

Word count: 747

Column	Situation	Action	Justification
brand	Same brand with multiple names. Some non-existent brands	Map to one brand. Remove non-existent brands	Categorising is easier as no duplicate categories
model	Nans	Drop nans	Model is the primary key
screen_size, price	Values are strings	Extract numbers	Numerical analysis can be performed if values are floats
color	Different shades of the same colour	Map to one colour	Shade of the same colour isn't important
harddisk, ram	Multiple units	Convert units and turn to float. Drop nans	Numerical analysis can be performed if values are floats and same unit
cpu	Some specific CPU models, mostly just CPU family (i.e. intel i5)	Map all CPUs to their CPU family	Data can now be categorised by CPU family
OS	Same OS with multiple names	Map to one OS	Categorising is easier as no duplicates
special_features	Same special features with multiple names. High nan %	Map to one name. Leave nans	Categorising is easier as no duplicates. Laptop doesn't need special features, so no need to drop nans
graphics	Some GPU names but mostly Integrated/Dedicated	Move GPU names to GC. Fill nans with Integrated or Dedicated based on GC	Specific GPUs should be in GC column
graphics_coprocessor (GC)	High nan %. Many names for same GPU	Drop nans. Map to one name	Knowing the GPU is vital information. Categorisation is easier
cpu_speed	High nan %, multiple units	Drop column	CPU speed alone is not a good indicator of CPU performance. This combined with a high nan% makes the column useless
rating	High nan %	None	Rating is non-critical information with a nan implying the laptop has no reviews

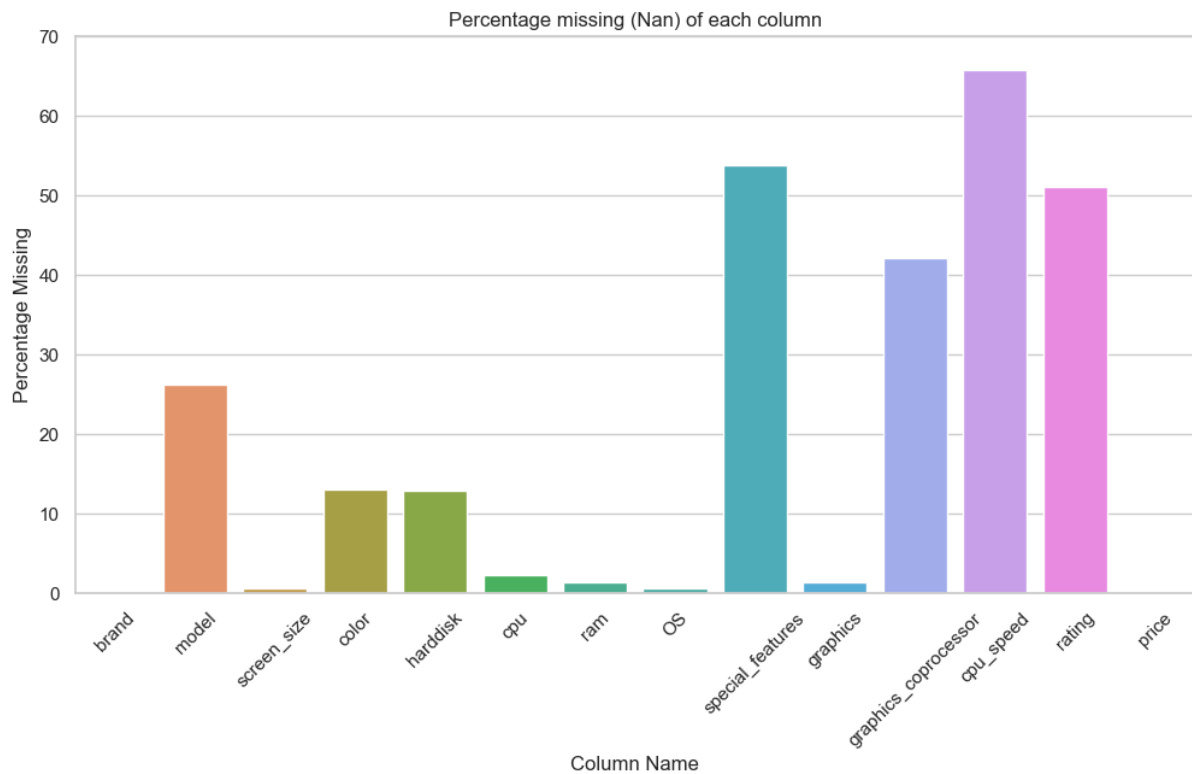


Figure 1

Customer 1 needs a new gaming laptop running Windows, since this is the OS with the most support for games. To play the latest games smoothly at high quality, a good CPU and GPU are required as well as sufficient RAM. Furthermore, lots of storage space is required to be able to store a diverse library of games as modern games can be larger than 100GB. Therefore, these 4 variables should be maximised whilst remaining within the customer's price range.

UserBenchmark's GPU dataset^[1] measures the relative performance of a GPU compared to an Nvidia RTX 2060S. Looking at the benchmarks of Nvidia's 10-40 series (figure 2) shows that performance generally improves with each series, however a GPU isn't always better just because it's in a later generation. It also worth noting that the 'Ti' version of a GPU is always better than the regular version.

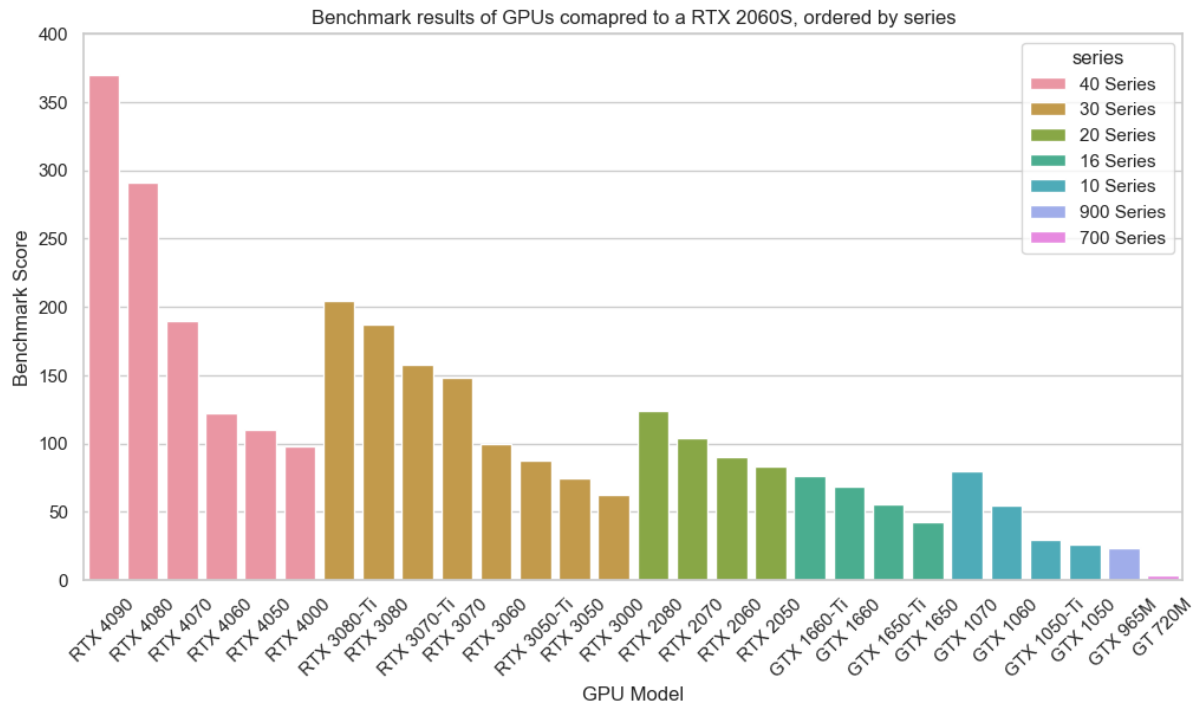


Figure 2 – Benchmark of most common gaming GPUs

UserBenchmark also has a CPU dataset^[2] which measures the effective speed of a CPU relative to an Intel i9-9900K. Most laptops in the amazon laptop dataset do not have a specific CPU model and as a result the only way to compare CPUs is by looking at the average benchmark of a laptop CPU's family. Figure 3 shows that Core i9 is on average best, with Ryzen 9 a small step behind. Otherwise, a Ryzen is generally better than it's Intel Core equivalent.

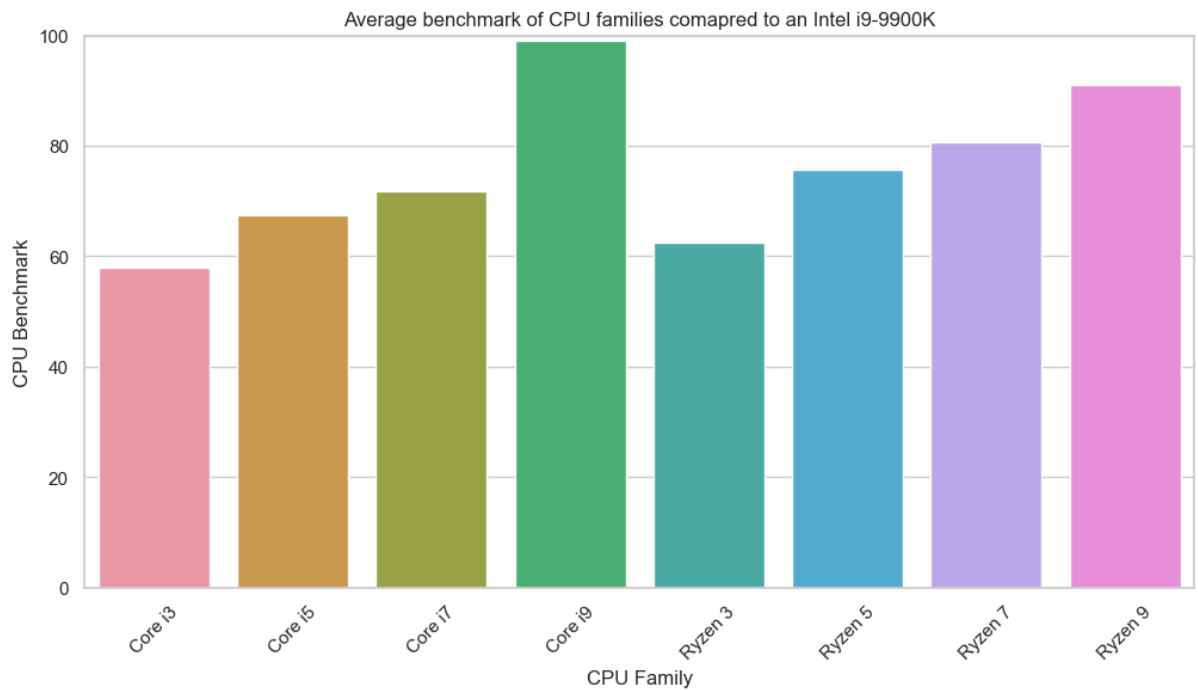


Figure 3

Using the follow equation, each of the 4 important variables is given equal important and provides a value on the total effectiveness of a laptop at gaming.

$$Gaming\ Score = RAM * 2 + \frac{Storage}{20} + GPU\ benchmark + CPU\ benchmark$$

$$Gaming\ Score\ Efficiency = \frac{Gaming\ Score}{Price}$$

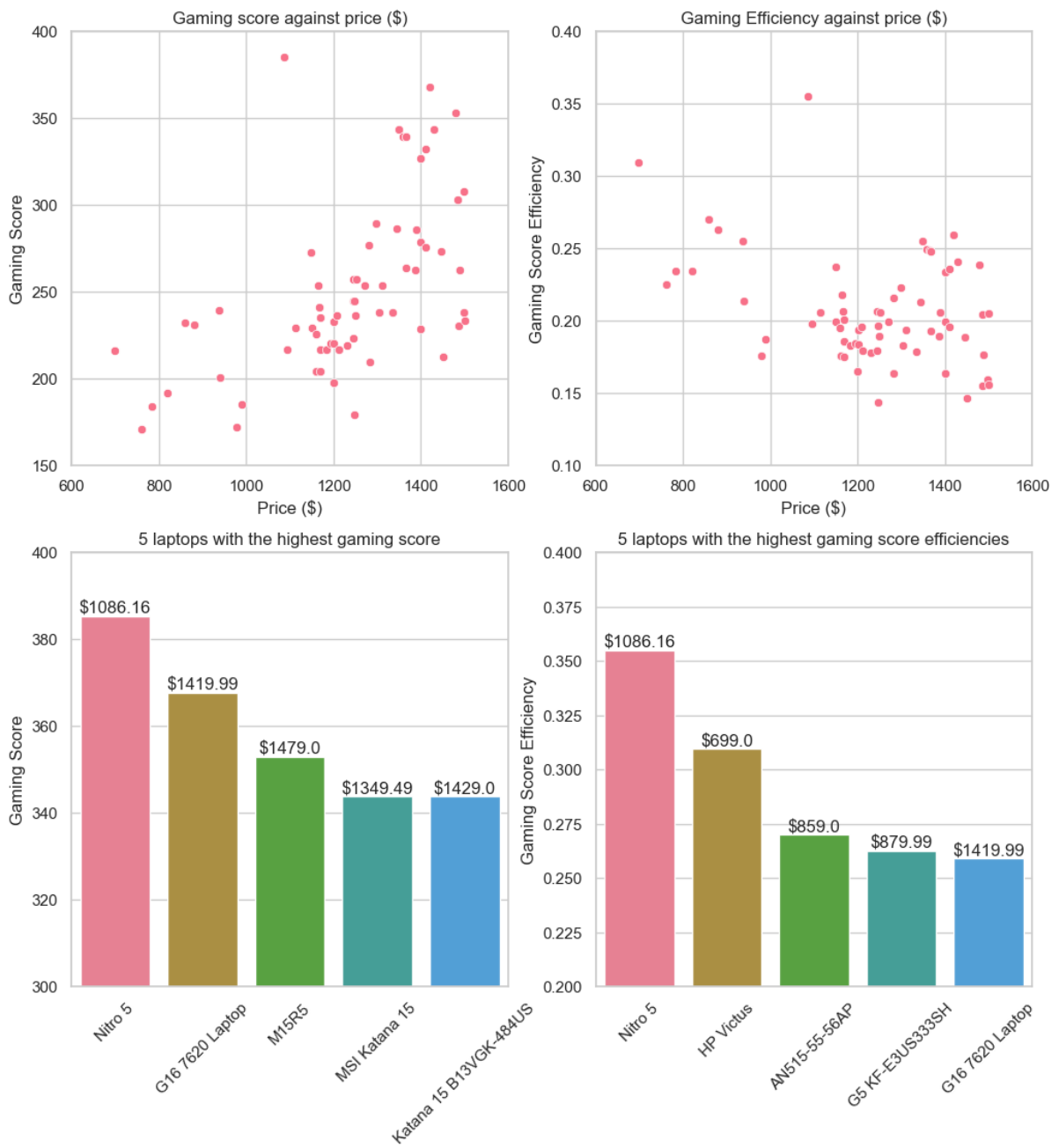


Figure 4

Taking a mix of the most powerful and most value-for-money laptops shown in figure 4, the final recommendations are:

1. G16 7620 Laptop
2. MSI Katana 15
3. Nitro 5
4. AN15-55-56AP
5. HP Victus

Customer 2 is a businessman who travels a lot. They want a laptop with; a small screen so that it is easy to fit in a bag, an above average rating to ensure the laptop is reliable and a fingerprint reader to protect sensitive business documents. With these constraints in mind, they would then like a laptop with as much storage as possible.

Looking at figure 5, any laptop must have a screen size of less than 15 inches and a rating greater than 4.1 as these are the mean values. Therefore, the laptops of interest are in the top left of the lefthand graph, the same graph revealing that there is no correlation between rating and screen size.

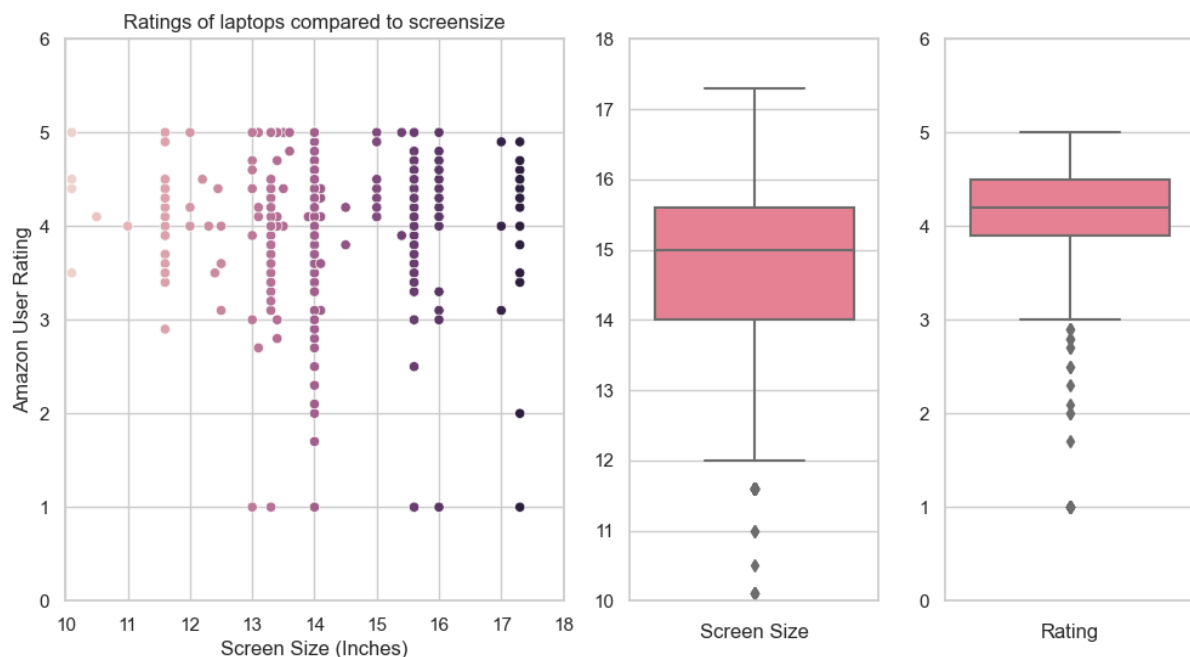


Figure 5

Figure 6 shows the laptops with fingerprint readers which meet these constraints. Taking the 5 laptops with the highest storage, the final recommendations are:

1. Lenovo Thinkpad T14
2. Dell Precision
3. XPS 13 9320
4. HP Envy
5. Acer Swift X Laptop

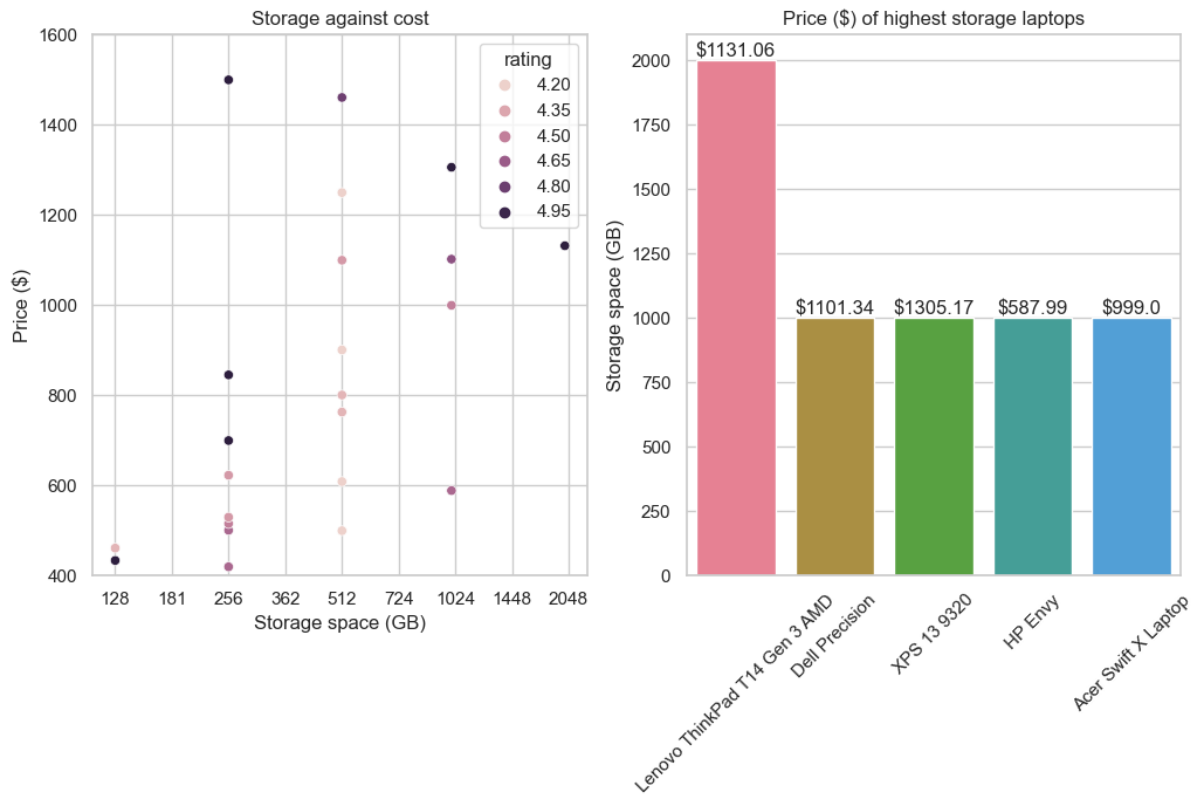


Figure 6

[1] – UserBenchmark. (2023) *GPU_UserBenchmarks*. Retrieved December 3rd, 2023, from https://www.userbenchmark.com/resources/download/csv/GPU_UserBenchmarks.csv

[2] – UserBenchmark. (2023) *CPU_UserBenchmarks*. Retrieved December 4th, 2023, from https://www.userbenchmark.com/resources/download/csv/CPU_UserBenchmarks.csv