

**CMPSC 390**  
**Data Analytics**  
**Fall 2017**

**Lab 2: Google Analytics.**

**Save this lab assignment to: [labs/lab2](#)**

Google Analytics is 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. In the second part of the lab you are asked to explore the issues that arise with using tools such as Google Analytics as it relates to privacy of health data.

## Objectives

To learn how to setup and use metric reports in Google Analytics. To become proficient with the tool kit and learn to create reports. The latter part of this lab is to determine how the misuse of data may impact unsuspecting individuals.

## 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.

## Part 1: Google Analytics

In order to complete this lab, your website has to be correctly linked and configured with the Google Analytics online toolkit so that data regarding your website's traffic, can be collected and displayed using the toolkits reports. The reports offered by the Google Analytics framework provide convenient metrics to be used to understand the trends of the data 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. Google Analytics provides default reports where these displays have been added, but users are encouraged to create their own reports where particular graphical displays of metrics have been added.

### A tutorial to create a customized report

In this lab, you will study the default reports given by Google Analytics with the intention of creating your own tutorial to instruct beginners on how to create a customized report to respond to the specific questions given below. Using screenshots and clear and concise language, your tutorial should explain in detail how to exhibit visual displays of analysis. In addition, you are to state how the metrics you incorporate into your report are able to correctly respond to the particular questions below. Finally, you are to explain how to understand the visualizations to answer these below questions concerning a time-window of a month prior to the current date. You may use any text editor to produce your tutorial.

1. How many people have used the web site since last month?
2. Where were these users from (country and city)?
3. Which pages did they read?
4. What was the highest bounce rate?
5. How many visits were there?
6. Number of unique visitors?
7. Number of page views?
8. Average visit duration?
9. Percentage of new visits?
10. Percentage of search traffic?
11. How did this traffic come to the website?

## Questions

After you have created your tutorial, please respond to the following questions in a new report document (this document should be separate from your tutorial document).

1. For a web site such as [www.amazon.com](http://www.amazon.com), which one metric is the most important to determine the business performed daily by the site? Why?
2. For a web site such as [www.facebook.com](http://www.facebook.com), which one metric is the most important for determining the general ease of use of the site. Why?
3. What metric would you suggest is often included in a report but is, actually, not very informative or provides no real value to an analysis? Why?

## Part 2: Health information privacy in using Google Analytics

Read an article titled “Privacy Implications of Health Information Seeking on the Web” by Timothy Libert. In your report document write at least half-page discussion clearly addressing the points below. You must support your responses by including specific examples mentioned in this article or cite other scholarly references to back up your claims.

- Describe privacy issues surrounding health data.
- How do tools such as Google Analytics exacerbate these privacy issues?
- Why is it difficult to protect health information privacy online?
- What steps could be taken to appease the harm caused by health information privacy issues and to avoid future problems?

## Important Details

**Lab directory structure:** Make sure you have placed your submission materials for this lab into `labs/lab2` directory in your Bitbucket repository (`cs390f2017-billb`).

**Note: Please remember to include your name on everything you submit for the class.** Although the instructor collects your work from Bitbucket, each work must be graded outside of the Bitbucket directory and so without adding your name, the instructor will be unable to award you credit for your work.

## Required Deliverables

Submit electronic versions of the following deliverable through your Bitbucket repository (`cs390f2017-billb`) by correctly using appropriate Git commands, such as `git add`, `git commit -m "your message"` and `git push`. When you have finished, please ensure that the Bitbucket Web site has your pushed work. Please see the instructor if you have any questions about assignment submission.

1. Your tutorial document instructing how to setup a customized report that respond to specific questions of web traffic interest.
2. Your responses to the three questions concerning the metrics of other websites.
3. Your report responding to the questions concerning the privacy article.