

A10: Laptop Data Graphs with ggplot2

The data set used in this exercise was downloaded from Kaggle at the following URL:

<https://www.kaggle.com/datasets/jacksondivakarr/laptop-price-prediction-dataset?resource=download>. A brief description of each included column is provided below. Produce the graphs described below using ggplot2 and deliver the results as an HTML webpage generated from an R Markdown or Quarto file. Use headers or text to differentiate each component of the assignment. Make sure to include both the code and the resulting graphs in your submission. When reading in the data, make sure all character columns are treated as factors.

brand: laptop brand name

name: name of laptop

price: price in US Dollars×100 (divide by 100 to get price)

spec_rating: specification score (0 to 100)

processor: processor name

CPU: central processing unit (CPU) specs

Ram: amount of installed RAM

Ram_type: type of RAM

ROM: size of hard disk

ROM_type: type of hard disk (SSD or Hard-Disk)

GPU: installed graphics processing unit (GPU)

display_size: size of display in inches

resolution_width: resolution in width dimension in pixels

resolution_height: resolution in height dimension in pixels

OS: operating system

warranty: length of warranty in years

***The following tasks from A3 are required to complete this exercise.**

T1: Write code to recode the Ram factor levels as follows and convert to a numeric type (Original → New):

"12GB" = "12", "16GB" = "16", "2GB" = "2", "32GB" = "32", "4GB" = "4", "64GB" = "64", "8GB" = "8".

T2: Write code to recode the ROM factor levels as follows and convert to a numeric type (Original → New):

"128GB" = "128", "1TB" = "1000", "256GB" = "256", "2TB" = "2000", "32GB" = "32", "512GB" = "512", "64GB" = "64".

T3: Write code to create a field that indicates whether the machine has a NVIDIA GPU.

T4: Write code to create a single column that differentiates between Intel and AMD processors. Any other manufacturer should be coded as "Other".

T5: Write code to create a single column that differentiates between i3, i5, i7, and i9 Intel processors. All other processors should be coded as "Other".

G1: Create a histogram showing the distribution in prices of only computers with i3, i5, i7, and i9 processors (not differentiated). (4 Points)

G2: Create a grouped kernel density plot to compare the distributions in prices of only computers with i3, i5, i7, and i9 processors. (4 Points)

G3: Create a grouped boxplot that compares the distribution of prices for all computers with i3, i5, i7, or i9 processors. All other computers should be excluded. (4 Points)

G4: Create a grouped violin/boxplot that compares the distribution of prices for all computers with i3, i5, i7, or i9 processors. All other computers should be excluded. (4 Points)

G5: Create a grouped boxplot that compares the distribution of prices for all computers with i3, i5, i7, or i9 processors. Use the fill color to differentiate computers that do and do not have an NVIDIA graphics card. All other computers should be excluded. (4 Points)

G6: Create a scatterplot with price mapped to the x axis and spec_rating mapped to the y-axis for only computers with i3, i5, i7, or i9 processors. Differentiate the processor type using the point color. (4 Points)

G7: Add to graph G6 by using point size to show the amount of RAM. (4 Points)

G8: Create a scatterplot with the display_size mapped to the x-axis and price mapped to the y-axis for only computers with i3, i5, i7, or i9 processors. Differentiate the processor type using the point symbol shape. (4 Points)

G9: Create a boxplot with brand mapped to the x axis and price mapped to the y axis for only computers with i3, i5, i7, or i9 processors. Differentiate the processor using fill color. (4 Points)

G10: Create a boxplot with brand mapped to the x axis and price mapped to the y axis for only computers with i3, i5, i7, or i9 processors. Differentiate the processor using fill color. Only include the following brands: Acer, Asus, Dell, HP, Lenovo, and MSI. Add horizontal lines showing the mean price for each processor type using just this subset of data. (4 Points)