# Video Streaming Websites Comparison Based on Usability Evaluation

Megat Irfan Zackry Bin Ismail
1st author's affiliation
School of Computing, Faculty of Engineering
Universiti Teknologi Malaysia, Johor Bahru, Malaysia
megatirfanzackry@graduate.utm.my

## ABSTRACT
This study used website performance evaluation tools including Pingdom, GTMetrix, and Website Grader to evaluate a few online video streaming websites which include Netflix, Hotstar, and Apple Tv+. In this study, I found that all three website performs very well and excels in their own way and are similar in terms of usability.

## KEYWORDS
Video Streaming Website, URL, GTMetrix, Pingdom, Website Grader, ANOVA.

## 1. INTRODUCTION
Since the global pandemic started in year 2019, most cinemas can't operate well, and it became harder to maintain profitable and some of them even went under. During that time, most people changed from buying tickets to a cinema, into purchasing subscription from online video streaming platform. This was a booming time for online streaming industry as they provide easy, accessible, and cheaper option to watch a movie.

This billion-dollar industry [1], however needs to compete to attract as many people as possible into subscribing their own services. One of the ways to seem more appealing than others is by improving their website's web performance. Low performance can bring bad evaluation to the website company and be overthrown by company that provide better service, at their website. Many criteria can affect web performance. Some of the criteria are [2]:

1. Overall load time - Time required to render the website and download it. The latency, size of files and the number of HTTP requests by the file usually affects the load time. It can be reduced by the availability of preloaded file.

2. Time for a site to be usable - Means the time required for a user to be able to use it even when there are still some other assets being loaded in the background.

3. Smoothness and interactivity - Refers to how users feel when using the site and how smooth it is when browsing the site, or ability to click and open button and pop-ups.

4. Perceived performance – Refers to how fast it took for the website to respond to users request and what the website can do to prevent user to think there is something wrong with the website thinking it is broken when loading a large size of the file.

5. Performance measurements - Refers to the actual speed when no request being made by the user.

## 2. METHODOLOGY
This study was conducted by using the software-based usability evaluation method and ANOVA test. Three video streaming websites usability scores were collected using three usability testing tools. The selected website and its URL are as shown in Table 1.

**Table-1:** URL for each Video Streaming Website

| Video Streaming Website | URL |
|---|---|
| Netflix [3] | https://www.netflix.com/ |
| Hotstar [4] | https://www.hotstar.com/ |
| Apple Tv+ [5] | https://tv.apple.com/ |

Three usability testing tools were used in this study to achieve best analysis and comparison in terms of many aspects. Below are the description and the reasons to use each of the testing tools

### 2.1 Pingdom [6]
Pingdom had almost the same report compared to GTMetrix except its method in determining the time to stop test. Pingdom determines the time to stop test with on load time, which will stop when processing on the page is finished while GTMetrix uses the fully loaded time that stops after 2 seconds of no internet activity, which is theoretically longer than load time.

### 2.2 GTMetrix [7]
GTMetrix is a web performance testing tool that gives you insights into how well any website loads and provides professional reports about the website's analysis includes the First Paint (FP) and First Contentful Paint (FCP). With seven global test regions, over 20 mobile devices and variants, and multiple connection speeds and Analysis Options, you can test your site and optimize it for any scenario [8]. This usability testing tool was chosen due to its complete report on various aspects of analyzing the websites' performance.

### 2.3 Website Grader [9]
This website evaluation tool provides analytic results such as security score, accessing the website via mobile score, search engine optimization (SEO) and performance score which was not provided by other two website performance analytics tools.

While analyzing the websites, I found that Pingdom's location is Asia-Tokyo, Japan, and GTMetrix is Vancouver, Canada. The Website Grader did not specify location due to the website's aspect evaluation does not differ based on the location.

After analyzing all the three video streaming websites on the usability testing tools, the result and highest score in the usability testing tool were generated to ease data presentation. The website with the highest score and the best video streaming websites was concluded based on the analysis from the three usability testing tools.

## 2.4 Analysis of Variance (ANOVA) test

The last method is to use the scores and results collected from the software-based usability evaluation tools and perform an Analysis of Variance (ANOVA) test which is a method of testing the equality of three or more population means by analysing sample variances [10]. It is also used to test the differences between the means of each website performance using the three previously mentioned website testing tools.

Three main results from each website testing tool that will be used are the performance score from Pingdom, overall GTMetrix page speed from GTMetrix, overall score from Website Grader. The ANOVA test is conducted using Microsoft Excel. Once all three websites are evaluated and analysed using the three website usability tools and ANOVA test, the website with the best and highest scores can be found and concluded as the best website in terms of usability among the three video streaming websites.

# 3. RESULT AND DISCUSSION

### A. Results of Pingdom Tool

The first website evaluator tools that was used was Pingdom. The video streaming websites were tested based on overall performance, page size, load time, requests, and grade.

Table 2 shows the result of the video streaming website based on the five factors that were evaluated by Pingdom. Table 3 shows the best video streaming websites with the highest score based on each usability criteria.

**Table-2:** Pingdom Tool For 3 Video Streaming Website

| Video Streaming Website | Performance | Page Size (MB) | Load Time (s) | Requests | Grade |
|---|---|---|---|---|---|
| Netflix | 88 | 2.00 | 0.923 | 44 | B |
| Hotstar | 96 | 1.37 | 0.506 | 73 | A |
| Apple Tv+ | 83 | 8.50 | 2.450 | 119 | B |

**Table-3:** Evaluation of Video Streaming Websites Based on Usability Criteria

| Website Analysis Criteria | Video Streaming Website | Highest Score Points |
|---|---|---|
| Performance | Hotstar | 96 |
| Page Size | Hotstar | 1.37 |
| Load Time | Hotstar | 0.506 |
| Requests | Netflix | 44 |
| Grade | Hotstar | A |

### B. Results of GTMetrix Tool

The second website evaluator tools that was used was GTMetrix. The video streaming websites were tested based on page speed grade, fully loaded time, total page size, and the total number of requests sent.

Table 4 shows the result of the video streaming website based on the four factors that were evaluated by GTMetrix. Table 5 shows the best video streaming websites with the highest score based on each usability criteria.

**Table-4:** GTMetrix Tool For 3 Video Streaming Website

| Video Streaming Website | Page Speed Grade | Fully Loaded time (s) | Total page size | Total number of requests |
|---|---|---|---|---|
| Netflix | A (96%) | 5.8 | 1.91 MB | 31 |
| Hotstar | B (81%) | 6.0 | 833 KB | 30 |
| Apple Tv+ | E (50%) | 16.1 | 73.3 MB | 202 |

**Table-5:** Evaluation of Video Streaming Websites Based on Usability Criteria

| Website Analysis Criteria | Video Streaming Website | Highest Score Points |
|---|---|---|
| Page Speed Grade | Netflix | A (96%) |
| Fully Loaded time | Netflix | 5.8s |
| Total page size | Hotstar | 833 KB |
| Total number of requests | Hotstar | 30 |

### C. Results of Website Grader Tool

The last website evaluator tools that were used is Website Grader. The video streaming websites were tested based on performance, SEO, mobile, security, and the overall score.

Table 6 shows the result of the video streaming website based on the five factors that were evaluated by Pingdom. Table 7 shows the best video streaming websites with the highest score based on each usability criteria.

**Table-6:** Website Grader Tool For 3 Video Streaming Website

| Video Streaming Website | Performance (30) | SEO (30) | Mobile (30) | Security (10) | Overall (100) |
|---|---|---|---|---|---|
| Netflix | 14 | 0 | 30 | 5 | 49 |
| Hotstar | 15 | 25 | 30 | 5 | 75 |
| Apple Tv+ | 7 | 30 | 20 | 10 | 67 |

**Table-7:** Evaluation of Video Streaming Websites Based on Usability Criteria

| Website Analysis Criteria | Video Streaming Website | Highest Score Points |
|---|---|---|
| Performance | Netflix | 15 |
| SEO | Apple Tv+ | 30 |
| Mobile | Netflix Hotstar | 30 |
| Security | Apple Tv+ | 10 |
| Overall | Hotstar | 75 |

### D. Results from ANOVA test

The ANOVA test is done by taking the three main results from each website testing tool which are the performance score from Pingdom, overall GTMetrix page speed from GTMetrix, overall score from Website Grader. The test is conducted with a significant value (α) of 0.05. Table 8 show the results that were collected.

**Table-8:** 3 main results from each website tools.

| Video Streaming Website | Pingdom | GTMetrix | Website Grader |
|---|---|---|---|
| Netflix | 88 | 96 | 49 |
| Hotstar | 96 | 81 | 75 |
| Apple Tv+ | 83 | 50 | 67 |

**Table-9:** ANOVA test Part 1

| Groups | Count | Sum | Average | Variance |
|---|---|---|---|---|
| Netflix | 3 | 233 | 77.6667 | 632.3333 |
| Hotstar | 3 | 252 | 84 | 117 |
| Apple Tv+ | 3 | 200 | 66.66667 | 272.3333 |

**Table-10:** ANOVA test Part 2

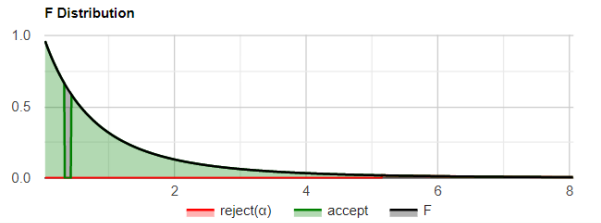
| Source of Variation | Sum of Square | df | Mean Square | F-value | P-value |
|---|---|---|---|---|---|
| Between Groups | 461.556 | 2 | 230.778 | 0.67765 | 0.54282 |
| Within Groups | 2043.33 | 6 | 340.556 | | |
| Total | 2504.89 | 8 | | | |

$$H_0: \mu 1 = \mu 2 = \mu 3$$

$$H_1: At\ least\ one\ mean\ is\ different.$$

From the tables above, there a few things I can conclude. The degree of freedom from the data is calculated to be numerator = 2 and denominator = 6. With significance value 0.05, The F-critical value = is 5.143253. According to the test performed, the F – value is 0.67765. P-value is also calculated to be = 0.54282.

**Figure 1:** F distribution graph to show rejection α




Since P-value 0.54282 is greater than significance level 0.05 and F-value statistics 0. 67765 is less than F-critical value 5.143253, I **accept $H_0$.** There is sufficient evidence to claim that the means of all three video streaming websites are considered to be equal.

Based on the 3 website evaluation tools, here are the top performing websites according to each website tool.

**Table 8:** Best performing websites based on each website evaluation tool

| Website Evaluation Tool | Best Performing Website | Overall Score/Grade |
|---|---|---|
| Pingdom | Hotstar | A (96%) |
| GTMetrix | Netflix | 96/100 |
| Website Grader | Apple Tv+ | 67/100 |

## 4. CONCLUSION

At the end of this study, I have found several interesting outcomes. Among three software-based tools that were mentioned, each video streaming websites earned themselves the best overall score for different tools. Each of the website perform well with their own and aspect. For example, Netflix, performs well in terms of their page speed on GTMetrix, Hotstar, were the best at their performance according to Pingdom, and Apple Tv+, even though received low score from both tools, managed to get the highest overall score at Website Grader.

I conclude that in terms of usability, all three website performs very well and excels in their own way. The ANOVA test that has been done suggest that all three websites are similar in terms of usability.

## 5. REFERENCES

[1] Julia Stoll, "Video streaming worldwide - statistics & facts," Video streaming worldwide - statistics & facts | Statista, 02-Mar-2021. Accessed 9 November 2021, https://www-statista-com.ezproxy.utm.my/topics/7527/video-streaming-worldwide/.

[2] MDN contributors, 2020, What is web performance, 30 Oct 2020 https://developer.mozilla.org/en-US/docs/Learn/Performance/What_is_web_performance.

[3] "Unlimited movies, TV shows, and more.," Netflix. https://www.netflix.com/

[4] "Watch TV Shows, Movies, Live Cricket Matches Online," Hotstar. https://www.hotstar.com/.

[5] "Apple TV," Apple TV. https://tv.apple.com/.

[6] Pingdom, 2021, SolarWinds Worldwide, Accessed 9 November 2021, https://tools.pingdom.com.

[7] GTMetrix, 2020, Carbon60, Accessed 9 November 2021, https://gtmetrix.com.

[8] G. L. Rand, "GTMetrix – Assuring Optimal Website Performance for All Your Users," GTMetrix Interview, 11-May-2020. Accessed 9 November 2021, https://www.websiteplanet.com/blog/gtmetrix-interview/. Accessed: 11-Nov-2021.

[9] Website Grader n.d, Google Lighthouse, Accessed 9 November 2021, https://website.grader.com

[10] D. Kiernan, "Natural Resources Biometrics," Chapter 5: One-Way Analysis of Variance | Natural Resources Biometrics, 16-Jan-2014. Accessed 9 November 2021, https://courses.lumenlearning.com/suny-natural-resources-biometrics/chapter/chapter-5-one-way-analysis-of-variance/

## 6. AUTHOR'S PROFILE

Name: Megat Irfan Zackry
Matrix No.: A20EC0205
Email: megatirfanzackry@graduate.utm.my
Course Code: SECV2113
Section: 03