2.1 Web Analytics Concept

Web analytics is the measurement and analysis of data to inform an understanding of user behavior across web pages.

Web analytics is the process of analyzing the behavior of visitors to a website. Web analytics is the process of collecting, analyzing, and reporting on website data to understand user behavior and optimize performance. The data is collected through **analytics tools** and platforms (for example, Google Analytics) that keep track of what happens on a website : how many people get there, where they came from, where they go after landing, how many make a purchase, how much they spend, and so on.

Web analytics is a term that applies to all forms of online measurement. There are many forms of web analytics that allow you to understand exactly how people engage with your website.

2.2 Web Analytics - Objectives, Importance and Benefits

2.2.1 Objective

The objective of web analytics is to serve as a **business metric** for promoting specific products to the customers who are most likely to buy them and to determine which products a specific customer is most likely to purchase. This can help improve the ratio of revenue to marketing costs.

In addition to these features, web analytics may track the **clickthrough** and **drilldown** behavior of customers within a website, determine the sites from which customers most often arrive, and communicate with browsers to track and analyze online behavior. The results of web analytics are provided in the form of tables, charts and graphs.

2.2.2 Importance

Website analytics provide insights and data that can be used to create a better user experience for website visitors.

Understanding customer behavior is also key to optimizing a website for key conversion metrics.

For example, web analytics will show you the most popular pages on your website, and the most popular paths to purchase.

With website analytics, you can also accurately track the effectiveness of your online marketing campaigns to help inform future efforts.

2.2.3 Benefits

Web analytics allows you to understand three things :

- How people find your website
- How people engage with your website
- What your website's strengths and weaknesses are

This is incredibly valuable information, which allows you to improve your website, as well as improve your digital marketing efforts. Without web analytics, we're just guessing what works and what doesn't.

Web analytics allows us to make informed decisions based on actual data. That's what sets digital marketing apart from other classical marketing activities.

The key benefits of analytics are that it enables you to :

- Get closer to the customer
- Accurately gauge user experience
- Gain insights from real customer experience
- Determine the likelihood that a given customer will repurchase a product after purchasing it in the past.
- Personalize the site to customers who visit it repeatedly.
- Monitor the amount of money individual customers or specific groups of customers spend.
- Observe the geographic regions from which the most and the least customers visit the site and purchase specific products.
- Predict which products customers are most and least likely to buy in the future.

Further, analytics tools reveal that you have a page with lots of traffic but very few conversions. Clearly, there is some work to do; but how do you make sense of this information ?

- Are people leaving because something on the page is broken ?
- Are they leaving because they are looking for something different altogether ?
- Is it because something on the page doesn't fill them with confidence ?
- Or is it because a crucial bit of information is missing ?

These are just a few of the many potential reasons behind your website visitors' actions, but without knowing how they experience a website or why they do what they do, your chances of optimizing and growing the business are limited. Web analytics do help in answering above raised queries.

2.3 Web Analytics Present - Future and Brief History

2.3.1 Brief History

Almost 32 years since its inception, the World Wide Web has come a long way.

1990 - The Beginning of The World Wide Web - The internet is essentially a constant dialogue of HTML code, flowing back and forth between a web user and a web server. 25th December 1990, Sir Tim Berners-Lee successfully implemented the first dialogue of this kind, creating the Internet.

1993 to 2003

Each time a certain HTML element is requested by a visitor, it is called a "hit" and is recorded into a log file. A hit may include text on a web page, an image, sound or a video file. However, at early stage, the Internet was comprised of mostly static pages limited to text and links. Therefore, when a page received a hit by a visitor, it was assumed that they were engaging with the entire contents of the page. As web usage increased, website owners became preoccupied with this metric and measurements. The generation of log file analysis led to the beginning of commercial web analytics, which was marked by the development of WebTrends in 1993.

In 1995, the very first log file analysis program was completely free of charge. Until this time, web analytics could only be understood by tech teams, but Analog made the reports generated through log files more comprehensible to online business owners, with clear documentation and visual graphs.

In 1996, Web-Counter, the first widely used hosted hit counter service was developed. This sparked the trend of odometer-style hit counters being displayed on website landing pages. Accrue, Omniture, and WebSideStory were founded.

Since 1997 to 2003 it was huge research time that was invested in the webpages analysis tools. As webpages began to include visuals and other elements besides text, it became clear that the number of hits a server accumulated no longer represented the number of pages requested. Javascript tagging became the new method of data collection to accurately report on diverse web traffic and trends. This is still the most widespread method of data collection today.

2004 - 2014

Web analytics became more established as an essential tool for web optimization, providing increasingly complex solutions that reported massive amounts of data. The Web Analytics Association (WAA), now known as the Digital Analytics Association, was founded.

Google analytics quickly became the most widely used web analytics service on the market. Focusing heavily on quantitative analysis, it tied in directly with Google's other web marketing offerings.

The launch of in-page analytics allowed website owners to see everything their visitors did on a webpage. Video session playbacks of visitor behavior delivered qualitative usability and conversion-based data, while heatmaps and form analytics provided online businesses with both quantitative statistics and qualitative behavioral data about website visitors.

Various organizations started offer solutions for businesses seeking conversion rate optimization, improved customer experience and increased user engagement through data sourced from analytics tools, such as **mouse tracking.**

Google launched Universal Analytics, meaning that users could be tracked across **multiple devices** and platforms through user IDs. Beyond that, offline behavior began to be monitored, and customer data became richer with the addition of demographic and other information.

Google analytics for mobile apps was also launched, supplying app analytics via SDKs for android and iOS. The world of web analytics grew to include the currently termed standard mobile analytics.

2015 - 2021

Google analytics incorporated machine learning into its app analytics, giving marketers smarter insights on the go. This enabled a streamlined mobile google analytics experience, showing more relevant metrics, with real-time monitoring. This was a turning point in the analytics industry.

By combining cognitive computing, machine learning and psychological research, organization started working on cloud platforms to solve critical digital questions for enterprises and enable them to provide top-quality customer experience. The new platform helps businesses better understand customer intent and behavior, and quantify impacts in order to more efficiently prioritize optimization tasks. This is made possible through the use of intuitive workflows and visualizations.

2.3.2 Web Analytics Future

Web and app design are heavily influenced by analytics. One wouldn't dream of designing own websites or apps without A/B testing and tweaking. Now, as the development of UX and UI is done hand in hand with derived insights from web analytics, this bond becomes ever stronger. The future is about optimizing this

relationship as well as optimizing the insights accrued, in order to deliver maximum satisfaction to both customer and enterprise. Machine learning and cognitive computing are the focus of current web analytics evolution, and analytics providers continue to innovate in the field.

The web analytics market is segmented by application (Online marketing, Mobile analytics, Content marketing, Social media management, E-mail marketing), Offering (Solution, Services), End-user vertical (Retail, Manufacturing, Information Technology, BFSI, Healthcare, Transportation and Logistics) and Geography.

Currently there are below classified groups of people who use web,

- **Ignore :** Don't collect or report on analytics data.
- **Basic :** Use data to measure WHAT is happening.
- **Intermediate :** Use data to measure what is happening + determine WHY.
- **Advanced :** Use data to measure what is happening + determine why + make ONE-OFF data-informed changes.
- **Elite :** Use data to measure what is happening + determine why + make ONGOING data-informed changes.

Looking at such classification definitely large groups users are keen about analysis of their webpages usage. Hence, web analytics is definitely the future of data analysis where one can track customer behavior and other insights by a matter of seconds.

2.4 Main Categories of Web Analytics

2.4.1 Off-Site and On-Site Web Analytics

The two main categories of web analytics are off-site web analytics and on-site web analytics.

Off-site web analytics

The term *off-site web analytics* refers to the practice of monitoring visitor activity outside of an organization's website to measure potential audience. Off-site web analytics provides an industrywide analysis that gives insight into how a business is performing in comparison to competitors. It refers to the type of analytics that focuses on data collected from across the web, such as **social media, search engines** and **forums.**

On-site web analytics

On-site web analytics refers to a narrower focus that uses analytics to track the activity of visitors to a specific site to see how the site is performing. The data gathered is usually

more relevant to a site's owner and can include details on site engagement, such as what content is most popular. Two technological approaches to on-site web analytics include **log file** analysis and page **tagging.**

Log file analysis, also known as ***log management,*** is the process of analyzing data gathered from log files to monitor, troubleshoot and report on the performance of a website. Log files hold records of virtually every action taken on a network server, such as a web server, email server, database server or file server.

Page tagging is the process of adding snippets of code into a website's HyperText Markup Language code using a **tag management system** to track website visitors and their interactions across the website. These snippets of code are called tags. When businesses add these tags to a website, they can be used to track any number of metrics, such as the number of pages viewed, the number of unique visitors and the number of specific products viewed.

2.4.2 Types of Web Analytics Based on Purpose of Web Analytics

Many aspects play a part in the success of your online marketing and sales strategies. Below are the types of analytics you should pay attention to when assessing the overall progress of your sales and marketing efforts,

- **Funnel analytics :** These analytics give you insight into how leads and customers travel through your sales and marketing funnels.
- **On-site engagement analytics :** These analytics answer questions you may have about user experience and usability on your website.
- **Customer analytics :** These types of analytics give you the opportunity to get feedback from your customers. With customer analytics, you can get information about possible customer service issues in your marketing efforts and unfulfilled customer needs.
- **Lifecycle analytics :** The sales acquisition process is made up of multiple steps with unique marketing and sales techniques deployed to meet the customers where they are at on the buyer's journey.
- **Marketing analytics :** Many common analytics tools like Moz and HubSpot are marketing analytics tools. These types of analytics gather pertinent data around the health of your marketing campaign efforts and compile them into a centralized dashboard.

2.5 Web Analytics Fundamentals - The Process

Web analytics is the collection, reporting, and analysis of website data. The focus is on identifying measures based on your organizational and user goals and using the website data to determine the success or failure of those goals and to drive strategy and improve the user's experience.

Basic Steps in Web Analytics

1. Collecting data

To start an analysis, you first need to collect the necessary data. Most web analytics tools insert JavaScript code into HTML text of websites so that they can capture data and store it in database tables. Data they capture may include webpage clicks, the device user accessed, geographic location of the visitor and so on.

Before collecting website data, the following questions must be asked.

- Is the data accurate ?
- Are all the required data are being collected ?

There are various methods for data collection, including Web Logs, JavaScript tagging, packet sniffing and more.

2. Analyzing data and processing data into information

There are various ways in which data can be converted into insights. This is a very important step when optimizing the website. Some of them are discussed below.

- There is a need to trend the metrics over time to understand the website behaviour, as the website can't compare itself to another website.
- Number of page visits could be good or bad depending upon the kind of website. Time has always been a good indicator of engagement.
- A high bounce rate means you are either bringing the wrong audience or you are not up to the mark.
- A visitor's intent can be identified by analyzing the keyword driven traffic.
- A visitor's requirement can be identified using the internal site search data.

This step involves transforming data into metrics by creating ratios from counts you obtained in the first step. For example, bounce rate is dividing the early-leavers count by total visitor. Though this metric is important to understand the success of the webpage, it needs to be combined with other metrics and information to generate actionable insights to develop a marketing/business strategy.

For example, bounce rate can be due to a slow loading website, dull content, bugs on the website etc. Additional information is necessary to understand the root cause and take action.

3. Tracking KPIs

Organization track metrics important to their business strategy which are commonly called Key Performance Indicators (KPIs). They provide history to key metrics so companies can measure how they are progressing over time. For example, conversion rate and cost per conversion are typical KPIs.

To measure the achieved goals, you must analyze the performance of the website, i.e. whether it is going up or down by creating key performance indicator. A good KPI should have the following attributes.

- Simple
- Relevant
- Timely
- Instantly useful

4. Formulating strategies and identifying actions to take

These are the online goals, objectives and standards for your organization or business. Usually, these strategies are related to making money, saving money or increasing marketshare.

Goals are what drive success of a given business objective. Goals should be all of the following,

- Actionable
- Measurable
- Understandable

Reviewing all the analytics information and their business goals, businesses need to decide what to do. Continuous analytics enables organization to test the results of their strategies and make changes accordingly. For example, **A/B testing** is used commonly to improve conversions by testing two different designs for a page.

This stage involves implementing insights to formulate strategies that align with an organization's goals. For example, search queries conducted on-site can help an organization develop a content strategy based on what users are searching for on its website.

5. Experimenting and testing

Businesses need to experiment with different strategies in order to find the one that yields the best results. For example, **A/B testing** is a simple strategy to help learn how an audience responds to different content. The process involves creating two or more versions of content and then displaying it to different audience segments to reveal which version of the content performs better.

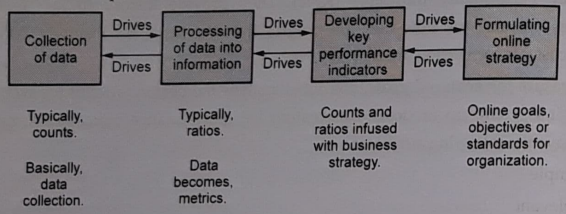

Fig. 2.5.1

2.6 Capturing Data for Web Analytics

There are two main technology approaches for collecting web analytics data : log-file analysis and page tagging as depicted in below Fig. 2.6.1.

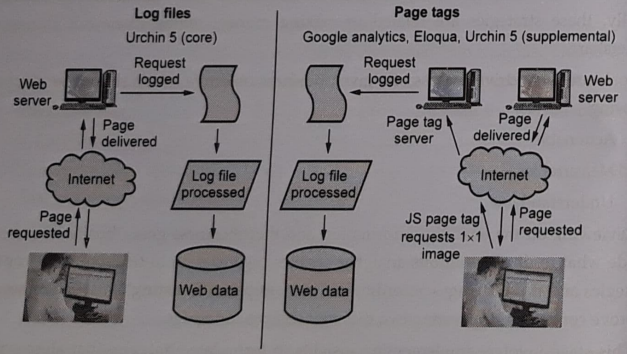

Fig. 2.6.1 Data capturing methods

Log-file analysis software reads the records, called log files, on the web server, which record all clicks that take place on the server. Web servers have always stored all the clicks that take place in a log file, so the software interprets data that have always been

available. A new line is written in a log file with each new request. For example, clicking on a link, an Ajax call, or submitting a form will each result in a new line being written.

Page tagging, on the other hand, sends information to a third-party server, where statistics can be generated. The browser executes JavaScript code that communicates with the tracking software, creating page tags.

Pixel tracking can be used to track e-mail campaigns. Here, a tiny, transparent pixel is placed in the e-mail. When you load the images in the e-mail, you will also load the tiny image that tracks your activity.

2.6.1 Log-File Analysis

- Log files are normally produced by Web servers, so the raw data are readily available. Page tagging, however, requires changes to the web site.
- Log files are very accurate - they record every click. Page tagging can be less accurate. If a user's browser does not support JavaScript, for example, no information will be captured.
- Log files are in a standard format, so it is possible to switch vendors and still be able to analyze historical data. Page tagging is proprietary to each vendor, so switching can mean losing historical data.
- Log files record visits from search engine spiders - useful for search engine optimization.
- Log files record failed requests, whereas page tagging only shows successful requests.

Log File-Based Tools Pros and Cons

Log file based approach Pros

- They (more) accurately reflect the actual load on your web servers - your IT department probably cares about this a lot more than your marketing department doe.
- They captures data very early in the process - as soon as you could possibly know someone is trying to view a page, they record it.

Log files based Cons

- They miss hits to cached pages (by browser, by proxy) - this can make for some rather nonsensical clickstreams.

- They are limited to data captured in the web server log file - this can be a fairly severe limitation if, for instance, you have rich meta data in the content of your pages and you want to use that meta data to group your content for analysis
- They capture a lot of useless data - I just went to the **Microsoft home page** and watched 65 discrete requests hit their web servers to render the page (images, stylesheets, Javascript include files, etc.); this is fairly typical and means you wind up pre-processing the log file to strip out all of the crud that you don't really care about
- It is difficult for them to filter out spiders/bots - there is a "long tail" of spiders crawling the web, so this is not simply a matter of knocking out Google's bot, Yahoo's bot, and Baidu's bot; there is an unmanageable, constantly changing list of known bots and spiders...and many bots mask themselves, which is extremely difficult to detect (this was actually the far-and-away biggest culprit with the client who spawned this post).

2.6.2 Page Tagging

In terms of page tagging, you should know the following :

- JavaScript makes it easier to capture more information (e.g., products purchased or screen size of a user's browser). You can use log-file analysis to capture this information, but it will involve modifying the URLs (uniform resource locators).
- Page tagging can report on events, such as interactions with a Flash movie, that log-file analysis cannot.
- Page tagging can be used by companies that do not run their own web servers.
- Page tagging service providers usually offer a greater level of support. This is because it is a third-party service, whereas log-file analysis software is often managed in-house.

Page Tag-Based Tools Pros and Cons

Page-tagging based approach Pros

- Because they are Javascript-based, they are able to capture lots of juicy supplemental data about the visitor and the content.
- Most (not all, mind you) spiders/bots do not execute Javascript, so they are automatically omitted from the data.
- The Javascript "forces" the page tag to fire...even on cached pages.

Page-tagging based approach Cons

- They requires the page tag Javascript to be deployed on every page you want tracked - even if you have a centrally managed footer that gets deployed to all pages. Chances are there are still some important corner case pages where this is not the case; and even if that is not the case now, that could happen in the future; we had a pretty robust system that was undermined when the design of a key landing page was completely overhauled and the page tag was nuked in the process.
- They do not record a hit to the page until the page has been at least partly delivered to the client - if you have visitors that bounce off of your site very quickly, you may never see that they hit the site at all.
- If Javascript is disabled by the client, then you have to put in some sort of clunky workaround to capture the traffic and what you capture will not be nearly as rich as what you capture for visitors who have Javascript enabled.

2.7 Data Sources for Web Analytics

The power of any analytics product is limited by the quality and diversity of its data sources. Web analytics commonly relies on below data sources,

- **Visitor data** includes
  - Data captured via javascript code snippets or cookies
  - Direct HTTP request data : Data sent by a web client (browser) to request a resource such as an image on a webpage
  - Application-level data sent with HTTP requests : This data is generated and processed by applications such as JavaScript, PHP and ASP.Net and includes how a visitor interacts with the web page. They are mostly collected by internal logs rather than web analytics services.
  - External data needed to analyze visitor data. External data is combined with on-site data to help understand website behavior data. Most common example is understanding the geolocation of users since IP addresses associated with geographic regions and internet service providers. This is a common feature offered by almost all web analytics software.
- **CRM :** If companies can connect visitors to leads in their CRM system, they can have an accurate view of where their revenue is coming from.
- **Search engines :** Search engines are a major source of traffic for companies and tools like google search console or bing webmaster tools can help companies better understand and optimize traffic from search engines.

- **Other search engine related data :** There are numerous data providers that cover different aspects of search engine traffic such as competitors' rankings, search engine friendliness of the company's web properties, etc.
- Server logs : Can provide additional, aggregated information about visitors that do not allow their data to be tracked.

Data	Type	Source
Page view	Measurement	HTTP request
Client profile/User-Agent (browser, OS)	Dimension	Application, HTTP request
User action (keyboard and mouse)	Measurement	Application
Geo location	Dimension	External
Visit or session	Measurement, Dimension	Application
Referrer (preceding webpage)	Dimension	HTTP request ("referrer" header)
Referral (channel identification)	Dimension	Application
Client profile (screen size, color depth)	Dimension	Application
IP address	Dimension	Network protocol
User profile	Dimension	External
Revenue or profit	Measurement	External

Table 2.7.1 Web analytics major data and source summary

Effect of GDPR on Data Collection

Countries are creating data privacy laws to protect their citizens' data from being abused. These laws include

- GDPR (General Data Protection Regulation) in Europe
- CCPA (California Consumer Privacy Act) in California, USA
- POPIA (Protection of Personal Information Act) in South Africa

Numerous countries either have published similar laws or are working on them. Most of these laws tend to be similar and we will focus on GDPR which has the largest geographic coverage.

In EU, GDPR is a regulation law on data protection and privacy. After GDPR, companies are required to acquire the user's consent to record their data. That is why

most EU websites nowadays feature a cookie disclaimer that asks users to accept cookies. If users do not accept cookies, none of their personal data can be stored. And most users are unlikely to accept cookies unless they are forced to accept them and GDPR prohibits companies from forcing users to accept cookies. However, this is not the end of your company's web analytics. We have seen 3 ways to mitigate the impact of GDPR :

- Users who consent to cookies can be analyzed to gain insights into user behavior. However, this provides only partial data, the cookie accepting users are likely to differ in their behavior when compared to cookie refusing users.
- Servers generate log files including non-individualized data such as visitor numbers per page. These can be used for high-level analytics as they do not constitute personal data.
- Data without personal information can be valuable as well. Vendors like Salesviewer use visitor data without relying on cookies by focusing on attributes like IP, browser characteristics, etc. to build a device fingerprint. Comparing this against their database, they can identify user's companies which they share with website owners. Their legal advisors certified them to be compliant to GDPR as well as national data privacy laws since their data processing is limited, pseudonymous and allows for opt-out.

2.8 Web Analytics Tools and Selecting Optimal Web Analytics Tools

2.8.1 Definition of Web Analytics Tool and its Task

Web analytics tools collect data to show you how visitors arrive at your website and what they do once they're there. These tools let you compare data over time to see patterns. This data also lets you measure performance against **benchmarks** and goals to see how your website is performing, where performance can be improved, and the effects of the actions you take to improve it.

Web analytics tools report important statistics on a website, such as where visitors came from, how long they stayed, how they found the site and their online activity while on the site. In addition to web analytics, these tools are commonly used for **product analytics, social media analytics** and **marketing analytics.**

Some of the things that website analytics tools can tell you include :

- How do people find your site? What do they do after they get there ?
- Which content on your site do people engage with ? When and how are they engaging with it ?

- Why do some people buy and others don't ? How can you get more of them to take action ?
- Some analytics programs also include data from social media and mobile apps, overlapping with the broader category of digital analytics.

2.8.2 Widely used Web Analytics Tools

The following are some of the most commonly used tools :

- **Google Analytics -** The 'standard' website analytics tool, free and widely used.
- **Piwik -** An open-source solution similar in functionality to Google and a popular alternative, allowing companies full ownership and control of their data.
- **Adobe Analytics -** Highly customizable analytics platform (Adobe bought analytics leader Omniture in 2009).
- **Kissmetrics -** Can zero in on individual behavior, i.e. cohort analysis, conversion and retention at the segment or individual level.
- **Mixpanel -** Advanced mobile and web analytics that measure actions rather than pageviews.
- **Parse.ly -** Offers detailed real-time analytics, specifically for publishers.
- **CrazyEgg -** Measures which parts of the page are getting the most attention using 'heat mapping'.
- **Clicktale -** Uses 'heat mapping,' keystrokes and mouse movement.
- **Optimizely -** Is a customer experience and **A/B testing** platform that helps businesses test and optimize their online experiences and marketing efforts, including conversion rate optimization.

2.8.3 Selecting Optimal Web Analytic Tool

In order to begin planning your web analytic tools, you need to first define the features that you are looking for, and then find a suitable service for each attribute.

As you can imagine choosing a web analytics tool is a critical process because whether you choose right or wrong you will be usually stuck with it for a while. Also because we tend to overstate the importance of historical web data (a post coming on this one of these days) it is quite likely that a quickie divorce will not be in the offing.

Acquiring the right web analytics tool will have a huge impact on the success of a company's digital marketing initiatives. By considering the below categories and carefully evaluating various web analytics tools, a marketer can obtain a user-friendly and

affordable tool that provides comprehensive insight into user behavior across a number of digital marketing channels.

Most Desired Features of Web Analytics Tool

1. **Traditional web analytics reports :**
   - **General usage reports,** such as visits, visitors, page views, bounce rates and top content.
   - **Visitor attribution,** such as country, browser and OS, screen resolution, flash version, etc.
   - **Traffic source analysis,** as visitors can come from many marketing channels. This enables measuring each channel separately and drilling down to a single keyword or banner placement.
2. **Data export :** For enabling the export of data to external tools in order to perform further analysis (Data Mining and BI Tools).
3. **Heat Maps, Scroll Maps and Click Maps :** A heat map (or heatmap) is a graphical representation of data where values are depicted by color. Heat maps make it easy to visualize complex data and understand it at a glance Heatmaps are a powerful way to understand what users do on website pages - where they click, how far they scroll, what they look at or ignore.

   Website heatmaps visualize the most popular (hot) and unpopular (cold) elements of a webpage using colors on a scale from red to blue.

   Heat map is an umbrella term for different heatmapping tools : Scroll maps, click maps, and move maps.

   Scroll maps show the exact percentage of people who scroll down to any point on the page : The redder the area, the more visitors saw it.

   Click maps show an aggregate of where visitors click their mouse on desktop devices and tap their finger on mobile devices (in this case, they are known as touch heatmaps). The map is color - coded to show the elements that have been clicked and tapped the most (red, orange, yellow).

   Move maps track where desktop users move their mouse as they navigate the page. The hot spots in a move map represent where users have moved their cursor on a page, and research suggests a correlation between where people are looking and where their mouse is - meaning that a move map gives an indication of where people might be looking as they go through your page.

4. **Funnel analysis :** For defining funnels and providing detailed conversion rate reports for all of the steps along each funnel. In most cases, being able to provide additional attributes for each step is necessary, in order to later enable the option of segmenting the data based on these attributes.
5. **Offline events :** In various cases, where some user flows may occur outside the site (on a different site, based on server side processing, from a downloadable client, or from a different browser or computer), we need a way to send events and transfer the user identification to the selected web/funnel analytics tool.
6. **AB (Split testing) and MVT (Multivariate testing) :** A tool that will allow us to prepare and run experiments for all the steps along the funnels, and present the results of these experiments.
7. **Visitor usage reports :** In order to be able to drill down through each user and see the overall usage.
8. **Real user monitoring :** A tool that will provide as much information as possible about the end user experience, such as: On average, how long does it take the page to load around the world, how many people leave the site before it is fully loaded, how long does it take to render the page on different browsers, etc.
9. **Voice of customer :** In order to be able to interact with the users and receive their feedback. This tool should also be able to specify who sees the surveys, when and where.
10. **Behavioral targeting :** For personalizing the content on the site, based on where the visitors came from and what they did on the site. This is something that I usually perform based on AB and MVT testing and after I have a clear understanding of my site and visitors usage.

Main categories used for web analytics tool evaluation

- **Features :** Each web analytics program offers a variety of features to help you understand your website visitors better. This section offers a top level view of several features that will be useful in measuring your website traffic.
- **Traffic :** Web analytics keep track of a large variety of very detailed information regarding who visits a website, what each visitor does on the site (what they click on, pages they view, etc.), and at what point they exit the website.
- **Referrals :** According to the Web Analytics Association, a referrer is "the page URL that originally generated the request for the current page view or object." Essentially, this is where your guest came from immediately before arriving on your website.

- **Report stat intervals :** Web traffic and other statistics are only useful if they are measurable, and that implies setting applicable date ranges. Whether you need yearly or monthly reports, or details down to the day or hour; it's important that your analytics provider is versatile with reporting dates and the accompanying stats.
- **Events :** According to the Web Analytics Association, an event is "any logged or recorded action that has a specific date and time assigned to it by either the browser or server."
- **Visitor details :** Web analytics programs keep track of each visitor to your site. This information can be used to identify target audiences, develop campaigns, or learn what might work better to increase conversions. Detailed geographic information about where the visitor is accessing the website from is also available in most cases.
- **File exporting :** After you compile stacks of data, what do you do with it ? Most web analytics programs offer a variety exporting options to meet your specific needs.
- **Tech support / help :** Web analytics solutions can be very complex. Choosing which one to use on your site is only the first of many steps, so look for a solution that offers product support for a period of time following the initial purchase.

2.9 Web Analytics - What is to Measure - The Metrics

2.9.1 Concept of Metric

Defining metrics is a vital part of web analytics process and tracking the right metrics can lead to the success of analytics projects.

There are three types of Web analytics metrics

1. **Count :** These are the raw figures captured that will be used for analysis.
2. **Ratio :** This is an interpretation of the data that are counted.
3. **KPI (key performance indicator) :** Either a count or a ratio, these are the figures that help you to determine your success in reaching your goals.

These are the most basic web metrics. They tell you how much traffic your web site is receiving. Looking at repeat and returning visitors can tell you about how your web site creates loyalty. As well as growing overall visitor numbers, a web site needs to grow the number of visitors who come back. An exception might be a support web site - repeat visitors could indicate that the web site has not been successful in solving the visitor's problem. Each web site needs to be analyzed based on its purpose.

In analysis, metrics can be applied to three different domains

1. **Aggregate :** All traffic to the Web site for a defined period of time.
2. **Segmented :** A subset of all traffic according to a specific filter, such as by campaign (PPC) or visitor type (new visitor vs. returning visitor).
3. **Individual :** The activity of a single visitor for a defined period of time.

2.9.2 Web Analytics Key Metrics

Type 1 - Website Traffic Metrics

Metrics related to website traffic can tell you how many people are coming to your site, where they're coming from, and what pages they're looking at.

- **Hit :** A request to the server (and a fairly meaningless number on its own).
- **Page :** Unit of content (so downloads and Flash files can be defined as a page).
- **Page views :** The number of times a page was successfully requested.
- **Visit or session :** An interaction by an individual with a Web site consisting of one or more page views within a specified period of time. This metric indicates the website's total number of visitors. Though this metric is the third most tracked metric of Google Analytics, it not very useful for marketers because it may count same people twice or more if they log multiple sessions.
- **Average session duration :** The average amount of time users spend on your website. This is another metric that is neither good nor bad because you can not increase this average by advertising or stuffing more keywords onto the page.
- **Percentage of new sessions :** Google explains this metric as "An estimate of the percentage of first-time visits.". This metric reveals important insights such as how fast your brand and website are growing. It is calculated as unique visitors/sessions.
- **Pages per session :** The average number of pages viewed during a session on your website. Measuring this KPI indicates the interest of an average user about your company.
- **Sessions by channel :** Your website is growing as new visitors come but how did they find you in the first place ? Google analytics enables you to group and track the performance of channels such as organic, paid, direct, social, email, referral, etc. Marketers can decide which channels they need to invest more or which channels are underperforming despite their budget.
- **Goal conversion rate :** Businesses mostly start web analytics with a goal such as increasing sign up number, purchase, or click to other pages. With Google Analytics, you can create your goals and start tracking your success rate.

- **Pageviews :** Pageviews is the total number of pages viewed on your website within a given time. It includes repeated views of a single page that's why it is considered as generous measurement.
- **Pageviews by page :** This metric is the total number of views for each page within a given time. It is a generous measurement same as pageviews.
- **Exit Rate :** Exit rate is a metric like bounce rate with one difference. A bounce is always an exit, but an exit doesn't have to be a bounce because the bounce rate is measured against the number of entrances while the exit rate is measured against a particular page.
- **Unique visitors :** The number of individual people visiting the web site one or more times within a period of time. Each individual is only counted once. Types of visitors can be categorized as follows :
  - **New visitor :** A unique visitor who visits the web site for the first time ever in the period of time being analyzed.
  - **Repeat visitor :** A unique visitor with two or more visits within the time period being analyzed.
  - **Return visitor :** A unique visitor who is not a new visitor.

This is the number of first time visitor over a period of time (organizations mostly prefers to use per month). This metric is important because you can measure the popularity of the website. By focusing on spikes and lags, you may identify what is causing trends.

Type 2 - User Behavior Metrics

Metrics related to user behavior can show you how visitors engage with your website, and tell you if your site is providing a **good user experience.**

- **Bounce rate :** This is the percentage of single-page site visits. This metric is neither good nor bad metric because it cannot identifies the web page's success. The user may get what he wants in one click.
- **Ad impressions :** Resources about the ad impressions metric, which measures how many times an ad placed on a website was viewed by potential site visitors (only some of whom actually click the ad and visit the advertising website).
- **Referrers and social media traffic :** Resources about the referrers metric in web analytics, which measures the number of visitors coming to a website via referring sites - other sites that link to the website. A special case of referral traffic is social media traffic - traffic to a website from social media sites like Facebook and Twitter.

- **Direct traffic in web analytics :** Resources about the direct traffic metric in web analytics, which measures the number of visitors entering a website directly by typing the URL or clicking on a bookmark, not via other websites.
- **Source and medium in web analytics :** Resources about possible source of traffic to a website (referrals from other sites, search traffic, direct, links from emails, etc.) and the source/medium metric in web analytics, which measures how many of a website's visitors came from each source and provides details (such as a specific URL the visitors originated from).

Type 3 - Audience Metrics

Knowing your audience can help you with everything from creating marketing campaigns to product development, and google analytics can give you a lot of information about the people who visit your site.

- **Breakdown by location :** Tracking the location of visitors and filtering them based on location help businesses improve their local marketing and SEO efforts for specific countries.
- **Breakdown by device :** Knowing the devices and browser of web visitors help businesses optimize and plan the resources they will use. For instance, checking the percentage of mobile users can influence businesses' decisions on spending money for mobile SEO. They can look at the average CTR of mobile users and decide whether to make their website more mobile-friendly.
- **Breakdown by browser :** Type of browser website users use can help businesses identify quick actions. For example, Chrome users can see the URL and name of images once they place their cursor on images (As seen on the left corner of the image below). And if Chrome is the dominant browser option among your audience, writing a related title for the image file can be an initiative to improve the user experience of Chrome users.
- **Mobile :** More than 50 % of all internet traffic is now on mobile devices, but that applies to the whole world, not your specific site. Knowing what type of devices your visitors use can help you decide if you need to prioritize a **mobile-friendly format.**

2.9.3 Web Analytics Metrics Explained Using User and Data Characteristics

2.9.3.1 Visit Characterization

The following help characterize the visit to a particular web site :

- **Entry page :** The first page of a visit.
- **Landing page :** The page intended to identify the beginning of the user experience resulting from a defined marketing effort.
- **Exit page :** The last page of a visit.
- **Visit duration :** The length of time in a session.
- **Referrer :** The URL (Uniform Resource Locator) that originally generated the request for the current page.
  - **Internal referrer :** A URL that is part of the same web site.
  - **External referrer :** A URL that is outside of the web site.
  - **Search referrer :** The URL has been generated by a search function.
  - **Visit referrer :** The URL that originated a particular visit.
  - **Original referrer :** The URL that sent a new visitor to the web site.
- **Click-through :** The number of times a link was clicked by a visitor.
- **Click-through rate :** The number of times a link was clicked divided by the number of times it was seen (impressions).
- **Page views per visit :** The number of page views in a reporting period divided by the number of visits in that same period.

These are the terms that tell you how visitors reach your web site and how they move through the web site. The way that a visitor navigates a web site is called a click path. Looking at the referrers, both internal and internal, allows you to gauge a click path that visitors take.

2.9.3.2 Content Characterization

The following help characterize how visitors move through the web site :

- **Page exit ratio :** Number of exits from a page divided by total number of page views of that page.
- **Single-page visits :** Visits that consist of one page, even if that page was viewed a number of times.
- **Bounces (single page-view visits) :** Visits consisting of a single page view.
- **Bounce rate :** Single page - view visits divided by entry pages.

When visitors view a page, they have two options: leave the Web site, or view another page on the Web site. These metrics tell you how visitors react to your content. Bounce rate can be one of the most important metrics that you measure! There are a few exceptions, but a high bounce rate usually means high dissatisfaction with a web page.

2.9.3.3 Conversion Metrics

Other metrics that apply to eMarketing tactics include the following :

- **Event :** A recorded action that has a specific time assigned to it by the browser or the server.
- **Conversion :** A visitor completing a target action.
- **Impression :** Each time an advertisement or a page is served.
- **Open :** Each e-mail that is deemed open. Usually if the images are loaded, an e-mail is considered open.

In order to test the success of your web site, you need to remember the TAO of conversion optimization :

- Track
- Analyze
- Optimize

Using your goals and KPIs, you'll know what metrics you will be tracking. You will then need to analyze these results and take appropriate actions. And the testing begins again!

2.10 Understanding Quick Stream Data quality - Streaming Analytics

2.10.1 Concept of Streaming Analytics

Stream processing analyzes and performs actions on real-time data though the use of continuous queries. Streaming Analytics connects to external data sources, enabling applications to integrate certain data into the application flow, or to update an external database with processed information.

Essential to stream processing is streaming analytics. Streaming Analytics is the ability to constantly calculate statistical analytics while moving within the stream of data. Streaming analytics allows management, monitoring, and real-time analytics of live streaming data.

Streaming analytics involves knowing and acting upon events happening in your business at any given moment. Since Streaming Analytics occurs immediately, companies

must act on the analytics data quickly within a small window of opportunity before the data loses its value. The data can originate from the **Internet of Things (IoT),** mobile phones, and mobile devices such as iPads, market data, sensors, Web clickstream, and transactions. Data that loses its value results in additional costs such as: operational, administrative, business risks, reputation damage, potential legal action, reduction in productivity, inability to make informed decisions, and reduces a company's competitive edge.

Streaming analytics taps into streams of GPS data from cars, continuously aggregates that data, and merges it in real time with the location information of customers. Each move a taxi driver makes and each move a car makes, streaming analytics calculates which cars are closest to a specific taxi driver based on any selection criteria chosen.

Real-time analytics makes it possible for organizations to capture live streams of data, process them very quickly and extract insights or perform operations on the data in real time or near-real-time. It is based on stream processing technology that can handle a very high throughput of event data.

There are two types of real time analytics,

- **On-demand analytics :** Providing data or computation results to users or applications in real time. For example, displaying the current product price to a user on an eCommerce site.
- **Continuous analytics :** Processing events on a continuous basis and streaming the results to end users, applications, or a data store. For example, showing live stock market data on a dashboard in a financial institution.

There are plenty more places where real-time data can make a difference on a business's bottom line,

- **Creating targeted pricing strategies :** If your business runs promotions on items, testing the right pricing is paramount to ensuring that customers buy your products. Streaming data can allow for more precise actions on price elasticity for each customer, timing of discounts, customized offering and sales channel
- **Detecting fraud in real time :** Access to real-time streaming data means you can respond quickly to any financial irregularities - so instead of writing off the costs of a fraudulent transaction, a company can flag it immediately.
- **Building customer loyalty and capturing market share :** Building more responsive relationships helps to gain customer trust and capture revenue. So companies that can propose and interact with their prospects in close to real time with a customized offering of content, pricing, and solution will lead to loyal and happy customers.

- **Finding operational efficiencies :** Real-time data analytics can continually monitor data integrity and let you respond automatically. Adoption of streaming can help eliminate manual processes that are susceptible to error, enable better data interoperability with other organizations, and increase speed-to-market by making data more actionable.

2.10.2 Assessing if Streaming Analytics is Right for Organization

Not all your problems will benefit from streaming analytics equally and getting started with real-time data can be overwhelming. There are plenty of ways to capture, ingest, and process data, and plenty of information to be gleaned from analyzing your company's data. Which data is the right data to gather and analyze ? What's the right way to prioritize the data you want to capture in real time, and which data can wait ? To decide if streaming analytics is right for you, it helps to consider the following :

- **Assess your current environment :** Identify which applications generate data in your organization and rank those data streams based on their importance. For example, in retail, the need for real-time applications would probably rank higher for website clickstreams compared to back office payroll, given direct revenue generation opportunity.
- **Map real-time analysis use cases to the data streams :** Decide which are your critical activities to improve top and bottom line, whether it be responding to customers, detecting faulty products or enhancing security.
- **Evaluate buy vs. build :** Do you have staff with the right skill sets to capture the maximum value from the technology? Do you have the resources to hire these experts? This will have an implication on time to value as you choose between an open source technology vs. a fully managed service.

2.10.3 Advantages of Streaming Analytics

- **Provides deeper insight through data visualization :** Visualization of vital company information can help companies manage their key performance indicators (KPIs) on a daily basis. KPI data is viewed in real time, which produces a single source of truth of real-time data that can provide a helicopter and granular view of a company at any given time. The data can improve sales, reduce costs, identify errors, and provide information to react faster to risks to mitigate them. Streaming analytics accelerates decision-making and provides access to business metrics and reporting.

- **Offers insight into customer behavior :** Streaming analytics allows companies to gain visibility into what customers are buying, not buying, customer preferences, and dislikes. This gives companies the ability to generate additional profit and retain existing customers. It allows companies to rapidly respond to customer needs and increase revenues through up-selling and cross-selling of goods and services.
- **Remain competitive :** Businesses can identify trends and benchmarks, develop white papers, use cases, and generate forecasts of their company and industry. This reduces internal and external threats and provides awareness of industry changes. This helps companies become innovative, remain competitive, and strengthen their brand.

2.10.4 Disadvantages of Streaming Analytics

- **Lack of experts :** An issue with Streaming Analytics is the lack of availability of experts in the field. There are only a small number of Data Scientists and a smaller number of companies that hire them. Streaming Analytics is still a recent technology and adoption is slow by most developers due to their lack of expertise. "The streaming application programming model is unfamiliar to most application developers," wrote Forrester analysts Mike Gualtieri and Rowan Curran in a Q3 2014 Forrester report on **Big Data and Streaming Analytics.**
- **Perform risk analysis :** Streaming analytics allows companies to view and analyze the latest media and industry news to keep abreast on the latest development in their industry. It also provides companies with data on customers and vendors allowing companies to take action when a risk or specific event occurs.
- **Securing data :** Streaming analytics allow companies to analyze internal and external threats that affect the company or industry. Companies can identify sensitive information that is not protected or that is not adequately protected and ensure state, federal and regulatory requirements are met.

2.11 Data Quality in Web Analytics

2.11.1 Data Quality Central Concept

Companies that deal with big data, use analytics tools in order to make short and long-term data-driven decisions and operate efficiently. It is crucial for those brands that their collected data is valuable and makes it possible to produce equally valuable business insights. Data quality is the ability of data to serve its purpose - to help create useful, reliable insights for your company. Information that leads to making the right business decision is high quality information. So, how do we evaluate data ? Some articles may

mention five, six or even ten factors you have to remember about. Here are some of the characteristics used to describe data's quality :

- **Completeness :** Incomplete data lacks some information that might be useful. Where there are gaps in datasets, there is the possibility of producing unreliable analysis and making the wrong decisions for your business, which could have awful outcomes.
- **Accuracy and reliability :** What if the data is not correct at all or comes from an uncertain source ? Misleading information - once again - leads to bad decisions. Without accuracy, your data may do as much bad as good.
- **Availability :** There are many people in the company that work with data and can use it to do their job better. If some data is not available for certain employees that may benefit from using it, it does not serve its purpose. Be sure that the experts who need particular information have access to it.
- **Timeliness :** In some cases, using historical data to make decisions is not a good idea. If your company needs to produce insights a short time after collecting data, you should learn more about real time big data analytics. Bear in mind that relying on outdated information may bring you to an inaccurate decision.
- **Granularity :** Data may give you knowledge about details or a general state of something. Many times, you cannot make a good decision operating only on general data.
- **Relevance :** Some data may not be useful. So, what is the point of storing it at all ? Information you find totally useless should not be taken under consideration during data analysis.

Why is data quality important ? Using false or inaccurate data, you won't be able to make the right decisions for your company and you'll have to waste your time and resources in order to solve the problems you created yourself instead of investing in development.

2.11.2 Low Quality Data

You should not base your judgment on unreliable information - if you don't know the source or you're not sure if you can rely on some data, you should fix the issue by eliminating unreliable sources. Incomplete data should also be eliminated, as it doesn't give you (or your analytics systems) a clear, real view of the situation. Some data may be ambiguous - easy to misinterpret - and therefore lead to bad conclusions.

Quite serious problems may be caused by duplicated data. Imagine having many profiles of the same customer in your database. Apart from confusing your staff and

taking up additional storage, this also leads to an inaccurate customer count and weakens marketing analysis, etc.

You need to remember about updating data - if you wish to improve your offer and marketing campaigns. Some of your customers may have been students just a few years ago, but today they might have two children and a dog - and completely different needs than before. You can use new information to recommend better suited products or optimize advertising.

2.11.3 The Data Quality Management Process

Data quality management is a process whose goal is to eliminate useless data in order to maintain the high quality of the information sets to be used for analytics.

1. Defining the required data quality

Ensuring high quality data starts with defining what the data should actually be like. You do this by establishing thresholds and rules - requirements for your information. Ideal data will be 100 % compliant with your data quality characteristics (accuracy, etc.). As you probably suspect, reaching 100 % for all the attributes is very difficult. Usually, a company decides which data and qualities are the most important.

2. Checking data accuracy

After setting rules, you need to have a look at your data and see if it meets the rules you have set. This process makes it possible to separate low quality information from high quality information in order to ensure good business insights.

3. Identify what causes low data quality

There has to be a cause for collecting low quality data. Have you ever wondered why the information gathered by your company is not good enough to produce useful business insights? Eliminating the sources of incomplete or unreliable data is the first step towards improving the decision-making process. Sometimes this can be done easily, by making the form you use for collecting data clearer. You can set a validation rule in your system so it doesn't accept data if it is wrong.

4. Monitor and control data

The data quality management process never ends – you need to review your data quality regularly to improve it. What you can't forget is that the business environment is changing all the time. Other data may become important or should be evaluated in a different way.

2.12 Web Analytics Best Practices

Web analytics can strongly support the qualitative research and testing finding. Some best practices to keep in mind related to this field are discussed below.

1. **Identifying the right resources :** Setting up web analytics is as much a technical task as it is a business task. The resource who is setting up analytics needs to know technical analytics implementation so that data is accurately recorded and measured. The resource also needs to be aware of how the data will be used. Without having the knowledge and ability to do both, the resource will be ineffective at setting up web analytics. So having just a single person implement web analytics is rare because of the mix of technical and business skills needed for the job. This is also the reason why small companies struggle with effective use of web analytics for business performance measurement
2. **Encourage a data-driven environment for decision making :** After collecting the relevant data to answer whether you have met (or fail to meet) your goals, find out what you can do to improve your KPIs. Are there high-value content (based on user feedback to the website) that is not getting any traffic? Find out why through user path analysis or engagement analysis of top sources for that page. Leverage the experimentation and testing tools to try out different solutions and find the best placement that generates the most engagement for that page.
3. **Strong understanding of business processes :** Learning how your business partners do their jobs gives you better insight into the types of data and analysis that will make them successful. It also helps you identify what types of measurement improvements will lead to improved reporting. The goal of a great analytics organization is to feed the business users exactly the right information at the right time to make informed decisions that will lead to higher conversions. When working with business partners, you must know what actions they will take, not just what data they need.
4. **Avoid only providing traffic reports :** Reporting about visits, pageviews, top sources, or top pages only skims the surface. Large numbers can be misleading; just because there is more traffic or time spent on site doesn't mean that there is success. Reporting these numbers is largely tactical; after all, what do 7 million visits have to do with the success of your program ?
5. **Create filters that prevent inaccuracies from internal traffic :** Your online marketing and sales activities and campaigns won't operate in a bubble whereby only certain web users will have their activity tracked. Marketing and sales managers, members of your marketing and sales teams, your company's president,

or other employees can and will like, comment, and share blog posts and social media content. They will also click through your website and may even make an online purchase. If you don't have filters set up in Google analytics to exclude internal traffic and interactions, your data will be greatly skewed and give you an inflated and inaccurate progress report of your sales and marketing activities.

6. **Always provide insights with the data :** Reporting metrics to your stakeholders with no insights or tie-ins to your business or user goals misses the point. Make the data relevant and meaningful by demonstrating how the website data shows areas of success and of improvement on your site.
7. **Avoid being snapshot-focused in reporting :** Focusing on visits or looking only within a specific time period doesn't capture the richer and more complex web experiences that are happening online now. Pan-session metrics, such as visitors, user-lifetime value, and other values that provide longer-term understanding of people and users, allow you to evaluate how your website has been doing as it matures and as it interacts with visitors, especially the returning ones.
8. **Consider seasonality, algorithm updates and bots :** While web analytics data is reliable for the most part, it isn't immune to inaccurate spikes and declines. The main causes of such sudden changes in data numbers include regular Google search algorithm updates and fraudulent bot traffic. If you come across a sudden spike or decline in your web analytics data, stop and look at the possible reasons behind them and exclude that data from the rest. Excluding off-season data will give you a more accurate view of how well your sales and marketing strategies are doing.
9. **Communicate clearly with stakeholders :** Be consistent in the information you provide, know your audience, and know the weaknesses of your system and disclose them to your stakeholders.
10. **Strict documentation :** This entire process of evaluation and truth finding is about putting together bits and pieces of information and figuring out the answer to a puzzle. To do so, documentation is a must. There are a lot of terms and concepts used in web analytics, all of which need to be defined for a team of analysts to accurately deliver results and gauge performance. A data dictionary and an events calendar, for instance, should be the starting points of documentation because they guide analysts to spot anomalies and refer to the gathered data for answers. An "analytics events" calendar can help analysts understand anomalies in the data due to sale events, code changes, or outages. Also, you might want to make an intranet or document repository available to your business partners that contains links to the data dictionary and other relevant documentation.

11. **Venue for sharing ideas and successes :** Great analysis is useful only if it's used to improve performance. To this end, successful organizations hold roundtables of business users or analysts to share successes and recommendations for improved measurement. Typically, this is done by functional area, such as marketing or merchandising. This also helps the analytics manager understand how the business users employ the data. Further, sharing successes with executives helps build support for additional analytics projects and resources. Building out a document repository for shared analysis is an important part of this process.
12. **A system to log requests :** Managing a huge backlog of requests is a common problem for Web analysts. Coming up with a system to log and prioritize those requests is critical. You must provide backlog visibility to your managers and stakeholders to ensure that you handle the most impactful requests first.; this can help you justify additional analytics resources. Outlook inbox is not the ideal system to log and manage critical requests. One solution is an online form, such as **Adobe FormCentral,** which is a great way to funnel requests into a spreadsheet that can be easily shared, sorted, and updated.
13. **Having a roadmap for improvements :** The road to improvement is a best practice for every business, more so for one that has everything to do with the online community. Changes on the internet are fast paced and regular. For web analytics to stay abreast of these changes, it has to have a roadmap for improvements and constant progress.

2.13 Three Layers of So What Test

To get best possible result from web analytics process one need to make right choices about the web metrics. A very simple test can allow you to figure out if the chosen metric is working as expected or not. It is called the **Three Layers of So What test.**

One should ask, every web chosen metric the question, "so what" three times.

Each question provides an answer that in turn raises another question (a "so what" again). If at the third "so what" you don't get a recommendation for an action you should take, you have the wrong metric. Drop it then.

This brutal recommendation is to force you to confront this reality: If you can't take action, some action (any action!), based on your analysis, why are you reporting data ?

The purpose of the "so what" test is to undo the clutter in your life and allow you to focus on only the metrics that will help you take action. All other metrics, those that fall into the *nice to know* or *the highly recommended* or the *I don't know why I am reporting this but*

it sounds important, would be definitely questioned in this test would be dropped if they are not needed. You need to face up to the reality that just because it sounds good, or looks good on paper during a board meeting, doesn't mean that the actual statistic is relevant or worth anything. If you can not act upon it, who cares ?

Asking "So What ?" three times to every key metric should eventually end in an actionable item.

Example

Let's check out how you would conduct the "so what" test with a couple of examples. You run a report and notice a trend for this metric.

Here is how the "so what" test will work,

"The trend of repeat visitors for our website is up month to month."

So what ? - 1

"This is fantastic because it shows that we are a more visited website now."

So what ? - 2

"We should do more of work for the website to leverage this trend." (Or – a specific action based on analysis of what caused the trend to go up.)

So what ? - 3

If your answer to that last "so what" is: "I don't know... isn't that a good thing... the trend is going up... or... I am not sure there is anything we can do... but it is going up right ?"

At this point you should understand that metric might give you wrong impression or it is not the metric you needed to be analyzed.

Obviously one size does not fit all when it comes to web metrics and there may be instances where certain rules do not apply, but they will be the exception and not the norm. Apply your 3 tests and see if you can actually do something with that wonderful key statistic you have been bragging about for the past 6 months.

2.14 Diving Deep in to the Web Analytics Concept - Understanding Basics, Revisiting Foundation Metrics

2.14.1 Understanding Web Analytics Basic

Web analytics is the collection, reporting, and analysis of website data. The focus is on identifying measures based on your organizational and user goals and using the website data to determine the success or failure of those goals and to drive strategy and improve the user's experience (UX).

Basic Terms Related to Web Analytics

1) Website traffic

The most critically important and fundamental of all web analytics is undoubtedly website traffic. Essentially, **this term refers to the number of visitors to your site on any given day, week, or month.** There are many proven methods and effective tactics for gaining more website traffic, including :

- Paid digital marketing campaigns
- Search Engine Optimization (SEO) efforts
- Posting fresh, original, keyword-targeted organic content via blogging
- Establishing a social media presence with engaged followers
- Public relations (Print or Digital)
- Word of mouth
- Email marketing
- Commenting on other relevant blogs or articles

Perhaps the most important nuance to this web analytics metric is "Unique Visitors". While website traffic is a calculation of all inbound visitors, **the number of unique visitors is the amount of new users who have spent time on your site.** The difference between this figure and the total amount of website traffic is thus the number of returning visitors or users who have been to your site more than once.

2) Website traffic source

While website traffic may be the most crucial web analytics concept, traffic sources provide critical context for understanding the value of your marketing efforts. In short, the website traffic source explains where your visitors are coming from on the Internet.

The four primary traffic sources are :

1. **Organic search :** Visitors coming from search engines (Google, Bing, etc.)
2. **Referral :** Web traffic from another website
3. **Direct :** A user types the exact domain URL into their computer browser
4. **Social :** Website traffic coming from social media networks (Facebook, Twitter, etc.)

If you can keep a finger on the pulse of this aspect of web analytics, **you can see where your visitors are coming from.** This is the first step to understanding what is successful online, and where you may need improvement.

3) Bounce rate

Bounce rate forms the bedrock of measuring the success of your website design, user interface, and incoming traffic. **A "bounce" occurs on your website when a user visits your site, but then immediately leaves.** The lower the bounce rate, the better. However, a bounce rate of zero is practically impossible, as some users are bound to visit by accident.

Bounce rate is a critical figure in web analytics, as it **provides great insight into the quality of your web presence.** A high bounce rate can be an indicator of several issues. Perhaps the visitor didn't find what they were looking for or expected in your website. Perhaps they didn't like the interface or the site loaded too slowly. Analyzing traffic source and bounce rate in a web analytics platform can help you discern what may be the source of a high bounce rate.

4) Top pages

The "Top pages" on your website **indicate which individual pages are the most important or valuable to your audience.** These are typically listed as the top 10 most visited pages in more web analytics platforms. It is important to note that this list can experience dramatic variation early on, so check back often. Once this list is established, your top pages can be used to understand where your traffic is visiting and spending time on your site. This can also tell you where to focus on improvement and which pages will be most impacted by changes to their design.

5) Exit pages

Exit pages are related to bounce rate in web analytics, but don't confuse the two. In fact, exit pages stand in direct contrast your bounce rate. A bounce occurs when a visitor immediately leaves your website, but your **exit pages are calculated by traffic that visits multiple pages on your site before leaving.** Exit pages are thus are a list of the top points of departure from your website. The last page your user visits before leaving can provide data and insight into potential issues with site design or content. Once your list of exit pages are determined, try to figure out why users are leaving from this location.

2.14.2 Revisiting Foundation Metrics

Critical to developing relevant and effective web analysis is creating objectives and calls-to-action from your organizational and site visitors goals, and identifying key performance indicators (KPIs) to measure the success or failures for those objectives and calls-to-action.

In a lot of ways, your website serves as the hub of your analysis efforts. You can use your website to distribute content, promote specific products and services to customers, communicate your brand's story and more.

Here's what web analytics comprises based on the type of data you should be tracking and how you can use them to improve web analysis efforts.

1. User analytics

This has a lot to do with knowing who your site visitors are in order for you to give them the best user experience possible. When you have data about their age, gender or location, it's easier for you to incorporate features you think they will need to help them navigate through your site.

For example, you can better deal with potential problems in language usage if you know where your users are located geographically. With this data, you can provide an option for your site's visitors to use content translated in their own language.

This part of web analytics can also tell you the type of device that visitors typically use to access your site. If most of them are mobile device users, it might be advantageous for you to optimize your site for mobile viewing or spend more on mobile advertising.

2. Content analytics

You can identify your low- and high-performing content with this metric. It goes without saying that your best content will get the most views, likes, shares, mentions and other engagement signals. You can then analyze what topics are most attractive for users, what type of keywords, CTAs and headlines work and even how the layout or design affects your content creation plans.

You can also identify which specific audiences are engaging with your content : are they average content consumers or are they looking for very technical and thorough information from your content ?

From these insights, you can focus on creating enhanced content that's tailored to your target market, helping you steer clear of misallocated resources and instead drive higher audience engagement for your site.

3. Behavioral analytics

Since every visitor on your site leaves a digital footprint, you have numerous opportunities to look for patterns in the way that visitors are using your site or what they're using it for - whether their purpose is to get in touch with you, check your offerings, make a purchase, share their experiences with your business, and so on.

Through behavioral analytics, you can track which pages on your site are performing and therefore should be optimized. It will also pinpoint on which pages they're spending the most or the least of their time. It could be that users are filling their shopping carts but once they get to the checkout page, they choose to abandon their cart because they have to fill out a lengthy checkout form.

The focus is on the series of events that show how users interacted on your site until they decided to buy or sign up for your product or service, or otherwise.

All things considered, your end goal in behavioral analytics is to be able to offer a smooth and seamless website experience for visitors through a user-friendly interface, clear navigational flow, fast-loading web pages and the like.

4. Traffic analytics

Traffic analytics lets you track which third party site or platform directed the highest traffic to your site, so you can boost your marketing efforts in that channel. If you're doing both Search Engine Optimization (SEO) and display ad campaigns but notice that your site traffic is mostly organic, then you'd do better to increase your budget in this department instead of paying for more space to run your ads.

Meanwhile, your social media reach can also be a good indicator and driver of your web traffic. Perhaps someone found your brand on LinkedIn, Facebook, or Twitter and visited your site or your blog afterward; in such case, make sure that you leverage these platforms for content and social marketing moving forward.

5. Acquisition analytics

Unlike traffic analytics - where the focus is on the source of traffic - acquisition analytics shows you whether those traffic sources are helping you acquire leads or customers for your site. If visitors from sites that link to your website download your content or install your app, it's a good indication that those sites are giving you high quality traffic that converts, which can serve as sales leads for the company in the future.

Conversely speaking, you may be getting a lot of traffic from other sources, but if they're not adding any value or generating revenue for your business, you might need to rethink your marketing efforts so you don't spend unnecessarily on campaigns that don't produce significant results.

2.14.3 Understanding Standard Reports Generated by Web Analytics

2.14.3.1 Web Analytics Dashboard

A web analytics dashboard is simply a reporting interface that displays the data derived from the monitoring of your website performance. This is usually done by tracking some specific metrics or key performance indicators such as online conversions, bounce rates, page views, referral traffics and so forth.

Web analytics tools or software are indispensable tools for every webmaster. This is because it allows smooth collection, collation, measurement, and analysis of data for their websites.

2.14.3.2 Web Analytics Standard Reports

There is more than one way to slice user data flowing in from a website. Web designers, digital marketing experts and other members of companies' web teams can make use of multiple types of reporting to determine how people are interacting with their sites, each pointing to different outcomes.

It can be noted that while it may be tempting to simplify web analytics to traffic reporting, consisting of statistics about website visitors, page views, or which parts of the site are the most important, there is ample room to go deeper and companies should do so. Opportunities for real improvement come from avoiding these superficial numbers and using reports that can explain more of the why behind customer behavior.

The following are seven types of analytics reports web developers can use to determine if they are on the right track with their campaigns and designs. They are based on the reports available in the Google Analytics interface, and apply to a wide range of companies across numerous industries.

Audience data

Audience reporting is one of the most fundamental types of research for marketing departments to undertake. Search engine journal notes that by grouping users into audience segments based on specific metrics, companies can better shape and direct their campaigns. Groupings can be based on what people do (such as using a certain feature of the website) or when they do these things (for example, activity on the site within a specific time period).

The metrics compiled in these reports can be tracked over time, meaning it's possible to determine the long-term relationship between a company and its site visitors. Lifetime value calculations fall into this category, allowing marketers to see how much a single member of the audience has spent with the company through all their sessions on the site.

User behavior

Rather than breaking the audience into slices and analyzing from that high-level view, web analytics tools' user behavior reports are based on the experience of using the website, and the types of activities that happen during a user's session. In Google analytics, that may mean creating flow charts of which actions lead to one another, determining how visitors are making use of search features and calculating the technical performance of the page.

Search engine journal adds that this category of analytics allows stakeholders to see heat maps to demonstrate which parts of their website are the best performing, as well as get a more detailed look at how their visitors are interacting with the company. Performance reports on the number of mobile vs. desktop visitors falls under this heading, as does tracking load times across different browsers or devices.

Conversion rates

Especially important in e-commerce, where the website is the main point of interaction between buyer and seller, conversion reports can determine how users are finding their way to companies' intended destinations and what factors could be stopping them. For example, if there is a page where users typically give up on a purchase and abandon their carts, this kind of report will highlight the trouble spot and empower web developers to correct the problem.

Since conversion reports can track user activity back and show how customers are finding the site, this branch of web analytics is essential for companies seeking to figure out which of their campaigns produce the most value. Identifying the most valuable combination of channels in the path to purchase is useful for marketers reporting to the C-suite and planning their next campaigns.

Real-time reporting

While a great deal of web analytics work is strategic in nature, there are tactical real-time insights available that help companies make adjustments in the moment. Google Analytics notes that it's possible to process audience data in real time, allowing companies to perform fast-paced remarketing operations.

Search engine journal points out a few practical uses for real-time metrics. For instance, when a new and potentially valuable piece of multimedia content has gone live, stakeholders can see whether people are viewing it, what path they followed to find the site, and whether they are then converting and performing the intended next steps. If the reality doesn't match the company's hopes, a quick adjustment to the social media posts and other promotional content could improve performance.

Acquisition patterns

When studying user acquisition information, companies compare the way users' paths to the web page affect their behavior once they are on the site. Google's console for acquisition metrics allows organizations to determine whether their social campaigns, search engine optimization efforts, and paid advertising reflect in the behavior of site visitors.

By combining a few different types of web analytics, acquisition reports allow digital marketers to adjust strategies. Rather than forcing their inbound marketing campaigns and user behavior calculations to exist in silos, stakeholders can draw explicit connections between the ways they generate leads and the resulting conversions. This correlation is important for allocating budget to impactful projects.

Advertising results

While today's online companies are intent on attracting customers through methods such as social media outreach and inbound marketing, traditional digital ads still have a part to play in outreach campaigns. Of course, as with any source of website traffic, the impact of ads should be subject to intensive study and analysis to determine which campaigns are driving results and which can be rethought or discontinued.

In the case of Google analytics, the integration between ads and analytics is especially close because Google is one of the primary sellers of online advertising. Companies that invest in Google ads campaigns can track the resulting customers who have clicked through, determining the types of behaviors they exhibit on the web page. If these users have a high bounce rate, something has gone wrong with the campaign, but if their conversion rate is abnormally high, the business can boost its investment.

User Flow Visualization

There are a few different views that show the paths visitors take on a website. Each of these reports has its own strategic use, as companies strive to create the clearest possible paths from an initial visit to a desired outcome, such as a purchase or subscription. When seeing how actual users are moving across pages, web developers may realize where their pages have room for impactful improvement.

Google analytics offers the ability to track individual users' courses through the site, providing an extremely granular view that can inspire changes to website features. Stakeholders can also choose to create content groupings, which show how audiences are interacting with specific sections of the website. Each of these breakdown styles can inspire adjustments to the site that will promote greater numbers to reach intended outcomes.

2.14.3.3 Google Analytics - A Widely used Web Analytics Tool - Standard Set of Reports

The tool allows you to **track websites, blogs and social networks.** In addition, it puts at your disposal predetermined and customizable reports.

- Google analytics comes with **78 standard reports** that each fall in to five categories.
- The **five categories** of standard reports include : real time reports, audience reports, traffic source reports, content reports, and conversion reports.
- **Real-time reports :** These four reports give you information about what is happening on your website RIGHT NOW. Where people are located in the world who are on your website right now, how they got to your website, etc. This is a fun report to look at, especially if your website has lots of traffic.
- **Audience reports :** These 13 reports give you the details about who is coming to your website (*that is, the your audience*), what languages they speak, where are they located in the world, are they NEW visitors to your site, or are they return visitors, what devices they use to visit (iPhone, iPad), etc.
- **Traffic sources report :** This is the biggest category of standard reports, and it covers all the different paths that people took to find your website. This is the category that tells you the following: did people type your URL directly into their browser, what words did they type into Google/Yahoo/Bing in order to find your website ? What landing page appeared after they typed the phrases into a search engine ? Are people visiting your website because of your monthly E-Newsletter or your daily Tweets and Facebook updates ? You can learn it all. (Some of the reports here can only be activated by linking your Google Analytics account to a Google Webmaster Tools account.)
- **Content report :** With 21 reports, this is the second biggest category, and it tells you all about your website's content : what content is most popular ? Where do people spend the most time ? Which is the biggest entrance page, and which is the biggest exit page ? What links are people clicking on ? What PDFs are people downloading ?
- **Conversion reports :** The 15 reports in this category are both very useful, and the hardest to activate and use of all the standard reports. Each one typically requires consulting to activate - but once activated, these reports show you the "outcomes" of your website. Basically, the idea is that you should ask yourself the following question : why does my website exist ? Some of the answers to that question can be addressed in this section. For instance, if you sell a wide variety of products on

your website, then you should enable e-commerce tracking so that you can see where your best customers are coming from, and how they behave. If you need people to download whitepapers or eBooks, you should have conversion tracking installed by setting up Goals to make sure that you can see which visits resulted in a PDF download. And for some websites that have a structured "conversion path," you'll want to define a sales funnel that you can measure against.

2.14.4 'Quality Content' Concept

This term gets banded around so much that it's lost all meaning over the years. Of course, it doesn't help that 'quality' is a highly subjective thing, but your content is going to fall short if it doesn't meet the following criteria,

- **Valuable :** Offers something of value that your audience can't get elsewhere.
- **Accessible :** Content that's there when users need it and provides a positive user experience every step of the way.
- **Compelling :** Titles that compel users to click through and content that compels them to take action.

Before you create any kind of content, ask yourself - what problem is this going to solve for our audience ? Also, bear in mind there's little value in repetition. Instead, create content that tells search engines and users that you have something unique to offer. You're not like the other brands in your industry; you're better.

Accessible content needs to be two things. First, content needs to be in the right place when users are looking for it - ranking for a search term in Google, visible to the right audience in Facebook, part of a targeted remarketing campaign, etc. If your content isn't within reach when your audience needs it, it's worthless to them.

Web analytics is often the go-to resource for measuring content quality and understanding user behavior. But it's also a great resource for evaluating user needs, including the content that users find desirable and valuable. It fact, users tell us what they like every time they visit our website. We just have to stop and listen.

Here are a number of ways web analytics can help you discover what content topics your readers care about.

1. Blog categories and tags

An effective website navigation scheme not only improves usability but also supports content analysis. Using a topic-driven navigation, you can gain insight into which topics are most popular and, thus, which valuable topics you can revisit or expand upon for new blog posts.

A common blog feature is to list links to blog post categories and tags. When users browse these category pages, they're informing us about topics that interest them (and consequently, when they ignore certain links, ones that don't).

In Google analytics or other analytics tools, evaluate page views for your category (topic) pages.

As with categories, also evaluate the popularity of blog tags. While category pages give you an idea of the broad topics users find interesting, tag pages give you insight into more specific areas of interest.

Category and tag page metrics are general and shouldn't be taken at face value. Instead, use them to start your research. In the example above, strategy could be toward the bottom of the list because of the placement of the link on the page or because the keyword is inappropriate. If posts are more popular than the category they fall under, you may be using the wrong keyword to describe the topic.

After using this method to uncover potentially valuable topics, you can dig deeper with some of the following methods to help validate your insights. You can also then use traditional metrics such as bounce rate and time on page to help answer more specific questions, such as "Do people who view this topic page find what they are looking for ?"

2. Top-performing blog posts

Viewing top-performing pages gives you insight into more specific topics of interest on a per-post basis.

Sort blog pages by page view. Take note of the topics, including related category and tag pages, and pay particular attention to combinations of topics. You may, for example, find that posts covering both *communication* and *social media* are very popular.

Open site explorer by SEOmoz can help you dig deeper through evaluating top-performing pages by search engine authority, inbound links, and social media shares (with a Pro Subscription). Discover which posts and topics rank well in search results (i.e., those that are most referenced by external sites) and which are most share-worthy.

3. Search (internal and external)

When visitors use the search box on your website, they're again telling you what topics interest them most right now.

However, visitors don't always know the right search terms to use, so don't take search terms completely at face value. Dig deeper.

In your analytics tool, view individual internal site search terms to evaluate whether the search results match the inquiry (i.e., does the user find what she or he is looking

for ?). This can be measured, in part, by whether visitors click on the search results or choose to leave your site and abandon their search.

The same goes for external organic search from Google and other search engines. What are the top topics (keywords) driving traffic to your site ? This approach is focused on prospective readers. View search terms to learn what topics are attracting new visitors. Where are the opportunities for increasing your readership with new, relevant content ?

For this exercise, focus less on the exact search terms and more on the topics they relate to. What topics are people searching for ?

4. Evaluate referral traffic by topic

In my experience, referral traffic is a metric that is often overlooked. Yet, it provides valuable, dual insight, revealing details on who your readers are as well as on which topics they care about most.

Who is referring traffic to your website ? Look past the common referring websites (Facebook, Twitter, affiliates, etc.) and note sources that are new or different. Who are they ? What is their business ? What topics do they cover ? Go for the long tail. Dig deep to discover sites that don't refer to you often but have shown an interest in your content.

Which blog topics are referrers linking to ? Once you have a sense of who your referring readers are, discover what content they are looking for. View top blog posts by page views and referring websites to find out which of your blog topics your readers are commonly linking to. What are new referrers linking to? More importantly, what referring links do people click on most frequently ?

5. Outgoing traffic

Track outgoing links to learn what external resources and related topics interest your readers. By default, most web analytics applications, such as Google analytics, don't track outbound links, but you can tag them manually or use third-party plug-ins (such as Google analytics for WordPress that will auto-tag outbound links so that you can track them.

Again, pay attention to topics. What external websites do your readers find desirable ? Do those websites cover topics that would be valuable and appropriate for your readers ?

6. Comment rate

User engagement is an important benchmark for evaluating content quality and one of the most effective ways to measure engagement with blog content is comment rate. Find out what percentage of readers comment on your blog posts. But don't stop there. Look at the comment rate by topic. What topics resonate most with your readers ?

Coupled with some of the aforementioned methods of evaluating popular posts, comment rate by topic can help you understand what topics your readers are most excited about.

7. Sharing rate

While you are evaluating comment rate, also take note of what percentage of readers share your posts via Twitter, Facebook, LinkedIn, Google+, email, and other social platforms. What blog categories and tags do users share the most ? Sharing rate can speak to the caliber of your topic. For example, a big retweet number gives your topic a high qualitative value, not just bragging rights.

2.14.5 Navigation Reports

As a important part web analytics below questions are the common questions one may get around navigation and they need to be answered,

1. Which individual navigation links are clicked the most ?
2. Which navigation areas are clicked the most ? This is usually related to the main section area, not individual links.
3. From which pages are visitors using each navigation link ?
4. For what percent of website visits is navigation used ?
5. In what order do website visitors use navigation links ?
6. Where do people go after viewing a blog post ?
7. Which navigation links lead to key website success milestones being accomplished ?
8. How many people use my contact form after viewing a product page ?
9. Is my website's navigation correctly structured ?

These questions can be answered by generating rich set of navigation reports that are generated by web analytics. Various navigation reports are discussed below,

1. Navigation summary

The navigation summary report allows to select a particular page to then see how people find that page and where people travel to next on your website. It gives a 'before' and 'after' snapshot of one's content. It's not complicated which makes it easy to interpret and one can quickly switch between the pages one wants to analyze.

2. Users flow

The users flow report takes navigation path analysis to the next level. It's a visualization of the different ways people travel through web content.

The visualization is made up of 'nodes' - the pages and other elements in the report and 'connections' - the paths linking the nodes together.

The nice thing about the way your data is visualized is that everything is related to everything else in the report. In other words, the sizes of the nodes and connections show you how many people view particular pages and take particular paths.

3. Page analytics

There was a report that was previously available inside the Google analytics interface called In-Page analytics. It over laid data from your reports directly onto your website which allowed you to browse your website and see details about how people engaged with your content.

4. User explorer

The user explorer report provides a user-by-user view of the navigation paths people take. While the other reports we've looked at all provide aggregate data, this report allows you to see how individual users experience your website. The report allows us to view the different types of actions each user takes on our website. We can see the pages they view, conversions, transactions and even events we've implemented to track custom actions, like watching videos and scrolling.

5. Goal flow

The goal flow report is very similar to the users flow report you wouldn't be the first person to get them confused. However, the goal flow report only shows you the navigation paths for the goals you've configured inside Google analytics. Once you've configured your goal you'll be able to see how people travel through each step you've defined in your goal funnel. This shows you where people drop-off at particular points, skip steps and loop-back to previous steps in the path.

The path to conversion is likely to be even more complicated than what you'll see in the report. Today, a linear path to conversion is the exception rather than the rule. So I recommend focusing on areas where people are dropping-off to identify opportunities to optimize those pages. For example, if lots of people are leaving a step to view your privacy policy, then it's probably a case of including some of those details within the page to reduce friction.

6. Segments

Applying segments to your reports can provide powerful insights into how people navigate through your website. You can start by creating basic segments, like a segment

for a particular landing page. Since this segment is applied to your reports, you can use it to understand the most important marketing channels driving people to the page and whether the page is meeting your objectives in terms of engagement and conversions. You can even apply the segment to the all pages report and the navigation summary. Once you have the basics down you can graduate to more advanced segments. You can create a sequence-based segment that requires people to view a series of pages in a particular order to be included in the segment.

You will be able to analyze how those people find your website and engage with your other content. And you could even create two sequence-based segments for different navigation paths to compare their performance.

7. Path analysis

If one is using a Google analytics App + Web property, then the path analysis report lets you understand how people travel through your website. You can see the pages people view and other actions people take as they engage with your content.

One can apply segments and filters to the report to narrow the focus of your analysis. For example, you can view paths for your paid traffic or for people who have already converted. The report also lets you select a node and view individual users who took the path you have selected. This lets you view a filtered version of the user explorer report.

Review Questions

1. *Define web analytics. What are objectives of web analytics ?* **(Refer section 2.1)**
2. *Classify web analytics.* **(Refer section 2.4)**
3. *Discuss web analytics process.* **(Refer section 2.5)**
4. *What are various data sources for web analytics.* **(Refer section 2.7)**
5. *Explain web analytics key metrics.* **(Refer section 2.9)**

❑❑❑