710	<2e-16 ***
SD0768	0.131928	0.2734996	0.483	0.629514
SD0800	0.131928	0.2734996	5.040	4.69e-09 ***
SD0256	0.3921646	0.272996	1.430	0.157454

Signif. codes: 0 '****' 0.001 '*' 0.01 '**' 0.1 ' ' 1

Residual standard error: 0.2780 on 1242 degrees of Freedom

Multiple R-squared: 0.8413 Adjusted R-squared: 0.8896

F-statistic: 181.3 on 58 and 1242 DF, p-value: < 2.2e-16

El análisis realizado mediante un modelo de regresión lineal múltiple permitió identificar los factores que más influyen en el precio de una laptop. Utilizando cinco variables clave:

- tipo de laptop (TypeName)
- memoria RAM (Ram)
- resolución de pantalla (Ppi)
- procesador (Cpu_brand)
- almacenamiento SSD

Se logró un ajuste del modelo del 80%, lo que significa que estas características explican gran parte de la variación en el precio.

Por ejemplo:

Las laptops tipo Notebook o Gaming entre otras, con pantallas de alta calidad (mayor PPI), con procesadores potentes como Intel i5 o i7 y con más memoria RAM y SSD, tienden a ser más costosas.

En conclusión, el modelo es estadísticamente significativo y confiable para analizar el comportamiento del precio de laptops según sus componentes. Esta información permite comprender cómo cada característica contribuye al valor de las computadoras.