

CMPSC 301 Data Analytics Spring 2020

Lab 1: Google Analytics and Response to Marketing Analytics 27th January 2020



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 platform that provides 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 summarize your understanding of these tools. You are also asked to discuss some of the ethical implications that may result from the collected information of visitors to a website. In the second part of the lab you are asked to read an article at [1] and explore the issues which follow the theme of the speaker about marketing research.

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/X44RZ5Be>

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

The Google Analytics framework provides convenient metrics to be used to understand the Web traffic and its trends over the course of a user-defined time-frame. Histograms, charts, tables and similar graphical methods of displaying information are provided as convenient modules which may be inserted into any report. It is left to you, the analyst, to understand how to understand and apply the information provided from website analytics. To give you some experience, you are required to spend some time to get to know your Google Analytics setup and to answer some [questions in blue](#) concerning the use of this information.

In order to complete this lab, your website has to be correctly linked and configured with the Google Analytics online toolkit. Either you or others are encouraged to click around the website so that some website traffic is generated to show data in your final customized report.

Questions In Blue

After you have spent some time to familiarize yourself with your Analytics output (i.e., plots, histograms and statistics) please locate the Markdown file, `writing/Reflections.md` that contains questions to which you are to respond. You will edit this file using Markdown language and links to useful resources are offered in the `README.txt` file of your GitHub Classroom repository. The questions are provided below for your convenience.

Part 1: Reflection on Google Analytics

1. For a web site such as <https://www.amazon.com>, which one metric would the most important to be able to determine the amount of business performed daily by the site? Why?
2. For a web site such as www.facebook.com, which one metric is the most important to be able to determine how much time is spent on the site. Why?
3. 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?
4. **Hypothetical Question:** Imagine that you are an admin of a banking website that allows people to check their balance and to complete other major banking tasks at your bank. You have noticed that no-one has actually checked their savings balance in spite of the fact that over one hundred people have logged into the site during the last few days. You suspect that something is wrong with the website.

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 where problem(s) in the website are likely to be found.

5. **Ethical Question:** Read the article [1] in the `articles/` directory. Summarize the article in a paragraph.

In a new paragraph, describe the ethical implications of collecting information about patients who consult health-websites. In particular, how this information may be used by third parties which could harm the visitors to the web site.

Part 2: Reflection on the Speaker's Presentation

Based on the talk given by Ron Mattocks on Monday, 27th January, discuss the following:

1. How is data analytics is used in various marketing applications. Explain how tools such as Google Analytics could aid in marketing research. Please give concrete examples to support your arguments.
2. Reflect on the any negative impacts of using data analytics in marketing applications. Provide at least one citation to a primary source to validate your argument.

Important Details

Lab directory structure: Outside of your classDocs/ repository, you are to create a labs directory (mkdir labs in which you are to add the GitHub Classtoom repositories for each of your weekly labs (use this command mkdir labs/labsxx, where xx is the two digit lab number). For example, your first and second lab repository should be located in the paths, labs/lab01 and labs/lab02, respectively.

Add you name to your work: Please remember to include your name on everything you submit for the class.

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 the given starter file, you will see the following output, specifying the missing minimum length and commit requirements, among others.

```
✗ The reflection.md in writing has exactly 0 of the 'TODO' fragment
✓ The file reflection.md exists in the writing directory
✓ The reflection.md in writing has at least 4 of the 'heading' tag
✗ The reflection.md in writing has exactly 0 of the 'Your Name' fragment
✗ The reflection.md in writing has at least 400 word(s) in total
✓ The reflection.md in writing has at least 2 of the 'list' tag
✗ The repository has at least 3 commit(s)

--- FAILURES ---

✗ The reflection.md in writing has exactly 0 of the 'TODO' fragment
  → Found 7 fragment(s) in the reflection.md or the output
✗ The reflection.md in writing has exactly 0 of the 'Your Name' fragment
  → Found 1 fragment(s) in the reflection.md or the output
✗ The reflection.md in writing has at least 400 word(s) in total
  → Found 154 word(s) in total of file reflection.md
✗ The repository has at least 3 commit(s)
  → Found 2 commit(s) in the Git repository

Passed 3/7 (43%) of checks for lab1_solution!

> Task :grade FAILED
```

Figure 2: When you initially run GatorGrader, you will see which checks need to be completed. In this image, all checks need attention.

Command to launch GatorGrader

With Docker Desktop installed, you will 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 2. The command is to be run from within the root of the GitHub Classroom repository and will function to check that the objectives of your assignment have been achieved.

First, to ensure that the following 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 following command in the terminal:

```
bash
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.

Required Deliverables

This portion of the assignment invites you to submit an electronic version of the following deliverable through your GitHub Classroom lab repository. Note: this repository is the one which you clone from the above link.

1. Your reflection document file: `writing/reflection.md` responding to the questions of the two parts discussed above.

When you have finished, run GatorGrader to ensure that your work was correctly completed. Push your work using the above three *git* commands, and then ensure that the GitHub web site has your pushed work by visiting your repository at the repository site (find your work repository site by checking your repositories at www.github.com). Please see the instructor if you have any questions about assignment submission.

References

- [1] T. Libert, “Privacy implications of health information seeking on the web,” *Communications of the ACM*, vol. 58, no. 3, pp. 68–77, 2015.