

SAS FINAL PROJECT: Laptop Sales Analysis

Erick Tinoco & Atharva Borole

October 25, 2024

CIS 5250-01

Dataset/Datasource URL: <https://www.kaggle.com/datasets/owm4096/laptop-prices>

INTRODUCTION

In our modern society, laptops have become vital tools for various activities, including work, education, and leisure. As more individuals depend on these devices, it's crucial for businesses to understand which laptops are favored by consumers and which features they prioritize. Gaining this knowledge can greatly impact a company's performance in a competitive environment. In an article (2024), I. Mitic stated that the Global laptop sales totaled \$127.6 billion in 2023, with 114.8 million units shipped worldwide in the first two quarters of the year. The market was dominated by six large PC manufacturers: Lenovo, HP, Dell, Apple, ASUS, and Acer.

In another article (2024), Ketaki Joshi mentioned that in 2023, the global laptop market was valued at USD 170 billion. Between 2023 and 2032, this market is estimated to register the highest CAGR of 4.4%. The global laptop market is expected to reach USD 247.7 billion by 2032. This growth is driven by increasing demand in educational and corporate settings. Factors influencing this growth include technological progress, economic factors, consumer preferences, and business requirements. These statistics highlight the growing importance of laptops in various sectors. Our project aims to analyze a dataset to determine which laptop is most sought after by users.

To reach our objective, we will employ SAS, a powerful tool for analyzing and visualizing data. By creating visual representations of the information, we can enhance our understanding of consumer preferences. Our aim is to identify not just the top-selling laptop but also the essential specifications that are important to buyers. This will involve looking at factors like processing

speed, battery life, design, portability, and price to determine what features make a laptop attractive.

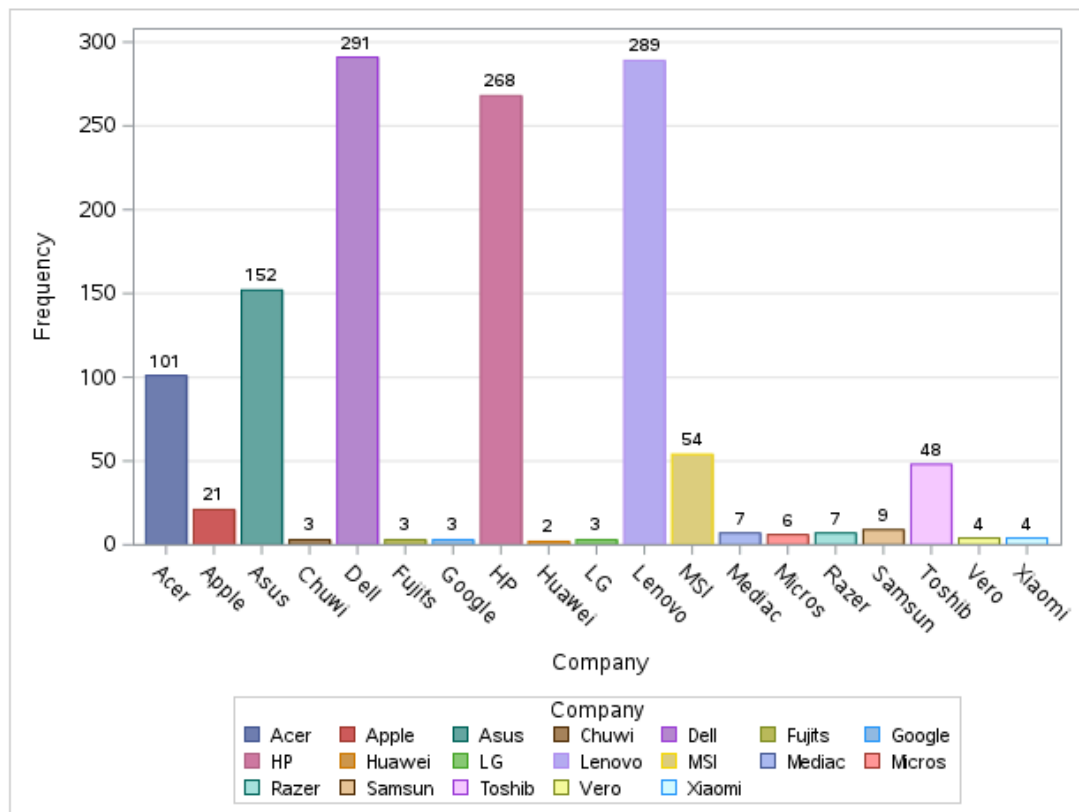
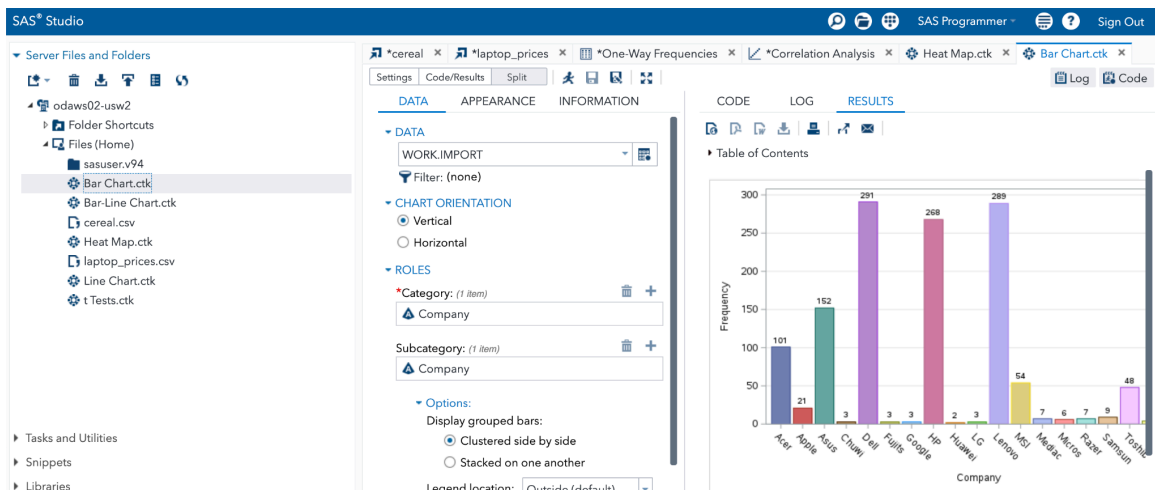
Once we identify the most popular laptop, we will also investigate which models are struggling with sales. This part of our analysis is crucial, as it can reveal key reasons why certain products might not resonate with consumers. For example, we might discover that certain laptops lack features that buyers prioritize, such as long battery life or strong performance, or that they are priced too high compared to competing models. By understanding these trends and consumer preferences, companies can make informed decisions to enhance their product offerings and refine their marketing strategies. This could involve making adjustments to design, features, or pricing to better align with what customers are looking for. Ultimately, our goal is to provide valuable insights that will guide manufacturers and retailers in meeting customer needs more effectively, ensuring that their products not only attract interest but also lead to higher sales and greater customer satisfaction.

By the end of our project, we hope to explain what draws customers to certain laptops. We want to help businesses make better decisions in a competitive market, resulting in products that more effectively meet what consumers want. Our research will benefit manufacturers while also providing useful information to consumers about the choices that suit their needs.

Company	Product	Type/Name	Screen	RAM	OS	Weight	Price (€)	Screen	ScreenW	ScreenH	Touchscreen	IPS/Panel	Retina/Display	CPU_company	CPU_freq	CPU_model	PrimaryStorage	SecondaryStorage	PrimaryStorageType	SecondaryStorageType	GPU_company	GPU_model
Apple	MacBook Pro	Ultrabook	13.3	8	macOS	1.37	1339.69	Standard	2560	1600	No	Yes	Yes	Intel	2.2	Core i5	128	0	SSD	No	Intel	Iris Plus Graphics 640
Apple	MacBook Air	Ultrabook	13.3	8	macOS	1.34	898.94	Standard	1440	900	No	No	No	Intel	1.8	Core i5	128	0	Flash Storage	No	Intel	HD Graphics 6000
HP	250 G6	Notebook	15.6	8	No OS	1.86	575	Full HD	1920	1080	No	No	No	Intel	2.5	Core i5 7200U	256	0	SSD	No	Intel	HD Graphics 620
Apple	MacBook Pro	Ultrabook	15.4	16	macOS	1.83	2537.45	Standard	2880	1800	No	Yes	Yes	Intel	2.7	Core i7	512	0	SSD	No	AMD	Radeon Pro 455
Apple	MacBook Pro	Ultrabook	13.3	8	macOS	1.37	1803.6	Standard	2560	1600	No	Yes	Yes	Intel	3.1	Core i5	256	0	SSD	No	Intel	Iris Plus Graphics 650
Acer	Aspire 3	Notebook	15.6	4	Windows 10	2.1	400	Standard	1366	768	No	No	No	AMD	3	A6-Series 9420	600	0	HDD	No	AMD	Radeon R5
Apple	MacBook Pro	Ultrabook	15.4	16	Mac OS X	2.04	2159.97	Standard	2880	1800	No	Yes	Yes	Intel	2.2	Core i7	256	0	Flash Storage	No	Intel	Iris Pro Graphics
Apple	MacBook Air	Ultrabook	13.3	8	macOS	1.34	1158.7	Standard	1440	900	No	No	No	Intel	1.8	Core i5	256	0	Flash Storage	No	Intel	HD Graphics 6000
Aus	ZenBook UX430UN	Ultrabook	14	16	Windows 10	1.3	1495	Full HD	1920	1080	No	No	No	Intel	1.8	Core i7 8550U	512	0	SSD	No	Nvidia	GeForce MX150
Acer	Swift 3	Ultrabook	14	8	Windows 10	1.6	770	Full HD	1920	1080	No	Yes	No	Intel	1.8	Core i5 8250U	256	0	SSD	No	Intel	UHD Graphics 620
HP	250 G6	Notebook	15.6	4	No OS	1.86	393.9	Standard	1366	768	No	No	No	Intel	2.5	Core i5 7200U	500	0	HDD	No	Intel	HD Graphics 620
HP	250 G6	Notebook	15.6	4	No OS	1.86	344.99	Full HD	1920	1080	No	No	No	Intel	2	Core i3 8000U	300	0	HDD	No	Intel	HD Graphics 520
Apple	MacBook Pro	Ultrabook	15.4	16	macOS	1.83	2439.97	Standard	2880	1800	No	Yes	Yes	Intel	2.8	Core i7	256	0	SSD	No	AMD	Radeon Pro 555

Visualization 1

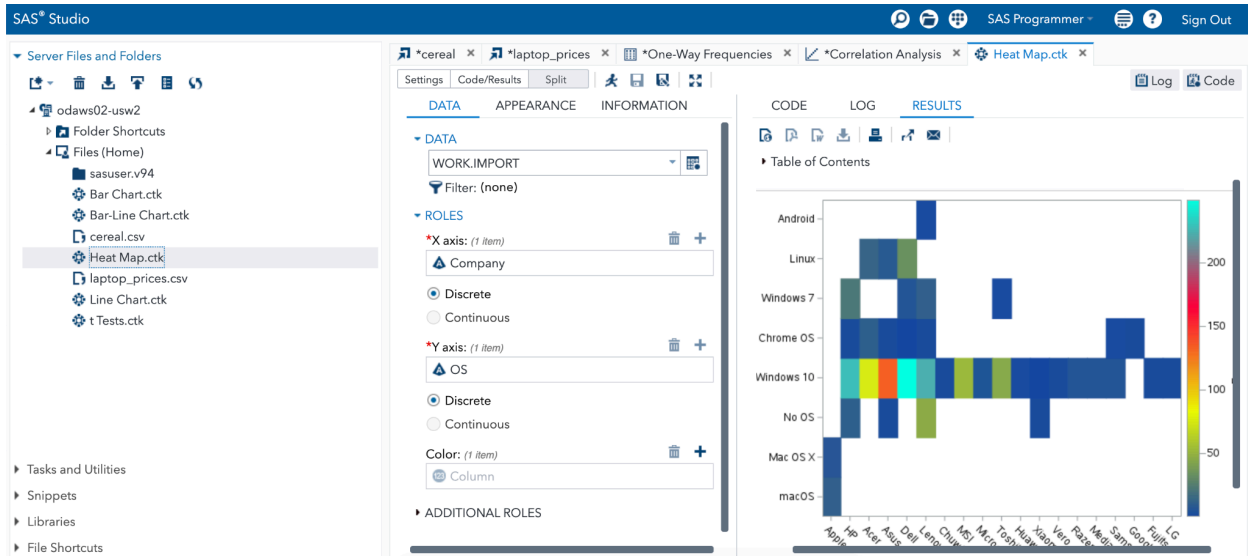
Which brand is the most bought?

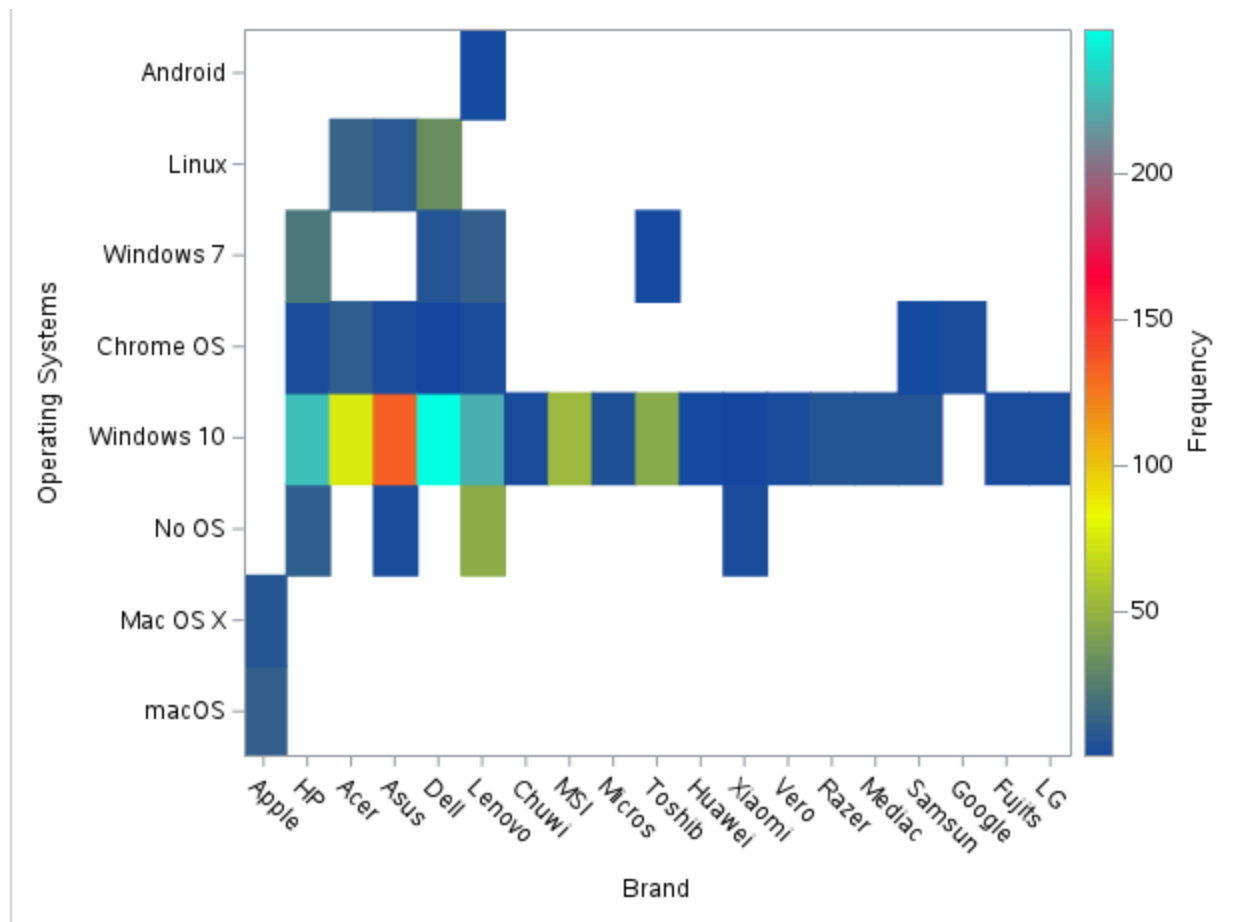


The bar chart highlights that Dell, Lenovo, and HP dominate the market, with Dell leading with 291 mentions, Lenovo at 289, and HP at 268, suggesting that their wide range of products plays a key role in their popularity. These companies are known for offering a variety of laptops and computers that cater to different customer needs, from budget-friendly models for students and casual users to high-performance devices for businesses and professionals. This broad appeal allows them to reach a wide audience and maintain a strong presence in the market. On the other hand, Apple, which is mentioned only 21 times, targets a more specific group of customers who prefer high-end, premium products. This focus on luxury devices limits its frequency in comparison to companies with a more diverse product lineup that covers various price points. Additionally, brands like Chuwi and Fujitsu, with only 3 mentions each may struggle to gain the same level of attention due to their smaller market share, limited product availability, or regional focus. Similarly, companies such as Huawei and Google, with just 2 and 3 mentions respectively, may not prioritize the traditional laptop market as heavily, resulting in fewer models and lower visibility compared to major brands like Dell and HP. Overall, according to the chart, we think that having a variety of products that meet different user needs and price ranges significantly boosts a company's market presence, while brands with a more specialized or limited approach may find it harder to compete at the same level.

Visualization 2

Within the brand, what type of Operating System is the most popular?



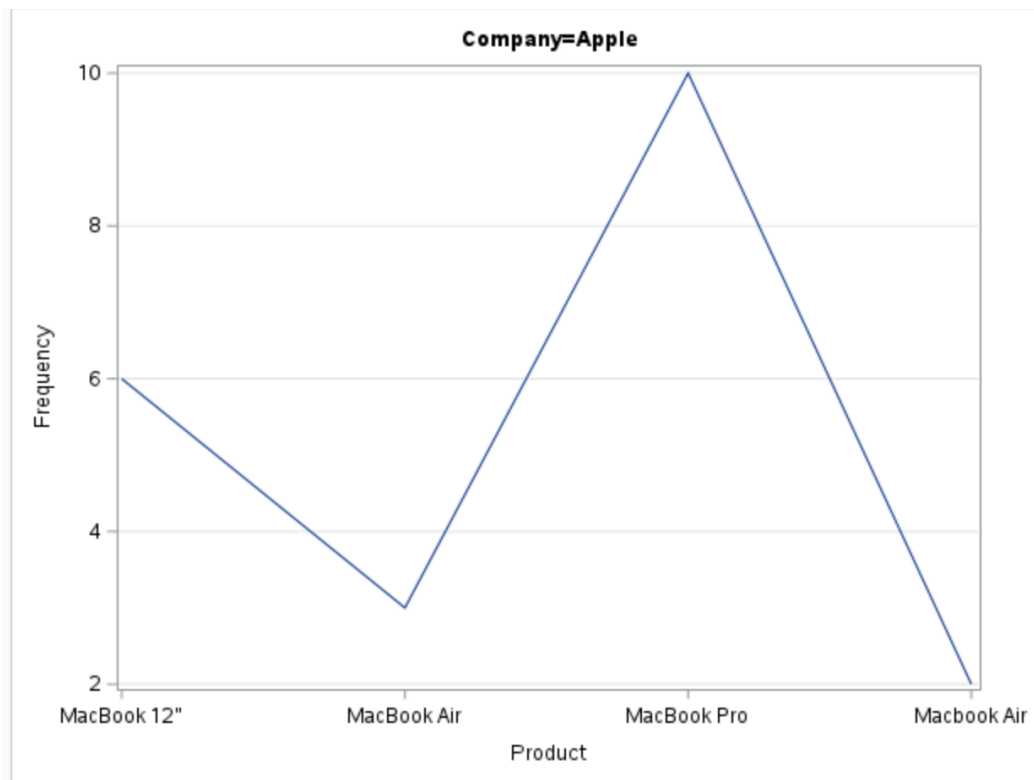
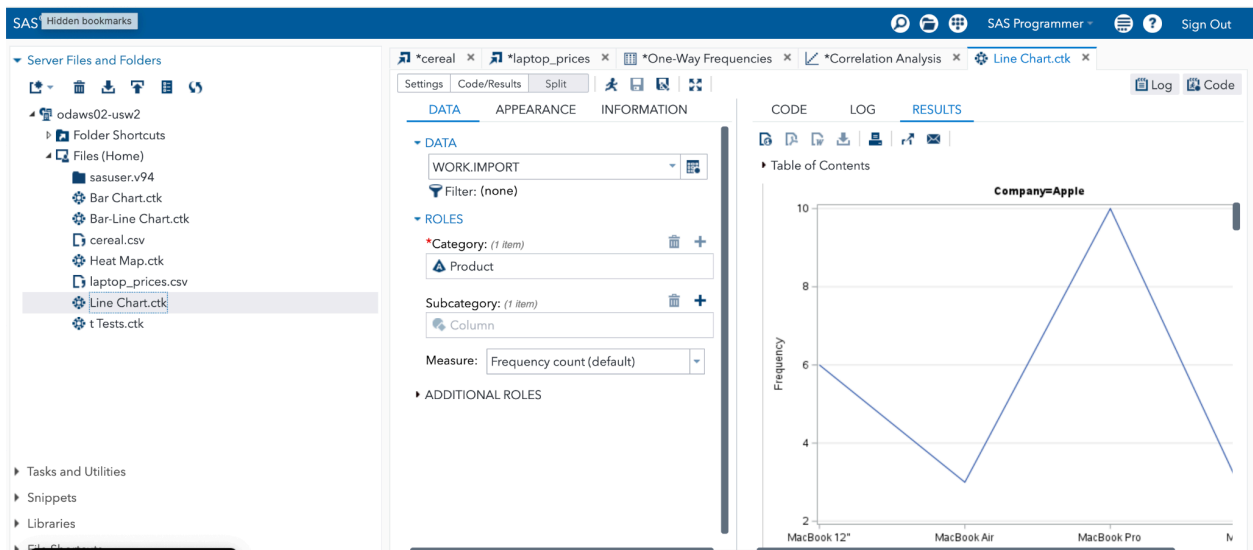


The graph shows that Windows 10 is the most popular operating system among the brands listed, with nearly all using it, except for Apple and Google. Dell is the largest user of Windows 10, followed closely by HP and ASUS. Its popularity stems from its versatility and reliability, offering strong security features like Secure Boot and Windows Hello. Windows 10 also has a user-friendly interface that adapts to different preferences, making it accessible for a wide range of users. It allows for easy access to apps and files across various devices, enhancing the overall user experience. Regular updates help keep the system secure, and its compatibility with many types of hardware improves usability. For businesses, Windows 10 provides solid security and identity protection, making it a preferred choice for professional environments. Chrome OS is the second most widely used operating system, adopted by seven brands on the list. In contrast,

macOS and macOS X are exclusive to Apple, as they are proprietary systems. Android is used only by Lenovo but unlike Apple, is not owned by the company. Other brands, such as LG, Fujitsu, Razer, Mediac, Vero, Huawei, Microsoft, MSI, and Chuwi, primarily choose Windows 10 for its flexible features and strong support. Additionally, HP, Acer, Lenovo, and Xiaomi offer laptops without any operating system, highlighting the variety of options available in the market.

Visualization 3

Within the most recognized brand, which model laptop is the most popular?

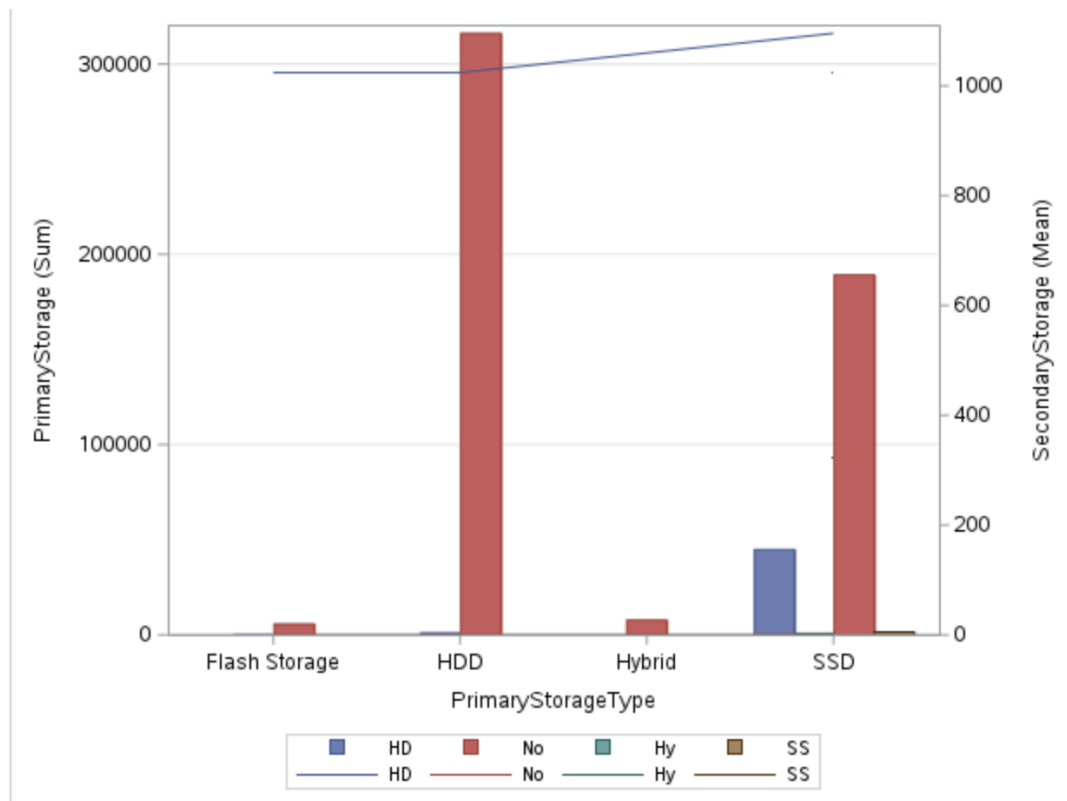
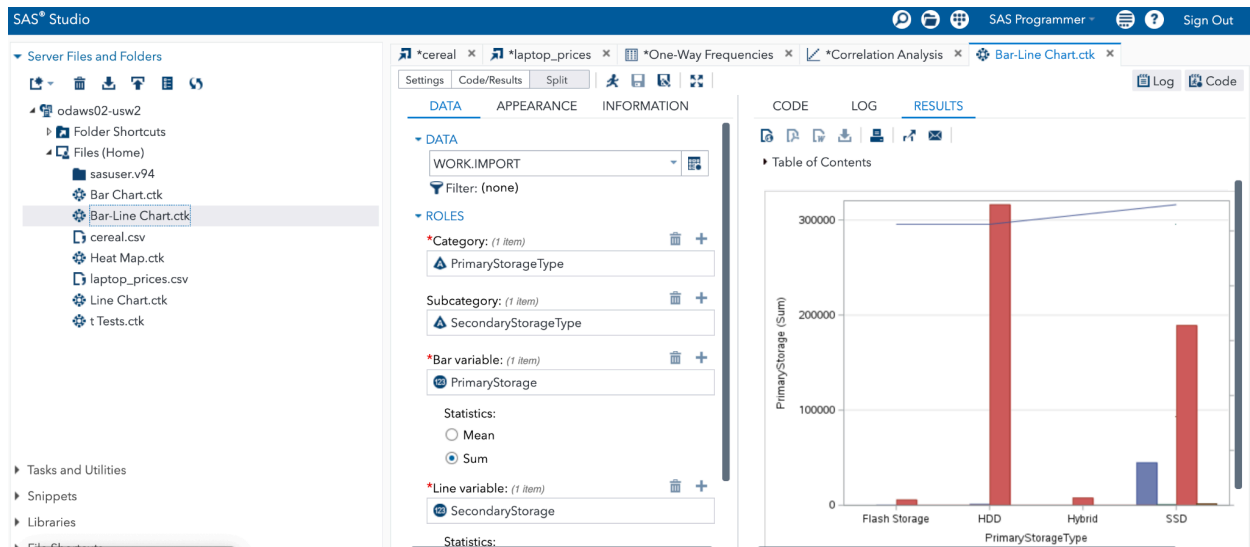


When it comes to brands in our database, Apple is the most recognized brand that most people are familiar with. When it comes to Apple, it seems that the MacBook Pro is more popular than

the MacBook Air. One aspect that is obvious between the two are display sizes. The MacBook Pro comes with a 14-inch or 16-inch display while the MacBook Air can be 13.6 inch or a 15-inch display. The MacBook Pro models are more powerful, have better screens and faster processors. When it comes to CPU, MacBook Pros use the M3 Pro which is a great choice for creative professionals and the M3 Max which is a great choice for graphics-intensive workflows like 3D modeling and high-res video editing since the CPU can support up to 128 GB of unified memory. On the other hand, MacBook Airs are lighter, cheaper, with a fan less design since the CPU is not as intensive. Generally speaking, MacBook Pros are for people who work in content creation trades such as photo and video editings, or for developers who run a lot of VMs, containers, or who process data sets that are larger than 1 TB. The MacBook Air can perform the previous stated tasks but it takes them longer due to the quality of the CPU's. In the end, college students, business professionals and avid gamers would lean towards buying a MacBook Pro based on the specifications rather than a MacBook Air. People who tend to buy the MacBook Air are not in need of a fast CPU nor a large storage capacity.

Visualization 4

Within the laptops in the database, which type of primary storage is most used?



Simply looking at the graph, the most common primary storage found in all the laptops in the database is the hard disk drive. On the other hand, the least found primary storage in the database was flash storage. In general terms, HDDs are the traditional computer storage units which are physically larger than flash since the internal composition is different. The composition consists of a spinning magnetic disk with a mechanical read/write head. Flash is typically much smaller in size and the storage capacity is even smaller. The storage capacity of flash can range from 8 MB to 1 TB. When people hear flash storage, USB thumb drives come to mind. However, flash runs on your phone and on a growing number of laptops and enterprise storage arrays. In total, flash could be considered an SSD type of storage, they are smaller and faster than their HDD counterparts but they can degrade over time since they are electrically programmed where Flash memory cells must be erased before new data can be written/ This deterioration caused by this protocol can limit the lifespan of its rewritable cells. In total, people want to rely on trustworthy primary storage that can hold a large amount of data longer. The cost of longer wait times outweighs storage types such as flash that are fast but deteriorate faster and the lifespan is smaller.

Data Description

Company	Laptop Manufacturer
Product	Brand and Model
TypeName	Laptop Type(Notebook, Ultrabook, Gaming, ...etc)
Inches	Screen Size
RAM	Total amount of RAM in laptop (GBs)
OS	Operating System installed
Weight	Laptop weight in kilograms
Price_euros	Price in Laptop in Euros (Target)
Screen	Screen definition (Standard, Full HD, 4K Ultra HD, Quad HD+)
ScreenW	Screen width (pixels)
ScreenH	Screen height (pixels)
Touchscreen	Whether or not the laptop has a touchscreen
IPSPanel	Whether or not the laptop has an IPSpanel
RetinaDisplay	Whether or not the laptop has retina display
CPU_Company	
CPU_freq	Frequency of laptop CPU (Hz)
PrimaryStorage	Primary storage space (GB)
PrimaryStorageType	Primary storage type (HDD, SDD, Flash Storage, Hybrid)
SecondaryStorage	Secondary storage space if any (GB)
SecondaryStorageType	Secondary storage type (HDD, SSD, Hybrid, None)
GPU_company	
GPU_model	

One-Way Frequency

The screenshot shows the SAS Studio interface. On the left, the 'Tasks and Utilities' pane lists various statistical procedures, with 'One-Way Frequencies' selected. The center pane shows the 'DATA' and 'ROLES' tabs. Under 'DATA', the variable 'WORK.IMPORT' is selected. Under 'ROLES', 'PrimaryStorage' is selected as the analysis variable. The right pane shows the 'RESULTS' tab, which contains a 'Table of Contents' and a table titled 'The FREQ Procedure'.

PrimaryStorage	Frequency	Percent	Cumulative Frequency	Cumulative Percent
8	1	0.08	1	0.08
16	10	0.78	11	0.86
32	43	3.37	54	4.24
64	15	1.18	69	5.41
128	175	13.73	244	19.14
180	5	0.39	249	19.53
240	1	0.08	250	19.61
256	508	39.84	758	59.45
500	124	9.73	882	69.18
508	1	0.08	883	69.25
512	136	10.67	1019	79.92
1024	240	18.82	1259	98.75
2048	16	1.25	1275	100.00

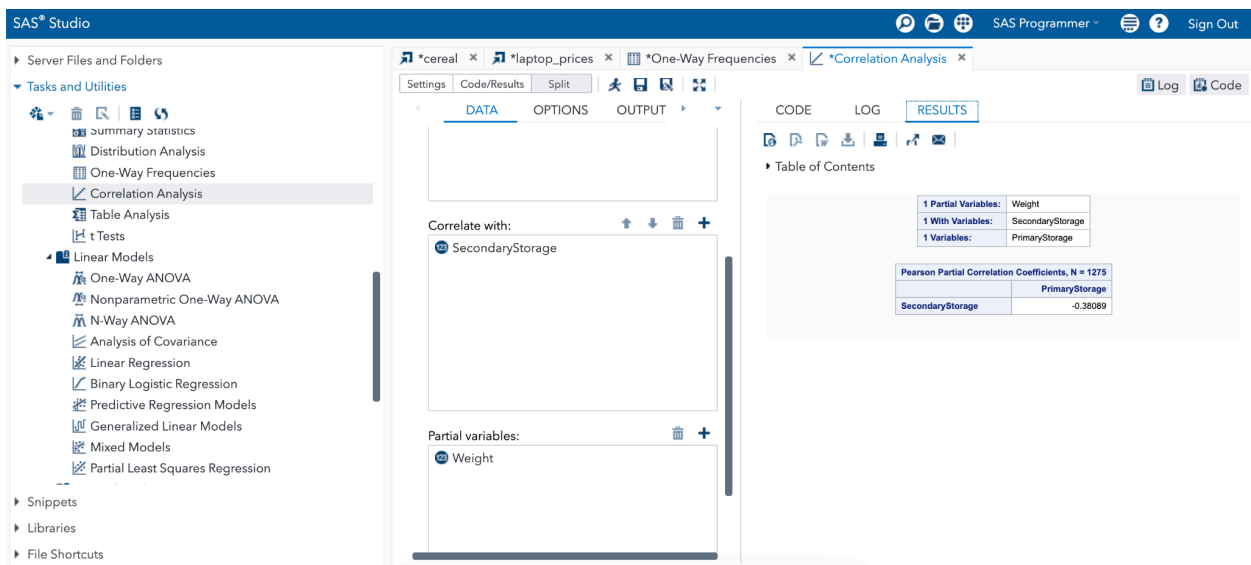
The FREQ Procedure

PrimaryStorage	Frequency	Percent	Cumulative Frequency	Cumulative Percent
8	1	0.08	1	0.08
16	10	0.78	11	0.86
32	43	3.37	54	4.24
64	15	1.18	69	5.41
128	175	13.73	244	19.14
180	5	0.39	249	19.53
240	1	0.08	250	19.61
256	508	39.84	758	59.45
500	124	9.73	882	69.18
508	1	0.08	883	69.25
512	136	10.67	1019	79.92
1024	240	18.82	1259	98.75
2048	16	1.25	1275	100.00

This frequency table summarizes the distribution of various storage capacities, measured in megabytes, across a dataset. The PrimaryStorage column is displaying the different storage

values it holds such as 8MB, 16MB, and 128 MB. The frequency column is showing how many times this storage size occurs in the dataset. This high frequency suggests that nearly 40% of the dataset consists of devices with 256 MB of storage. This cumulative information reveals that most of the data (over 98%) lies within the range of 8 MB to 1024 MB, with a large concentration specifically between 256 MB and 512 MB. The table thus provides a clear view of the distribution and helps identify common storage capacities within the dataset.

Correlation Analysis



1 Partial Variables:	Weight
1 With Variables:	SecondaryStorage
1 Variables:	PrimaryStorage

Pearson Partial Correlation Coefficients, N = 1275	
	PrimaryStorage
SecondaryStorage	-0.38089

In the correlation analysis, there are three variables in consideration: Weight, PrimaryStorage, and SecondaryStorage. The analysis shows us that there is a relationship between all three variables. In essence, if the primary storage is a Hard Disk Drive, the weight of the laptop will increase. When the weight is getting heavy, the secondary storage should not increase the weight of the laptop because it defeats the purpose of portability. Henceforth, the secondary storage decreases. The negative correlation (-0.38089) suggests that, controlling for Weight, as PrimaryStorage increases, SecondaryStorage tends to decrease. The strength of this relationship is moderate.

Statistical Summary

Hypotheses

Null Hypothesis: The Hard Disk Drive (HDD) is the most popular storage type

Alternative Hypothesis: The Hard Disk Drive (HDD) is not the most popular storage type

Type of T-Test

For this test, we will conduct a one sample T-test. After researching many articles about the topic, the majority of the articles state that the HDD is the most sold primary storage in the world. That statement will be considered as our well known value. On the other hand, the data extracted from our dataset contains data about storage which will serve as our sample mean.

Calculation of T-Test

- **One-sample t-test:**

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

On the numerator, the X represents the sample mean while the u represents the population mean.

On the denominator, the s represents the sample standard deviation while the n simply represents the sample size.

WE ARE GOING TO USE A TWO TAILED TEST because we are trying to reject the null hypothesis.

Degrees of freedom

Since our test is a One-sample t-test, the formula for degrees of freedom is as follows:

$$df = n - 1$$

Similar to the formula of a one-sample t-test, the n represents the number of cases while df is an acronym for degrees of freedom.

Variable: PrimaryStorage

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.772771	Pr < W	<0.0001
Kolmogorov-Smirnov	D	0.291487	Pr > D	<0.0100
Cramer-von Mises	W-Sq	19.43823	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	107.9968	Pr > A-Sq	<0.0050

Variable: PrimaryStorage

N	Mean	Std Dev	Std Err	Minimum	Maximum
1275	444.5	365.5	10.2371	8.0000	2048.0

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
444.5	424.4	464.6	365.5	351.9	380.3

DF	t Value	Pr > t
1274	43.42	<.0001

Summary Statistic	Explanation
Mean	Represents the sample mean
Standard Deviation	A measure of how spread out the numbers in a set of data are
Min	In the range of the data, this is the minimum value
Max	In the range of data, this is the maximum

	value
N	The sample size

From the image, these are the elements we need to focus to reach our result:

- **N is for sample size:** 1,275
- **Mean:** 444.5
- **Standard Deviation:** 365.5
- **Standard Error:** 10.2371
- **Degrees of Freedom (DF):** 1,274
- **T-value:** 43.42
- **P-value (Pr > |t|):** <0.0001

The analysis starts with the T-value. The T-value is 43.42. This is a very high t-value, suggesting a big difference between the sample mean and the null hypothesis mean (likely 0 or some other reference value).

The P-value is <0.0001. This indicates that the result is statistically significant. Since the p-value is much smaller than common significance levels (e.g., 0.05), you can reject the null hypothesis with strong confidence.

The result of the t-test is statistically significant, suggesting that the mean value of the variable "PrimaryStorage" differs from the hypothesized value (which is likely 0 or another reference

value, depending on the test setup). Despite the non-normality of the data, the large sample size makes the t-test result reliable.

References

“Laptops by the Numbers: 2024 Market Share & More.” *Fortunly*, fortunly.com/articles/lap-top-market-share/. Accessed 24 Oct. 2024.

Joshi, Ketaki. “Laptop Statistics by Companies, Shipments, OS, Market Size and Facts.” *Coollest Gadgets*, 9 Aug. 2024, www.coollest-gadgets.com/laptop-statistics.

Jones, J. (2024). *MacBook Pro vs. Macbook Air: How to decide which Apple Laptop is best for you*. ZDNET.
<https://www.zdnet.com/article/macbook-air-vs-macbook-pro-how-to-decide/>

GeeksforGeeks, J. P. (2024, September 8). *Difference between hard drives and flash drives*.
<https://www.geeksforgeeks.org/difference-between-hard-drives-and-flash-drives/>