

# Introduction to Web Analytics

Dr. Barry Shepherd  
Institute of Systems Science  
National University of Singapore  
Email: barryshepherd@nus.edu.sg

© 2018 NUS. The contents contained in this document may not be reproduced in any form or by any means, without the written permission of ISS, NUS, other than for the purpose for which it has been supplied.

## What is Web Analytics?

- Web Analytics
  - The measurement, collection, analysis and reporting of web data for purposes of understanding and optimizing web usage (Wikipedia)
  - The study of visitor, traffic and navigation patterns to discover the success of a particular website
  - GOALS ~ increase website traffic, improve the user experience (UX), ensure website meets its business goals

What most people  
think of as Web  
Analytics

Web Analytics Tools: Google Analytics, WebTrends, ....

# On-site vs Off-Site Web Analytics

---

- On-Site Web Analytics
  - Analysis of visitors behavior once they are at your website – analysis of webserver logs
  - Example tools include Google Analytics, WebTrends etc
- Off-Site Web Analytics
  - Analysis of your *potential* website audience and how your website compares to others
  - Main Tools and Methods:
    - **Panels:** Companies such as comScore and Nielsen recruit panels of users who install monitoring software to measure their web activity. Panelist demographics are known but they are mostly home users.
    - **ISP Data:** Companies such as Hitwise aggregate anonymous data provided by ISPs – the ISP web log data. Covers all visitor types: home, work, mobile, educational, public access etc. This data is anonymous hence demographics are not available.

## The Bigger Picture: Web Mining

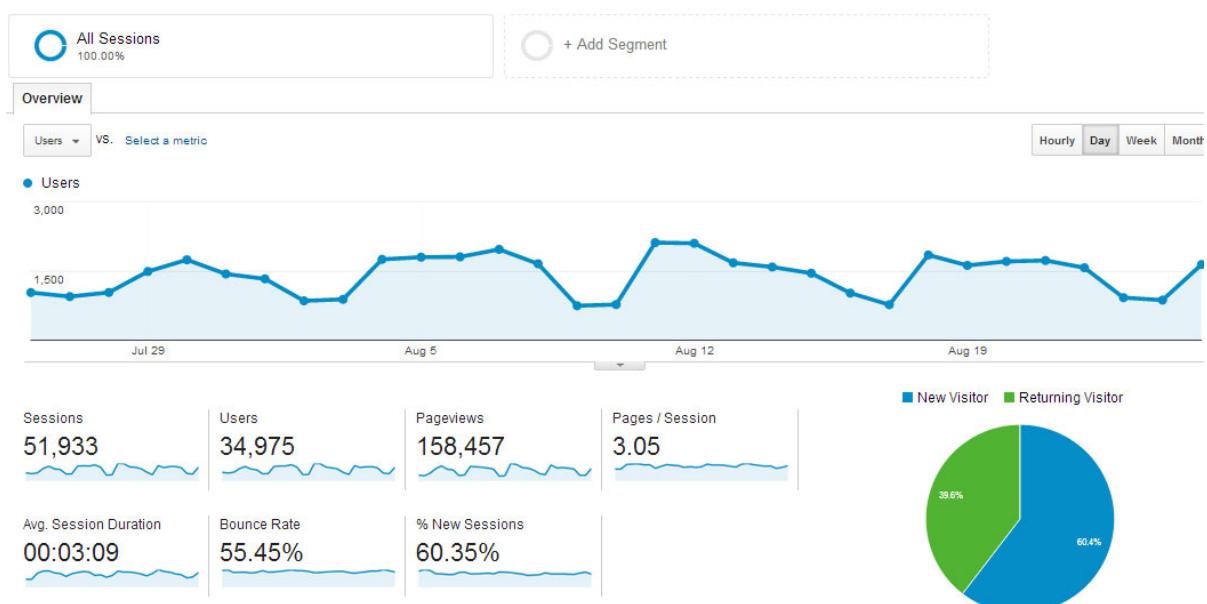
---

- Web Usage Mining
  - Understanding and modeling how users interact with the web, determining what their interests are and what they are looking for
  - Focus on: website optimisation, improved search, behavioral modeling, content personalisation, making recommendations, ....
- Web Content Mining
  - What can we learn from the content that is on the web?
  - Focus on web crawling & sampling, text mining
- Web Structure Mining
  - What can we learn about the interconnection between websites and between entities on the web?
  - Focus on graph & network mining, e.g. social network mining

# Agenda for the Elective

- Day1: Web Site Analytics (Barry)
  - Website Analytics methods and tools
  - Google Analytics Workshop
- Day2: Web Usage Mining (Barry)
  - Mining Web Server Logs (page-view and search query logs)
  - Mining Online User Behaviors (Search relevance prediction, Ad click prediction, Content Recommendation)
- Day3 & 4 & 5: Web Structure Mining (Dr Li Xiaoli, I2R)
  - Graph & Network Mining concepts
  - Link Analysis, Social Network Mining
  - Day5 (pm) will be presentations for Xiaoli's assignment project

## On-Site Web Analytics: How is it done?



ISS Website Users (shown using Google Analytics)

# On-Site Web Analytics: How is it done?

- Log File Analytics
  - Whenever a webserver delivers a webpage to your browser this is recorded into a web server log
    - Typical fields = Date-time, IP-address of client, Browser Type, URL requested.....
  - Every served page (page-view) is logged – log size can hence be huge
  - Usually performed in-house – you buy and maintain the analytics software
- Page Tagging
  - You insert some small tracking code (*usually JavaScript or Flash*) onto every important web page in your website
  - When the users browser displays the webpage it executes the tracking code which send a message back to the Web Analytics Server
    - Typical fields = Page-ID, Date-time, IP address of client, Browser Type, Screen Resolution ....
  - Can be performed in-house, but usually provided as a third-party service

## Advantages of log file analysis

- Pros
  - No changes to website required, the raw data is already available
  - Easy to switch web analytics tool, the data is on the company's own servers, in a standard, not proprietary, format
  - The web server records every transaction it makes, e.g. serving PDF docs - does not rely on the visitors browser cooperating - does not require JavaScript to be enabled
  - Search engine spiders get logged, can be useful information for SEO
  - Doesn't slow page loading, no external server calls (e.g. DNS lookups)
- Cons
  - Search engine spiders and robots get logged – they must be detected in the logs and removed for normal analysis
  - Web caching isn't logged; if a person revisits a page that can be retrieved from the browser cache, then the web server gets no request. The person's path through the site is hence lost. Caching can account for up to one-third of all page-views, omission can seriously skew many site metrics.
  - Can be more expensive to set up Web Log Analytics tools: can require specialist knowledge of web logs
  - Web logs can periodically be purged

# Advantages of Page Tagging

- Pros
  - Available to companies who don't have access to their own web servers
  - Also records visits to cached pages
  - Can record events which do not involve a request to the web server, e.g. interactions within Flash movies, partial form completion, mouse events, visitor screen sizes, even (potentially) price of goods the visitor purchased.
  - The page tagging service manages the process of assigning cookies to visitors; with log file analysis the server has to be configured to do this
- Cons
  - Require JavaScript to be enabled in the visitor's browser
  - Requires an additional DNS look-up by the user's computer to determine the IP address of the collection server. Delay or failure in completing a DNS look-up may result in data not being collected

# Web Analytics Tools Survey

## Self-hosted software [\[ edit \]](#)

### Free / Open source (FLOSS) [\[ edit \]](#)

This is a comparison table of web analytics software released under a [free software license](#).

Name	Platform	Supported databases	Tracking Method
Analog	C	Logfile-based	Web log files
AWStats	Perl	Logfile-based	Web log files
Open Web Analytics	PHP	MySQL	JavaScript or PHP pagetag
Piwik	PHP	MySQL	JavaScript or PHP pagetag or Web log files
W3Perl	Perl	Logfile-based	Web log files
Webscraper	C	Logfile-based	Web log files

From: [http://en.wikipedia.org/wiki/List\\_of\\_web\\_analytics\\_software](http://en.wikipedia.org/wiki/List_of_web_analytics_software)

# Web Analytics Tools Survey

## Proprietary [ edit ]

This is a comparison table of web analytics proprietary software.

Name	Company	Platform	Tracking Method	
Mint	Mint	PHP	Cookies via JavaScript	
Sawmill	Flowerfire Inc	Windows/Linux/BSD/POSIX	Web log files	Start
Splunk	Splunk Inc.	Windows/Linux/BSD/Solaris	Web log files	
Tealeaf CX	IBM	Windows/Linux	Network traffic monitor	
Unica NetInsight	IBM	Windows/Linux/Solaris	Web log files & Cookies (with or without JavaScript)	
Urchin	Google	Windows/Linux/BSD	Cookies & Logs	

# Web Analytics Tools Survey

## Hosted / Software as a service [ edit ]

This is a comparison table of hosted web analytics software as a service.

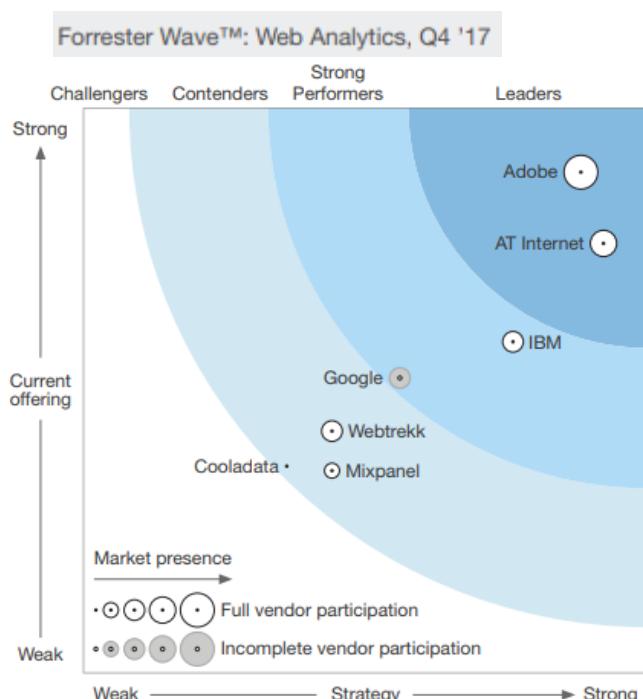
Name	Company	Tracking Method	Latest stable release	Price in USD
Adobe Analytics	Adobe Systems	Cookies via JavaScript	15	Negotiable
Bing Webmaster Tools	Microsoft	Cookies via JavaScript	N/A	Free
Analyzer	AT Internet	Cookies via JavaScript	N/A	Negotiable
Bango Mobile Web Analytics	Bango plc	Mobile ID and cookies	4.0	From \$49/month
Data Workbench	Adobe Systems	Cookies via JavaScript	6.51	Negotiable
Google Analytics	Google	Cookies via JavaScript	N/A	Free (Standard), \$150,000 Annual (Premium)
Kissmetrics	Kissmetrics	Mobile libraries; Cookies via JavaScript	N/A	\$159-\$599 Monthly Plans
LiveChat	Livechat Software	Cookies via JavaScript	N/A	from \$36/month
Logentries	Logentries	Web log files; Mobile libraries; JavaScript	Continuous	Free
Mapmyuser.com	Mapmyuser, LLC	Cookies via JavaScript	N/A	Free
Mixpanel	Mixpanel	Cookies via JavaScript	N/A	Free (Start)/Monthly Plans
Nullstack Analytics	Nullstack Labs	Javascript, Cookies via Server	14	Free / Monthly Plans
Piwik Cloud	Piwik PRO	User ID, fingerprint and Cookies	2.14.0	\$29+/month
Quantcast	Quantcast Corporation	Cookies via JavaScript	N/A	Free
SimilarWeb	SimilarWeb Ltd.	Cookies via JavaScript	N/A	Free / Monthly plans (from \$199)
StatCounter	StatCounter	Cookies via JavaScript	N/A	Free - \$5/month ... \$119/month
Webtrekk Q3	Webtrekk	Cookies via JavaScript	N/A	From \$202/month
Webtrends	Webtrends	Cookies via JavaScript	N/A	N/A
Woopra	iFusion Labs LLC	Cookies via JavaScript	1.2	Free - \$499.95+/month
Yahoo! Web Analytics	Yahoo!	Cookies via JavaScript	Not available anymore	Free

# Enterprise Web Analytics Tools

Vendor	Product evaluated
Adobe	Adobe Analytics
AT Internet	Analytics Suite
Cooladata	Cooladata
Google	Google Analytics 360 { Google Analytics for premium users
IBM	IBM Watson Customer Experience Analytics
Mixpanel	Mixpanel
Webtrekk	Webtrekk Analytics

## Vendor Inclusion criteria

1. The web analytics solution has sparked client inquiries and/or the vendor has web analytics technologies that put it on Forrester's radar. Forrester clients often discuss the vendor's web analytics products through inquiries; alternatively, the vendor may, in Forrester's judgment, warrant inclusion or exclusion in this evaluation because of web analytics technology trends or market presence.
2. The vendor provides a dedicated web analytics software solution. In other words, the vendor offers a software solution that has been specifically built to deliver web analytics functionality. This functionality is core to the solution and is not simply an add-on to other analytical functionality, such as interaction analytics from session-based replay data, customer analytics, or insights-driven optimization (e.g., A/B testing or online testing). The web analytics solution is offered to and used as a standalone software tool by enterprise customers.
3. The software solution has a complete set of functionality for advanced enterprise web analytics needs. Evaluated vendors provide a complete set of advanced web analytics functionality, including in-browser instrumentation for collecting visitors' contextual and behavioral data, the provision and management of an extensive set of out-of-the-box metrics and dimensions, and customizable self-service dashboards.
4. The vendor has an enterprise user base. Ten or more enterprises are users of the vendor's web analytics solution. Forrester defines enterprise-sized customers as firms with at least \$1 billion in annual revenue.



<https://www.adobe.com/content/dam/acom/en/experience-cloud/offer-assets/54658.en.exp.report.forrester-wave-web-analytics-q4-2017.pdf>



INSTITUTE OF SYSTEMS SCIENCE

ATA/BA-DAT/IntroToWebAnalytics/V6

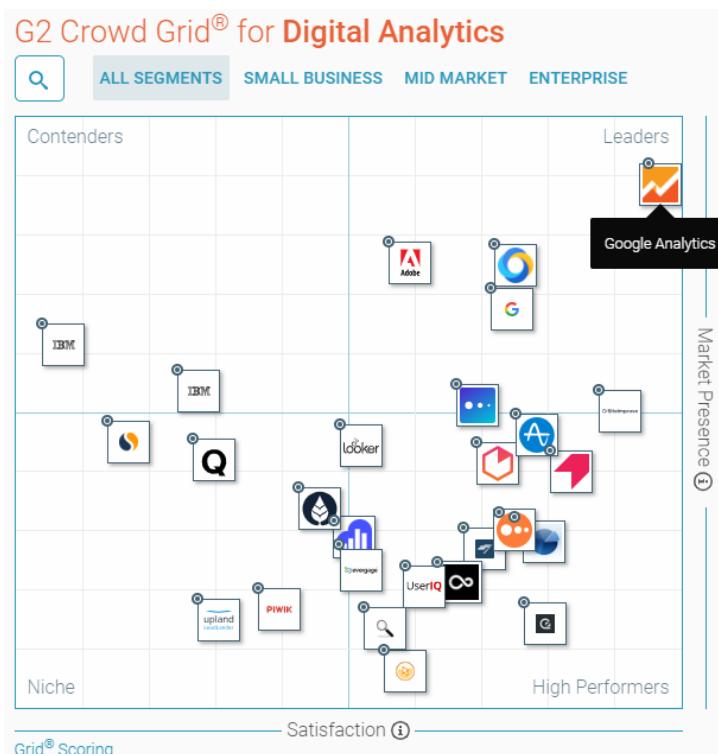
© 2018 NUS. All rights reserved.

Page 13

# Digital Analytics Tools

"Digital analytics software tracks website visitors and measures web traffic."

<https://www.g2crowd.com/categories/digital-analytics>



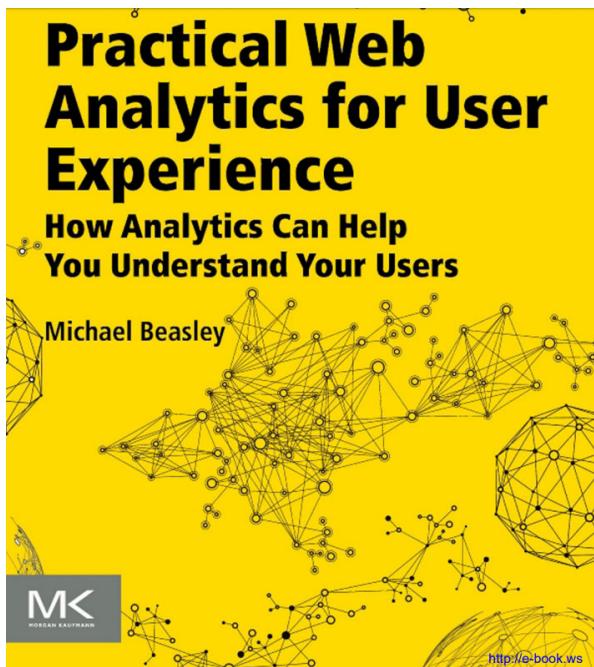
INSTITUTE OF SYSTEMS SCIENCE

ATA/BA-DAT/IntroToWebAnalytics/V6

© 2018 NUS. All rights reserved.

Page 14

# Useful Reference



<http://freepdf-books.com/download/?file=3724>

Good introduction to the main concepts

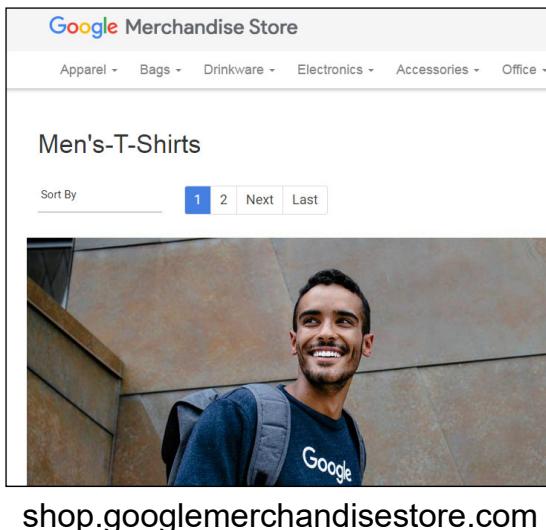
Uses Google Analytics – but some GA functionality has changed since publication

## On-Site Web Analytics: Main Objectives

- How popular is your website – how much traffic?
  - Basic site usage & traffic metrics. **GOAL = Monitor website health**
- What types of people visit your website?
  - Demographics & interests
  - **GOAL = Reach the right people / tailor your website to your audience**
- How did they get to (find) your website ?
  - Search engine, referral, bookmark. **GOAL = Increase traffic / monitor campaigns**
- What do they do on your site?
  - How long do they stay? Most/least popular pages? Is it easy to use?...
  - **GOAL = Ensure the website meets the customer/visitor needs**
- Does your website achieve its business goal(s)?
  - Measure Conversions, downloads, sign-ups etc.
  - **GOAL = Ensure the website meets your needs**

# Getting Started with Google Analytics

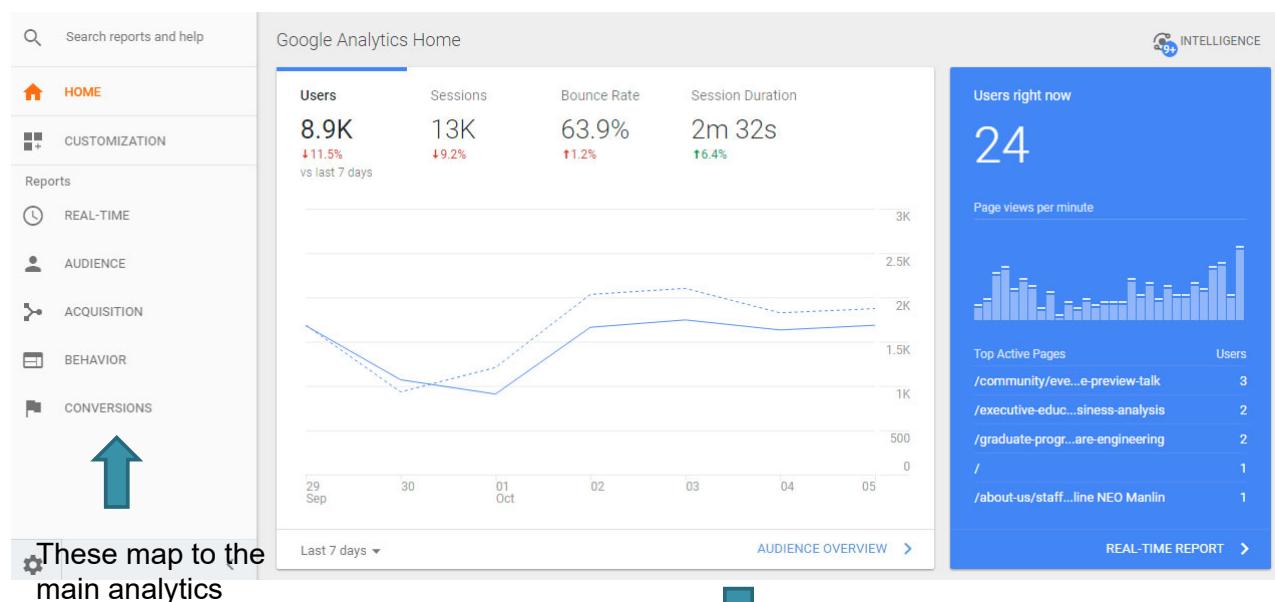
- Create your own Google Analytics (GA) demo account
  - This will give you access via GA to Google's e-commerce store traffic data
  - To set up the account: <https://analytics.google.com/analytics/web/demoAccount>



The GA demo account allows you to explore traffic data from a real e-retail site

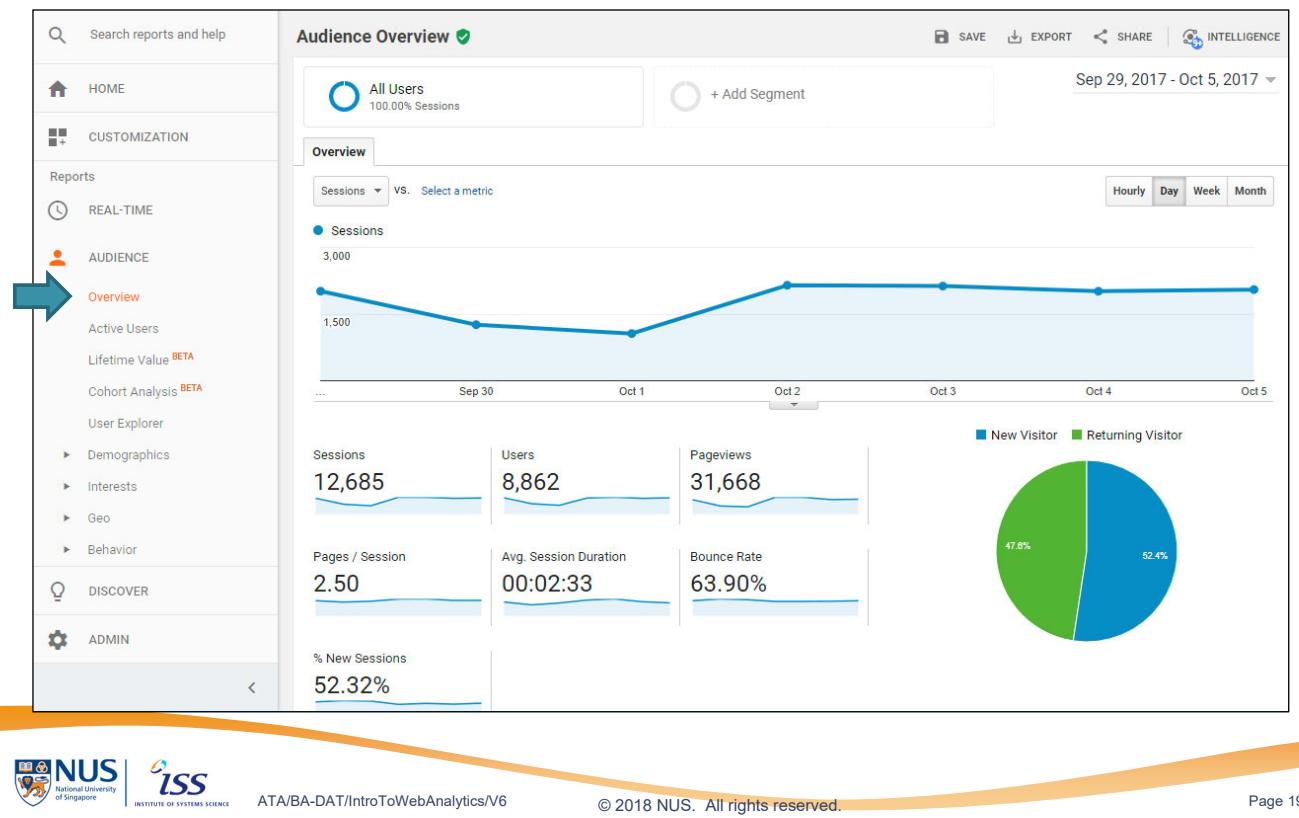
# Getting Started with Google Analytics

- Once you have the account created you should see something like this.....



Lots more info when scrolling down

# Basic Website Metrics (Audience)



# Basic Website Metrics (Audience)

## Page View

A request for a file, or an event such as a click, that is defined as a page in the setup of the web analytics tool. An execution of the page tagging script

- Number of Page Views
- Page View Duration

Difference between time of the request for a page and the time of the next request. Not recorded if there is no next request

- Number of Hits

A request for any file from the web server. Available only in log analysis. Overestimates site popularity since webpage usually comprise many files

Note: All metrics require you to define a date range

## Visit (Session)

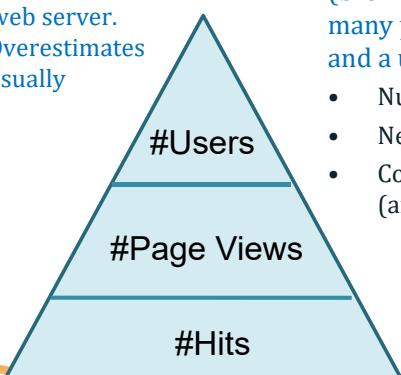
A series of page views from the same user during the same web “session”

- Number of Visits
- Average Visit Duration
- Average Page Depth (pages per visit)
- Bounce Rate (%visits to single page only)

## Visitor (User)

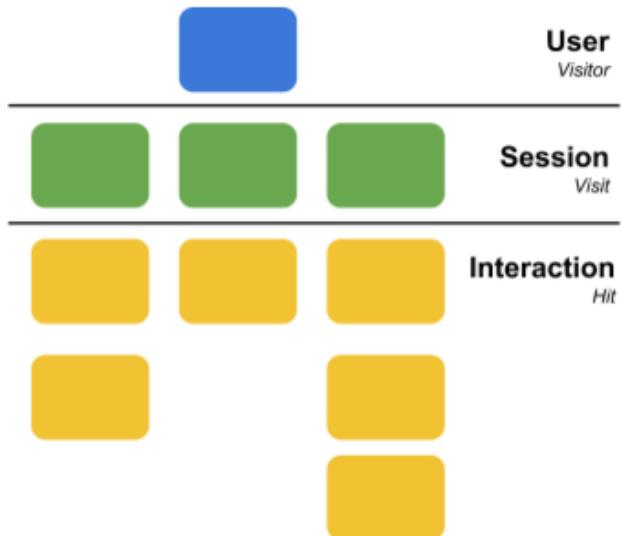
Usually identified by cookies OR the unique combination of IP address and User Agent (browser). Not the same as unique person – many people can share the same computer and a user may have many machines.

- Number of Visitors
- New vs Returning Visitors (%new visits)
- Conversions & Conversion rate (any action on the site you define as a goal)



# The GA Data Model

- **User (visitor)**—the client that visits the site, such as the browser or mobile phone operated by a person.
- **Session (visit)**—the period of time during which the visitor is active on the site.
- **Interaction (hit)**—the individual activities that send a GIF request (hit) to the Analytics servers. These are typically characterized by a pageview, but can include:
  - a pageview
  - an event (e.g. click on a movie button)
  - a transaction
  - a social interaction



## Basic Website Metrics: Time Dimension

- A snap-shot of the metrics on any given day is of limited use. More important is to look at the trends over the recent past.
  - Day-of-week, week-of-month, month-of-year, season – these may all be expected to impact site usage in (potentially) predictable ways
  - BUT look for unexpected changes in any of the metrics, this may indicate something on the website or elsewhere has changed and perhaps needs fixing, e.g:

Symptom	Possible Cause
<b>decrease</b> in visit duration, conversions, ..	Removal of popular pages or products
<b>increase</b> in visit duration, conversions, ..	Addition of new pages or products
<b>increase</b> in pages viewed, visit duration, ..	Positive changes in website structure
<b>increase</b> in repeat page views, looping (navigational difficulty)	Negative changes in website structure
<b>increase</b> in bounce rate	Negative changes to homepage, poorer targeting /marketing
<b>decrease in</b> bounce rate	Improvements to homepage, better targeting/marketing
<b>increase</b> in #visitors, %new visitors	Better marketing

# Users (Visitors)

- Audience Overview can also show #users on each day
- But how many are repeat visitors ~ regular users of the website?



## Active Users

- Users who have visited the website at least once in the recent past
- E.g. if you visited in the last 14 days then you are a 14-day active user



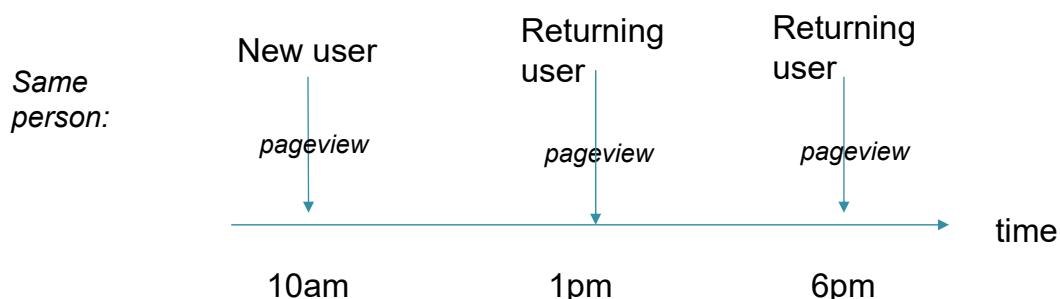
# Counting Users: issues (1)

- Metrics such as page-views are additive, you can add up the days
- Unique Users cannot be easily added over days, weeks, months
- E.g. A hotel has two rooms (Room A and Room B). How many unique visitors were there over the 3 days period?

	Day 1	Day 2	Day 3	Total
Room A	John	John	Mark	2 Unique Users
Room B	Mark	Jane	Jane	2 Unique Users
Total	2	2	2	?

# Counting Users: issues (2)

- New visitors + Repeat visitors do not equal total visitors
- If a user makes their first visit to the website then they are (correctly) a new visitor.
- But if they revisit the site again that day or week – then they appear to be a returning user on the same day or week



# Counting Users in GA

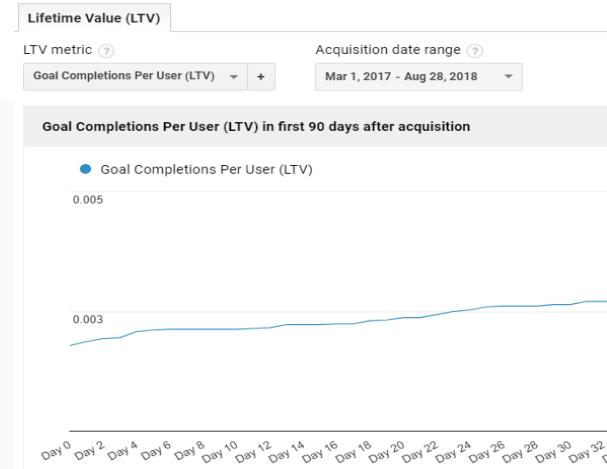
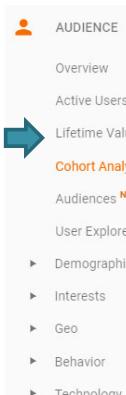
- GA will automatically select which of the following two calculations to use depending on the report being viewed.
- Pre-Calculated Data
  - A set of unsampled, pre-aggregated data tables, updated on a daily basis.
  - Fast - Enables you to quickly select reports for any date range
  - Used exclusively in reports when the only dimension is a time frame
  - Derived by counting sessions
- On-The-Fly Calculations
  - Takes much longer and is used when you choose to view users over any non-date dimension
  - Uses cookies to identify unique users  
(warning: counting unique cookies overestimates unique users potentially by up to 25% or more)

## Lifetime Value Report

- Plot LTV for users in the 90 days after their acquisition
- LTV defined by a metric

### Available metrics

- Appviews Per User (LTV)
- Goal Completions Per User (LTV)
- Pageviews Per User (LTV)
- Revenue Per User (LTV)
- Session Duration Per User (LTV)
- Sessions Per User (LTV)
- Transactions Per User (LTV)



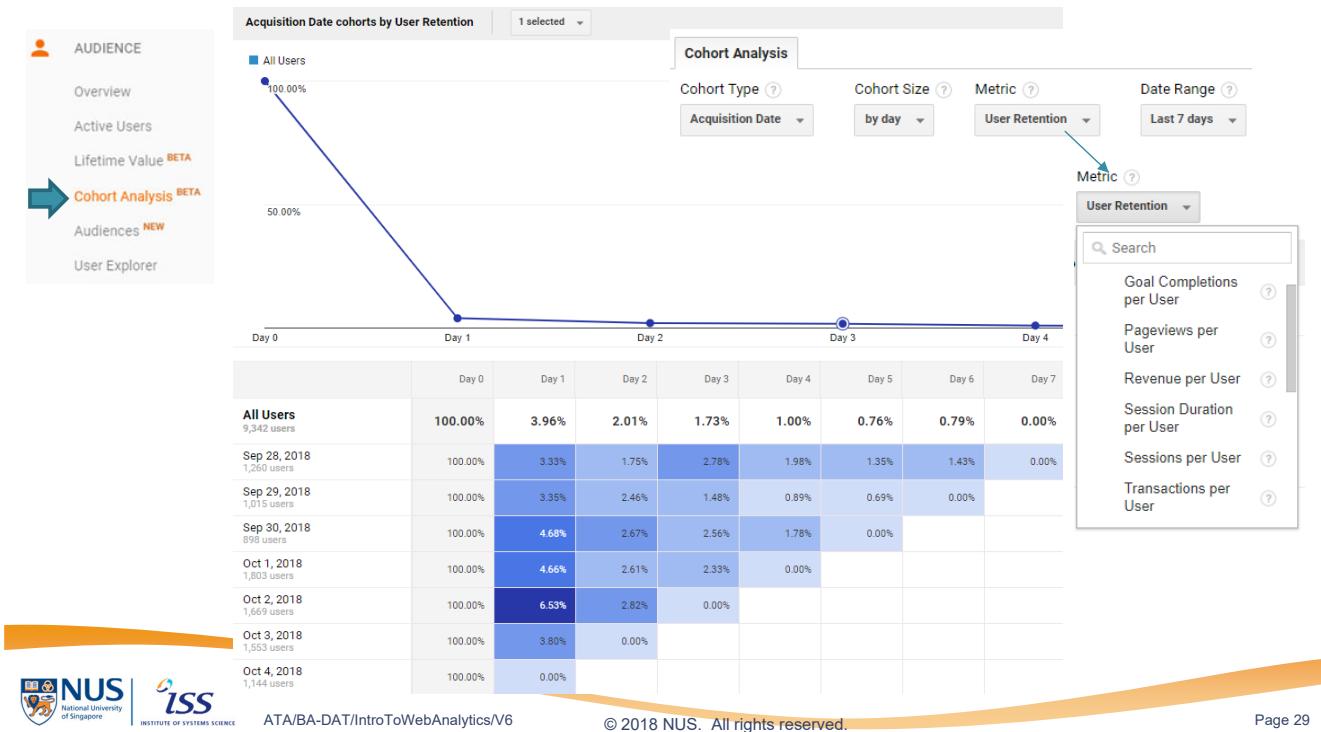
### Example Scenarios

- Compare LTV for users acquired through email or paid search
- Compare LTV of users acquired through different methods. E.g. organic search vs social vs email

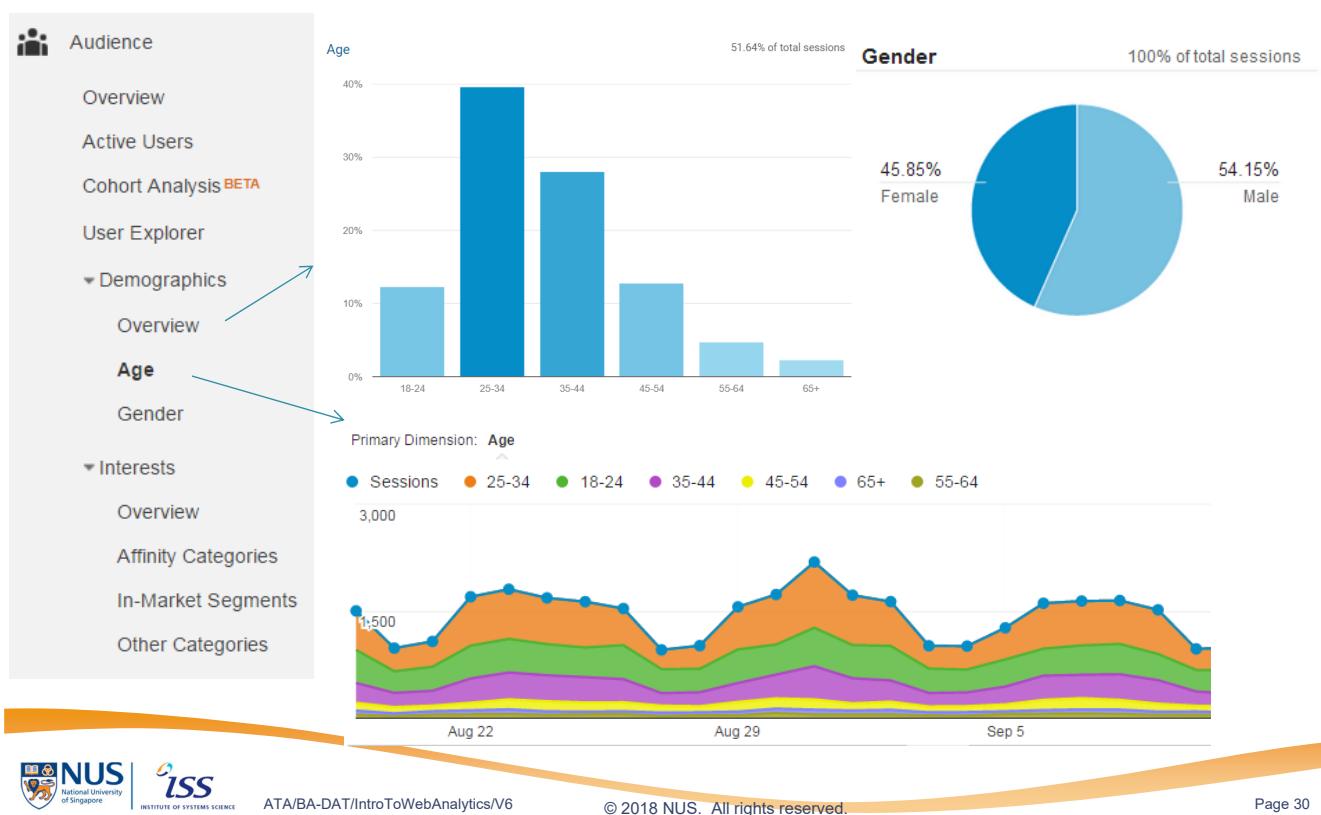
Acquisition Channel	Users	Goal Completions Per User (LTV)	Goal Completions (LTV)
	62,030 % of Total: 100.00% (62,030)	<0.01 % of Total: 100.00% (~0.01)	210 % of Total: 100.00% (210)
1. Organic Search	32,151 (52.32%)	<0.01 (50.97%)	56 (26.67%)
2. Direct	13,785 (22.43%)	<0.01 (74.30%)	35 (16.67%)
3. Display	7,787 (12.67%)	<0.01 (142.81%)	38 (18.10%)
4. Referral	4,141 (6.74%)	<0.01 (113.07%)	16 (7.62%)
5. Social	1,612 (2.62%)	0.02 (635.40%)	35 (16.67%)
6. Paid Search	887 (1.44%)	0.03 (890.81%)	27 (12.86%)

# Cohort Analysis

A cohort is a group of users who share a common characteristic that is identified in this report by an Analytics dimension. E.g. all users with the same Acquisition Date. Only Acquisition date is supported currently, but there are many metrics to choose from.



## Who is the Audience? Demographics in GA



# Interests

Affinity Category (reach)	69.32% of total sessions	In-Market Segment	65.80% of total sessions
3.86% Technophiles		3.76% Consumer Electronics/Mobile Phones	
3.41% Movie Lovers		3.45% Financial Services/Investment Services	
3.30% TV Lovers		3.25% Travel/Hotels & Accommodations	
2.94% Shutterbugs		2.72% Employment	
2.87% Mobile Enthusiasts		2.24% Education/Post-Secondary Education	
2.85% News Junkies & Avid Readers/Entertainment & Celebrity News Junkies		2.17% Computers & Peripherals/Computers/Laptops & Notebooks	
2.72% Travel Buffs		2.08% Travel/Air Travel	
2.38% Avid Investors		1.77% Real Estate/Residential Properties	
2.35% Music Lovers/Pop Music Fans		1.74% Software/Business & Productivity Software	
2.26% Political Junkies		1.72% Autos & Vehicles/Motor Vehicles/Motor Vehicles (Used)	

Affinity Categories	Lifestyles similar to TV audiences, for example: Technophiles, Sports Fans, and Cooking Enthusiasts
In-Market Segments	Product-purchase interests
Other Categories	Provides the most specific view of your users. For example, Affinity Categories includes Foodies, while Other Categories includes Recipes/Cuisines/East Asian

In-market means that a user is researching products and are actively considering buying

## Obtaining Demographics & Interests

- Inferencing:** Whenever you visit a website that has partnered with the Google Display Network this is stored in a cookie on your machine. Different sites are associated with different demographics hence the list of sites you have visited can be used to infer gender and age etc. Note that cookies identify a web browser on a specific computer, not a specific person
- Volunteered to 3<sup>rd</sup> Party Sites:** some sites might provide Google with demographic and interests information that people share on certain websites, such as social networking sites
- Volunteered to Google:** demographics derived from Google profiles

# Obtaining Demographics & Interests

Once you update Analytics to support Advertising Reporting Features, Analytics collects Demographics and Interests data from the following sources:

Source	Applies to	Condition	Result
Third-party DoubleClick cookie	Web-browser activity only	Cookie is present	Analytics collects any demographic and interests information available in the cookie
Android Advertising ID	App activity only	You update the Analytics tracking code in an Android app to collect the Advertising ID	Analytics generates an identifier based on the ID that includes demographic and interests information associated with users' app activity
iOS Identifier for Advertisers (IDFA)	App activity only	You update the Analytics tracking code in an iOS app to collect the IDFA	Analytics generates an identifier based on the IDFA that includes demographic and interests information associated with users' app activity

Demographics and interests data may only be available for a subset of your users, and may not represent the overall composition of your traffic: Analytics cannot collect the demographics and interests information if the DoubleClick cookie or the Device Advertising ID is not present, or if no activity profile is included.

<https://support.google.com/analytics/answer/2799357?hl=en>

*By enabling the Advertising Features, you enable Google Analytics to collect data about your traffic via Google advertising cookies and anonymous identifiers, in addition to data collected through a standard Google Analytics implementation.*

## Affinity Categories

- These are marketing segments that you can buy



Search reports and help		Affinity Category (reach) <small>?</small>	Acquisition			Behavior		
			Sessions <small>?</small> <small>▼</small>	% New Sessions <small>?</small>	New Users <small>?</small>	Bounce Rate <small>?</small>	Pages / Session <small>?</small>	Avg. Session Duration <small>?</small>
Overview			7,327 % of Total: 57.76% (12,685)	51.87% Avg for View: 52.32% (-0.86%)	3,766 % of Total: 56.74% (6,637)	65.06% Avg for View: 63.90% (1.81%)	2.54 Avg for View: 2.50 (1.54%)	00:02:31 Avg for View: 00:02:33 (-1.34%)
Active Users								
Lifetime Value <small>BETA</small>								
Cohort Analysis <small>BETA</small>								
User Explorer								
▶ Demographics		1. Shoppers/Value Shoppers	5,581 (4.45%)	48.75%	2,721 (4.19%)	62.16%	2.66	00:02:39
▼ Interests		2. Lifestyles & Hobbies/Green Living Enthusiasts	4,819 (3.85%)	47.91%	2,309 (3.55%)	61.53%	2.66	00:02:45
Overview		3. Lifestyles & Hobbies/Business Professionals	4,707 (3.76%)	47.59%	2,240 (3.45%)	61.27%	2.80	00:02:53
Affinity Categories		4. Technology/Technophiles	4,547 (3.63%)	47.99%	2,182 (3.36%)	60.55%	2.77	00:02:44
In-Market Segments		5. Travel/Travel Buffs	4,451 (3.55%)	49.76%	2,215 (3.41%)	65.11%	2.50	00:02:28
Other Categories		6. Lifestyles & Hobbies/Art & Theater Aficionados	4,091 (3.26%)	50.87%	2,081 (3.20%)	63.31%	2.58	00:02:38
▶ Geo		7. Shoppers/Shopaholics	3,994 (3.19%)	50.40%	2,013 (3.10%)	65.30%	2.54	00:02:33
▶ Behavior		8. Sports & Fitness/Health & Fitness Buffs	3,911 (3.12%)	50.12%	1,960 (3.02%)	60.52%	2.72	00:02:50
▶ Technology		9. News & Politics/News Junkies/Entertainment & Celebrity News Junkies	3,815 (3.04%)	55.28%	2,109 (3.24%)	68.91%	2.29	00:02:08
▶ Mobile		10. Travel/Business Travelers	3,660 (2.92%)	47.76%	1,748 (2.69%)	60.00%	2.77	00:02:52
▶ Custom								

# Geo: Identifying User Location

**Audience**

- Overview
- Active Users BETA
- Cohort Analysis BETA
- Demographics
- Interests
- Geo
  - Language
  - Location
- Behavior
- Technology
- Mobile
- Custom
- Benchmarking

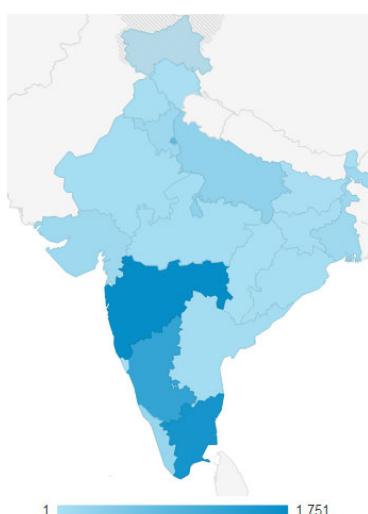
Users Flow

- Usually based on reverse-IP technology, your IP address is used to infer your location
- In general, reverse-IP is very accurate at country-level (~ 99%), less accurate at city and postal code level (~ 90%+)

Country	Acquisition			Behavior
	Sessions	% New Sessions	New Users	Bounce Rate
	<b>49,701</b> % of Total: 100.00% (49,701)	<b>57.21%</b> Avg for View: 57.06% (0.28%)	<b>28,432</b> % of Total: 100.28% (28,359)	<b>55.60%</b> Avg for View: 55.60% (0.09%)
1. Singapore	32,833 (66.06%)	49.84%	16,365 (57.56%)	56.39%
2. India	7,147 (14.38%)	60.08%	4,294 (15.10%)	49.90%
3. China	3,651 (7.35%)	90.93%	3,320 (11.68%)	49.82%
4. United States	1,344 (2.70%)	77.31%	1,039 (3.65%)	58.93%
5. Malaysia	446 (0.90%)	73.09%	326 (1.15%)	67.49%
6. Hong Kong	370 (0.74%)	66.76%	247 (0.87%)	51.08%
7. Indonesia	316 (0.64%)	64.87%	205 (0.72%)	57.91%
8. Taiwan	247 (0.50%)	79.35%	196 (0.69%)	67.21%
9. (not set)	247 (0.50%)	61.54%	152 (0.53%)	57.49%
10. Myanmar (Burma)	246 (0.49%)	52.03%	128 (0.45%)	40.65%

## User Location Drill-down

- GA allows drill down to the state or city level, e.g. clicking on India yields:



Region	Acquisition			Behavior		
	Users	New Users	Sessions	Bounce Rate	Pages / Session	Avg. Session Duration
	<b>676</b> % of Total: 5.86% (11,527)	<b>553</b> % of Total: 6.59% (8,391)	<b>966</b> % of Total: 5.62% (17,197)	<b>64.60%</b> Avg for View: 59.99% (7.67%)	<b>2.20</b> Avg for View: 2.56 (-13.96%)	<b>00:02:14</b> Avg for View: 00:02:47 (-19.90%)
1. Tamil Nadu	185 (25.97%)	143 (25.86%)	269 (27.85%)	60.22%	2.38	00:02:20
2. Karnataka	119 (17.38%)	91 (16.46%)	180 (18.63%)	60.00%	2.35	00:01:48
3. Maharashtra	67 (9.77%)	59 (10.67%)	93 (9.63%)	65.59%	1.77	00:01:31
4. Telangana	67 (9.77%)	55 (9.95%)	80 (8.28%)	73.75%	1.70	00:00:45
5. Delhi	55 (8.02%)	45 (8.14%)	102 (10.56%)	66.67%	1.83	00:02:53
City	Acquisition			Behavior		
	Users	New Users	Sessions	Bounce Rate	Pages / Session	Avg. Session Duration
	<b>676</b> % of Total: 5.86% (11,527)	<b>553</b> % of Total: 6.59% (8,391)	<b>966</b> % of Total: 5.62% (17,197)	<b>64.60%</b> Avg for View: 59.99% (7.67%)	<b>2.20</b> Avg for View: 2.56 (-13.96%)	<b>00:02:14</b> Avg for View: 00:02:47 (-19.90%)
1. Chennai	153 (21.89%)	110 (19.89%)	215 (22.26%)	58.14%	2.46	00:02:17
2. Bengaluru	119 (17.02%)	91 (16.46%)	179 (18.53%)	60.34%	2.35	00:01:48
3. Hyderabad	75 (10.73%)	63 (11.39%)	88 (9.11%)	73.86%	1.77	00:01:06
4. New Delhi	36 (5.15%)	28 (5.06%)	64 (6.63%)	68.75%	1.95	00:03:05
5. Mumbai	33 (4.72%)	29 (5.24%)	43 (4.45%)	62.79%	2.07	00:02:08
6. Coimbatore	31 (4.43%)	26 (4.70%)	44 (4.55%)	65.91%	2.20	00:02:44
7. (not set)	28 (4.01%)	23 (4.16%)	45 (4.66%)	68.89%	1.53	00:02:10
8. Kolkata	21 (3.00%)	17 (3.07%)	29 (3.00%)	51.72%	2.72	00:07:07
9. Kochi	18 (2.58%)	15 (2.71%)	34 (3.52%)	55.88%	3.82	00:03:32
10. Pune	17 (2.43%)	14 (2.53%)	19 (1.97%)	52.63%	1.79	00:01:18

## New versus Returning Users

Audience

Overview

Demographics

Interests

Geo

Behavior

New vs Returning

Frequency & Recency

Engagement

Technology

Mobile

Custom

Benchmarking

Users Flow

- The best ratio is business goal specific. E.g. online newspaper versus (say) online funeral services!

Acquisition			
User Type	Sessions	% New Sessions	New Users
	<b>46,987</b> % of Total: 100.00% (46,987)	<b>55.72%</b> Site Avg: 55.64% (0.13%)	<b>26,179</b> % of Total: 100.13% (26,145)
1. New Visitor	<b>26,179</b> (55.72%)	100.00%	<b>26,179</b> (100.00%)
2. Returning Visitor	<b>20,808</b> (44.28%)	0.00%	0 (0.00%)

Behavior				Conversions	Goal 1: Registration
Bounce Rate	Pages / Session	Avg. Session Duration	Registration (Goal 1 Conversion Rate)		
<b>53.98%</b> Site Avg: 53.98% (0.00%)	<b>3.16</b> Site Avg: 3.16 (0.00%)	<b>00:03:14</b> Site Avg: 00:03:14 (0.00%)	<b>0.00%</b> Site Avg: 0.00% (0.00%)		
54.89%	2.76	00:02:18	0.00%		
52.83%	3.66	00:04:25	0.00%		

## Recency & Frequency

- Histograms of #visits per user and #days since last session
  - Expectations should be in line with website business domain

Audience

Overview

Demographics

Interests

Geo

Behavior

New vs Returning

Frequency & Recency

Engagement

Technology

Mobile

Custom

Benchmarking

Users Flow

Distribution

Count of Sessions	Days Since Last Session
46,987	% of Total: 100.00% (46,987)

Days Since Last Session	Sessions	Pageviews
0	37,563	114,275
1	1,938	6,125
2	952	3,407
3	709	2,327
4	511	1,875
5	452	1,589
6	440	1,533
7	347	1,496
8-14	1,207	4,717
15-30	1,232	4,765
31-60	749	2,877
61-120	503	1,986
121-364	312	1,224
365+	72	230

# Engagement

- Average Visit Duration & Page Depth (number of pages users visit in a session). Typically long-tail distributions - most users visit few pages and spend little time!

**Audience**

Overview

- ▶ Demographics
- ▶ Interests
- ▶ Geo
- ▼ Behavior
  - New vs Returning
  - Frequency & Rec...
  - Engagement**
- ▶ Technology
- ▶ Mobile
- ▶ Custom
- ▶ Benchmarking
- Users Flow

All Sessions  
100.00%
+ Add Segment

Distribution

Session Duration Page Depth

Sessions	Pageviews	
<b>46,987</b> % of Total: 100.00% (46,987)	<b>148,426</b> % of Total: 100.00% (148,426)	
<hr/>		
Session Duration	Sessions	Pageviews
0-10 seconds	28,796	33,688
11-30 seconds	3,379	9,051
31-60 seconds	1,906	6,019
61-180 seconds	3,785	15,781
181-600 seconds	4,414	29,262
601-1800 seconds	3,626	32,629
1801+ seconds	1,081	

Page Depth	Sessions	Pageviews
1	10,317	10,317
2	2,564	5,128
3	1,354	4,062
4	729	2,916
5	497	2,485

ISS
INSTITUTE OF SYSTEMS SCIENCE

ATA/BA-DAT/IntroToWebAnalytics/V6

© 2018 NUS. All rights reserved.

Page 39

## Technology & Mobile

- Determine which browsers & O/S to optimise your website for
- Determine what mobile to optimize for

**Audience**

Overview

Active Users BETA

Cohort Analysis BETA

- ▶ Demographics
- ▶ Interests
- ▶ Geo
- ▶ Behavior
- ▼ Technology
  - Browser & OS**
  - Network**
- ▼ Mobile
  - Overview**
  - Devices
- ▶ Custom
- ▶ Benchmarking
- Users Flow

**Browser & OS**

Service Provider	Users
	11,527 % of Total: 100.00% (11,527)
1. singnet pte ltd	1,886 (15.25%)
2. singtel mobile	973 (7.87%)
3. m1 ltd	729 (5.89%)
4. starhub mobile	600 (4.85%)
5. starhub-ltd-ngnbn-services	561 (4.54%)
6. m1 net ltd	544 (4.40%)
7. starhub cable vision ltd	544 (4.40%)

**Mobile-Overview**

Device Category	Users
	11,527 % of Total: 100.00% (11,527)
1. desktop	6,574 (57.31%)
2. mobile	4,516 (39.37%)
3. tablet	381 (3.32%)

**Devices**

Mobile Device Info	Acquisition
	4,897 % of Total: 42.48% (11,527)
1. Apple iPhone	2,231 (45.56%)
2. Apple iPad	233 (4.76%)
3. (not set)	141 (2.88%)
4. Samsung SM-G950F Galaxy S8	136 (2.78%)
5. Samsung SM-N950F Galaxy Note8	131 (2.68%)
6. Samsung SM-G935F Galaxy S7 Edge	89 (1.82%)
7. Microsoft Windows RT Tablet	88 (1.80%)
8. Samsung SM-G965F Galaxy S9+	79 (1.61%)
9. Samsung SM-G955F Galaxy S8+	74 (1.51%)
10. Samsung SM-G930F Galaxy S7	68 (1.39%)

ISS
INSTITUTE OF SYSTEMS SCIENCE

ATA/BA-DAT/IntroToWebAnalytics/V6

© 2018 NUS. All rights reserved.

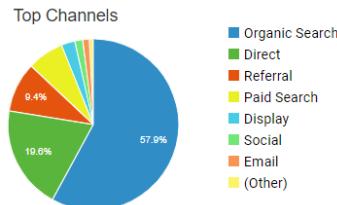
Page 40

# Where did your Audience come from?

ACQUISITION

- [Overview](#)
- [All Traffic](#)
- [Channels](#)
- [Treemaps](#)
- [Source/Medium](#)
- [Referrals](#)
- [▶ Google Ads](#)
  - [▶ Search Console](#)
  - [▶ Social](#)
  - [▶ Campaigns](#)

- Which channel yields the most traffic?
- Which channel yields the most profitable traffic?
  - Consider conversion rates, engagement rates, bounce rates ....
- Optimise your website to increase (profitable) traffic - SEO



- Referral = via a link in another website
- Direct traffic = they typed in your URL or had it bookmarked
- Display = they clicked on your ad
- Social = they came via social media

Default Channel Grouping	Users	New Users	Sessions	Bounce Rate	Pages / Session	Avg. Session Duration
1. Organic Search	6,970 (57.92%)	4,585 (54.60%)	10,114 (58.81%)	55.84%	2.79	00:03:05
2. Direct	2,363 (19.64%)	1,954 (23.27%)	3,190 (18.55%)	68.46%	2.04	00:02:04
3. Referral	1,135 (9.43%)	621 (7.39%)	1,978 (11.50%)	55.92%	2.80	00:03:28
4. Paid Search	834 (6.93%)	674 (8.03%)	998 (5.80%)	72.44%	1.98	00:01:16
5. Display	309 (2.57%)	302 (3.60%)	320 (1.86%)	87.19%	1.18	00:00:14
6. Social	179 (1.49%)	133 (1.58%)	225 (1.31%)	61.78%	2.43	00:02:36
7. Email	152 (1.26%)	71 (0.85%)	217 (1.26%)	64.98%	2.33	00:02:59

## Channel Traffic Drill-Down

- Click on “Organic search” to see the top keywords that led people to your website
  - These show the topics users are most interested in - ensure they get covered adequately in your site. What keywords yielded less traffic than expected? This could indicate that the website doesn’t have enough content in this area (e.g. not getting indexed by Google)\*

Default Channel Grouping	Keyword	Users	New Users	Sessions	Bounce Rate	Pages / Session	Avg. Session Duration
1. Organic Search		7,260 % of Total: 56.75% (12,792)	4,720 % of Total: 49.42% (9,550)	10,524 % of Total: 56.02% (18,786)	55.82% Avg for View: 61.16% (-8.73%)	2.73 Avg for View: 2.45	00:03:03 Avg for View: 00:02:37 (16.75%)
2. Direct	1. (not provided)	7,080 (97.11%)	4,596 (97.37%)	10,230 (97.21%)	56.16%		
3. Referral	2. nus iss	22 (0.30%)	15 (0.32%)	33 (0.31%)	15.15%		
4. Paid Search	3. (not set)	14 (0.19%)	5 (0.11%)	25 (0.24%)	44.00%		
5. Display	4. iss nus	10 (0.14%)	2 (0.04%)	21 (0.20%)	19.05%		
6. Email	5. nus-iss	4 (0.05%)	3 (0.06%)	4 (0.04%)	50.00%		
7. Social	6. iss	3 (0.04%)	3 (0.06%)	4 (0.04%)	0.00%		
8. (Other)	7. iss nus singapore	3 (0.04%)	2 (0.04%)	3 (0.03%)	0.00%	4.33	00:01:38
	8. https://www.iss.nus.edu.sg/	2 (0.03%)	1 (0.02%)	2 (0.02%)	0.00%	4.00	00:06:10

Many search queries are “not provided” due to privacy and other issues but lessons may still be learned from those that are

Most Google searches are performed via HTTPS, which causes the keyword dimension to automatically be set to (not provided).

# Search Query Analysis via Search Console

## About Search Console

The Search Console reports in Analytics provide information about the performance of your organic-search traffic. You can see data like user queries and the number of times your site URLs appear in search results (impressions), along with post-click data about site engagement like bounce rate and ecommerce conversion rate. This combination of data helps you optimize your site for the most profitable traffic.

**ACQUISITION**

- Overview
- All Traffic
- Google Ads
- Search Console** (highlighted)
- Landing Pages
- Countries
- Devices
- Queries

- The queries that led users to your site can also be investigated using the Search Console

Search Query	Clicks	Impressions	CTR	Average Position
	2,892 % of Total: 45.34% (6,378)	79,611 % of Total: 73.61% (108,152)	3.63% Avg for View: 5.90% (-38.40%)	37 Avg for View: 31 (19.11%)
1. nus iss	512 (17.70%)	709 (0.89%)	72.21%	1.0
2. iss nus	101 (3.49%)	154 (0.19%)	65.58%	1.0
3. chong yoke sin	96 (3.32%)	296 (0.37%)	32.43%	1.3
4. ssg funding	47 (1.63%)	123 (0.15%)	38.21%	1.1
5. iss	28 (0.97%)	600 (0.75%)	4.67%	14
6. dr chong yoke sin	26 (0.90%)	69 (0.09%)	97.69%	1.4
7. nus-iss	25 (0.86%)	40 (0.05%)	1.0	1.0
8. nus iss courses	23 (0.80%)	28 (0.04%)	1.0	1.0
9. institute of systems science	21 (0.73%)	36 (0.05%)	1.0	1.2
10. pmp exam singapore	21 (0.73%)	67 (0.08%)	31.34%	1.0

The number of times any URL from your site appeared in search results viewed by a user, not including paid Google Ads search impressions.

The average ranking of your website URLs for the query or queries. For example, if your site's URL appears at position 3 for one query and position 7 for another query, the average position would be 5 ( $3+7/2$ ).

Note: Organic Search keywords (see previous slide) show the keywords used to find your site via all organic search engines. The queries shown here in Search console are only those from the Google search engine

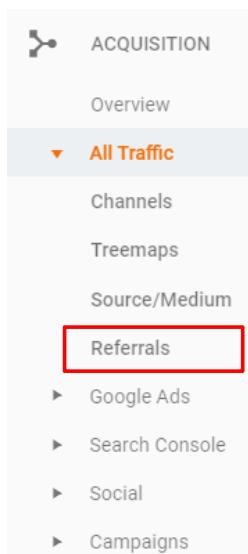
## Channel Traffic Drill-Down

- Clicking on Referral shows the sites that provide most referrals.....

Default Channel Grouping	Source	Acquisition			Behavior		
		Users	New Users	Sessions	Bounce Rate	Pages / Session	Avg. Session Duration
1. Organic Search		1,135 % of Total: 9.85% (11,527)	621 % of Total: 7.40% (8,391)	1,978 % of Total: 11.50% (17,197)	55.92% Avg for View: 59.99% (-6.80%)	2.80 Avg for View: 2.56 (9.35%)	00:03:28 Avg for View: 00:02:47 (24.22%)
2. Direct							
3. Referral	1. nus.edu.sg	279 (24.01%)	161 (25.93%)	430 (21.74%)	53.26%	2.46	00:03:42
4. Paid Search	2. scale.nus.edu.sg	214 (18.42%)	153 (24.64%)	488 (24.67%)	59.22%	2.88	00:01:43
5. Display	3. tms-iss.nus.edu.sg	153 (13.17%)	7 (1.13%)	237 (11.98%)	41.35%	3.55	00:06:06
6. Email	4. provost.nus.edu.sg	45 (3.87%)	29 (4.67%)	61 (3.08%)	44.26%	3.10	00:02:55
7. Social	5. mail.google.com	36 (3.10%)	13 (2.09%)	59 (2.98%)	47.46%	2.69	00:04:31
8. (Other)	6. myskillsfuture.sg	31 (2.67%)	20 (3.22%)	50 (2.53%)	52.00%	3.16	00:04:25
	7. forums.hardwarezone.com.sg	25 (2.15%)	23 (3.70%)	25 (1.26%)	84.00%	1.72	00:00:18
	8. ssg.gov.sg	25 (2.15%)	15 (2.42%)	31 (1.57%)	61.29%	3.32	00:05:52
	9. isc2.org	24 (2.07%)	18 (2.90%)	25 (1.26%)	56.00%	2.56	00:02:03
	10. goodagile.com	22 (1.89%)	11 (1.77%)	23 (1.16%)	65.22%	1.96	00:00:47

Can also get same info via the Referrals tab

# Referral Traffic - Questions To Ask



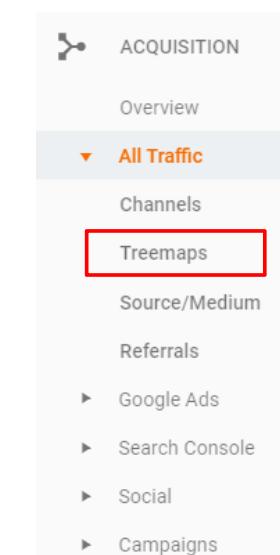
The sidebar shows the following navigation structure under 'ACQUISITION':

- Overview
- All Traffic (highlighted)
- Channels
- Treemaps
- Source/Medium
- Referrals (highlighted with a red box)
- ▶ Google Ads
- ▶ Search Console
- ▶ Social
- ▶ Campaigns

- Which websites refer most traffic to you & which websites refer users with the highest conversion rates?
  - This can give insights into the interests of your most valuable users (where else do they visit?)
- Do the referral sites send users to different landing pages?
- Are the referral sites those that you think your users will visit?
  - E.g. your website sells cameras but most referrals are from health-food websites!
  - If not then how to get referrals in relevant sites?

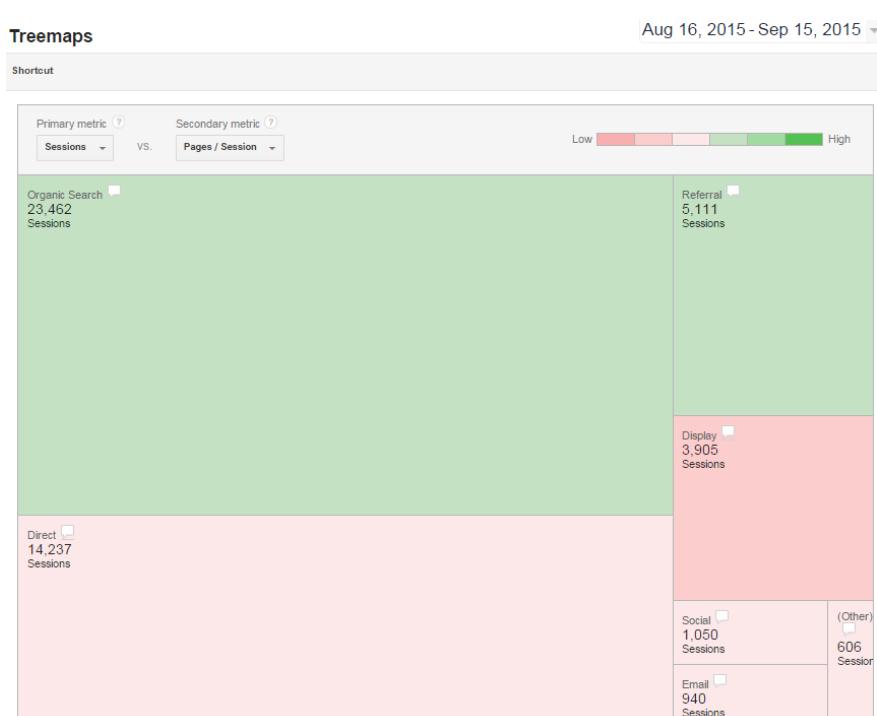
## Traffic Analysis: TreeMap

You can view any of the summary numbers as a tree map



The sidebar shows the following navigation structure under 'ACQUISITION':

- Overview
- All Traffic (highlighted)
- Channels
- Treemaps (highlighted with a red box)
- Source/Medium
- Referrals
- ▶ Google Ads
- ▶ Search Console
- ▶ Social
- ▶ Campaigns



# Traffic Analysis: Source/Medium

Source / Medium <span style="font-size: small;">?</span>	Acquisition			Behavior			
	Users <span style="font-size: small;">?</span> <span style="color: #ccc;">↓</span>	New Users <span style="font-size: small;">?</span>	Sessions <span style="font-size: small;">?</span>	Bounce Rate <span style="font-size: small;">?</span>	Pages / Session <span style="font-size: small;">?</span>	Avg. Session Duration <span style="font-size: small;">?</span>	
	12,792 % of Total: 100.00% (12,792)	9,555 % of Total: 100.05% (9,550)	18,786 % of Total: 100.00% (18,786)	61.16% Avg for View: 61.16% (0.00%)	2.45 Avg for View: 2.45 (0.00%)	00:02:37 Avg for View: 00:02:37 (0.00%)	
1. google / organic	7,057 (52.43%)	4,586 (48.00%)	10,218 (54.39%)	56.22%	2.70	00:03:01	
2. (direct) / (none)	2,681 (19.92%)	2,267 (23.73%)	3,484 (18.55%)	70.87%	1.93	00:01:55	
3. google / cpc	844 (6.27%)	671 (7.02%)	997 (5.31%)	71.31%	2.01	00:01:15	
4. Gter / banner	445 (3.31%)	439 (4.59%)	457 (2.43%)	88.62%	1.18	00:00:15	
5. scale.nus.edu.sg / referral	261 (1.94%)	196 (2.05%)	612 (3.26%)	62.09%	2.62	00:01:38	
6. nus.edu.sg / referral	248 (1.84%)	128 (1.34%)	368 (1.96%)	51.09%	2.45	00:03:23	
7. straitstimes.com / referral	192 (1.43%)	186 (1.95%)	201 (1.07%)	80.60%	1.33	00:00:25	
8. tms-iss.nus.edu.sg / referral	168 (1.25%)	5 (0.05%)	271 (1.44%)	40.59%	3.63	00:06:12	
9. bing / organic	147 (1.09%)	93 (0.97%)	227 (1.21%)	38.33%	3.55	00:04:08	
10. edm / email	134 (1.00%)	69 (0.72%)	189 (1.01%)	66.14%	2.11	00:02:25	

**Source** = the origin of your traffic, such as a search engine (e.g. *google*) or a domain (*example.com*).  
**Medium**= the general category of the source, e.g. organic search (*organic*), cost-per-click paid search (*cpc*)

## Traffic from Google Marketing Campaigns

❯ ACQUISITION

- Overview
- ▶ All Traffic
- ▼ Google Ads
- Accounts
- Campaigns
- Treemaps
- Sitelinks NEW
- Bid Adjustments
- Keywords
- Search Queries
- Hour of Day
- Final URLs
- Display Targeting
- Video Campaigns
- Shopping Campaigns

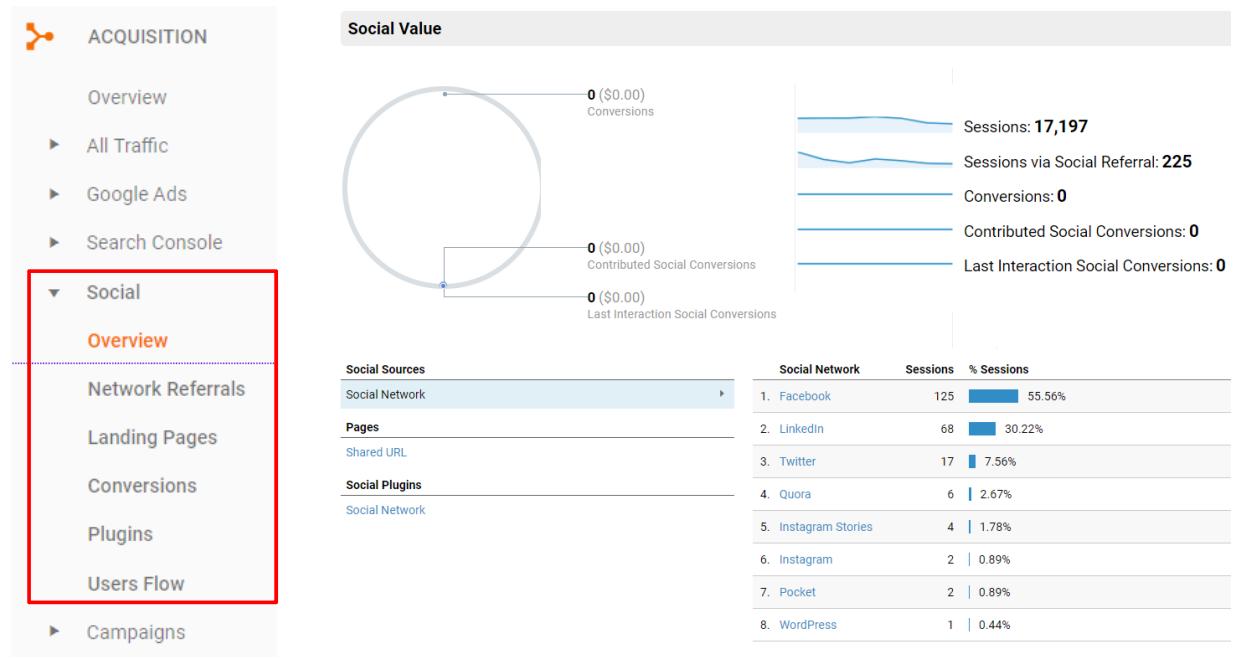
Campaign / Campaign ID	Acquisition				Behavior		Conversions	Goal 1: Registration <span style="font-size: small;">▼</span>	
	Clicks <span style="font-size: small;">?</span> <span style="color: #ccc;">↓</span>	Cost <span style="font-size: small;">?</span>	CPC <span style="font-size: small;">?</span>	Sessions <span style="font-size: small;">?</span>	Bounce Rate <span style="font-size: small;">?</span>	Pages / Session <span style="font-size: small;">?</span>	Registration (Goal 1 Conversion Rate) <span style="font-size: small;">?</span>	Registration (Goal 1 Completions) <span style="font-size: small;">?</span>	Registration (Goal 1 Value) <span style="font-size: small;">?</span>
	1,108 % of Total: 100.00% (1,108)	\$1,409.21 % of Total: 100.00% (\$1,409.21)	\$1.27 Avg for View: \$1.27 (0.00%)	296 % of Total: 1.35% (21,996)	79.39% Avg for View: 66.66% (19.10%)	1.71 Avg for View: 2.15 (-20.61%)	0.00% Avg for View: 0.00% (0.00%)	0 % of Total: 0.00% (0)	\$0.00 % of Total: 0.00% (\$0.00)
1. Display - Remarketing 396032808	749 (67.60%)	\$408.16 (28.96%)	\$0.54	220 (74.32%)	80.91%	1.72	0.00%	0 (0.00%)	\$0.00 (0.00%)
2. Search 395590488	329 (29.69%)	\$989.23 (70.20%)	\$3.01	59 (19.93%)	76.27%	1.68	0.00%	0 (0.00%)	\$0.00 (0.00%)
3. Display 395594688	30 (2.71%)	\$11.82 (0.84%)	\$0.39	7 (2.36%)	71.43%	1.43	0.00%	0 (0.00%)	\$0.00 (0.00%)
4. (not set)	0 (0.00%)	\$0.00 (0.00%)	\$0.00	10 (3.38%)	70.00%	1.70	0.00%	0 (0.00%)	\$0.00 (0.00%)

### Link/unlink Google Ads and Analytics

Linking your Google Ads account to your Analytics property lets you see the full customer cycle, from how they interact with your marketing (e.g., seeing ad impressions, clicking ads) to how they finally complete the goals you've set for them on your site (e.g., making purchases, consuming content).

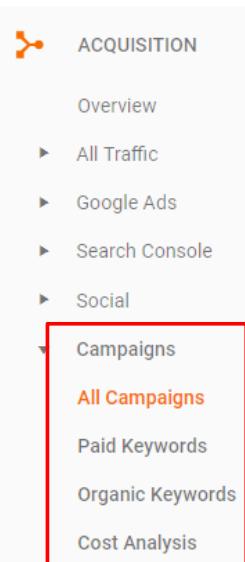
# Traffic from Social Media

- Identify where to increase your social effort



# Traffic from Custom Campaigns

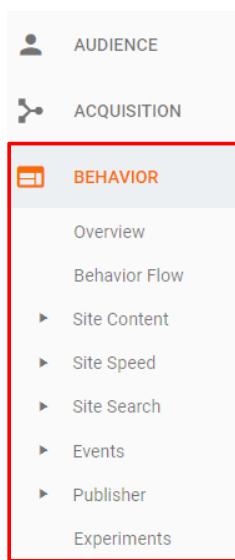
3rd Party campaigns etc....



Campaign	Acquisition			Behavior		
	Users	New Users	Sessions	Bounce Rate	Pages / Session	Avg. Session Duration
	533 % of Total: 4.62% (11,527)	414 % of Total: 4.93% (8,391)	673 % of Total: 3.91% (17,197)	74.59% Avg for View: 59.99% (24.33%)	1.81 Avg for View: 2.56 (-29.23%)	00:02:02 Avg for View: 00:02:47 (-27.10%)
1. Gter GDipSA banner	308 (55.90%)	302 (72.95%)	319 (47.40%)	87.15%	1.18	00:00:14
2. DID_sep	100 (18.15%)	49 (11.84%)	129 (19.17%)	64.34%	2.34	00:03:43
3. BM_Aug2018	28 (5.08%)	5 (1.21%)	53 (7.88%)	56.60%	2.25	00:02:44
4. DID_Sep	26 (4.72%)	15 (3.62%)	30 (4.46%)	90.00%	1.33	00:00:28
5. Gter GDipSA bann	19 (3.45%)	18 (4.35%)	19 (2.82%)	89.47%	1.26	00:00:20
6. AT_Sep	9 (1.63%)	3 (0.72%)	11 (1.63%)	45.45%	3.73	00:02:50
7. BDLT	7 (1.27%)	1 (0.24%)	14 (2.08%)	85.71%	1.21	00:00:39
8. EAMC	7 (1.27%)	3 (0.72%)	17 (2.53%)	47.06%	2.59	00:03:13
9. Learning_Day_2018	6 (1.09%)	1 (0.24%)	6 (0.89%)	66.67%	2.67	00:00:41
10. SFS-Cybersecurity	6 (1.09%)	3 (0.72%)	6 (0.89%)	83.33%	2.00	00:02:56

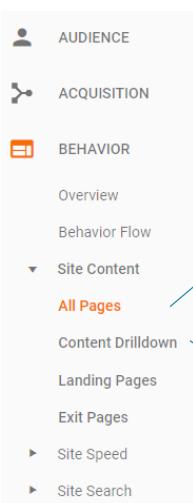
To monitor traffic from a custom campaign you must add campaign parameters to the referring URL (e.g. the URL link to your website in an email or the URL link in a banner ad)

# Audience Behavior: How do visitors use the site?



- Where do users go on the website, how long do they stay?
  - Are the pages you think important the ones most visited?
- Which pages get most & least traffic?
  - Is low traffic because users can't find the page OR page is no value?
  - Is high traffic because page is useful OR page easy to find (many links)?
- What pages do users use to enter & exit the site?
  - If the ratio of entrances to unique page-views for a page is high then the page is a landing page => needs welcome info & navigation aids
- Which Pages have high Bounce & Exit Rates and why?
  - High Bounce or Exit rate does not always imply bad page design
    - Page could do its job well, e.g. product page (high bounce), checkout page (high exits)
    - Page might be indexed by search engine inappropriately or referenced incorrectly from another website => users with no interest may arrive

## Audience Behavior: What Pages are Viewed?



All pages in descending order...

Page	Pageviews	Unique Pageviews
1. /	2,996 (6.80%)	2,398 (6.96%)
2. /graduate-programmes/programme/detail/graduate-diploma-in-systems-analysis	2,000 (4.54%)	1,686 (4.90%)
3. /graduate-programmes/programme/detail/master-of-technology-in-enterprise-business-analytics	1,825 (4.14%)	1,379 (4.00%)
4. /graduate-programmes/programme/detail/master-of-technology-in-intelligent-systems	1,362 (3.09%)	1,018 (2.96%)
5. /community/events/event-details/2018/10/06/default-calendar/graduate-diploma-in-systems-analysis-programme-information-session	1,113 (2.53%)	1,000 (2.90%)
6. /executive-education/discipline/detail/analytics-and-intelligent-systems	999 (2.27%)	537 (1.56%)
7. /graduate-programmes/programme/detail/master-of-technology-in-software-engineering	880 (2.00%)	719 (2.09%)
8. /executive-education/course-exams-finder/course-finder	761 (1.73%)	513 (1.49%)

Or grouped in a hierarchy, for drill-down...

Page path level 1	Pageviews
1. /executive-education/	21,048 (47.78%)
2. /graduate-programmes/	8,439 (19.16%)
3. /about-us/	4,276 (9.71%)
4. /community/	3,964 (9.00%)
5. /	2,996 (6.80%)
6. /stackable-certificate-programmes/	973 (2.21%)
7. /collaboration/	582 (1.32%)
8. /graduate-programmes	469 (1.06%)
9. /stackable-certificate-programmes	404 (0.92%)
10. /executive-education	258 (0.59%)

Page path level 2	Pageviews
1. /course/	13,936
2. /discipline/	4,722
3. /course-exams-finder/	761
4. /skillsfuture-series	534
5. /funding/	462
6. /executive-education	183
7. /funding	134
8. /terms-and-conditions-of-nus-iss-short-courses	87
9. /nus-iss-course-planner	78
10. /nus-iss-and-learner-s-commitment-and-respon	56

# Landing and Exit Pages

- High bounce rate is not always bad if the user gets what they are looking for first time!

**AUDIENCE**

**ACQUISITION**

**BEHAVIOR**

Overview

Behavior Flow

Site Content

All Pages

Content Drilldown

Landing Pages

Exit Pages

Site Speed

Site Search

Landing Page	Sessions	% New Sessions	New Users	Bounce Rate	Pages / Session	Avg. Session Duration
1. /	2,031 (11.81%)	38.01%	772 (9.19%)	28.85%	4.28	00:04:49
2. /graduate-programmes/programme/detail/graduate-diploma-in-systems-analysis	1,275 (7.41%)	61.80%	788 (9.38%)	77.80%	1.66	00:01:30
3. /community/events/event-details/2018/10/06/default-calendar/graduate-diploma-in-systems-analysis-programme-information-session	957 (5.56%)	75.03%	718 (8.55%)	73.77%	1.88	00:01:07
4. /graduate-programmes/programme/detail/master-of-technology-in-enterprise-business-analytics	709 (4.12%)	30.89%	219 (2.61%)	67.28%	1.99	00:03:05
5. /graduate-programmes/graduate-programmes	387 (2.25%)	44.96%	174 (2.07%)	25.32%	3.34	00:03:46
6. /graduate-programmes/programme/detail/master-of-technology-in-intelligent-systems	368 (2.14%)	27.99%	103 (1.23%)	65.22%	2.12	00:02:52
7. /graduate-programmes/programme/detail/master-of-technology-in-software-engineering	346 (2.01%)	32.08%	111 (1.32%)	59.83%	2.41	00:03:36
8. /executive-education/course/detail/nicf-certified-scrummaster/	323 (1.88%)	62.23%	201 (2.39%)	69.04%	1.92	00:02:21

Page	Exits
1. /graduate-programmes/programme/detail/graduate-diploma-in-systems-analysis	1,340 (7.79%)
2. /graduate-programmes/programme/detail/master-of-technology-in-enterprise-business-analytics	945 (5.50%)
3. /	822 (4.78%)
4. /community/events/event-details/2018/10/06/default-calendar/graduate-diploma-in-systems-analysis-programme-information-session	764 (4.44%)
5. /graduate-programmes/programme/detail/master-of-technology-in-intelligent-systems	631 (3.67%)
6. /graduate-programmes/programme/detail/master-of-technology-in-software-engineering	431 (2.51%)
7. /about-us/getting-to-nus-iss	366 (2.13%)
8. /about-us/staff/detail/7.5.5/CHONG Yoke Sin	283 (1.65%)
9. /executive-education/course/detail/working-with-data-cloudops-and-things-using-python/startup-and-sme	266 (1.55%)
10. /executive-education/course/detail/nicf-certified-scrummaster/	258 (1.50%)



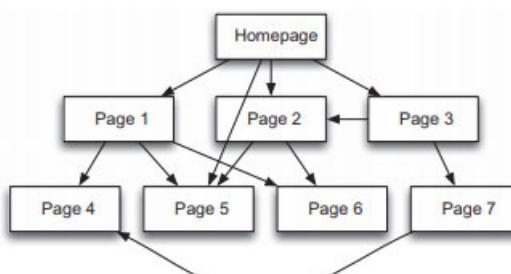

ATA/BA-DAT/IntroToWebAnalytics/V6

© 2018 NUS. All rights reserved.

Page 53

## Click Path Analysis

- What routes do users take through your website?
- What events do they generate along the way?
- Are there typical paths associated with different visitor segments?
- Is there an identifiable path from page type to page type?
  - Maybe easier to detect patterns looking only at page categories (will cover in workshop)
- Identifying click-patterns is also useful for
  - Ad Promotion, Fraud Detection, ...
- BUT first we must determine which web-pages belong to the same user session....



# Sessionisation

- Determining the sequence of webpages that were visited by the same person during a single visit
  - “Visit” is a conceptual term referring to a continuous period of interaction with the website. E.g. I may be logged into Gmail all day but actually only interact with it once in the morning and once in the afternoon (e.g. after lunch) – two sessions conceptually
- Issues:
  - (1) finding all pages viewed by a single user ~ typically use cookies, unless logged-in
  - (2) finding those pages viewed in a single visit ~ typically use heuristics
    - E.g. The time spent on a page cannot exceed 30mins (time-out threshold)
    - E.g. The total session duration may not exceed a fixed threshold
    - E.g. All page-views (except the first) must be referred from another page in same session (referrer based heuristics, h-ref)
  - Usually a “big data” task => Map-Reduce frequently used
  - Personalised Sessionisation: E.g. Compute the average time a user spends on each page on the website and set the page time-out threshold according to this

## Sessionisation: Examples

Time	IP	URL	Ref
0:01	1.2.3.4	A	-
0:09	1.2.3.4	B	A
0:19	1.2.3.4	C	A
0:25	1.2.3.4	E	C
1:15	1.2.3.4	A	-
1:26	1.2.3.4	F	C
1:30	1.2.3.4	B	A
1:36	1.2.3.4	D	B

Session 1	0:01	1.2.3.4	A	-
	0:09	1.2.3.4	B	A
	0:19	1.2.3.4	C	A
	0:25	1.2.3.4	E	C
Session 2	1:15	1.2.3.4	A	-
	1:26	1.2.3.4	F	C
	1:30	1.2.3.4	B	A
	1:36	1.2.3.4	D	B

Sessionisation  
using Time-  
Based Heuristic

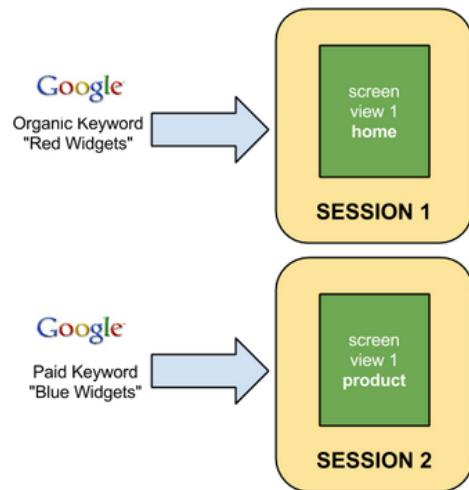
Time	IP	URL	Ref
0:01	1.2.3.4	A	-
0:09	1.2.3.4	B	A
0:19	1.2.3.4	C	A
0:25	1.2.3.4	E	C
1:15	1.2.3.4	A	-
1:26	1.2.3.4	F	C
1:30	1.2.3.4	B	A
1:36	1.2.3.4	D	B

Session 1	0:01	1.2.3.4	A	-
	0:09	1.2.3.4	B	A
	0:19	1.2.3.4	C	A
	0:25	1.2.3.4	E	C
Session 2	1:15	1.2.3.4	A	-
	1:30	1.2.3.4	B	A
	1:36	1.2.3.4	D	B

Sessionisation  
using h-ref  
heuristic

# Sessionisation in Google Analytics

- There are two methods by which a session ends:
- Time-based expiry
  - After 30 minutes of inactivity
  - At midnight
- Campaign change:
  - A campaign is what drove you to the website ~ a search engine, referring website, or campaign tagged URL (but not direct traffic)
  - If a user arrives via one campaign, leaves, and then comes back via a different campaign then this counts as a new session



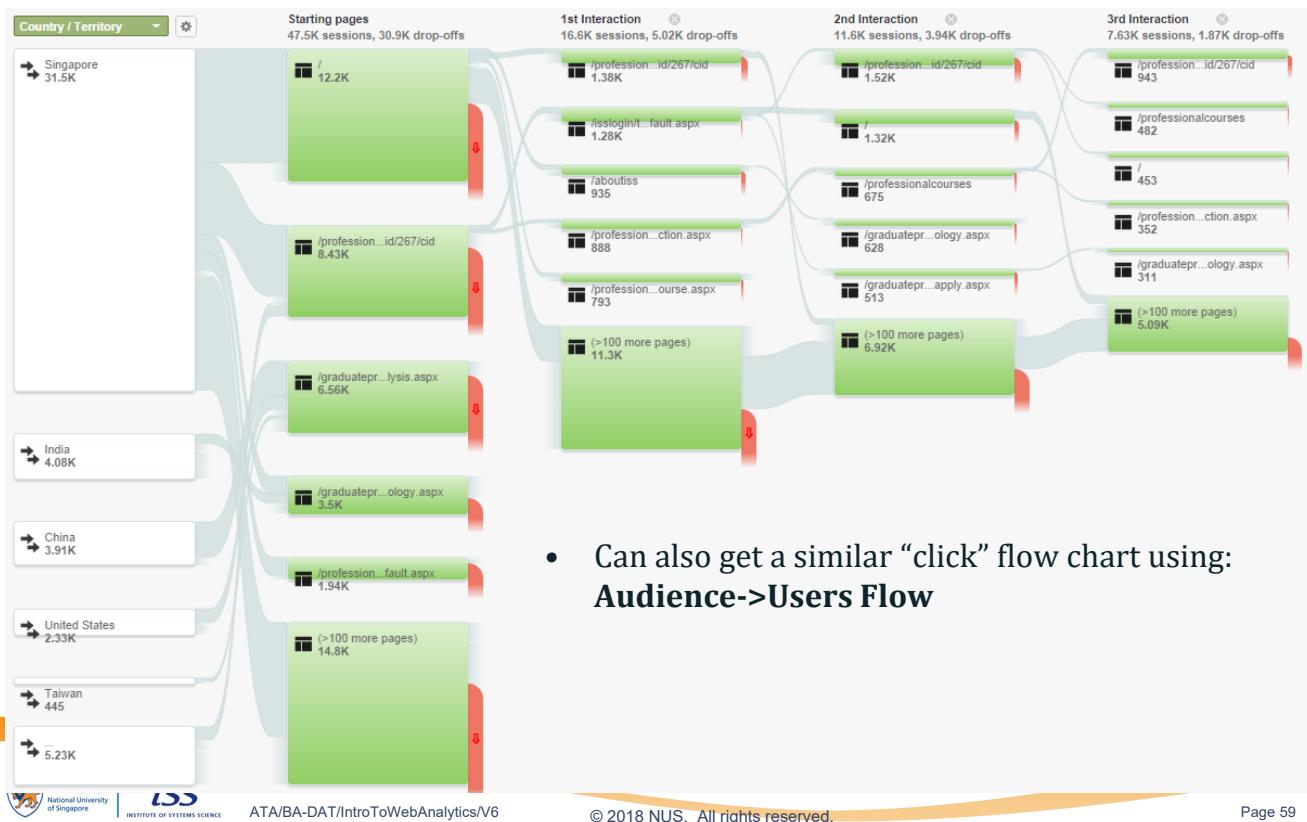
E.g. even if the time between two page-views on a website is small the session will change if the refer was outside of the site (e.g. you did a second search)

## Behavior Flow

- What are the common page routes thro the website?

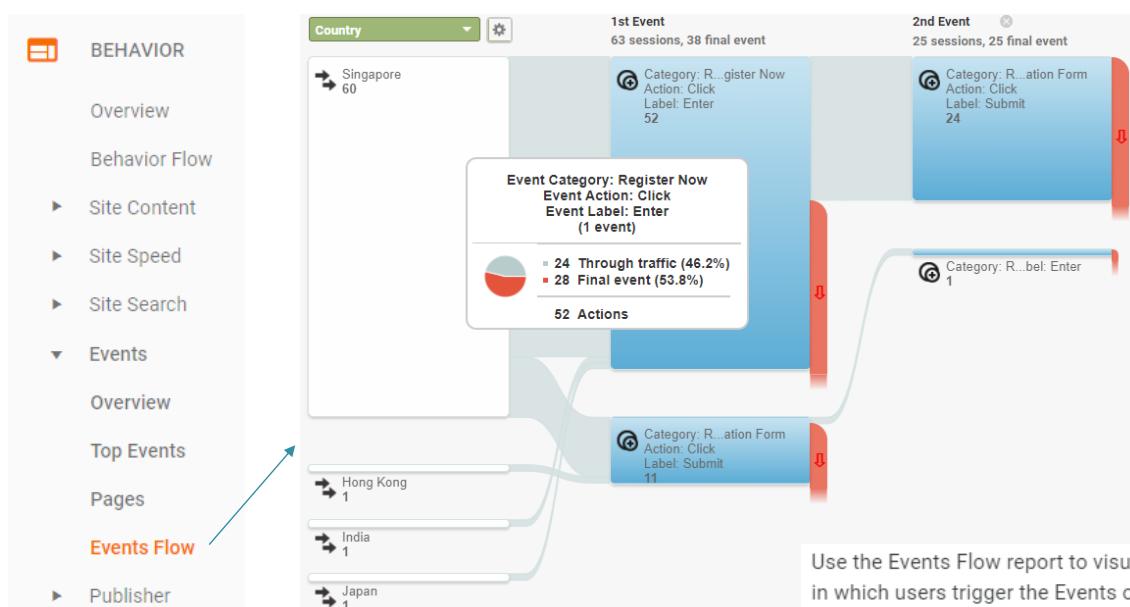


# Users Flow



- Can also get a similar “click” flow chart using: **Audience->Users Flow**

# Event Flow



*This may require more tracking code to be added to the website*

Use the Events Flow report to visualize the order in which users trigger the Events on your site. Events let you track special content like videos, downloadable items, and embedded gadgets. The Events Flow report can help you discover which Event content keeps users engaged with your site, and see the paths users take from one popular Event to the next.

# Click-Path Analysis: Navigation Summary

- Behavior->Site Content->  
->All pages-> Navigation Summary
- Select a page and see what pages users visited before and after

The screenshot shows a navigation analysis interface. At the top, there are tabs: 'Explorer', 'Navigation Summary' (which is selected), and 'In-Page'. Below the tabs, it says 'Group pages by: Ungrouped' and 'Current Selection: /graduate-programmes'. There are buttons for 'Show rows: 10' and a search icon.

Below this, there are four sections: 'Entrances' (Jan 1, 2018 - Oct 3, 2018: 32.02%), 'Exits' (Jan 1, 2018 - Oct 3, 2018: 23.77%), 'Previous Pages' (Jan 1, 2018 - Oct 3, 2018: 67.98%), and 'Next Pages' (Jan 1, 2018 - Oct 3, 2018: 76.23%).

On the left, there's a table titled 'Previous Page Path' showing the most common previous pages leading to the selected page. On the right, there's a table titled 'Next Page Path' showing the most common pages visited from the selected page.

At the bottom left are the NUS and ISS logos. In the center, it says 'ATA/BA-DAT/IntroToWebAnalytics/V6'. At the bottom right, it says '© 2018 NUS. All rights reserved.' and 'Page 61'.

## Site Navigation Issues: Site Search Analysis

The screenshot shows the NUS ISS website. At the top, there are links for myEMAIL, IVLE, LIBRARY, MAPS, CALENDAR, SITEMAP, and CONTACT. There is also a 'Login | Register as New User' button and a 'COURSE FINDER' search bar.

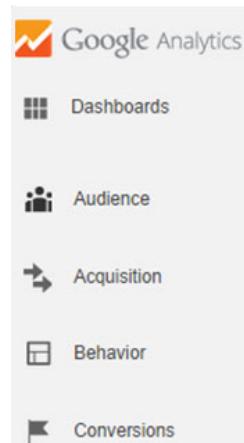
The main menu includes HOME, ABOUT ISS, GRADUATE PROGRAMMES, PROFESSIONAL COURSES, EVENTS, CONSULTING, CAREER SERVICES, and a red-highlighted 'SITE SEARCH' section which is expanded to show 'Overview', 'Usage', 'Search Terms', and 'Pages'.

To the right, there is a list of navigation issues:

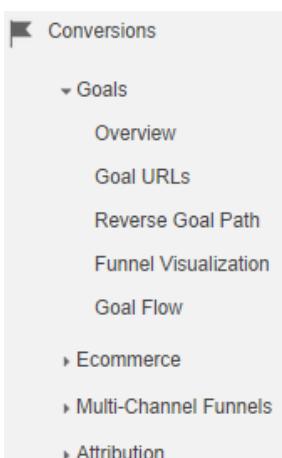
- What terms are most searched by users using your website search box?
  - Tells exactly what users are interested in
  - It can also indicate which content the user is having problems finding on your website => potential bugs in your website design
- Which site searches fail the most?
  - This could be information that your website is not providing

# On-Site Web Analytics: Main Objectives

- How popular is your website – how much traffic?
  - Basic site usage & traffic metrics. GOAL = Monitor website health
- What types of people visit your website?
  - Demographics & interests, GOAL = Reach the right people
- How did they get to (find) your website ?
  - Search engine, referral, directly etc... GOAL = Increase traffic
- What do they do on your site?
  - How long do they stay? Most/least popular pages? Is it easy to use?...
  - GOAL = Ensure the website meets their needs
- Does your website achieve its business goal(s)?
  - Measure Conversions etc.
  - GOAL = Ensure the website meets your needs



## Setting Website Goals



- What is your website for? What is its business/other goal?
- What do you want users to do when they visit?
- Distinguish the user's goals from the website goals
- Typical website goals (aka conversions)
  - E.g. make an on-line purchase, signup for newsletter, complete a survey, download a pdf file, view a video, download a coupon, get driving directions etc.
- What can be measured?
  - Reaching a page
  - Interacting with a page (watching a movie, using a widget, downloading a file etc.)
  - Remaining on the website for X minutes or visiting Y pages
  - Goal Conversion Rate ~ Percentage of users visiting the site that convert

# Defining Goals in Google Analytics

- Create goals using the Admin tab

+ NEW GOAL	
Import from Gallery	
Goal	↓
ITIL Mailer OSA	Id
ITIL RCV	Goal ID 7 / Goal Set 2
ITSM Mailer Registration	Goal ID 8 / Goal Set 2
Registration	Goal ID 6 / Goal Set 2
SHM	Goal ID 1 / Goal Set 1
Goal ID 11 / Goal Set 3	


  
**1** Goal description
 

Name	<input type="text"/>
Type	
<input type="radio"/> Destination ex: thanks.html	
<input type="radio"/> Duration ex: 5 minutes or more	
<input type="radio"/> Pages/Screens per session ex: 3 pages	
<input type="radio"/> Event ex: played a video	

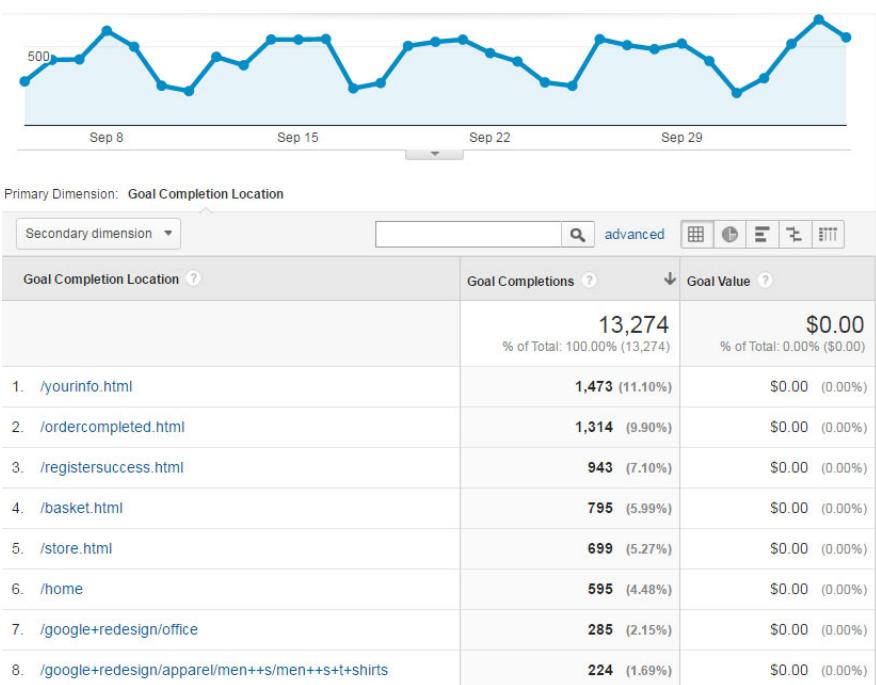
Next step
Cancel

- Ecommerce Tracking Goals
    - Transaction data: product name, SKU, category, price, quantity, tax, shipping
    - Requires more JavaScript tracking code
    - Should be tracking this anyway, but adding to GA enables tie in with website stats

# Goal Reports in Google Analytics

# Goal URL's Report

Useful if a goal can be achieved from many different pages, e.g. set a goal as "whitepaper download" but have multiple whitepapers to download each from different pages in the website



## Reverse Goal Path

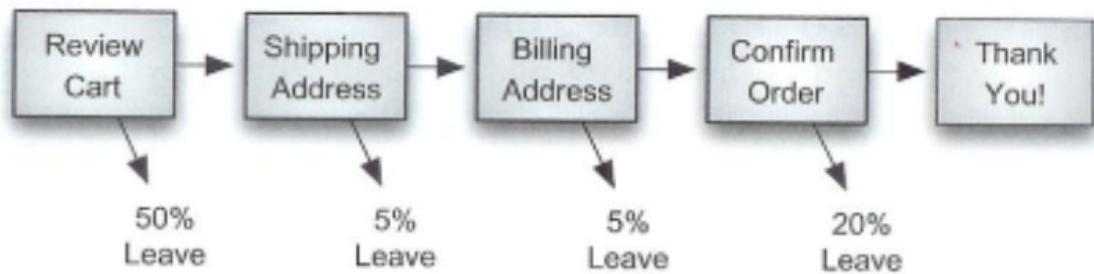
- Show the 3 pages prior to the goal page
- What pages did a user complete a goal on + the 3 pages prior to this

### Reverse Goal Path

Goal Completion Location	Goal Previous Step - 1	Goal Previous Step - 2	Goal Previous Step - 3	Goal Completions
1. /ordercompleted.html	/revieworder.html	/payment.html	/yourinfo.html	414 (3.12%)
2. /ordercompleted.html	/ordercompleted.html	/revieworder.html	/payment.html	341 (2.57%)
3. /ordercompleted.html	/revieworder.html	/payment.html	/payment.html	312 (2.35%)
4. /yourinfo.html	/basket.html	/store.html	/registersuccess.html	174 (1.31%)
5. /registersuccess.html	/signin.html	/home	(entrance)	165 (1.24%)
6. /ordercompleted.html	/revieworder.html	/payment.html	/register.html	96 (0.72%)
7. /yourinfo.html	/basket.html	/basket.html	/basket.html	87 (0.66%)

# Funnel Visualisation Report

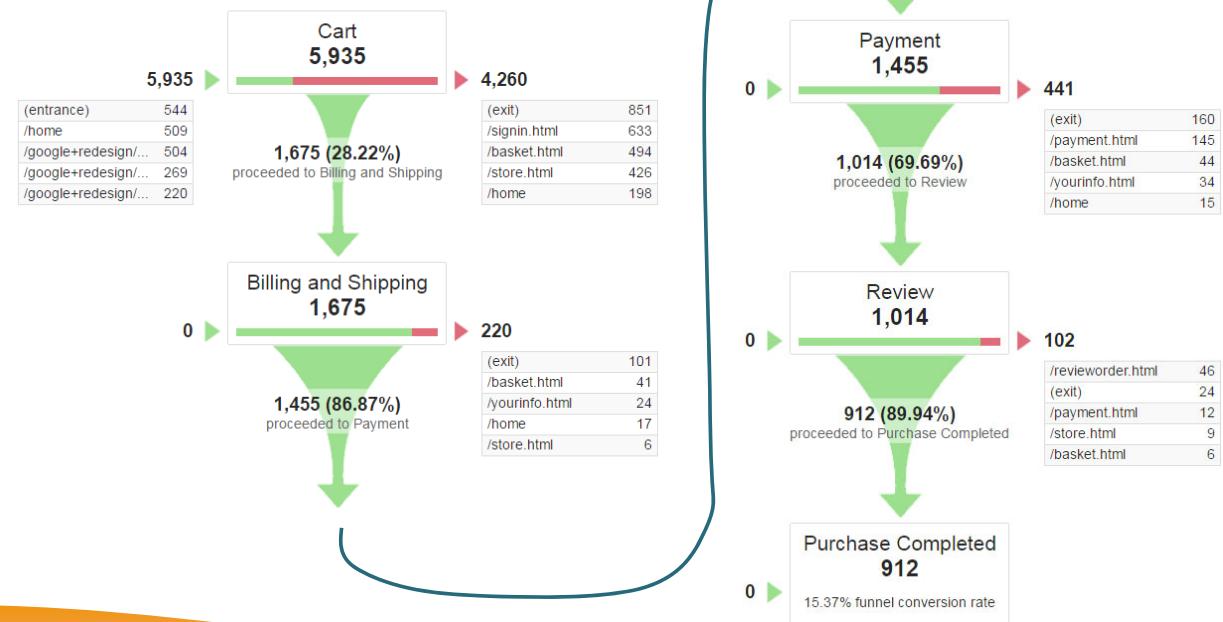
- Useful when the goal involves a sequence of page visits
- See how many people enter & complete, also how many quit before the end or enter from another (side) route



# Funnel Visualisation Report

## Purchase Completed

This Goal was completed in 912 sessions | 15.37% funnel conversion rate

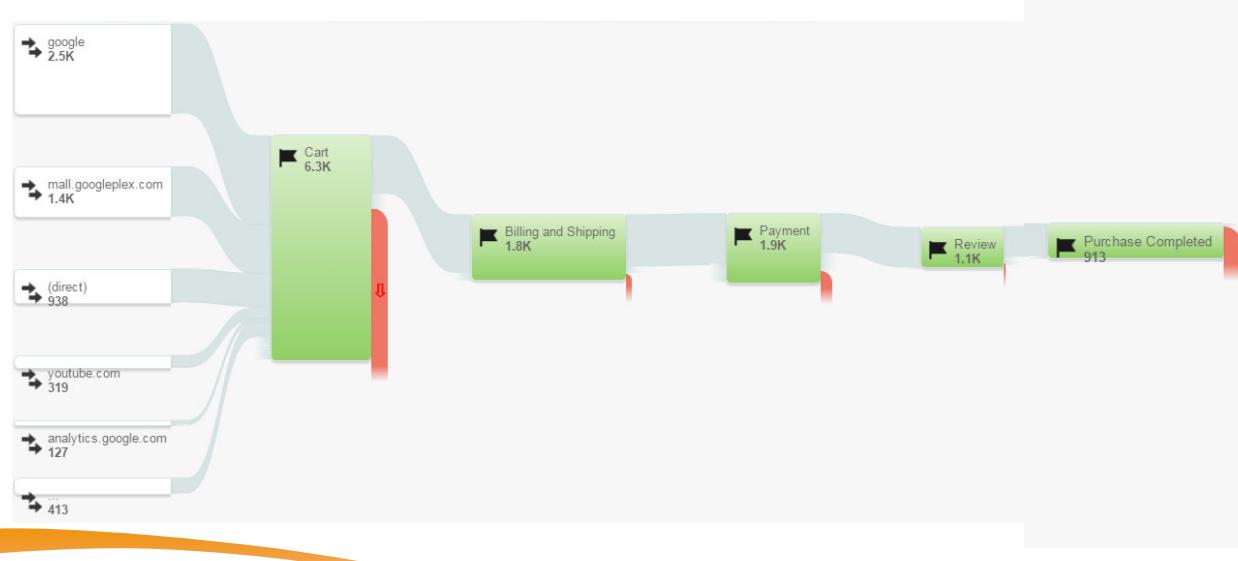


# Funnel Visualisation – Another Example

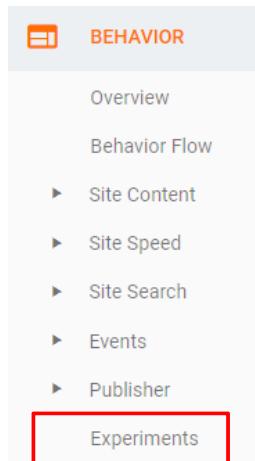


## Goal Flow

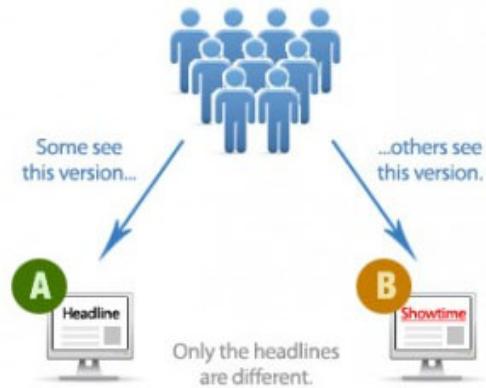
- The Goal Flow report is an interactive graphic that visualizes the path users traveled through a funnel towards a Goal. This report can help you see how people are navigating and complete a Goal



# Site Optimisation: A/B Testing



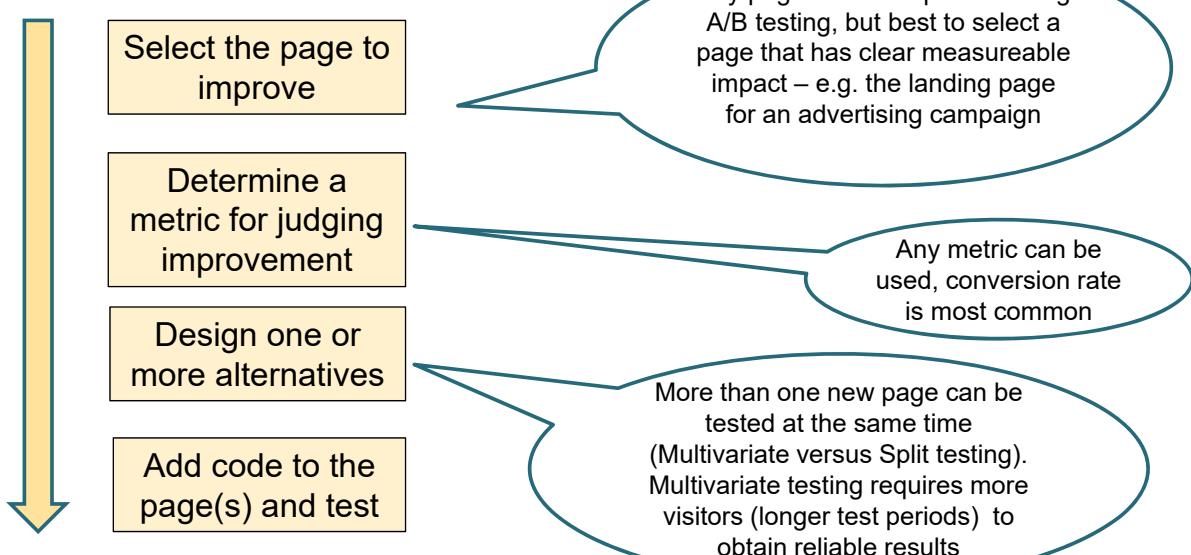
- You want to test a new design for a page – how to proceed?
- Create the new page (call it B), show it to some of the visitors to the website instead of the original page (call this A)
- After a while see which page (A or B) had more conversions (or other success measure)



## Essentials

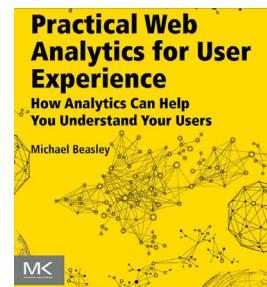
- Each visitor is randomly assigned to a variation, either A or B.
- The visitor always sees either only A or only B (even for repeat visits)
- Record which variation this visitor saw.

# Designing an A/B Test



# A/B Experimental Design Questions

- What split? (what percentage of users should see A versus B) ?
  - Normally, A = 50%, B = 50%
- How Long to Run?
  - Depends on complexity of test (number of pages being tested) and website traffic.
  - Ideally until any significant difference is observed, but clearly if there is no difference then need a cut-off time limit. You need to estimate your website traffic in advance to set a test completion date or else run until sufficient visitors have been logged.
  - A good rule of thumb for simple split test is at least 5000 visitors to the website\*
  - For Multivariate testing a good rule of thumb is to add 2000 visitors per day for a one month test for every additional page being tested. Hence a 4-way test would require about 8000 daily visitors over a month to complete\*



See chapter 13

## Testing Significance

- Most methods use the z-score, students t-test or welches t-test

$$z = \frac{\bar{x} - \Delta}{\frac{\sigma}{\sqrt{n}}}$$

$\bar{x}$  is the sample mean.  
 $\Delta$  is a value to be tested.  
 $\sigma$  is the population standard deviation  
 $n$  is the size of the sample



Look up the significance level of the z-value in the standard normal table

- E.g.

$$ZScore = \frac{p_{variationB} - p_{variationA}}{\sqrt{SE_{variationA}^2 + SE_{variationB}^2}}$$

$$SE = \sqrt{\frac{p(1-p)}{sampleSize}}$$

Where P ~conversion rate

# Testing Significance: An Example

- Original page (A): 200,000 unique visitors with 10,000 converting
- Test page (B): 100,000 unique visitors tested with 4000 converting
- Lift of A over B  $\sim 25\%$
- If A had only **8,100** visitors converting => conversion rate of 4.05%
- Lift of A over test B =  $(4.05 * 100) / 4.00 = 1.25\%$

**Group A is a clear winner (99.9%\*)**

Group A: **5%** conversion

Group B: **4%** conversion

	Visitors:	Goals:
<b>Group A:</b>	200000	10000
<b>Group B:</b>	100000	4000

**Calculate**

**Sorry, you have no clear winner**  
We estimate that you'll need 1583527 more visitors\*

Group A: **4.05%** conversion

Group B: **4%** conversion

	Visitors:	Goals:
<b>Group A:</b>	200000	8100
<b>Group B:</b>	100000	4000

**Calculate**

See <http://www.usereffect.com/split-test-calculator>

## A/B Testing in Google Analytics

Behavior

Overview

Behavior Flow

Site Content

Site Speed

Site Search

Events

AdSense

**Experiments**

In-Page Analytics

**Content Experiments - Create a new experiment**

1 Choose an experiment objective

Name for this experiment  
**Untitled experiment**

Objective for this experiment  
Select a metric - OR - [Create a new objective](#)

Percentage of traffic to experiment  
100% 

**Advanced Options**

Distribute traffic evenly across all variations  
**OFF**

Set a minimum time the experiment will run  
2 weeks 

Set a confidence  
95.0% 

This controls how many people who visit your site see one of your test pages, including your original page. Everyone else will see your original page. If you want quicker results, you might want to include a higher proportion of your visitors in the experiment. However, if your experiment is drastic and risky, you might want to include only a small proportion of your site's visitors.

Enable this option to assign an equal amount of traffic to each variation for the life of the experiment. When this option is disabled, Content Experiments follows the default behavior of [adjusting traffic dynamically based on variation performance](#).

A/B testing in GA is being replaced by Google Optimize (Google's custom A/B testing and personalization tool), which launched in 2016. the interface and capabilities are relatively similar

Based on “multi-arm bandit” experiment analysis. See:  
<https://support.google.com/analytics/answer/2844870>

# Google Analytics Academy

## Analytics Academy Courses

 <b>Digital Analytics Fundamentals</b>	 <b>Google Analytics Platform Principles</b>	 <b>Ecommerce Analytics: From Data to Decisions</b>	 <b>Mobile App Analytics Fundamentals</b>	 <b>Google Tag Manager Fundamentals</b>
Learn the core principles of digital analytics including how to create a Google Analytics account, collect data, and navigate reports.	Learn how the Google Analytics Platform collects, configures, and processes the business data you need for reporting and analysis.	Discover how to use Enhanced Ecommerce reports in Google Analytics to make informed ecommerce business decisions.	Learn how to use Google Analytics to attract new users to your mobile app and increase revenue.	Discover how Google Tag Manager can simplify the tag implementation and management process for marketers, analysts, and developers.
<a href="#">LEARN MORE</a>	<a href="#">LEARN MORE</a>	<a href="#">LEARN MORE</a>	<a href="#">LEARN MORE</a>	<a href="#">LEARN MORE</a>

<https://analyticsacademy.withgoogle.com/>



ATA/BA-DAT/IntroToWebAnalytics/V6

© 2018 NUS. All rights reserved.

Page 79

# Google Analytics Community

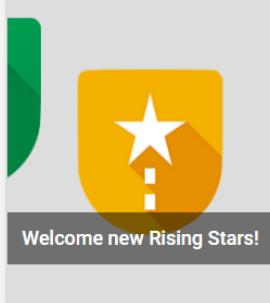
Welcome to the Official Google Analytics Community 3.1K members online now

People like you are finding answers, learning about new features, getting inspired, and helping others. Whether you are new to Analytics or an expert, your voice matters.

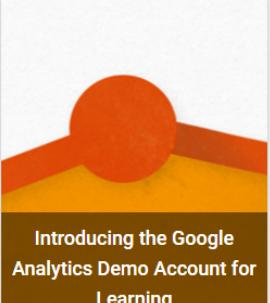
To complete your community registration, please accept the [Google Terms of Service](#) and the [Community Guidelines](#). X

[Register](#)

FEATURED CONTENT



Welcome new Rising Stars!



Introducing the Google Analytics Demo Account for Learning



Common Google Analytics pitfalls (and how to avoid them) - by Top Contributor

[https://www.en.advertisercommunity.com/t5/Google-Analytics/ct-p/Google\\_Analytics#](https://www.en.advertisercommunity.com/t5/Google-Analytics/ct-p/Google_Analytics#)



ATA/BA-DAT/IntroToWebAnalytics/V6

© 2018 NUS. All rights reserved.

Page 80