GPU Popularity Trends Over Time Based on Steam's Hardware Survey

Alexandru Nemțoc, *Student, UPT,* and Paul Stănilă, *Student, UPT*

I. Introduction

THIS application provides a simple and easy-to-use tool for scraping the results from Steam's Monthly Hardware Survey page [1] by downloading every available snapshot of the page that it can find on TheWaybackMachine [2]. It also provides a convenient way of storing the collected data into a very popular format for datasets, namely CSV. Finally, it provides the tools to generate graphs and views of the collected data in order to better understand the relations and the popularity trends of the different series of GPUS which were present in the Steam Hardware Survey. The data provided by this tool is an important point of reference for both the average consumer that is looking to buy a new graphics card, as well as for the game developers looking to release a new game as it provides insight on not only the current and the previous generations of graphics cards, but it also makes it easy to visualize the past and current popularity trends within one or multiple generations of GPUS.

II. Data Acquisition

The data used in this project was scraped from the *Steam Hardware Survey* page [1] (Figure 1). One of the problems with this approach is that Valve only shows the data for the current month and for the four previous months. This issue was solved with the help of a digital library called *TheWaybackMachine* [2] which contains snapshots of Valve's Hardware Survey page all the way back to 2008. TheWaybackMachine uses crawlers in order to save snapshots of some of the most popular pages on the internet, but since the Hardware Survey page wasn't that popular at the beginning, there are a couple of gaps in the data for some of the earlier years.

Steam's Hardware & Software Survey page currently contains 7 tables of data, all of the data found in this survey is presented as a percentage:

- The first table, *"Overall Distribution of Cards"*, contains data related to the DirectX version of each card.
- The second table, *"All Video Cards"*, represents the individual data for each graphics card present in the survey. Only the most popular cards are shown in the table, the less popular cards being grouped under a single entry called *"Other"*.
- The remaining tables contain only the data for the GPUS which support a specific version of DirectX. Currently there are 5 such tables, namely *"DirectX 12 GPUS, DirectX 11 GPUS, DirectX 10 GPUS, DirectX 9 Shader Model 2B and 3.0 GPUS, and DirectX 9 Shader Model 2.0 GPUS"*.

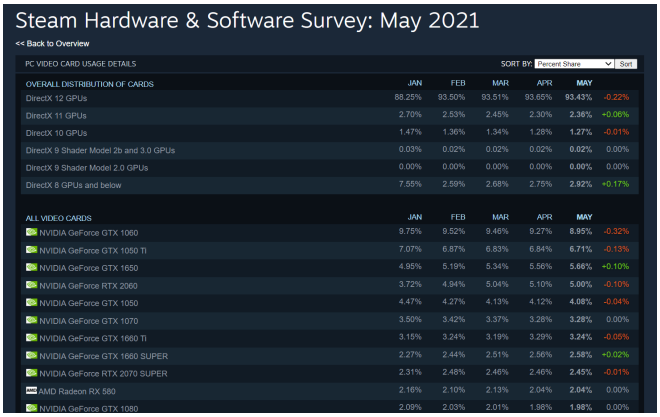

Fig. 1. Steam Hardware & Software Survey Page for GPUS: May 2021

Since the data for the older models of GPUS is no longer relevant, only the data presented in the *"All Video Cards"* table is used.

A crawler was used in order to extract the data from the webpage. The crawler starts by connecting to the WaybackMachine page of the very first snapshot which contains the Steam Hardware Survey data. Once the page is loaded, the crawler makes a copy of it onto the local disk and then it moves to the next snapshot. Figure 2 shows the console output of the crawler application. The link of the page containing the next snapshot is present in the downloaded page HTML file. The process of downloading the snapshot pages lasts until the crawler can no longer find the link to the next page, meaning that it had reached the last available snapshot.

In most cases, each month will have multiple snapshots associated with it. The next step is to only keep the most recent snapshot for each month. This is done automatically within the same project. Once the latest snapshot for a given month is found, a copy of it is saved into a folder called *latest*.

After downloading and selecting only the relevant snapshots, the data from each table needs to be extracted. This is done with the help of a Java library called *jsoup* [3]. Starting with the oldest snapshot, each file is parsed by the program and for each new GPU found in the table, a new entry is created into the data list. The data list consists of GPU entities. Each GPU entity has a name, a brand, and a map containing the monthly popularity of that GPU. The map is a *TreeMap<String, TreeMap<Month, String>>* where the first String, the key of the large map, represents the year, and the second String, the value of the small map, represents the popularity in that given month.

```
"C:\Program Files\Java\jdk1.8.0_291\bin\java.exe" ...
23:36:59.503 [main] INFO crawler.SnapshotDownloader - Downloading snapshots...
23:36:59.506 [main] INFO crawler.SnapshotDownloader - 1: 14.12.2008 -> http://web.archive.org/web/20081214203623/http://store.steampowered.com/hwsurvey/videocard/
23:37:01.930 [main] INFO crawler.SnapshotDownloader - 2: 16.12.2008 -> http://web.archive.org/web/20081216014343/http://store.steampowered.com:80/hwsurvey/videocard/
23:37:02.982 [main] INFO crawler.SnapshotDownloader - 3: 18.12.2008 -> http://web.archive.org/web/20081218003004/http://store.steampowered.com:80/hwsurvey/videocard/
23:37:03.711 [main] INFO crawler.SnapshotDownloader - 4: 19.12.2008 -> http://web.archive.org/web/20081219111733/http://store.steampowered.com:80/hwsurvey/videocard/
23:37:04.653 [main] INFO crawler.SnapshotDownloader - 5: 24.01.2009 -> http://web.archive.org/web/20090124025412/http://store.steampowered.com:80/hwsurvey/videocard
23:37:05.811 [main] INFO crawler.SnapshotDownloader - 6: 01.02.2009 -> http://web.archive.org/web/20090201122050/http://store.steampowered.com:80/hwsurvey/videocard/
23:37:07.826 [main] INFO crawler.SnapshotDownloader - 7: 11.02.2009 -> http://web.archive.org/web/20090211194920/http://store.steampowered.com:80/hwsurvey/videocard
23:37:09.590 [main] INFO crawler.SnapshotDownloader - 8: 28.02.2009 -> http://web.archive.org/web/20090228052736/http://store.steampowered.com:80/hwsurvey/videocard
23:37:11.463 [main] INFO crawler.SnapshotDownloader - 9: 09.03.2009 -> http://web.archive.org/web/20090309041349/http://store.steampowered.com:80/hwsurvey/videocard/
23:37:12.362 [main] INFO crawler.SnapshotDownloader - 10: 15.03.2009 -> http://web.archive.org/web/20090315051823/http://store.steampowered.com:80/hwsurvey/videocard
23:37:13.593 [main] INFO crawler.SnapshotDownloader - 11: 30.03.2009 -> http://web.archive.org/web/20090330184938/http://store.steampowered.com:80/hwsurvey/videocard
23:37:14.988 [main] INFO crawler.SnapshotDownloader - 12: 09.04.2009 -> http://web.archive.org/web/20090409034224/http://store.steampowered.com:80/hwsurvey/videocard/
23:37:15.861 [main] INFO crawler.SnapshotDownloader - 13: 28.04.2009 -> http://web.archive.org/web/20090428050156/http://store.steampowered.com:80/hwsurvey/videocard
23:37:17.295 [main] INFO crawler.SnapshotDownloader - 14: 25.05.2009 -> http://web.archive.org/web/20090525091916/http://store.steampowered.com:80/hwsurvey/videocard
23:37:18.470 [main] INFO crawler.SnapshotDownloader - 15: 26.05.2009 -> http://web.archive.org/web/20090526092103/http://store.steampowered.com:80/hwsurvey/videocard/
23:37:19.417 [main] INFO crawler.SnapshotDownloader - 16: 21.06.2009 -> http://web.archive.org/web/20090621014253/http://store.steampowered.com:80/hwsurvey/videocard
23:37:20.995 [main] INFO crawler.SnapshotDownloader - 17: 26.06.2009 -> http://web.archive.org/web/20090626085407/http://store.steampowered.com:80/hwsurvey/videocard/
23:37:21.874 [main] INFO crawler.SnapshotDownloader - 18: 08.07.2009 -> http://web.archive.org/web/20090708052214/http://store.steampowered.com:80/hwsurvey/videocard
23:37:22.786 [main] INFO crawler.SnapshotDownloader - 19: 01.08.2009 -> http://web.archive.org/web/20090801060013/http://store.steampowered.com:80/hwsurvey/videocard
23:37:24.846 [main] INFO crawler.SnapshotDownloader - 20: 10.08.2009 -> http://web.archive.org/web/20090810225016/http://store.steampowered.com:80/hwsurvey/videocard/
23:37:26.314 [main] INFO crawler.SnapshotDownloader - 21: 28.08.2009 -> http://web.archive.org/web/20090828230812/http://store.steampowered.com:80/hwsurvey/videocard
23:37:27.516 [main] INFO crawler.SnapshotDownloader - 22: 31.08.2009 -> http://web.archive.org/web/20090831144303/http://store.steampowered.com:80/hwsurvey/videocard/
23:37:28.691 [main] INFO crawler.SnapshotDownloader - 23: 11.09.2009 -> http://web.archive.org/web/20090911132120/http://store.steampowered.com:80/hwsurvey/videocard/
23:37:30.979 [main] INFO crawler.SnapshotDownloader - 24: 13.09.2009 -> http://web.archive.org/web/20090913053903/http://store.steampowered.com:80/hwsurvey/videocard
23:37:32.461 [main] INFO crawler.SnapshotDownloader - 25: 02.10.2009 -> http://web.archive.org/web/20091002062332/http://store.steampowered.com:80/hwsurvey/videocard
23:37:33.888 [main] INFO crawler.SnapshotDownloader - 26: 12.10.2009 -> http://web.archive.org/web/20091012230327/http://store.steampowered.com:80/hwsurvey/videocard/
23:37:36.156 [main] INFO crawler.SnapshotDownloader - 27: 31.10.2009 -> http://web.archive.org/web/20091031192115/http://store.steampowered.com:80/hwsurvey/videocard/
23:37:37.592 [main] INFO crawler.SnapshotDownloader - 28: 24.11.2009 -> http://web.archive.org/web/20091124080459/http://store.steampowered.com:80/hwsurvey/videocard/
```

Fig. 2. Console output of the crawler application. The first number represents the snapshot index, the second item represents the snapshot date, and the final item represents the snapshot URL.

Once all snapshots have been parsed, the data can be written into a CSV file. The headers of this CSV file are as follows: id, GPU Name, GPU Brand, ”MON-YEAR”, where ”MON-YEAR” represents an individual column for each month for which we have collected data. The program does not currently have the option to remove the empty month columns, but this can be done manually. Removing columns which do not contain any data helps in generating more aesthetic charts.

III. Results

The results obtained by parsing and cleaning the data are presented into the following graphs.

Figure 3 represents the monthly popularity of each manufacturer over the analyzed period. In the original survey, GPUs which do not pass a minimum popularity percentage are not listed individually, but rather under a common entry named ”Other”. This entry represents the remaining data needed to reach a total of 100%. In order to generate Figure 3, the popularity percentage of each brand was multiplied by a certain value in order to reach a total sum of 100.

Since neither AMD nor Intel have any relevant GPUs the following graphs will focus entirely on Nvidia’s past and current lineup. Figures 4, 5, and 6 shows the popularity of Nvidia’s previous generations of GPUs, namely the 500, 600, and 900 series. The current generations of Nvidia products are shown in Figures 7, 8, 9, and 10. Figures 11, 12, and 13 represent the multi-generation popularity of a certain product model. Figure 11 shows the popularity of Nvidia’s model 60 lineup, figure 12 shows the popularity of Nvidia’s model 70 lineup, and figure 13 shows the popularity of Nvidia’s model 80 lineup.

Additionally, figures 14, 15, and 16 represent the all-time most popular GPUs, the all-time most popular mobile GPUs, and the all-time most popular Ti-series GPUs.

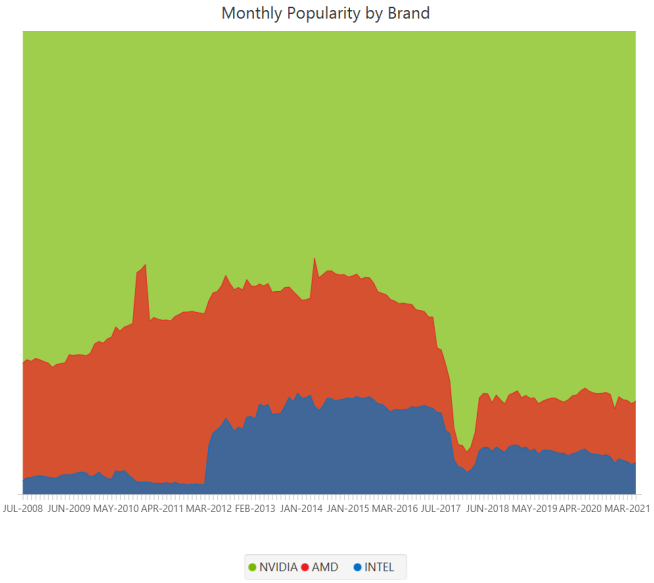

Fig. 3. Brand popularity graph showing Nvidia, AMD, and Intel GPUS.

IV. Conclusion

As we can see in Figure 3, Nvidia has always been on top when it comes to the GPU market share. Even when it was at its lowest point, it still had more than the other two GPU brands combined, and at its best, it had about 80% of the total GPU market share. This is mostly due to the lack of competition, as AMD had struggled to produce a decent graphics card for over a decade now, but things might change as its latest series of GPUS is quite promising. As for Intel, it never really was a GPU manufacturer, the only reason it is present in this graph is because of its integrated mobile units.

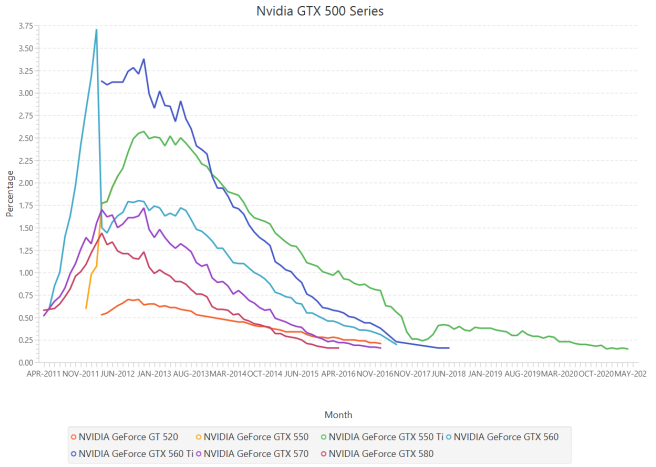

Fig. 4. Popularity graph of nvidia's GTX 500 series.

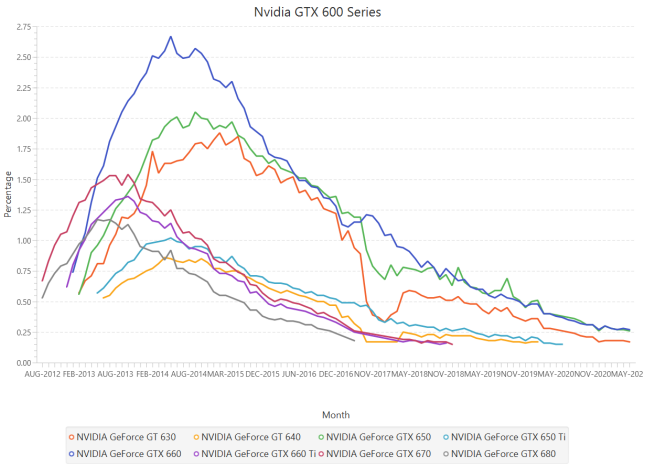

Fig. 5. Popularity graph of nvidia's GTX 600 series.

Figures 4 to 9 show us that even a decade ago, Nvidia's model 60 offered the best value for the money, a trend that was maintained all the way up the the RTX 2000 series, but things might change with the new RTX 3000 series, more specifically with the RTX 3070 offering unprecedented value.

Figures 11, 12, 13 prove that model 60 was the most popular model in every series released by Nvidia in the last decade. The popularity of these cards is double, if not triple, that of the second most popular model in the series. Occasionally, model 70 was almost, if not just as popular as the previous model, model 60, because the cost difference was not that significant, however, in recent years the price of a new graphics card went up significantly. Even if model 80 was the top of the line when it came to performance, it also came at a cost premium, making it an unlikely choice for most consumers, thus having the lowest overall popularity out of the three models that were analyzed.

Mobile GPUS were never really that popular because they were significantly slower than their PC counterparts, but still, the most popular models managed to take about 2% of the total market share within each generation (Figure 15).

In conclusion, Nvidia was and still is the dominating figure when it comes to graphics cards, AMD's most popular graphics card, the RX 580, only managing to get a mere 2% market share. Things might change in the future, since Intel will soon be launching its own dedicated GPU lineup, not to mention AMD delivering one of the most promising series of GPUs it had released in the past decade.

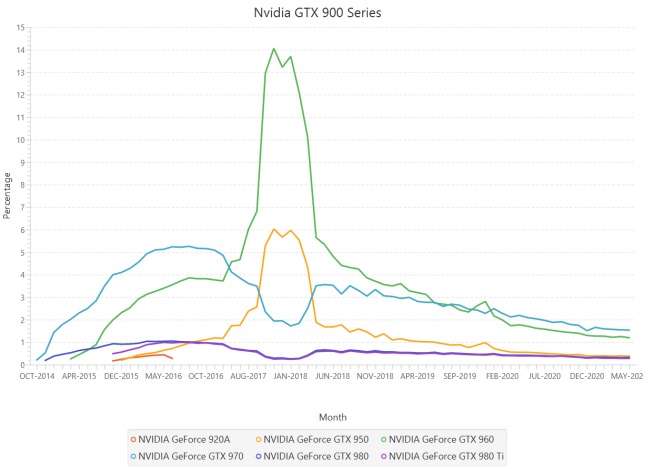

Fig. 6. Popularity graph of nvidia's GTX 900 series.

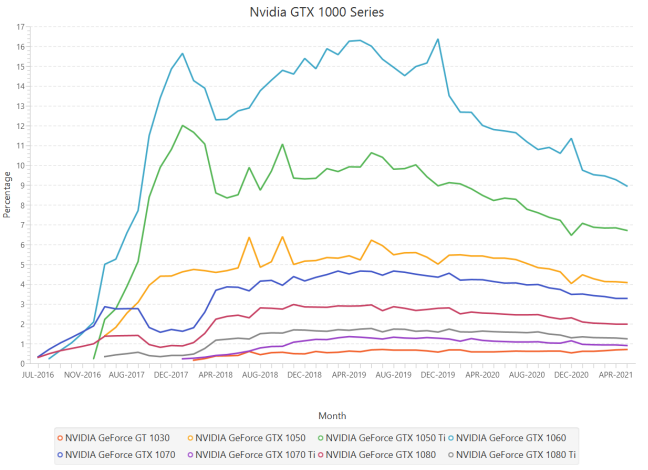

Fig. 7. Popularity graph of nvidia's GTX 1000 series.

REFERENCES

[1] Steam Hardware & Software Survey Page, https://store.steampowered.com/hwsurvey/videocard/

[2] Internet Archive - WayBackMachine, https://web.archive.org/

[3] jsoup: Java HTML Parser, https://jsoup.org/

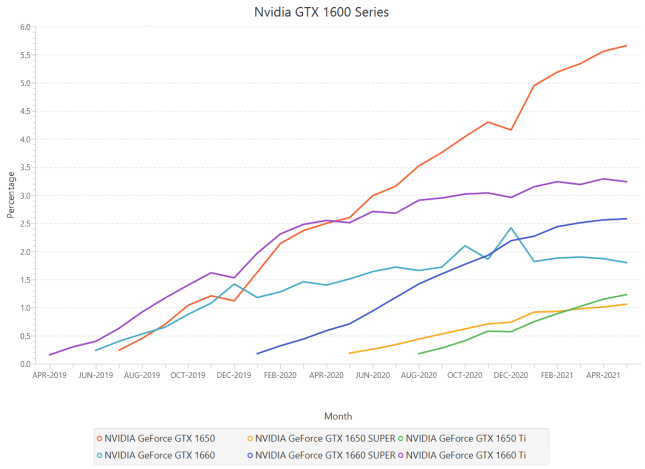

Fig. 8. Popularity graph of nvidia's GTX 1600 series.

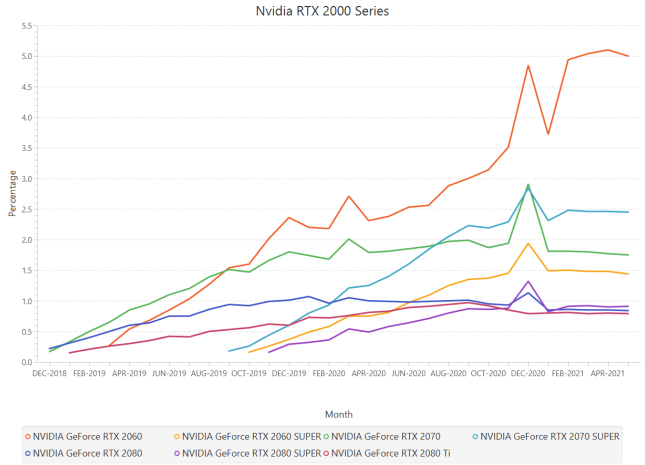

Fig. 9. Popularity graph of nvidia's RTX 2000 series.

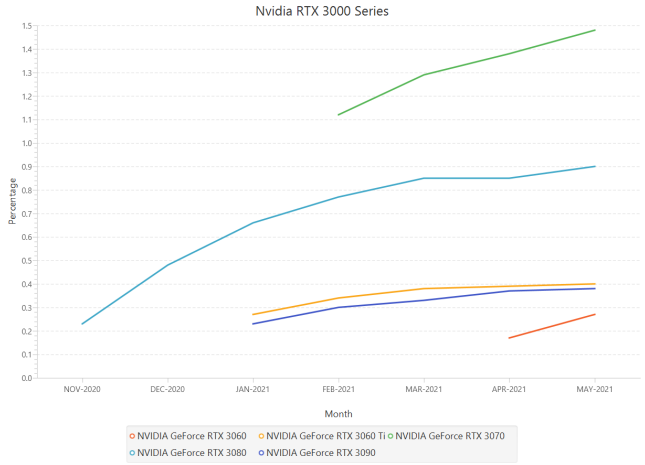

Fig. 10. Popularity graph of nvidia's RTX 3000 series.

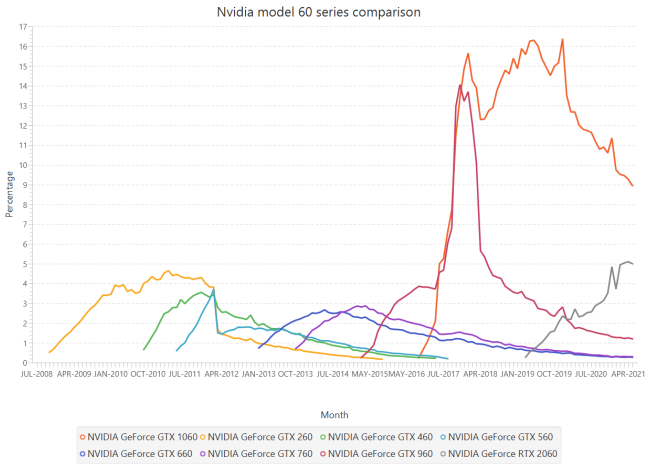

Fig. 11. Popularity graph of nvidia's model 60 lineup.

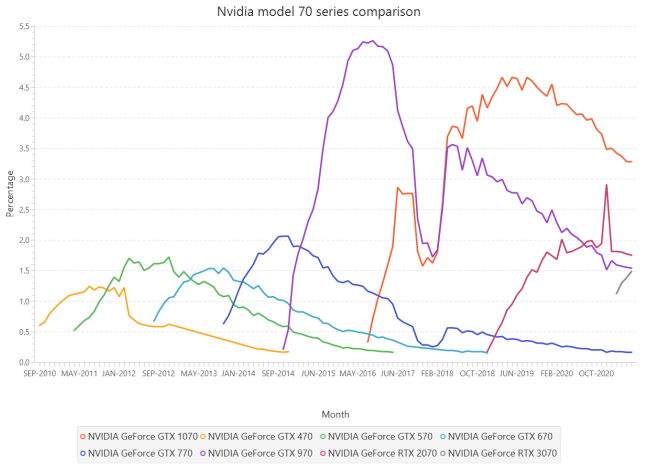

Fig. 12. Popularity graph of nvidia's model 70 lineup.

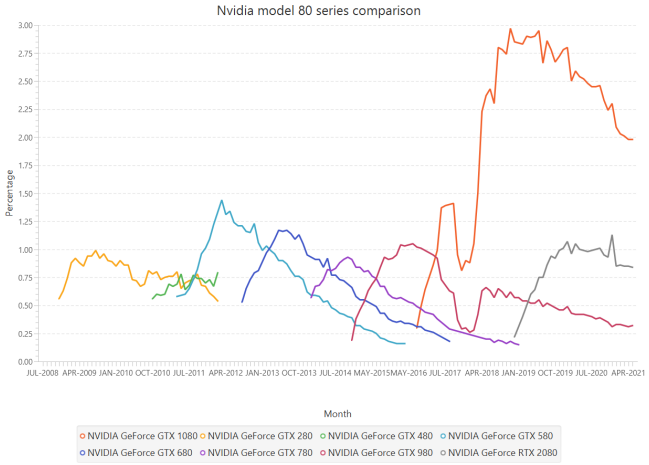

Fig. 13. Popularity graph of nvidia's model 80 lineup.

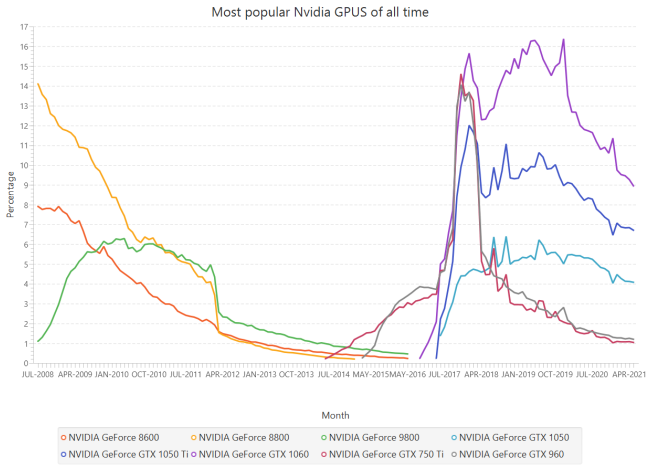

Fig. 14. Popularity graph of nvidia's all-time most popular GPUS.

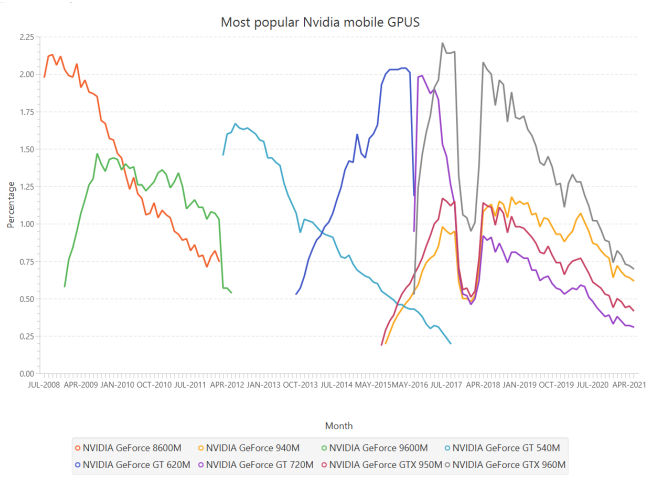

Fig. 15. Popularity graph of nvidia's most popular mobile GPUS.

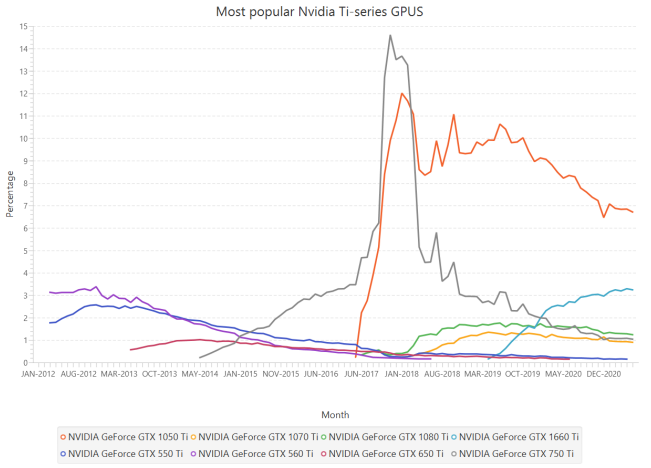

Fig. 16. Popularity graph of nvidia's most popular Ti-series GPUS.