



**MET CS688**

# ***WEB ANALYTICS AND MINING***

**ZLATKO VASILKOSKI**

WEB ANALYTICS – BASIC TERMS

# Web Analytics Overview

- Web analytics goal:
  - Improving the online experience of your customers and potential customers by data analysis from your business (and the competition).
  - Many examples you might be familiar with (real estate, pharmaceuticals, travel etc.). [Can you suggest few more?](#)
- Analytics program includes aspects of
  - Collecting relevant raw data
  - Understand significance contained in the data
- The focus is to understand the interaction with the customer
  - How the search keywords (and advertisements) influence the search process so that the business can have more visitors (potential customers).
  - Understand (measure) the user experience, behavior and satisfaction with the web site.
- Needed to achieve this
  - Infrastructure to collect and process data (technology)
  - Skills to analyze and interpret that data (qualified people)
- Common misunderstanding of comprehensive web analytics program
  - Focusing on collecting and reporting raw data without understanding their significance
- Skills needed: understanding statistics, mathematics.
  - Recently attempts to create tools to compensate for these skills.
- Fluidic and evolving subject in nature, roles are still being defined
  - Who is responsible for web analytics? Marketing or IT?

# Web Analytics Overview

- Goal - Illustrate how to establish a comprehensive analytics program.
- Good tools are needed for any implementation of analytics.
- Different tools are needed for different types of businesses.

# Example 1

- To better target, acquire, and retain customers, marketers need to use data analytics, content marketing, and customer engagement.
- For example, an office furniture retailer could increase its return on what it spends for advertising by use of data analytics in the following way.
- To acquire new customers, the retailer has to figure out better ways to find potential clients to target for its ads. To do that, however, the company needs huge amounts of **intent**—or **in-market**—data.
- **Intent data** is data collected about online users' activities—indicating some future action, or intent, such as ordering a product. This can be achieved by obtaining web site's key performance indicators using web (or free Google) analytics as described in the first part of this semester class notes.
- When potential customers is looking to furnish a new office may interact with the retailer's web site in a variety of ways, such as browse through architectural sites for design ideas, visit various office retail sites to evaluate items or perhaps even do some comparative analysis on a product review site. All these actions signal that the consumer is actively browsing, researching or comparing the types of products online furniture retailers sell. Customer's intents, preferences, and loyalties create impressions and Web/Google analytics tools can then capture that data enabling marketers to act on it.

# Example 2

- ENERGY STAR, a score from 0 to 100, is a measure of energy consumption performance. A score of 40 means performance worse than 40% of similar buildings nationwide. A score of 75 or higher makes you eligible for ENERGY STAR certification.
- Energy Management Associates, Inc. Help customers achieve ENERGY STAR certification rating by reviewing customer's portfolio, and providing calls with questions or suggestions, to raise customer's score and provide up to 20% savings opportunities.
- According to EMA's web site, customers are encouraged to provide a copy of their most recent utility invoice (12 to 14 month usage history) for each account (gas and electric). This data is contained in EMA's database that with help of **predictive data analytics**, can be used to enhance the detailed customer profile and predict the most likely factors affecting the ENERGY STAR score, by integrating an automatic analysis of the key performance indicators with historical data, by region, type of energy etc. and tied together with external data available on the web, such as location specific temperature and weather conditions.

# Example 3

- The use of **predictive data analytics** enables the retailer, insurance company, the travel agency or any similar eCommerce (click to order) business to predict the most likely purchase type from a given customer.
- This is typically achieved by leveraging historical data (sometimes tied together with a third-party data) to paint a detailed picture of the buyer profile. From that information, it can be determined:
  - Which customer will purchase
  - What product they will purchase
  - What message they will respond to
  - Which customers will focus on
- Predictive data analytics can recognize patterns and behaviors for more effective
  - messaging
  - plan persona-specific campaigns based on customer's habits and past preferences.

# Example 4

- Data analytics is also commonly used to:
  - Give real-time visibility of inventory and warehouse management systems.
  - Sifting through an abundance of social media information. Monitoring, analyzing and reporting on the voice of its customers on social media including Facebook, Twitter, and various blogs and forums.
  - To more accurately analyze the voice of the customer, predictive analytics (and NLP - Natural Language Processing) can be used to score data attributes and determine which social media posts are actionable and relevant thus filtering out the noise of irrelevant posts.

# Brief History of Web Analytics

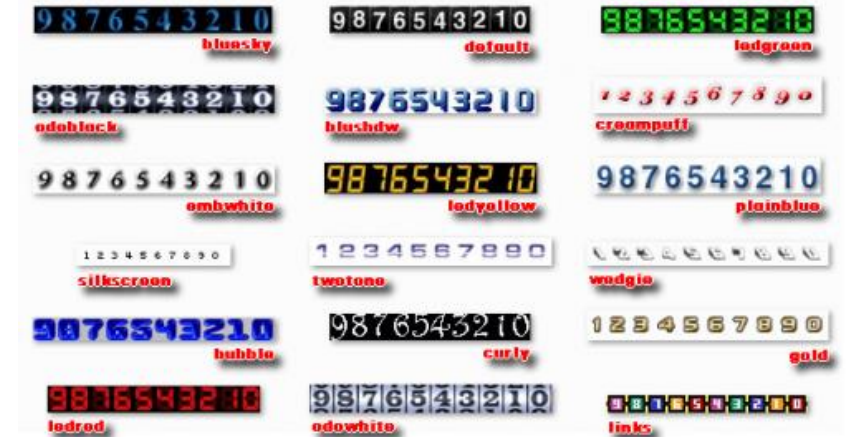
- Starting 1995, the internet community started seeing a **hit counter** on web sites. This plug-in counter communicated web sites' popularity, so almost everyone wanted it. However, such counters are not accurate. As the web-design industry matured, hit counters slowly vanished.





# Brief History of Web Analytics

- Every time someone locates a website the web server logs data.
- Before 1995, simple reports based on information that is automatically collected
  - filename
  - time, referrer (i.e., the website forwarding the request)
  - browser
  - operating system
  - computer data
- Later hit counter introduced on websites, but not accurate so they vanished.
- WebTrends usually associated with first commercial web analytics programs.
  - Commercial web-analytics programs appeared later, with a company called *Webtrends* leading the way. The *Webtrends* software package produced visualizations that appealed to many users.
  - Years later, however, the web site–analytics industry introduced software that was able to measure **click density**, or **site overlay**, and **heatmaps**.

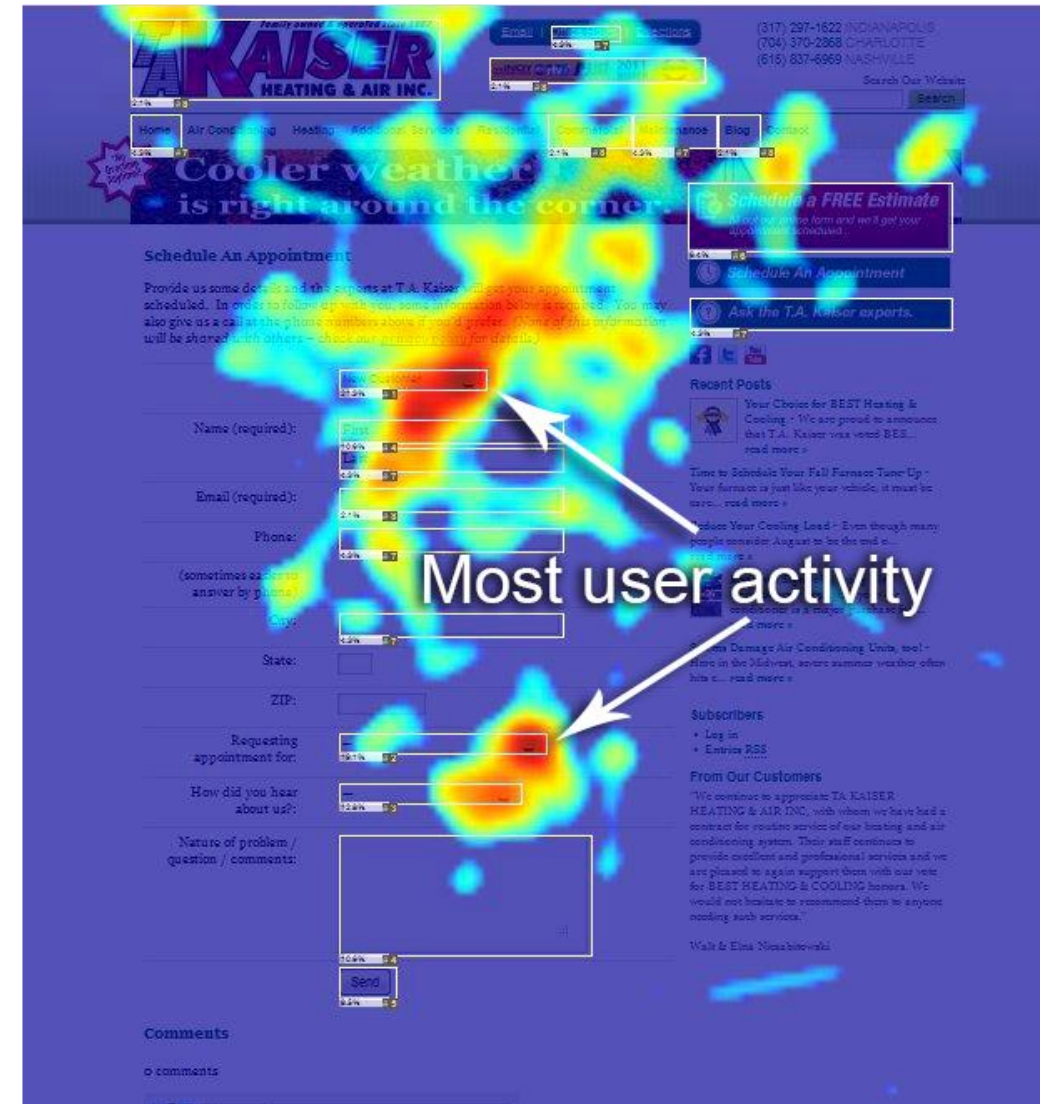


# Heat Map instead of publicly available Hit Counters

source: [http://www.shawnkanciruk.com/web\\_analytics\\_metrics/web-analytics.htm](http://www.shawnkanciruk.com/web_analytics_metrics/web-analytics.htm)



source: <http://zachhellermarketing.com/blog/2015/2/4/testing-with-heat-maps>



# Brief History of Web Analytics

- Several years later, introduced website analytics software that was able to measure click density or site overlay and heat maps.
  - Enabled to understand exactly, which links and where on a page, visitors were clicking.
  - Reports could described the consumer behavior by reporting number of clicks or overall percentage for the identified web pages.
- Large amount of data collected by vendors from every web site visit and click – **clickstream**.
  - It describes what data is collected.
  - It does not contain any insight into the significance of the data.
  - Too large to be useful.
- Big web analytics vendors:
  - Google, WebTrends, Coremetrics, Omniture, WebSideStory, IBM and Adobe.
  - Please have a look at SiteCatalyst a tool from Omniture (Adobe marketing cloud) to see the features and functionalities that can be found in modern tools.
- Ever changing and growing industry, gained much more popularity recently with Big Data.

# Addressing the “What” question

Important web analytics questions to address - “What happened?”

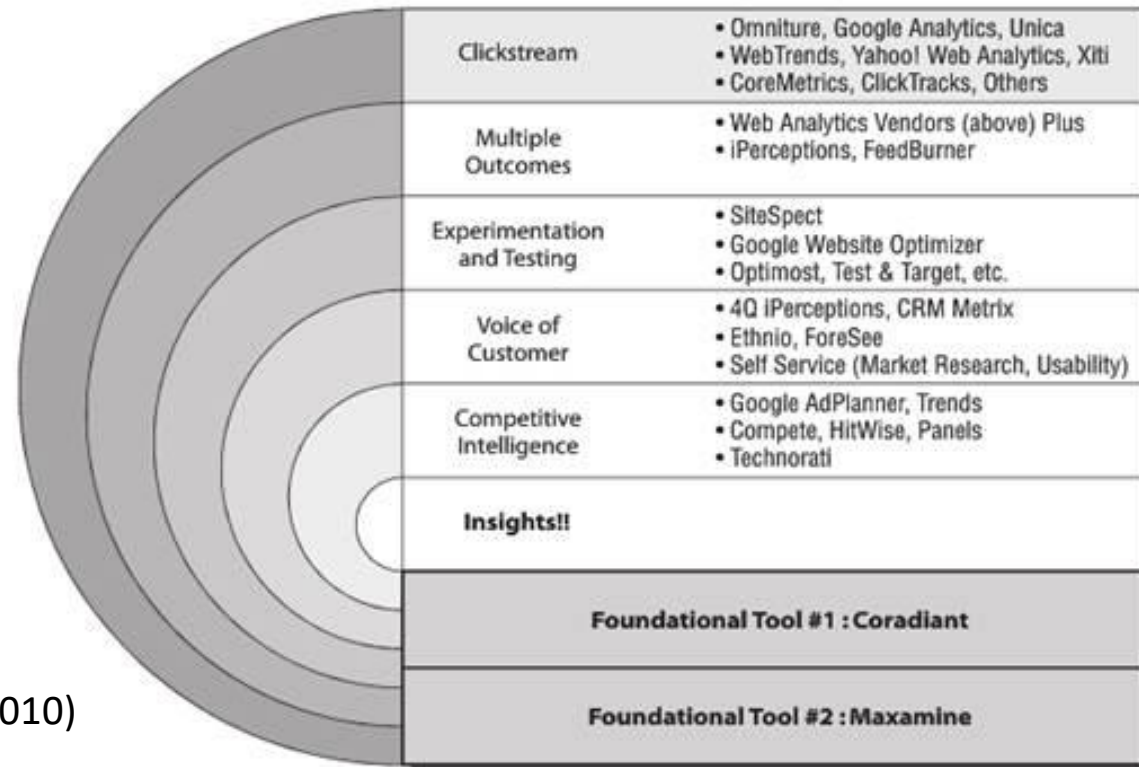
- What pages did people view on your website?
- What products did people purchase?
- What was the average time spent?
- What sources did they come from?
- What keywords and campaigns produced the clicks?

# Addressing the “Why” question

- Even more critical is to know “why” people do the things they do on your web site
  - This can be used to make intelligent decisions about your web presence.
- Important “why” questions to address
  - Analyze qualitative and quantitative data from our web site and a competitor’s web site
  - Focus on continual improvement of the customer’s online experience
  - Better translate data into desired outcomes, both on- and off- line
  - Decide whether we should go with channels like TV and radio advertising, rather than online advertising
  - Obtain a strategic advantage over competitors who focus only on clickstream
- Examples:
  - dropped shopping carts or
  - dropped registrations forms

# Web Analytics Tools

- Which web analytics tools to use?
- Depends on the size of your business and the available resources
  - Small Business – Click stream, outcomes, voice of the customer
  - Medium Business – All of the above plus Testing
  - Large Business – All of the above plus Competitive intelligence, Deep back-end analysis, Site structure gaps



From “Strategy and Tools”  
Webanalytics 2.0 by Kaushik (Wiley 2010)

# Goals, Context,

- It is Important to know
  - Goals of Analysis
    - Target that lets us measure success
    - Assumes understanding both the “what and the why”
    - Examples:
      - Knowing where visitors come from.
      - Understanding what they did in the website.
      - Did they meet the goal?
      - Did the consumers respond to the online marketing plan in the manner you expected?
      - Does the data analytics tool validate that?
  - Context Study
    - Website visitors are made up of groups of people that behave differently and have different objectives.
    - Their action or behavior can be analyzed with the web analytics tools.

# and Segments

- Segment (determine as much as possible features for classification)
  - Analyze particular groups (segments) of people that come to the website.
  - What this particular group has in common with other groups.
  - Relate visitors by geographical regions (east/west coast, Europe, Asia etc.)
  - Relate visitors by gender and their preferences.
  - Segment data by visitor's motivation.
  - Which group of people “browsed, shopped and purchased”.
  - Was this the group that you sent out an email marketing outreach newsletter?
  - If so, do we need to do more outreach via e-newsletters.



# Defining Basic Analytics Metrics

- In analytics it is essential to know which numbers are important and why.
- Note that the metric focuses on “Why”.
- Basic Analytics Metrics
  - Visits and Visitor Sessions
  - Referrals
  - Bounce & Exit Rate
  - Conversion Rate
  - Engagement
  - SEO, Social Media, Emails and Metrics

# Defining Basic Terminology

- **Visitor** - an individual (not necessarily a human) or device such as browser which accesses a Web site within a specific time period.
  - Unique visitor within a specific reporting period (no double counting).
- **Visit (Sessions)** – an interaction with a data source (example: text and/or graphics downloads) from a single browser (device) during a single session.
  - A visit can consist of a series of page views that a single visitor makes during a period of browsing activity. A visit ends after the visitor closes the browser, clears cookies, or is inactive for 30 minutes (customizable time period).
  - During each visit, users will engage in one or more interactions with the web site pages.
  - Analytics software will automatically track these interactions as “**pageviews**.” The pageview metric increases every time a page is viewed on your site. Other activity, like watching a video, mouse position, etc. can also be tracked. Such activities are better classified as “**events**” rather than pageviews.
  - **Cookie** (persistent or session) is a file on the user’s device that identifies the user’s unique browser.
  - Tracking code looks for cookies. If a cookie is deleted or blocked incorrectly counts unique visitors.

# Referrals - Where do visitors come from?

- **Referrals** indicate the place from which the user clicked to get to the current page.
- It is valuable to know how someone found your web site. Was it
  - through a search engine
  - positive review
  - social-media talk
  - email or e-newsletters
- Referrals are the lifeline for marketing advertisements.
- It's important to know which campaigns helped draw in new visitors or succeeded in getting loyal customers.

# Bounce and Exit Rate

- **Bounce rate** and the **exit rate** measure whether users find a web site or a web page useful.
- Bounces are counted for users who land on a page and leave immediately. They do not see the page content.
- Reasons can be site-design or usability issues or many other reasons.
- Typically expressed as percentage of single-page sessions.
- If the exit rate is high, the exit-rate metric can be meaningless, and it should not matter.
- Useful to find out if visitors are abandoning the site at a certain point in the middle of an e-commerce transaction.

# Bounce and Exit Rate

- **Bounce rate** and the **exit rate** measure whether users find a web site or a web page useful.
  - Bounce Rate: the percentage of single-engagement sessions. Determined by calculating the number of bounces over the total number of pageviews that started from that page.
  - Exit Rate: the percentage of exits on a page
    - **Monday:** Page A > Page B > Page C
    - **Tuesday:** Page B > Page A > Page C
    - **Wednesday:** Page A > exit
  - Report for Page A would show 3 pageviews and a 50% bounce rate since Tuesday's pageview did not start with Page A.

## Bounce Rate

$$R_b = \frac{T_v}{T_e}$$

$R_b$  = Bounce Rate  
 $T_v$  = Total one-page visits  
 $T_e$  = Total entrance visits

## Exit Rate

$$R_e = \frac{T_{\text{exits}}}{T_{\text{views}}}$$

$R_e$  = Exit Rate  
 $T_{\text{exits}}$  = Total exits from page  
 $T_{\text{views}}$  = Total visits to page

- Reference: <https://support.google.com/analytics/answer/2525491?hl=en>

# Bounce and Exit Rate

What are the Bounce and the exit rate considering the following single-session days

- **Monday:** Page B > Page A > Page C > Exit
- **Tuesday:** Page B > Exit
- **Wednesday:** Page A > Page C > Page B > Exit
- **Thursday:** Page C > Exit
- **Friday:** Page B > Page C > Page A > Exit

## Bounce Rate

$$R_b = \frac{T_v}{T_e}$$

$R_b$  = Bounce Rate  
 $T_v$  = Total one-page visits  
 $T_e$  = Total entrance visits

## Exit Rate

$$R_e = \frac{T_{\text{exits}}}{T_{\text{views}}}$$

$R_e$  = Exit Rate  
 $T_{\text{exits}}$  = Total exits from page  
 $T_{\text{views}}$  = Total visits to page

# Bounce and Exit Rate

What are the Bounce and the exit rate considering the following single-session days

- **Monday:** Page B > Page A > Page C > Exit
- **Tuesday:** Page B > Exit
- **Wednesday:** Page A > Page C > Page B > Exit
- **Thursday:** Page C > Exit
- **Friday:** Page B > Page C > Page A > Exit

*Bounce Rate:*

- Page A: 0% (one session began with Page A, but that was not a single-page session, so it has no *Bounce Rate*)
- Page B: 33% (*Bounce Rate* is less than *Exit Rate*, because 3 sessions started with Page B, with one leading to a bounce)
- Page C: 100% (one session started with Page C, and it lead to a bounce)

*Exit Rate:*

- Page A: 33% (3 sessions included Page A, 1 session exited from Page A)
- Page B: 50% (4 sessions included Page B, 2 sessions exited from Page B)
- Page C: 50% (4 sessions included Page C, 2 sessions exited from Page C)

## Bounce Rate

$$R_b = \frac{T_v}{T_e}$$

$R_b$  = Bounce Rate  
 $T_v$  = Total one-page visits  
 $T_e$  = Total entrance visits

## Exit Rate

$$R_e = \frac{T_{\text{exits}}}{T_{\text{views}}}$$

$R_e$  = Exit Rate  
 $T_{\text{exits}}$  = Total exits from page  
 $T_{\text{views}}$  = Total visits to page

# Conversion Rate

- In the context of studying goals and outcomes, this metrics is a significant one.
- The **conversion rate** (as a percentage), is defined as
$$\text{conversion rate} = \text{outcome} / \text{unique visits} * 100$$
- An example of an **outcome** could be something very simple like clicking on ads or coupons or subscribing to a newsletter.
- Should we use Unique Visitors (browsers) or Visits?
- Common for a unique visitor to visit the same page many times (purchase).
  - Can you think of an example?
- In this context conversion rate measures the process of converting a visitor into a buyer.
- Conversion rate can be calculated automatically by integrating analytics software with shopping carts.



# Engagement

- This is a qualitative metric hard to measure. The definition can be fuzzy.
- Most analytics software will track event and visit duration.
- This does not provide any information about the quality of engagement during that visit.
- What it matters is the time a visitor spends on a web site **with** engagement.
- The challenge is to distinguish between
  - Reading the information on a web page or
  - Looking for the information and not being able to find it.
- Example: Google Analytics tools allow us to research engagement
  - In-Page Analytics (visual assessment of how users interact with your web pages)
  - Behavior Flow analytics (visualizes the path users traveled from one page or event to the next).
- Both of these contribute to Engagement statistics.

# Review of Used Terms

Terms used:

- **Page** - a web document (HTML, text, video files etc.), excludes image files (hits).
- **PageView** - a page loaded by a browser when a visitor types the URL. GA logs a pageview each time the tracking code is executed.
- **Pages/visit** - Number of web pages viewed in a particular visit.
- **Bounce** - Single page viewed by the visitor who leaves promptly.
- **Average Time on Site** – Duration of site visit.
- **New Visits** – page accessed for the first time by a web browser (GA identification by the web browser. By setting a cookie).

# Keywords Used

- **Loyalty** (visitor behavior): Loyal visitors are frequently highly engaged with your brand, and a high number of visits indicates good customer/visitor retention. A high number of new visitors indicates strong visitor recruitment.
- **Recency** (visitor behavior): The frequency with which visitors return to your site can indicate their level of engagement with your brand and their readiness to buy.
- **Length of Visit** (visitor behavior): Length of visit is a measure of visit quality. A large number of lengthy visits suggests that visitors have interacted more extensively with your site. The graph allows you to visualize the entire distribution of visits instead of simply the average time on site taken from all visits. Keep in mind that the average time on site is skewed by visitors leaving browser windows open when they are not actually viewing or using your site. You can see whether a few visits are skewing the average time on site upward or whether most visits to your site have a long average time.
- **Depth of Visit** (visitor behavior): Depth of visit is a measure of visit quality. A large average number of pageviews per visit suggests that visitors interact extensively with your site. The graph allows you to visualize the entire distribution of visits instead of simply the average pageviews per visit. You can see whether a few visits are skewing your average upward or whether most visits to your site result in a high number of pages being viewed.
- **Browsers** (browser capabilities): Which browsers do your visitors use? Optimizing your site with the appropriate technical capabilities helps make your site more engaging and usable and can result in higher conversion rates and more sales.

# Web Analytics

- Collecting relevant raw web data .
- Understand significance contained in the data.
- Analysis is used for understanding and optimizing the web usage.
- Also measuring web traffic can be used as a tool for business and market research.
- Two categories of web analytics
  - **Off-site** web analytics measures on data gathered from sites other than your own (Internet as a whole) and measures potential website audience, including social media.
    - Used to understand how to market your site by identifying the keywords tagged to your site, either from social media or from other websites.
  - **On-site** web analytics measure the actual visitor traffic arriving on your website such as
    - Performance of your website in a commercial context
    - Web page associated with online purchases
    - Audience response to your marketing campaign

# Log Files vs. Page Tagging

- Two unique **on-site** methodologies used by analytics tools to collect web visitor-analytics data.

1. Log files approach
2. Page tags approach

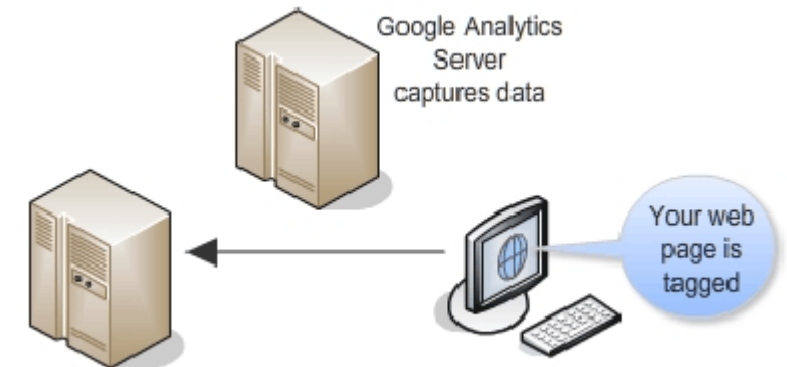
## 1. Log files approach (server-side data-collection methodology)

- All requests for web pages, images, PDFs made to your web server are captured.
- Data is collected by your web server and is independent of a visitor's browser.
- Advantages of log file approach is that it is not affected by firewalls or software that can block page tags. Also, it can differentiate between a complete download and a partial download.



## 2. Page tags approach

- Motivation to perform web analytics as an outsourced service + corrects the accuracy of log file.
- Involves the tagging of the web pages with special JavaScript.
- A remote server captures all the tagged pages and will report the analytics to you.
- Google Analytics uses this approach to collect web visitor-analytics data.



# Web Analytic Data Sources

Server log files

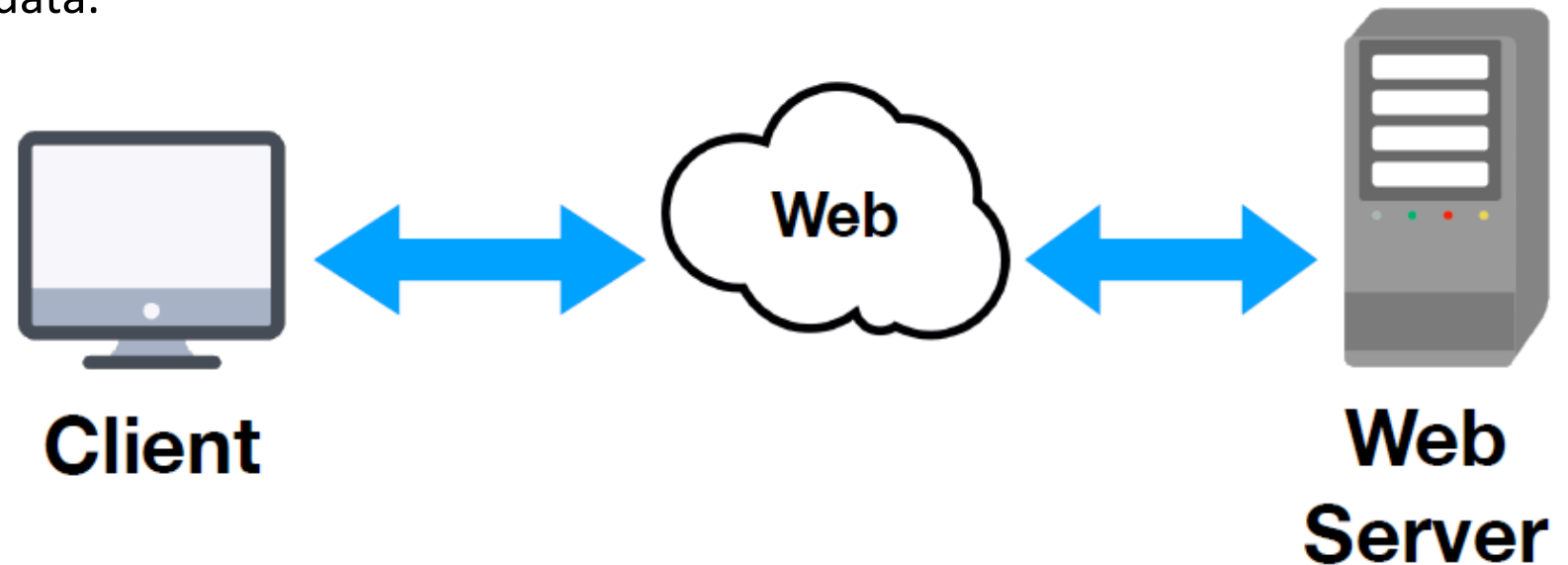
- Standard, easy to access
- Include search engine access info.

Page Tagging

- Cookies that can track mouse events, movie plays,....
- Inexpensive access to visitor data.

Click analysis

Visitor Geolocation analysis



# Page tags approach - Sample HTML page

source: <https://www.dummies.com/web-design-development/site-development/a-sample-web-page-in-html/>

```
<html>
<!-- Text between angle brackets is an HTML tag and is not displayed.
Most tags, such as the HTML and /HTML tags that surround the contents of
a page, come in pairs; some tags, like HR, for a horizontal rule, stand
alone. Comments, such as the text you're reading, are not displayed when
the Web page is shown. The information between the HEAD and /HEAD tags is
not displayed. The information between the BODY and /BODY tags is displayed.-->
<head>
  <title>Enter a title, displayed at the top of the window.</title>
</head>
<!-- The information between the BODY and /BODY tags is displayed.-->
<body>
  <h1>Enter the main heading, usually the same as the title.</h1>
  <p>Be <b>bold</b> in stating your key points. Put them in a list: </p>
  <ul>
    <li>The first item in your list</li>
    <li>The second item; <i>italicize</i> key words</li>
  </ul>
  <p>Improve your image by including an image. </p>
  <p></p>
  <p>Add a link to your favorite <a href="https://www.dummies.com/">Web site</a>.
  Break up your page with a horizontal rule or two. </p>
  <hr>
  <p>Finally, link to <a href="page2.html">another page</a> in your own Web site.</p>
  <!-- And add a copyright notice.-->
  <p>&#169; Wiley Publishing, 2011</p>
</body>
</html>
```



# Log Files Approach - Sample Server Log

source: <http://www.satsig.net/logfile.htm>

65.26.149.185 - - [04/Nov/2002:01:51:53 +0000] "GET /ivsats.htm HTTP/1.1" 200 9430 "http://search.dogpile.com/taxis/search?q=Satellite+Internet+Access+Dish&format=clone&brand=dogpile&attrib=rs" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)"

65.26.149.185 - - [04/Nov/2002:01:51:53 +0000] "GET /901-342s.jpg HTTP/1.1" 200 8600 "http://www.satsig.net/ivsats.htm" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)" 65.26.149.185 - - [04/Nov/2002:01:51:54 +0000] "GET /pas1rkuh.gif HTTP/1.1" 200 4189 "http://www.satsig.net/ivsats.htm" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)"

65.26.149.185 - - [04/Nov/2002:01:51:54 +0000] "GET /nss7kwas.jpg HTTP/1.1" 200 6271 "http://www.satsig.net/ivsats.htm" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)" 65.26.149.185 - - [04/Nov/2002:01:51:54 +0000] "GET /asiak2.gif HTTP/1.1" 200 6560 "http://www.satsig.net/ivsats.htm" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)"

65.26.149.185 - - [04/Nov/2002:01:51:54 +0000] "GET /ab2\_eu3.gif HTTP/1.1" 200 6635 "http://www.satsig.net/ivsats.htm" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)" 66.32.2.122 - - [04/Nov/2002:01:52:55 +0000] "GET /ssazelm.htm HTTP/1.1" 304 - "http://www.google.com/search?hl=en&lr=&ie=UTF-8&oe=UTF-8&as\_qdr=all&q=satellite+signal+meter+aim" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.0; Q312461)"

66.32.2.122 - - [04/Nov/2002:01:52:55 +0000] "GET /sf-95-3.gif HTTP/1.1" 304 - "http://www.satsig.net/ssazelm.htm" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.0; Q312461)" 24.43.169.115 - - [04/Nov/2002:01:53:02 +0000] "GET /ssazelm.htm HTTP/1.1" 200 11623 "http://www.google.ca/search?q=%22Free+to+Air%22%2Bsatellite+dish&hl=en&lr=&ie=UTF-8&oe=UTF-8&start=10&sa=N" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; .NET CLR 1.0.3705)"

24.43.169.115 - - [04/Nov/2002:01:53:03 +0000] "GET /sf-95-3.gif HTTP/1.1" 200 3536 "http://www.satsig.net/ssazelm.htm" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; .NET CLR 1.0.3705)" 64.130.130.17 - - [04/Nov/2002:01:55:13 +0000] "GET /ssazelm.htm HTTP/1.0" 200 11857 "-" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.0)" 64.130.130.17 - - [04/Nov/2002:01:55:14 +0000] "GET /sf-95-3.gif HTTP/1.0" 200 3536 "-" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.0)"



# Web Page Analytics Tools

- Google Analytics
- Adobe Analytics (Omniure SiteCatalyst)
- You can read more about Google Analytics/Adobe
- Analytics features and comparison: [https://](https://www.educba.com/google-analytics-vs-adobe-analytics/)
- [www.educba.com/google-analytics-vs-adobe-analytics/](https://www.educba.com/google-analytics-vs-adobe-analytics/)
- Your custom-made tool

# Google Analytics

- Standard (and free) tool for generating web analytics reports.
- It uses a **page-tag** approach for capturing analytics data,
  - Special JavaScript needs to be installed (tagged) in each web pages.
- It provides the means to track traffic, analyze it, and report useful metrics, such as site visits, marketing goals achieved, ad revenues generated etc.
- It is used in organizations of all sizes.
- You can install Google Analytics on a personal website and begin using it
  - to track visitors
  - create goals
  - measure website success

# Introduction to Google Analytics

- Google Analytics is a page-tag approach to capturing analytics data.
- Free tool that has the following functional uses:
  - Identify visitors or groups of visitors that became customers.
  - See which pages drive the most pageviews on your site. The Top Content report can answer questions you have about your most or least effective pages.
  - Review data for online advertising campaigns by tracking landing-page quality and conversions (goals) and pick out your best-performing ads.
  - Grant other people in your company or organization designing or viewing access to reports. You can also authorize an entire account, with administrator privileges, for anyone you want.
  - See live analytics as customers are interacting with your web site.
  - Study the location your visitors are coming from and what devices they are using (e.g., smartphones or desktops).
  - Leverage more than 80 reports with customizable templates.

# Introduction to Google Analytics

- Examples of Google Analytics metrics:
  - Daily visitors to your site
  - Average conversion rate
  - Top visited pages
  - Average visit time on the web site
  - How often visitors come back
  - Average visit's page depth
  - Geographic distribution of visitors
  - Bounce rate

# How Does Google Analytics Work?

- Data about visitors is collected from desktop computers and devices and it is stored in large Google databases.
- When we log into Google Analytics, we open up such a reporting database and review the data collected by the analytics.
- This process does not involve the active participation of the web server that hosts the web site.
- Depending on how much traffic you have, there will inevitably be some variance in how long it takes from the time a visitor views a web page on your company server to when Google reports the analytics data.

# How Does Google Analytics Work?

- According to Google, there are four main components to the Google Analytics system:
  - data collection,
  - configuration,
  - data processing, and
  - reporting.



# How Does Google Analytics Work?

## Collection

- You can collect data from any digitally connected environment, including a kiosk or a point-of-sale (POS) system.
- To track a web site, Google Analytics uses a small piece of JavaScript code, which tracks user engagement.
- You must place this piece of code on every page of the web site.
- When a user arrives at your web site, this JavaScript code will collect various pieces of information about how the user engages with your site. This process will be introduced in the next section.

## Processing

- JavaScript collects information about the web site, the browser or device, and the referring source and sends it to the Google servers for processing.
- This is the “transformation” step in which the raw data becomes useful information.
- For example, during data processing, Google Analytics will categorize users' devices as mobile or non mobile.

# How Does Google Analytics Work?

## Configuration

- In this step, Google Analytics applies any configuration settings you have identified to the raw data.
- For example, you may have used a filter to exclude data from your own internal users, as such data would produce inflated results.
- Once the data is processed, it is inserted into Google's analytics database. Such recording of data is permanent and can't be changed

## Reporting

- The last component of the Google Analytics platform is reporting.
- A simple-to-use web interface found at [www.google.com/analytics](http://www.google.com/analytics) makes it possible to retrieve data from your organization's Google Analytics account.



# Defining Basic Analytics Metrics

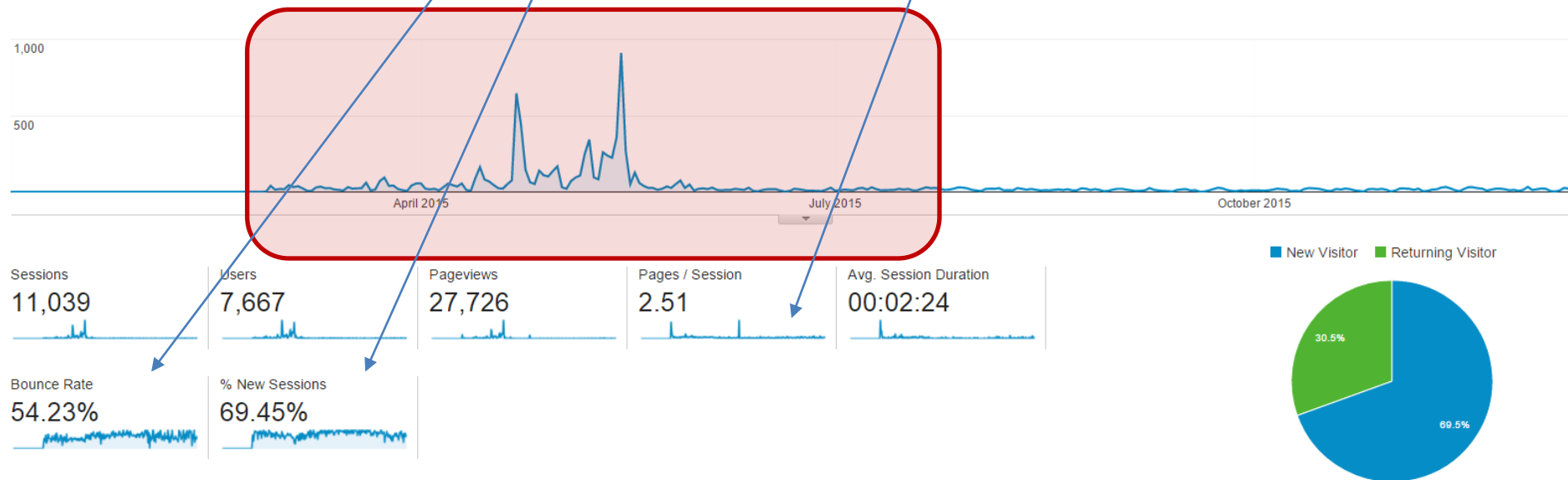
- In analytics it is essential to know which numbers are important and why.
- Note that the metric focuses on “Why”.
- Basic Analytics Metrics
  1. Visits and Visitor Sessions
  2. Referrals
  3. Bounce & Exit Rate
  4. Conversion Rate
  5. Engagement
  6. SEO, Social Media, Emails and Metrics

# Google Analytics - Case Study

- Data from a conference hosted by BU
- Promotes an annual conference for professionals practicing project management principles
- Event was held in mid-May 2015.

# Audience

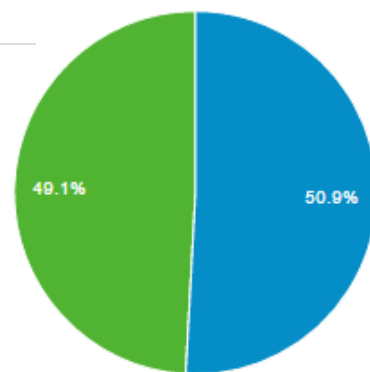
- Most traffic between April 1 and May 31
- Very low bounce rate of 54%
- Two-thirds of visitors are new
- The average visitors stays quite a long time (2.5 minutes)



# Audience

- Most traffic from U.S. (68%) followed by Russia (6%) and Canada (3%)
- Nearly even distribution of men and women
- Most traffic from desktop devices (82%) followed by mobile (14%)

Country	Sessions	% Sessions
1.  United States	7,557	68.46%
2.  Russia	710	6.43%
3.  Canada	319	2.89%
4.  Brazil	203	1.84%
5. (not set)	166	1.50%
6.  India	148	1.34%
7.  United Kingdom	117	1.06%
8.  Saudi Arabia	76	0.69%
9.  China	72	0.65%
10.  Peru	72	0.65%



Device Category	Acquisition		
	Sessions ? ↓	% New Sessions ?	New Users ?
	11,039 % of Total: 100.00% (11,039)	69.48% Avg for View: 69.45% (0.04%)	7,670 % of Total: 100.04% (7,667)
1. desktop	8,885 (80.49%)	71.20%	6,326 (82.48%)
2. mobile	1,771 (16.04%)	61.77%	1,094 (14.26%)
3. tablet	383 (3.47%)	65.27%	250 (3.26%)

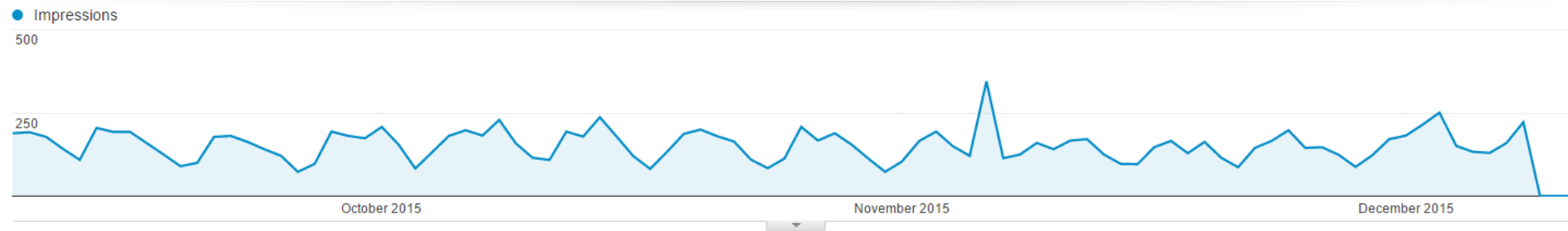
# Acquisition (top 10) – how users found the site

- Most new user traffic is “direct”; however, direct traffic also represents email marketing in this example
- Google and Bing organic traffic are also major sources
- Inbound links (back to your site from another Web site) from 4 websites also provide major sources of traffic

Source / Medium ?	Sessions ?	% New Sessions ?	New Users ? ↓	Bounce Rate ?	Pages / Session ?	Avg. Session Duration ?
	<b>10,500</b> % of Total: 95.12% (11,039)	<b>67.96%</b> Avg for View: 69.45% (-2.15%)	<b>7,136</b> % of Total: 93.07% (7,667)	<b>52.05%</b> Avg for View: 54.23% (-4.02%)	<b>2.59</b> Avg for View: 2.51 (3.02%)	<b>00:02:30</b> Avg for View: 00:02:24 (4.57%)
1. (direct) / (none)	3,845 (36.62%)	69.83%	<b>2,685</b> (37.63%)	53.91%	2.45	00:02:37
2. google / organic	1,509 (14.37%)	69.78%	<b>1,053</b> (14.76%)	52.22%	2.53	00:02:04
3. girlsguidetopm.com / referral	616 (5.87%)	86.36%	<b>532</b> (7.46%)	38.31%	2.85	00:02:02
4. pmi.org / referral	408 (3.89%)	85.29%	<b>348</b> (4.88%)	51.72%	2.32	00:02:25
5. pm4girls.elizabeth-harrin.com / referral	409 (3.90%)	75.06%	<b>307</b> (4.30%)	34.96%	3.21	00:02:24
6. MET Current Students / email	401 (3.82%)	62.84%	<b>252</b> (3.53%)	56.86%	2.36	00:02:21
7. MET Prospective Students / email	209 (1.99%)	64.59%	<b>135</b> (1.89%)	62.68%	1.96	00:01:58
8. bing / organic	174 (1.66%)	74.71%	<b>130</b> (1.82%)	41.95%	3.08	00:01:54
9. projectmanagement.com / referral	167 (1.59%)	76.05%	<b>127</b> (1.78%)	43.11%	2.99	00:02:52
10. MET / banner	185 (1.76%)	52.43%	<b>97</b> (1.36%)	34.59%	4.95	00:06:02

# Acquisition – how users found the site

- Organic search traffic mainly from “project management conferences” keyword



Primary Dimension: Query Other ▾

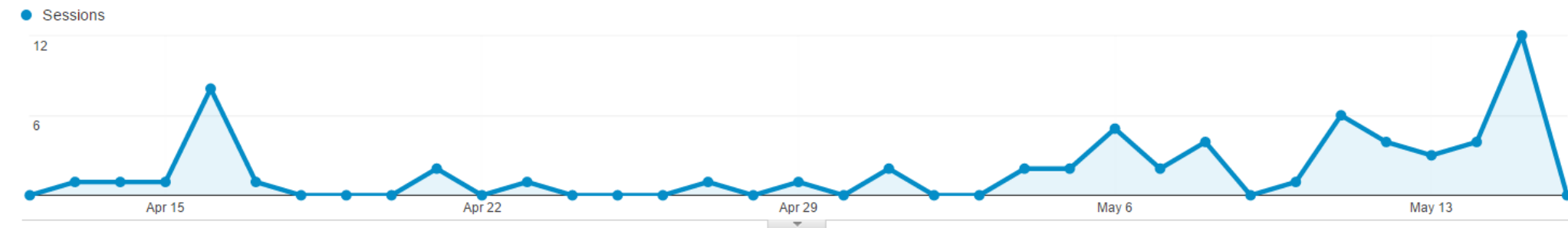
Secondary dimension ▾

advanced

Query	Impressions ?	Clicks ?	Average Position ?	CTR ?
	14,029 % of Total: 100.00% (14,029)	699 % of Total: 100.00% (699)	21 % of Total: 100.00% (21)	4.98% Avg for View: 4.98% (0.00%)
1. (not set)	6,954 (49.57%)	426 (60.94%)	22(102.93%)	6.13%
2. project stakeholder management pdf	254 (1.81%)	44 (6.29%)	4.2 (16.16%)	17.32%
3. project management conferences 2016	295 (2.10%)	38 (5.44%)	10 (46.23%)	12.88%
4. project management conference	914 (6.52%)	26 (3.72%)	8.3 (36.39%)	2.84%
5. project management conferences 2015	595 (4.24%)	21 (3.00%)	8.6 (37.98%)	3.53%
6. project management conference 2016	133 (0.95%)	15 (2.15%)	12 (53.83%)	11.28%
7. project management conferences	396 (2.82%)	13 (1.86%)	9.8 (44.26%)	3.28%
8. control stakeholder engagement	187 (1.33%)	12 (1.72%)	9.7 (43.84%)	6.42%

# Acquisition – how users found the site

- Paid search (Google advertising) generated 46 clicks



Primary Dimension: **Keyword** [Ad Content](#)

Plot Rows

Secondary dimension

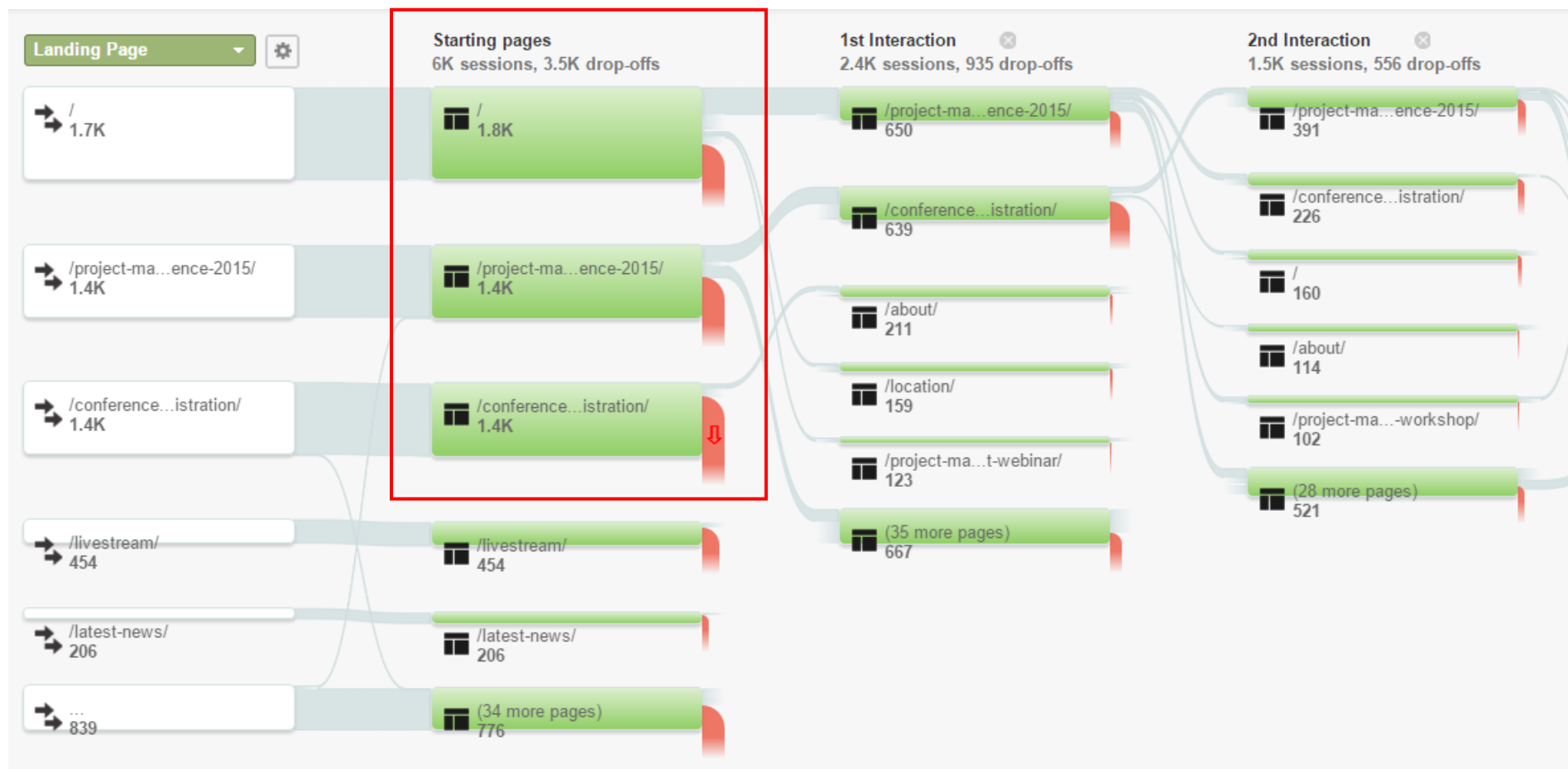
Sort Type: Default

advanced

<div></div> <div>Keyword</div>	Acquisition				Behavior		Conversions <div>All Goals</div>		
	Clicks	Cost	CPC	Sessions	Bounce Rate	Pages / Session	Goal Conversion Rate	Goal Completions	Goal Value
	<div>46</div> <div>% of Total: 100.00% (46)</div>	<div>\$180.93</div> <div>% of Total: 100.00% (\$180.93)</div>	<div>\$3.93</div> <div>Avg for View: \$3.93 (0.00%)</div>	<div>64</div> <div>% of Total: 1.07% (5,999)</div>	<div>37.50%</div> <div>Avg for View: 46.41% (-19.19%)</div>	<div>3.11</div> <div>Avg for View: 2.60 (19.69%)</div>	<div>101.56%</div> <div>Avg for View: 89.20% (13.86%)</div>	<div>65</div> <div>% of Total: 1.21% (5,351)</div>	<div>\$3.00</div> <div>% of Total: 0.72% (\$418.00)</div>
1. +project +management +conference	32 (69.57%)	\$113.40 (62.68%)	\$3.54	46 (71.88%)	34.78%	3.15	97.83%	45 (69.23%)	\$3.00(100.00%)
2. +project +management +webinars	10 (21.74%)	\$48.08 (26.57%)	\$4.81	13 (20.31%)	46.15%	3.23	107.69%	14 (21.54%)	\$0.00 (0.00%)
3. +project +management +pdus	3 (6.52%)	\$15.82 (8.74%)	\$5.27	3 (4.69%)	33.33%	2.67	133.33%	4 (6.15%)	\$0.00 (0.00%)
4. +project +management +workshops	1 (2.17%)	\$3.63 (2.01%)	\$3.63	2 (3.12%)	50.00%	2.00	100.00%	2 (3.08%)	\$0.00 (0.00%)

# Behavior

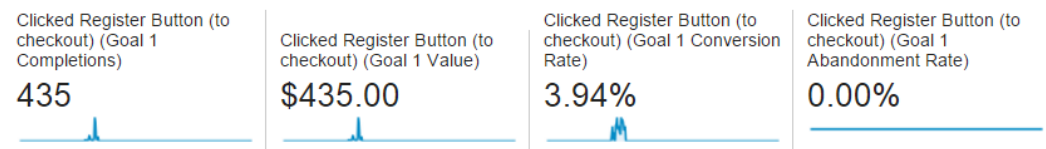
- Nearly equal # of visitors enter the site's main conference and registration pages as the homepage





# Goals / Conversions

- Best sources of conversions are “direct” (70%), email marketing and Google organic search traffic



## Goals

Goal Completion Location

Source / Medium

Source / Medium

Clicked  
Register Button  
(to checkout)  
(Goal 1  
Completions)

% Clicked Register  
Button (to checkout)  
(Goal 1 Completions)

Source / Medium	Clicked Register Button (to checkout) (Goal 1 Completions)	% Clicked Register Button (to checkout) (Goal 1 Completions)
1. (direct) / (none)	305	70.11%
2. google / organic	24	5.52%
3. MET Current Students / email	18	4.14%
4. MET Prospective Students / email	15	3.45%
5. pm4girls.elizabeth-harrin.com / referral	10	2.30%
6. MET / banner	9	2.07%
7. projectmanagement.com / referral	8	1.84%
8. myeventguru.com / referral	6	1.38%
9. Consolidated PMIP Lists since 2012 / email	4	0.92%
10. Email Blast / Email	4	0.92%

A

- A