 Due by 3:00pm, 21st September 2021

CMPSC 301
Data Analytics
Fall 2021

Lab 2: Google Analytics
and Response to Marketing Analytics

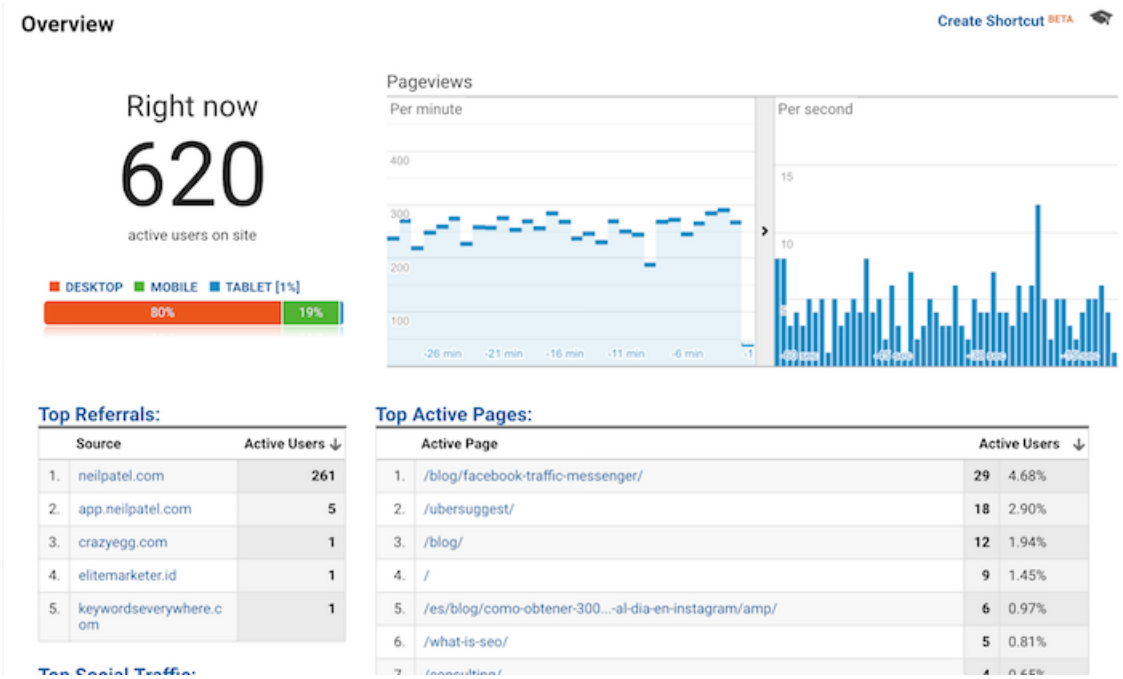

Figure 1: Google Analytics allows admins to study the traffic and activities on webpages. Information includes the number of viewers online, and the number of documents that they are viewing currently, as shown.

Objectives

Google Analytics, shown in Figure 1, is one of the most popular enterprise web analytics platforms providing rich insights into website traffic and marketing effectiveness. In the first part of the lab you are invited to investigate the available tools on Google Analytics and to summarize your understanding of the function of these tools.

You are then asked to further develop your website by adding two reflection pieces which center on ethical premises. You are also asked to attach your own analytics functionality to your website because you and your colleagues will be visiting the websites of others in the class later on to respond to the ethical reflection pieces. During this website research, your Google Analytics will be tracking traffic for you to use for the last part of the lab.

Reading Assignment

Please review class slides and your class notes. You can also find useful information in the Google Analytics Community by performing online research. Please take some time to gain experience

with using Markdown to write your work. See *Mastering Markdown* `https://guides.github.com/features/mastering-markdown/` for more details about Markdown.

GitHub Starter Link

https://classroom.github.com/a/3xahBLim

To use this link, please follow the steps below.

- Click on the link and accept the assignment.
- Once the importing task has completed, click on the created assignment link which will take you to your newly created GitHub repository for this lab.
- Clone this repository (bearing your name) and work on the lab locally.
- As you are working on your lab, you are to commit and push regularly. You can use the following commands to add a single file, you must be in the directory where the file is located (or add the path to the file in the command):
  - `git commit <nameOfFile> -m ''Your notes about commit here''`
  - `git push`

  Alternatively, you can use the following commands to add multiple files from your repository:
  - `git add -A`
  - `git commit -m ''Your notes about commit here''`
  - `git push`

Google Analytics

Asking visitors to a website to complete a survey to gain insights of their experience on a website has become an unusually onerous request, and internauts typically avoid such surveys with an air of annoyance. Developers of websites are still eager to know the users' impressions of their website and so analytical tools have arisen to allow the actions of users to be studied without their interruption on the website.

The Google Analytics (GA) framework provides convenient metrics to be used to study data concerning the Web traffic on a web site. Some of these can be viewed by different types of plots as shown in Figure 2. For a user-defined time-frame, trends of visits and site usage may be tracked, studied and plotted to gain some understanding of how the website is behaving for its visitors. The study of plots such as, histograms, charts, tables and similar graphics, allows one to conveniently track website activities and to make decisions without having to delve into massive amounts of raw numbers from click-based data.

Part 1: Visualization Techniques

In this lab, you are asked to, first, explore some of the visualizing techniques that GA provides. Please answer the following questions-in-blue using **clear and meaningful** language. Your responses should be well thought-out and be about four or five sentences in length and will be placed in the file, `writing/part1.md`.

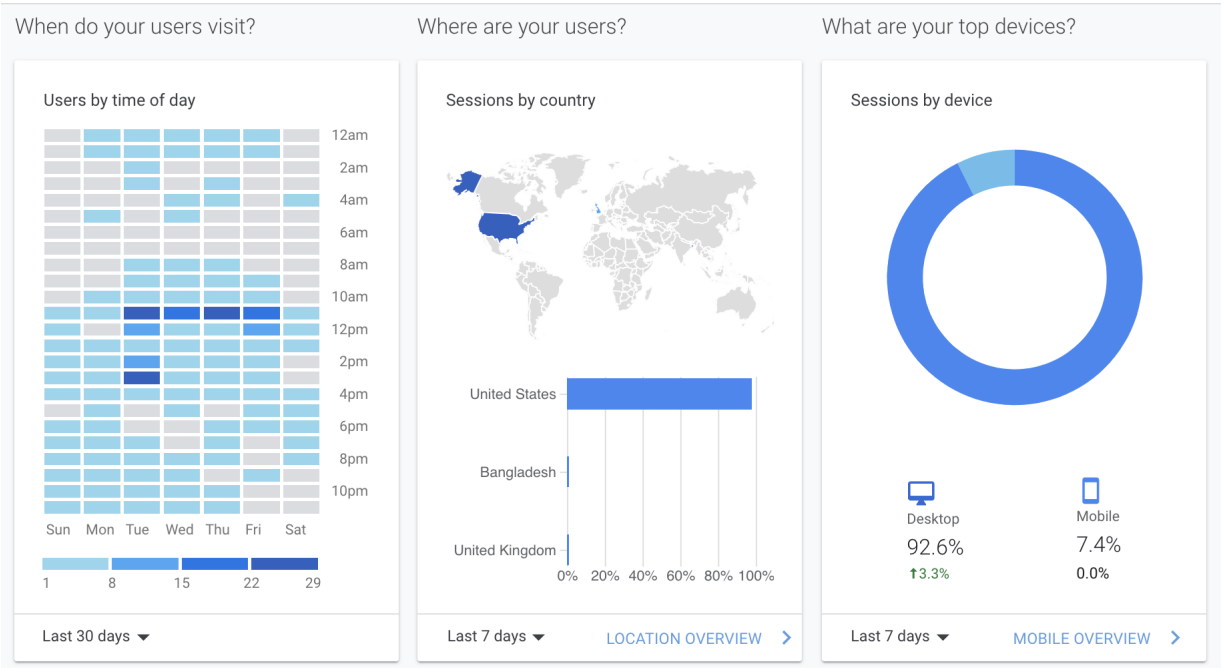

Figure 2: There are a number of different plots available on the dashboard of Google Analytics. One may ask oneself which are plots are more and less important for particular types of website analyses.

1. **Business**: For e-commerce web site such as `https://www.amazon.com`, which **one** (1) metric (plot, statistic or other) would be the most important in making an educated guess about the amount of *potential business activity* reported by the site? Why?

2. **Social media**: For a web site such as `https://www.facebook.com`, which **one** (1) metric (plot, statistic or other) would be the most important in making an educated guess about the amount of *usage* reported by the site? Why?

3. **Campus Calendar**: For an Allegheny College campus web site such as `https://sites.allegheny.edu/registrar/fall-2021-academic-calendar/` that shows important dates for the college, which **one** (1) metric (plot, statistic or other) would be the most important in making an educated guess about the amount of *usage* reported by the site? Why?

4. Reflect on the metrics and plots that you chose for the previous three questions. Assuming that your responses were not the same for each type of web site, discuss how it is that the goals of the website appear to favor particular metrics for monitoring the use of the site.

5. What metric would you suggest is often included in a report but may be interpreted differently depending on the goals of the web site? Argue why this metric has such different meanings for particular sites?

6. **Hypothetical Question**: Imagine that you are an administrator of the website of a large

bank. The site allows account-holders to check balances (checking and savings) and complete other important banking tasks.

You have noticed that no-one has checked their savings balance in spite of the fact that thousands of people have logged into the site during the last few days. You suspect that something is wrong with the website somewhere.

Using at least three different metrics available to you from Google Analytics (monitoring your banking website), describe plots and /or statistical evidence that you could show to the website developers to help them locate the problem in the website. Please justify your selections.

Part 2: Next Week's Lab; Ethics, Discord-Out and Responses

In class, you were to build a sample website to be connected to Google Analytics in order to learn more about the functionality of this analysis suite. In order to complete this part of the lab, your website must be configured correctly with your GA Tracking ID. If you want, you can create a separate website for the purpose of this lab, however, please consider being **creative**, **polite** and **respectful** in your website design and message. Remember, you could always re-purpose your site later to help you beginning your professional /career life. The design of the site is left to you to create.

What to do with your website

Your website **must** be completed with two additional pages where you address the below two questions-in-blue of responsible use. Your responses will be placed in the file, `writing/part2.md`

1. **Do you own one's "Personal Data"?** Data privacy, or the notion that one is able to control how one's data is collected and used, has become hot-topic in our information age. This topic concerns the rights of the individual in light of how one's data is collected and then reapplied toward some (often) unknown purpose.

   For example, in the business model of `https://www.youtube.com`, a user owns a channel in which s/he has left video content for other users to watch. As others click on the videos, they leave meta-data information about themselves, their interests, their locations and similar. YouTube is able to collect, process and apply this information to generate a profit in some way. In your opinion who do you suppose owns (should own) this data concerning these clicks?

2. **Victims of Breach?** Imagine using a website for a long time and having little idea about what types of information is being collected about you and your actions on the site.

   You have used this site to buy merchandize using your credit card several times. One day, you read an article in popular journalism that discusses a breach of the website's data (i.e., data was stolen from its servers by unknown malicious individuals).

   You quickly check your email to determine whether the website administration has tried to contact you to warn you or provide instructions on what to do about any incurred damages. You find no such communications to you on this topic from the website's administration.

Discuss issues of responsibility and liability for the breach. Ultimately, who (in your opinion) is responsible for any potential damages caused by the breach?

Reporting and Writing From Three Websites

Share your thoughts: On the day of the next lab, you are to place the link to your website in the course' Discord channel to allow your colleagues to visit your site and read your responses to the above two questions-in-blue.

You are to visit three (3) websites from the Discord channel to read the reflection pieces of your colleagues.

1. In your responses document (in your `writing/part2.md` document), you are to comment on one or both of the reflection pieces on each website that you visit.
2. Give the name and link of the three (3) websites you chose. Briefly state the premise of the reflection piece(s) by the author and respond to how you agree or disagree with the premise. **Please be respectful of the work of your colleagues in your critical assessment.**
3. Next, check the website analysis of your own page. Write a brief piece of one or two paragraphs about the traffic to your website and the metrics you used to measure your user's traffic and activity. For instance, which plots were most useful to you in determining your website's traffic? Was there any particular metrics in the analysis suite that you found especially helpful concerning the your traffic analysis?

GatorGrader

We will be using Docker and GatorGrader tools in our labs to run and check programs for correctness. Since you are not writing any programs for this lab, GatorGrader will be used to determine that your submission satisfies the minimum requirements, allowing the instructor to grade the content of your writing for other important details. For example, if you run GatorGrader with this project, then you will see the output of Figure 3. In the figure, you will note any submission errors to address to correctly commit your work.

MacOS and Linux Commands to launch GatorGrader

If you have Docker Desktop installed, you could run the following code from your terminal or command prompt. The first time you run the command, you will see the output shown in Figure 3. The command is to be run from within the root of the GitHub Classroom repository and will function to check the that the objectives of your assignment have been achieved.

If you are using MacOS or Linux, then you would need to ensure that the following (MacOS or Linux) command will work correctly, you must create the cache directory by running the command `mkdir $HOME/.dockagator` and push ENTER. Then, to see if your submission satisfies the minimal requirements, you can run the below command in the terminal

```
bash
```

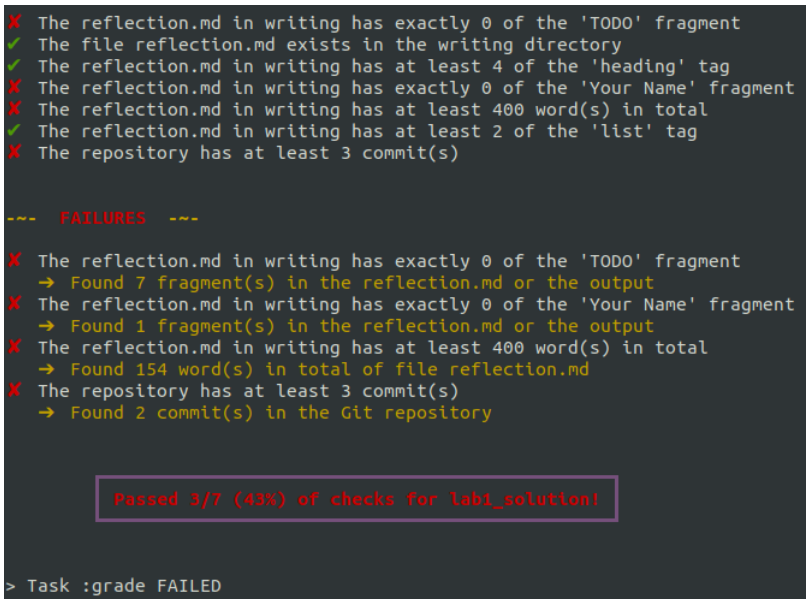

Figure 3: When you initially run GatorGrader, you will see which checks need to be completed. In this image, all checks need attention.

```
docker run --rm --name dockagator \
  -v "$(pwd)":/project \
  -v "$HOME/.dockagator":/root/.local/share \
  gatoreducator/dockagator
```

Please note, you may not need the word `bash` in your command and so if the above command does not run, try using the same command without the `bash` in it. Please see your `README.md` file for more details on this tool.

Windows Commands

If you are running your program on a Windows Operating System, you should run the following command instead, replacing the word `user` with the username of your machine:

```
docker run --rm --name dockagator -v  "%cd%":/project -v
"C:\Users\user/.dockagator":/root/.local/share gatoreducator/dockagator
```

The above line may need to be edited to be on one line when you enter it in your command prompt. Please see your TL's if you have trouble.

Using The Project GitHub Site to Check Submission

If you are having trouble with the above commands with Docker Desktop then you are advised to use the repository webpage to check your submission. Look for the green submission checkmark in

the GitHub website for your repository (right-side, above the listing of the files in your project). A red 'X' indicates a problem and a click will lead you to the submission report.

Summary of Deliverables

Below are the required deliverables for your work.

1. Submit all written work using Markdown. You will be editing supplied files in your GitHub Classroom repository.
2. Edit File `writing/part1.md`: Your responses to the reflection pieces from Part 1, as described above.
3. Complete your website, announce the link in the class Discord channel.
4. Edit File `writing/part2.md`: Your responses to the reflection pieces from Part 2, as described above.

Please ask questions as necessary.