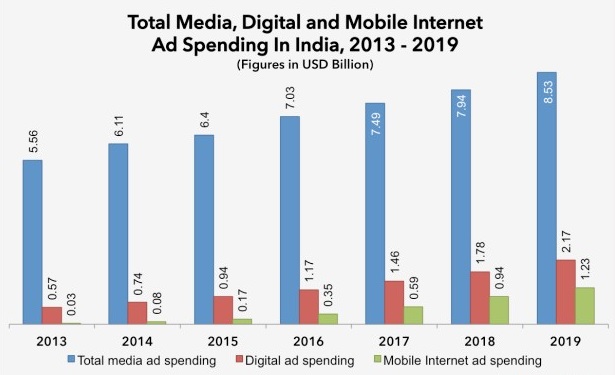
**A Comprehensive Guide to Digital Marketing and Analytics Every Data Science Professional Must Read**

## Introduction

One of the biggest challenges of breaking into the field of digital analytics is that the landscape of digital marketing is extremely complex. It’s a hard task finding professionals who know the best of both worlds – digital marketing and data science. There is a serious shortage in the supply of adequate talent, while spending on digital marketing continues to rise unabated.

This spending is quite prevalent in the developed economies. But here’s the good news – developing countries are starting to catch up, and are not far behind the curve. Check out the below chart and see for yourself the growth in the digital marketing spend share over time in India:



**Figure source : eMarketer**

I have seen positions open for years in the digital analytics space because of the shortage of such niche talent coupled with the immense growth of this field. Over the last few months, I have spent some time trying to understand the digital marketing landscape and how it integrates with our field of data science. In this article, I have shared my compiled notes that should simplify this complex world of digital marketing and help you understand how you can use your data science tools in this terrain.

Excited yet? Good. Let’s start with the most basic question and take it from there!

## Table of Contents

* What makes Digital Media better than Traditional Media?
* As a Digital Analytics Professional, why should I even care about Digital Media?
* Sub-channels of Digital Media
* Consider a Real-World Scenario
* Google Search Network vs. Google Display Network
* How does Google choose ads for every search and their Placements?
* Adwords (now Google Ads) vs. Adsense
* DoubleClick for Publishers (DCP) vs. DoubleClick for Advertisers (DCA)
* DoubleClick Search vs. DoubleClick Bid Management System
* Cookies, Pixels and Tag Containers
* DSP vs. SSP vs. Ad Exchange
* Data Management Platform
* Identity Resolution
* Digital Analytics Tools

## What makes Digital Media better than Traditional Media?

Marketing is all about getting 4 things right – reaching out to the right **customer** with the right **product** at the right **time** through the right **channel**. Marketing also needs a lot of testing to see which combination works for these 4 factors, so we need a channel that takes minimum time to hit the market. Traditional media channels like cable TV, flyers, radio, physical banners, etc. present a lot of challenges getting all of these 4 things right. To name a few:

1. Once printed or aired, an advertisement cannot be changed.

2. High time to market. This is actually a big challenge for any dynamic industry. A lot of offers require a quick response time (like Amazon Prime sales) and advertisers don’t wish to announce the sale beforehand to make sure customers regularly come to their website to check out offers.

3. No way to accurately measure viewed ads. All we know is the response from ad campaigns.

4. Limited ways to reach out to a specific audience.

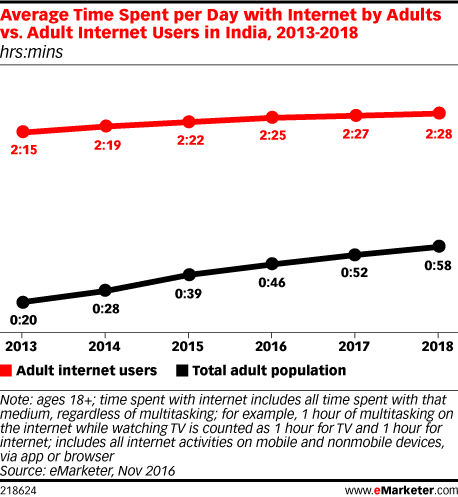
5. Even if we manage to reach the right customers, we might not really reach out at the right time. For instance, if we advertise during specific TV shows that are viewed by our target audience, we don’t really know whether the target customer is looking at the TV at the time when our ad is being played. Similarly, if we send a mail to a specific prospect who is traveling at the time, it’s a lost opportunity.

6. Given a low response from these channels, most of these traditional methods have a high cost per acquisition.

Digital marketing is able to overcome all the challenges mentioned above.”How” will become clearer as we cover other related topics in this article. With internet penetration reaching 60%+ around the globe and rapidly increasing (see chart below), digital media is now an ideal place to catch your audience’s attention.

Today, we can reach out to 4 billion+ internet users through digital media. Another survey (below) indicates 2 hr. 28 mins spent per day by adult internet users in India. Noting that average time spent by the total adult population is 1 hour, total internet user average should be somewhere between 1 hr to 2.5 hrs (say, 2 hours). An estimate of the ad opportunity available per day on the internet is 4 Bn\*2 hours = 8 Bn hours/day.

**Even if we assume only 1% of this internet time can be monetized with ads, we are still talking about 80 Million hours/day!**



**Figure source : eMarketer**

This exercise would have given you a sense of the magnitude of opportunity in digital marketing. And as an analyst, you would have definitely noticed the incremental curve for both internet users and the average time spent on the internet. Given the scale and precision of this channel, digital media presents multiple methods for advertising. Let’s broadly classify these sub-channels of digital media.

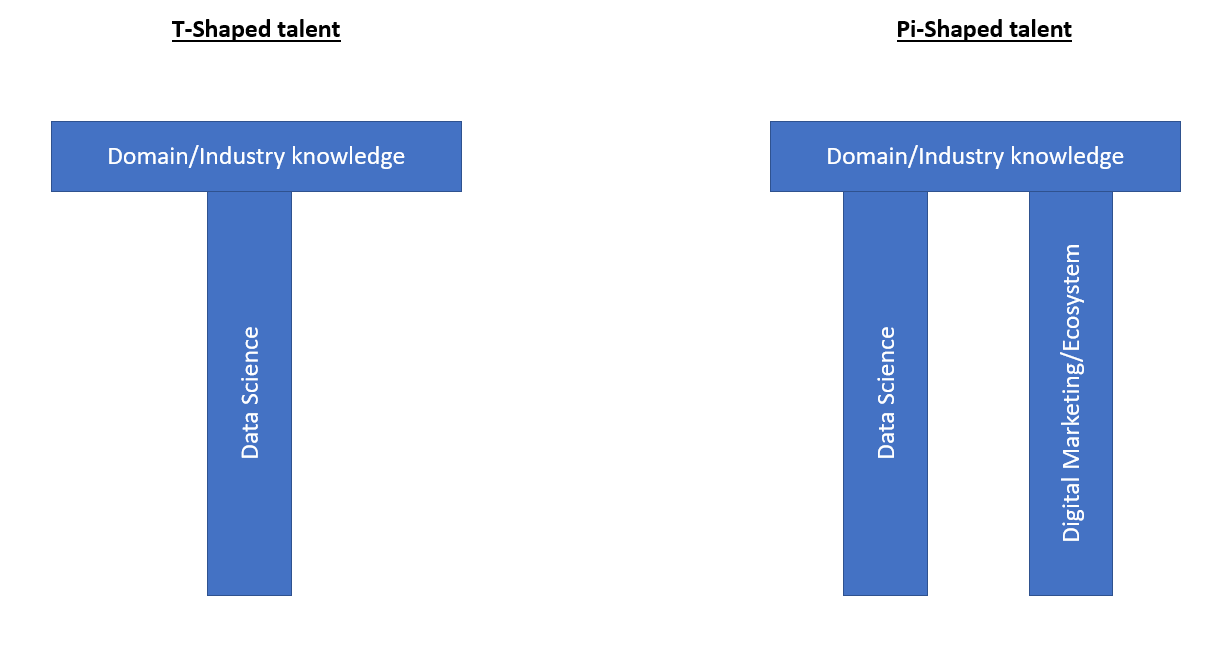
## As a Digital Analytics Professional, why should I even care about Digital Media?

A data scientist can contribute literally in any domain and any industry. For instance, there are data science roles in creating new medicine, intelligent voice assistants like Google Home, targeting strategy, etc. For each of these roles, a data scientist might need minimal knowledge of the domain, but can still make strong contributions with the tools we have. For instance, to contribute towards building a new medicine, you need not know all the biology behind it. All you need to know is an objective function and the parameters you need to optimize. In these scenarios, you can easily break down the work into what the business person needs to tell and what you need to know already.

However, digital analytics is slightly different. Given the fast changing dynamics in the industry, we are getting better data everyday. As an analyst, you need to decide:

* What data is useful?
* How to best capture this data for accuracy?
* How to use the tool(s) you have created with data science?

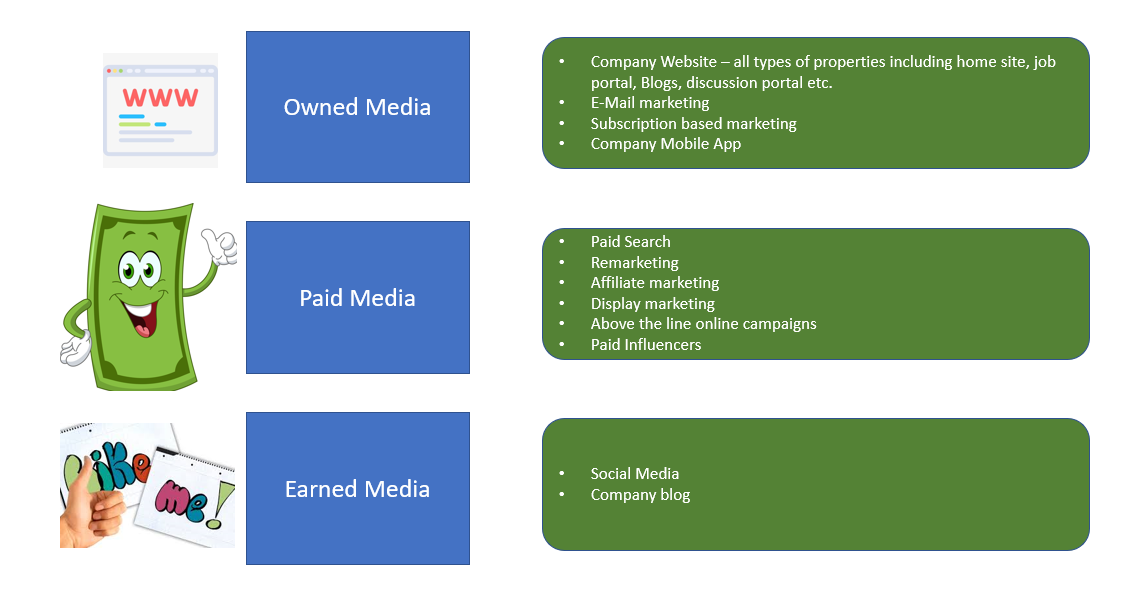
The barrier between a business person and an analyst becomes very blurry in this space. Here is how companies describe this new skill set they need for digital analytics:



The T-shaped talent suffices for most data scientist roles in different domains. However, digital analytics professionals need the Pi-shaped talent to be a rockstar. If you already understand the world of data science, then this article will help you grasp the second aspect of the Pi-shaped talent requirement, the digital ecosystem, and open up a land of opportunities for you. In no way does this article capture everything under the digital marketing umbrella, but this is as good a place as any to get your journey started.

## Sub-channels of Digital Media

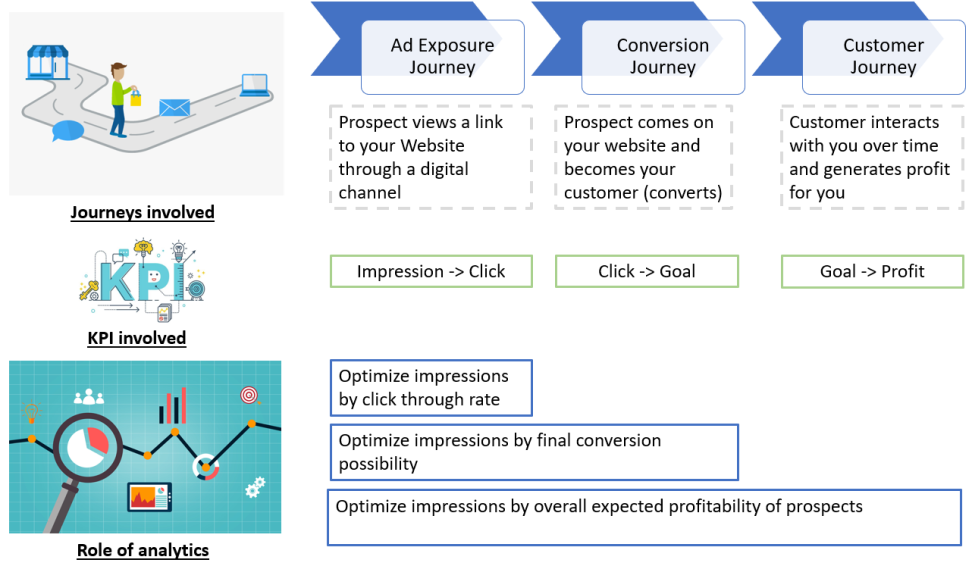
We generally talk about 3 types of digital media – owned media, paid media, and earned media. Here’s a description of each:



Data science has a strong application in each of these 3 types of media. Before we get into the actual applications of data science, let’s first broadly break down the journey of a prospect into stages.

**The Different Stages of Customer Acquisition**

There are three journeys involved in customer acquisition.



*“Our end objective of a successful campaign is to show ads to only those prospect that have a strong propensity of being a profitable customer.”*

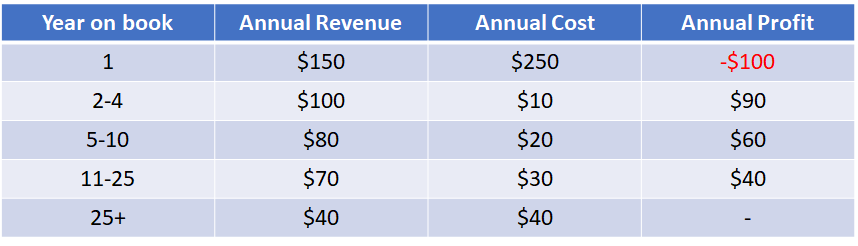
Looks simple, right? But in the real world, it really isn’t! Here’s why:

* The **ad exposure journey** is normally managed by digital media/ad serving companies like Google, Facebook, etc.
* The **conversion journey** can be managed by either your in-house team(s) or outsourced digital analytics firms like Adobe Site-catalyst and Google
* Finally, the **customer journey** is managed by your in-house analytics team as this is highly confidential company information.

As you might have guessed, it becomes very difficult to collaborate between all the three stages to reach ensure a successful campaign. Let’s look at a a real-life scenario that will make this challenge a lot more clearer.

## Consider a Real-World Scenario

Company X sells life insurance and acquires 50% of customers through paid digital media. A typical life insurance starts low on profit because of all the overhead costs involved, and gradually increases to maximum profit within 2 to 4 years. Year 5 onward, good customers start leaving the insurance to pick up better deals in the market, and hence the profit starts to decline. See this illustrative table that shows year wise profit:



Just relying on the average can so often be deceiving. It seems like every acquisition is worth (lifetime value) $1,130 to your company, but all customers are different and as analysts, we try to find the most profitable segment. Now suppose you found that there is one attribute, total number of insurance policies held by the customer till date, that can strongly separate highly profitable customers from others. Here is how simple this rule can be:

For customer with #insurance policy <= 2 => Customer lifetime value  = $4000

For customer with #insurance policy > 2 => Customer lifetime value  = -$50

Now, here is the challenge. To stay compliant with the regulatory authority, you cannot deny a policy just because of profitability reasons. Hence, the only way you can craft your acquisition portfolio is if you can control who is looking at your ads. For a simplistic view, let’s say 50% of your acquisition comes from your website directly and the remaining comes through Google’s Paid Search. Acquisition coming directly to your site is hard to control as these prospects are already down the conversion funnel.

So all you can do is control who looks at your paid search ads. Controlling paid search targeting looks straight forward – just tell Google to show ads to those people with #insurance policy <= 2. But there are a few challenges in doing this:

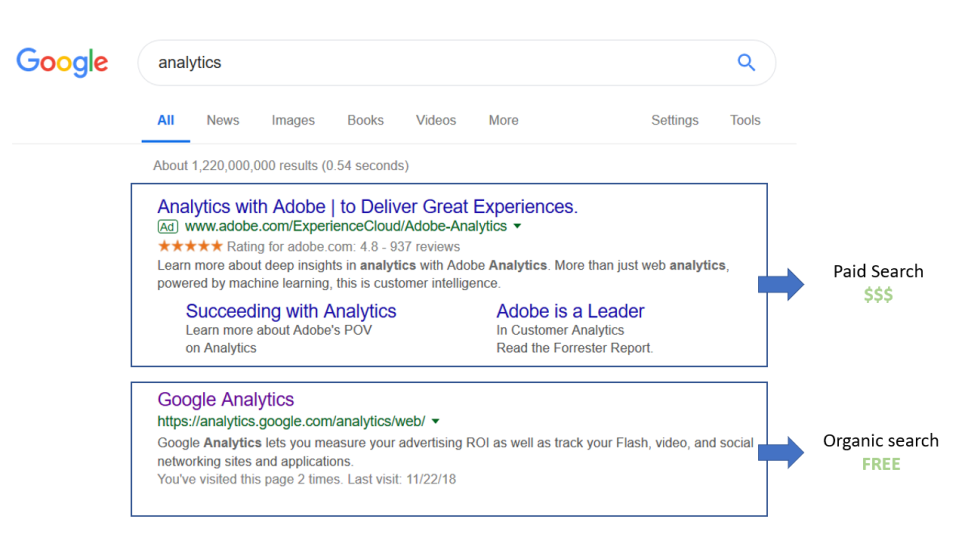
1. This rule is your intellectual property and you will definitely not want to give this gem to an aggregator that is being approached by other insurance competitors
2. Google might not really have this one variable to apply your rule
3. Google will not be open to sharing their variable on impression level to help you make this rule with the variable that is available to Google

So you see the challenge – the variable that is available for real-time targeting by ad server companies is not available with you to link profitability. Similarly, companies cannot make their profitability linked information available to Google to do specific targeting. Even though such direct targeting is difficult with paid media, we still have a lot of ways to get around this challenge.

We will talk more about these targeting audience methods later in this article. Before we move on to targeting audience methods on digital media, let’s get some jargon out of the way. The best way to understand these jargon terms is to put them in contrast against each other.

## Google Search Network vs. Google Display Network

When a common person thinks about Google, we think about the Search Engine (well, at least I do!). When we search for a keyword, we see a list of links. Advertisers have to pay for sponsored ads that appear at the top of the rankings. The list of links that come in organic search are free of cost and the rank is determined by Google’s proprietary Page Rank algorithm. In the picture below, Adobe has paid for the sponsored ad, whereas the Google Analytics link below it is an organic search.



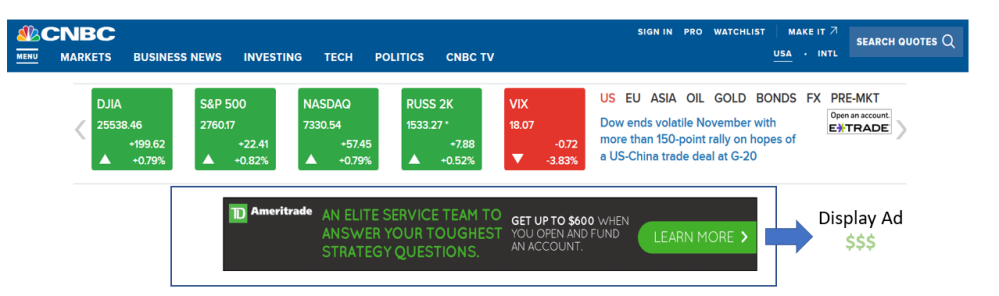
The Google Search network is far bigger than just the search engine. It currently includes:

* **Google Properties:** Google Search, Google Play, Google Shopping, Youtube and Google Maps
* **Other partners:** Amazon, The New York Times, W3Schools and Google custom search sites

**What about Google Display Network (GDN)?**

In addition to the search network, Google also partners with 2 million sites that can publish ad banners on their site. GDN has a huge coverage of about 90% internet users. The biggest challenge with GDN is the low response rate because these visitors are not really looking out for your product. These visitors are simply looking at news, or the weather or are watching a video, etc.

So, we classify such a visitor as a top-of-the-funnel prospect as he/she has a long way to go before making a purchase. Another challenge with this display network is attribution. Because a customer might view your ad today and make a purchase significantly later, attribution of these purchases becomes very subjective. GDN is very successful for brand awareness and multi-touch acquisitions (which requires a customer to see your product’s ad multiple times before making the final purchase). You might see GDN ads on your favorite sites if they partner with Google. Here’s an example:

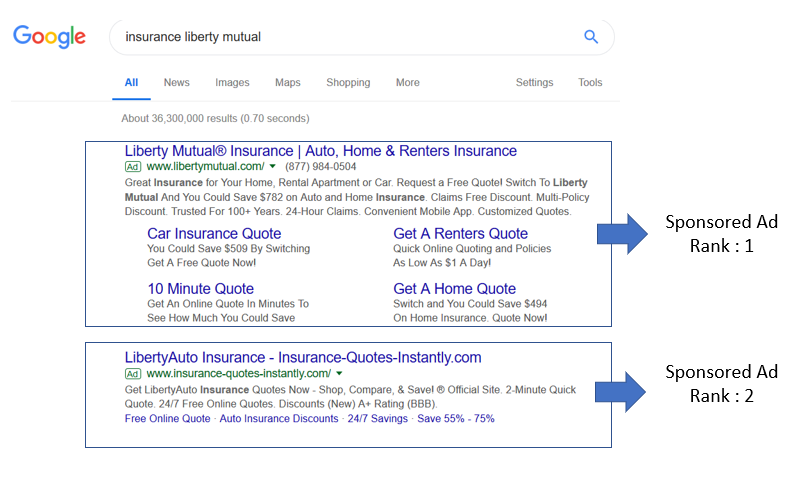


You should now have an understanding of Google’s core business of advertising. This is the majority revenue generator for Google. Every company wants to partner with Google to advertise their product/service. But how does Google choose which ads have to be displayed and when? Let’s review the Google Ad auction briefly before we jump to the digital marketing jargon.

## How does Google Choose Ads for Every Search and their Placements?

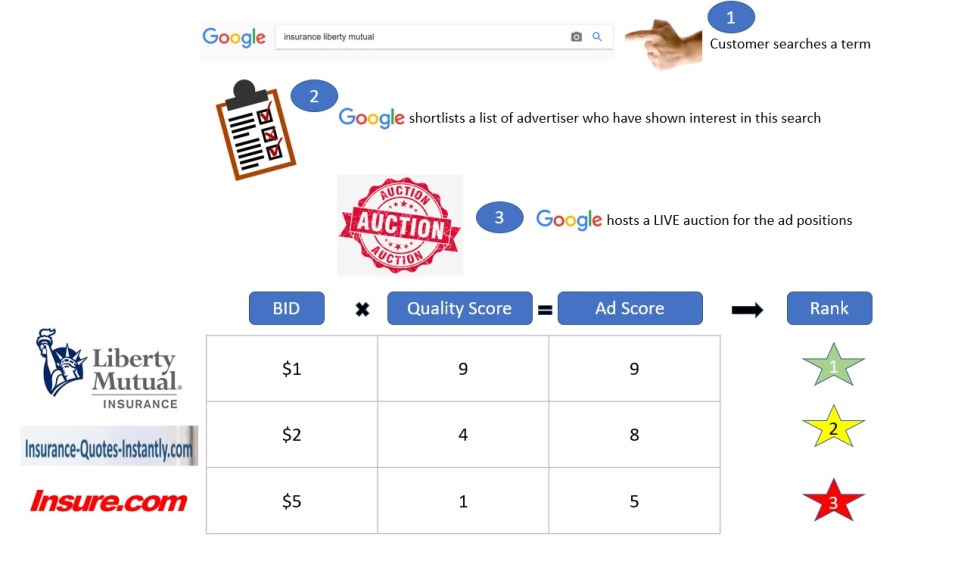
You’ll be amazed to know how Google does this. Let’s try to understand the logic through an illustrative example.

I searched for the term ***“insurance liberty mutual”***on Google Search. At this point, Google hosts a live auction for this one instance ad inventory. This auction will be conducted in mili/microseconds and the winners will get the top positions on the sponsored ad slots. Is a high bid all you need to get the top spots? Of course not. Because my search was specifically for liberty mutual insurance, no other competitor should ideally get the top spot. Otherwise Google’s customer will have a hard time getting to the relevant links (the core skill Google offers). Following is the result for an actual Google search of the term:



As you can see, “insurance-quote-instantly” also participated in this auction but did not win against liberty mutual. So how does Google bring in this dimension of relevancy to remove monetary bias? Google has a concept of quality score for each rank which is kind of a page rank score on organic search. This quality score is then multiplied by the bid to calculate the ad score. This ad score is finally used to rank order the ads for a search instance.

Here is a simplistic view of what goes on behind the scenes and how advertisers are charged for their ads:





Couple of things you should notice in this Google Auction model:

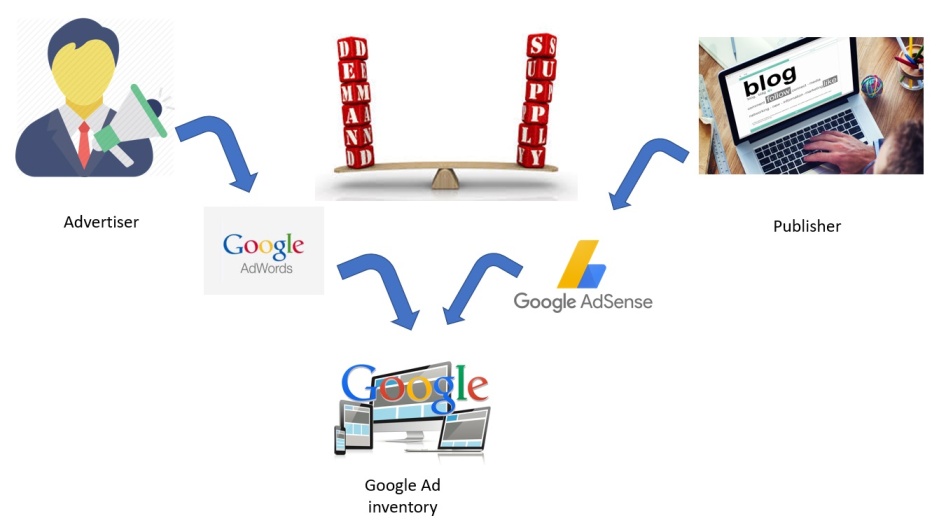
1. You don’t need to always bid very high to win an ad rank auction. For instances where people are searching for you specifically, you will win the bid with minimal cost.
2. You always pay the minimum amount you need to win the bid war. Even though Liberty Mutual would have bid millions of dollars in this auction, they would have paid only $0.89
3. It is easy to defend your proprietary keywords with relatively low price because of the high quality score

**CPC vs CPM bidding** – Quick note of the types of bidding we do on both search and display networks. CPC bidding is when the advertiser wants to bid on how much they are willing to pay for a click. CPM bidding is based on per 1000 impressions. It really depends on the type of objective an advertiser is looking to achieve. If the objective of the advertiser is branding, all that matters is impressions. However, if the advertiser wants customers to progress in the conversion funnel, they will generally bid on CPC.

One exception here are multi-view purchases, where a customer generally takes a decision after looking at an ad many times. In such cases, the advertiser might bid on CPM even when the objective is conversion.

## Adwords (now Google Ads) vs. Adsense

Simply put – Adwords is used by advertisers to post ads on Google’s display and search networks. Adsense is used by publishers to monetize their content. Adwords manages the demand side of ad inventory whereas Adsense manages the supply side. The below picture illustrates this concept well:



Even though Google does the heavy lifting for both the publisher and advertiser in terms of ad serving, it still provides many tools to match the right ad to the right placement. This is where analytics plays such a crucial role. Google delivers Ads to Ad Spaces with publishers through 3 methods:

1. **Contextual Targeting:** Google Bot crawls through the page and uses Natural Language Processing (NLP) to categorize the page by analyzing the article topic. Only text is analyzed and images/videos are simply ignored
2. **Placement Targeting:** Bidding based simply on placement. Here, advertisers hand pick publishers where they wish to show their ads, and simply bid for placements
3. **Interest-based Advertising:** Finally, an advertiser can target ads based on inferred interest. This one is specifically targeted based on the visitor and not the content. For instance, if a user visits an airline website and then comes to Analytics Vidhya. Even though flight booking ads are not really relevant for AV content, they can still come up based on Interest-based Advertising

What levers do ***Publishers*** have to subset in the pool bidding for their ad space?

1. **Exclusion of specific category ads:** For instance, a baby blog might not really like ads from dating sites. These categories are generally very sensitive, like political views, adult content, etc.
2. **Exclusion of specific bidders/domains:** You can exclude your competitors
3. **Create audience:** Publishers can also create audience if they see very strong patterns. For instance, you might have a tie up with an airline aggregator and will want to show airline ads to those people who already visited this aggregator, in order to nudge their buying behavior. You can also create complex variables
4. **Choosing placements and allowed sizes:** You can choose the exact placement and size of your ads. These can be above or below the fold. Above the fold does not need any scroll downs, whereas below the fold only appears on the visitor’s screen if he/she scrolls down. You can also choose if you want to include ads that are smaller than the size you defined in the auction
5. **Text vs Visual Ads:** You can also choose whether you want only text to appear in the ad window, or visuals like image/video. You can further choose the background color of the ad slot and other formatting options

What levers do ***Advertisers*** have to subset in the pool bidding for their ad space?

We will cover this topic in our section on Google Analytics later. For now, just note that you have levers on Demographic and Behavioral attributes you can use to target ad inventory. Further, you can choose keywords to make a precise selection of the auction you wish to participate in.

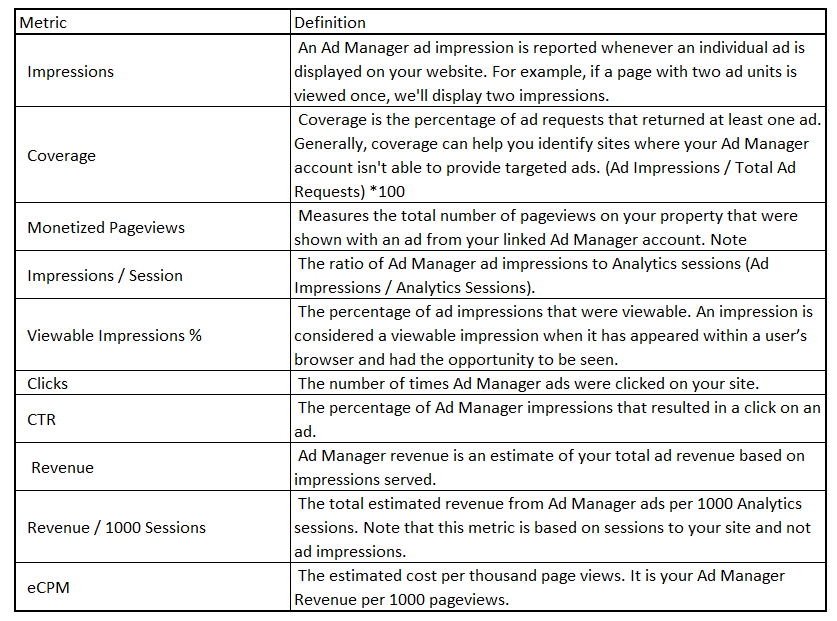
## DoubleClick for Publishers (DFP) vs. DoubleClick for Advertisers (DFA)

DoubleClick is an ad-serving company that was bought by Google for $3.1 billion in 2008. In simple words, DoubleClick for Publishers makes it easier for publishers to monitize their content. It also provides effective tracking of how your content is performing in context of advertisements.

DoubleClick for Advertisers, on the other hand, helps advertisers optimize their Search and Display campaigns. Google recently rebranded DFP as Google Ad Manager. DFP comes out much stronger than AdSense when you wish to manage your ads beyond Google Display Network – including Affiliated Brands, Real-time Ad exchanges, etc. Let’s see a few examples of how we analysts use the ad tracking done by DFP.

Suppose you own a travel blog and are using DFP to manage your ad inventory. There are 4 primary sections in your blog – Food, Destinations, People and Latest Deals. On each blog page, you have 3 banners – Leaderboard, Skyscraper and Square. The below picture will help you visualize each of these positions on your blog page:

The below table covers the key metrics that DoubleClick publishes by default:

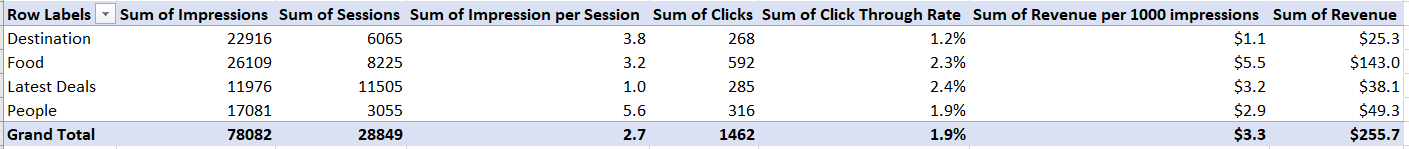


If you use DFA with Google Analytics 360, you will get endless dimensions. However, Google Analytics 360 requires you to part with a significant amount, not something everyone can afford. So let’s stick to some of the basic dimensions and what we can do with them. Here are a few dimensions which you can leverage to analyze your site’s performance with respect to ad revenue:

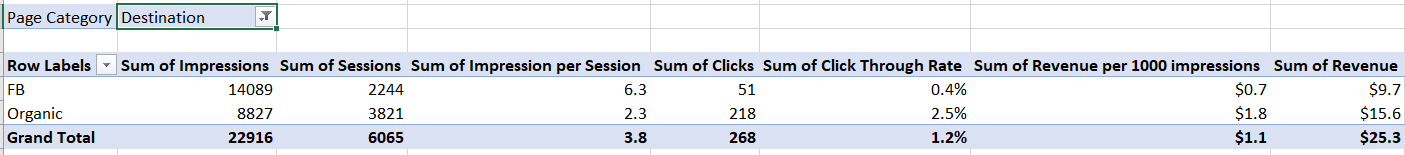
1. Page name
2. Page category
3. Ad slot number
4. Source/Medium

These dimensions are just examples of dimensions available to you for slicing and dicing information. Let’s take a look at a dimension metric view to learn more.

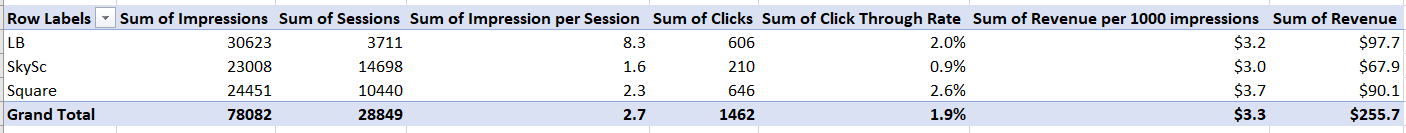
Starting with a basic view to see how each category is performing:



Clearly, the food section brings in the majority of the revenue (50%+) even though the destination category gets most of the impressions. The destination category has both click through rate (CTR) and revenue per 1000 impressions at the lowest value, indicating our ads are not being optimized well for this category. Seems like quite a huge opportunity, right? Breaking down the destination category further by the traffic sources gives the following results:



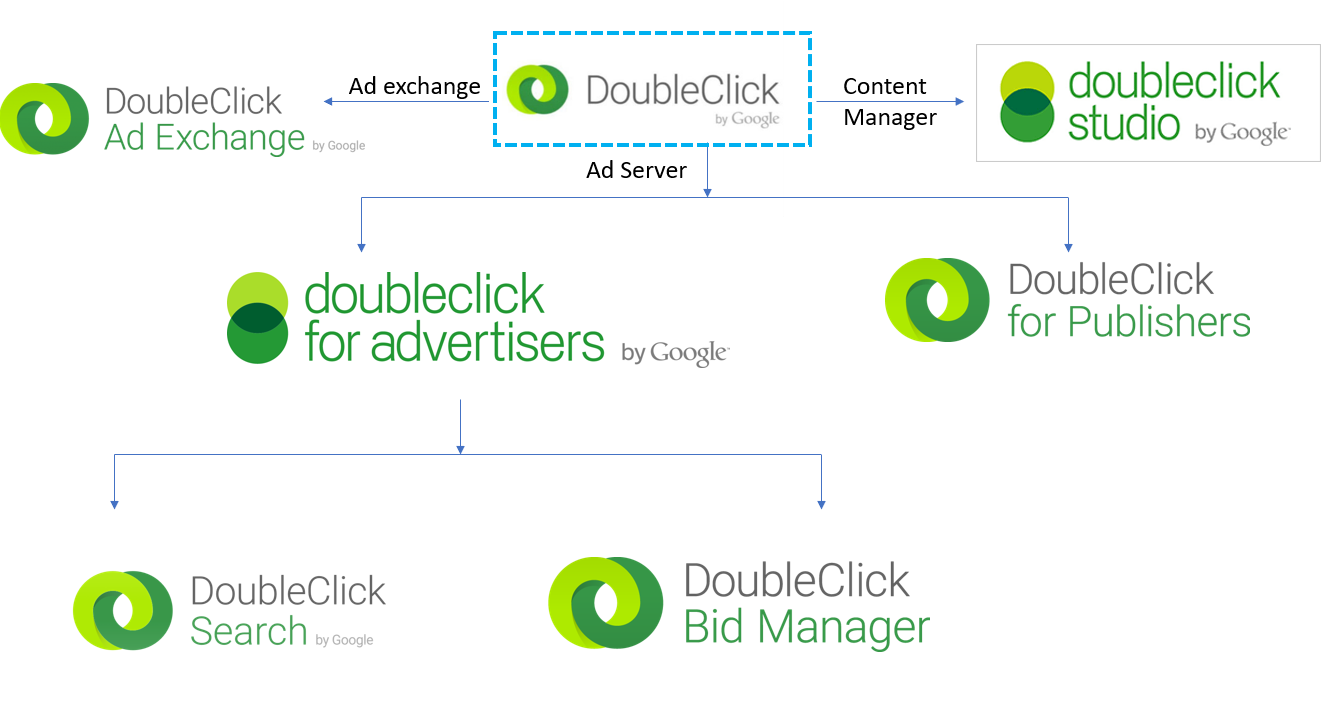
The above table shows that FB is the major source of traffic for our destination category. However, this source performs sub-optimal on both CTR and Revenue. This narrows down our search further. Are we presenting different information than what customers coming from Facebook are expecting? We can look out for this information by checking the bounce rate for this audience. We will hold that discussion for later. Let us also review how each banner type is performing across pages:



Leaderboard (LB) has the maximum number of impressions, which makes sense as it comes on the top and should appear during most visits. Skyscraper has a low CTR and revenue per 1000 impressions, indicating these banners might not be completely visible (definitely scope of improvement there). We can further deep dive into this analysis with DoubleClick for effective targeting, but we’ll keep that for a future article.

## DoubleClick Search vs. DoubleClick Bid Management System

DoubleClick is a suite of product provided by Google. The below diagram will help explain the types of services DoubleClick provides in the world of digital media:



DoubleClick Search is primarily used by advertisers to manage their ads on multiple Search Engines, including Google Search Network. DoubleClick Bid Manager is used to manage Display Ads across the Google Display Network and Real-Time Bidding platforms. DoubleClick Ad Exchange is like the New York Stock Exchange where ad inventory is bought and sold. DoubleClick also provides campaign management services like DoubleClick Studio.

## Cookies, Pixels and Tag Containers

Cookies are small text files that are placed on a visitor’s machine through websites they browse. A cookie contains some key information that can be used when the visitor returns. For instance, a lot of sites use cookies to save ID and passwords. Others use them to refill the checkout cart when the visitor returns. These are primarily first-party cookies.

Services like DoubleClick, LiveRamp, etc. place a 3rd party cookie that they use to track user activity across the web space. An important thing to note here is that you can only read those cookies that you have placed, because all cookies have unique properties. For example, Amazon cannot read a cookie that Wells Fargo has placed in a visitor’s browser.

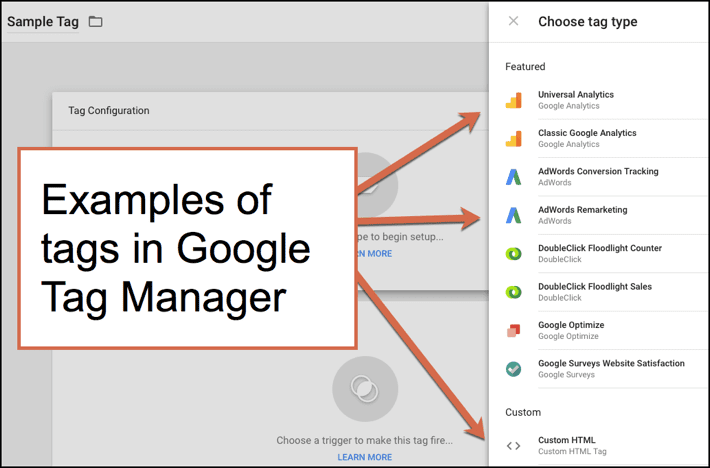
Pixels (tags) are typically an invisible single pixel. These pixels fire, or a JAVA code executes (both mean the same thing), when the webpage is loaded and capture important information about the visitor. Pixels can also place a new cookie in a visitor’s browser or check if there is an existing cookie already there. Websites need to embed this Java script in the site source code, which looks something like this:

Tag containers can contain multiple pixels or tags. One tag can trigger another set of tags and so on. Hence, containers are used to make conditional decisions that can determine if a set of pixels should fire or not.

**Google Tag Manager (GTM):** By now, you would have realized that the digital world is all about maintaining these tags. Why? So you can measure campaigns and create new audiences for prospecting.

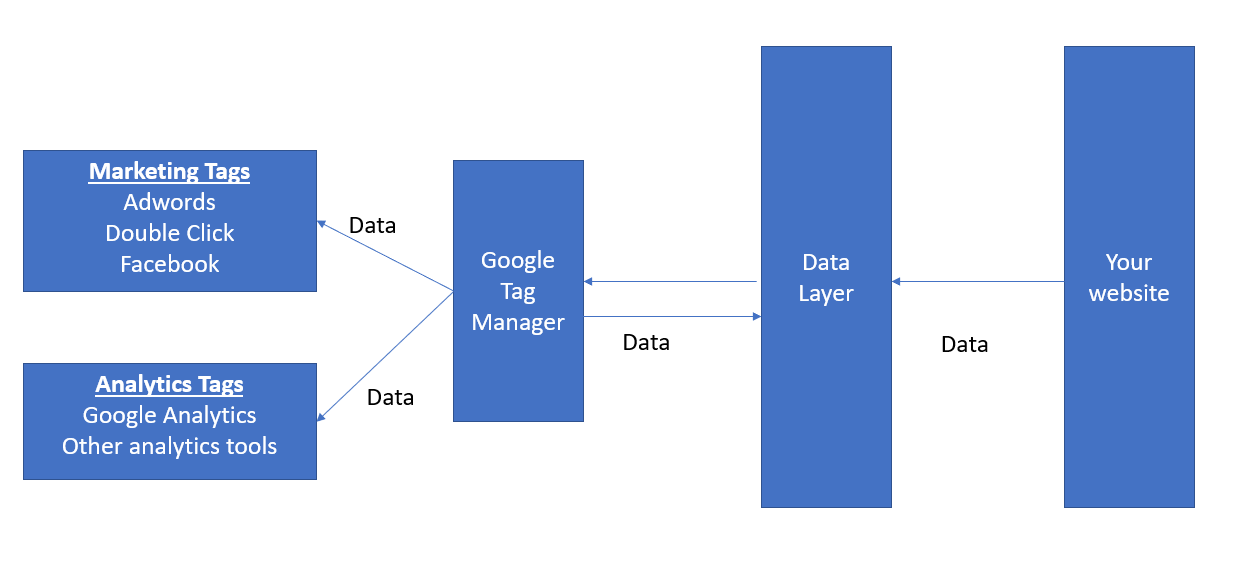
Here is a simple example – Tom runs a food blog. He actively markets on paid search using Adwords. He also re-markets his customer to return to his blog if the customer has not come back for some time. He additionally tracks his online traffic using Google Analytics. He even leverages DoubleClick to target display ads beyond Google Network. Imagine the number of tags Tom might have to put on his site – one for Adwords, one for re-marketing, one for DoubleClick floodlight and some custom DMP/DSP tags for specific re-marketing campaigns.

Handling so many tags within the source code of the site is extremely risky for Tom’s blog. Google Tag Manager is a solution to this problem. Google Tag Manager provides a single Java script which Tom needs to put in his source code which will allow him to manage all the tags to his site straight from the GTM interface. Additionally, Tom does not need to inform his IT team when making small changes to the tags because tags and the site source code are two separate entities superficially linked by Google Tag Manager. Below is a list of tags that Google Tag Manager can manage for you:



**Data layer** is another key concept you should know. This component is what makes Google Tag Manager such a powerful tool. Simply put, you can think of it as a bridge for data between your site and GTM. GTM can do a two-way communication with the data layer. GTM will then make this data available for all the tags sitting in it.

Think of the additional capability this adds to the dynamics! Now your marketing tags and analytics tag can directly pull segments of visitors you have created with on-site and DMP data. This can help you achieve the same level of targetability on off-site as that of on-site. The below schematic will make the process clear:



**Quick note on Floodlight tags:**DoubleClick uses floodlight tags to note visitor activities and sales. They provide two types of tags – FL Counter and FL Sales. FL Counter is primarily used to link some conversion to ad exposure. Hence, it is used to track subscriptions after an ad click from a visitor. FL Sales is primarily to store transactions with a dollar value. It can help you optimize your campaign on total ad spends instead of maximizing conversions.

## DSP vs. SSP vs. Ad Exchange (Demand Side Platform vs Supply Side Platform vs Ad Exchange)

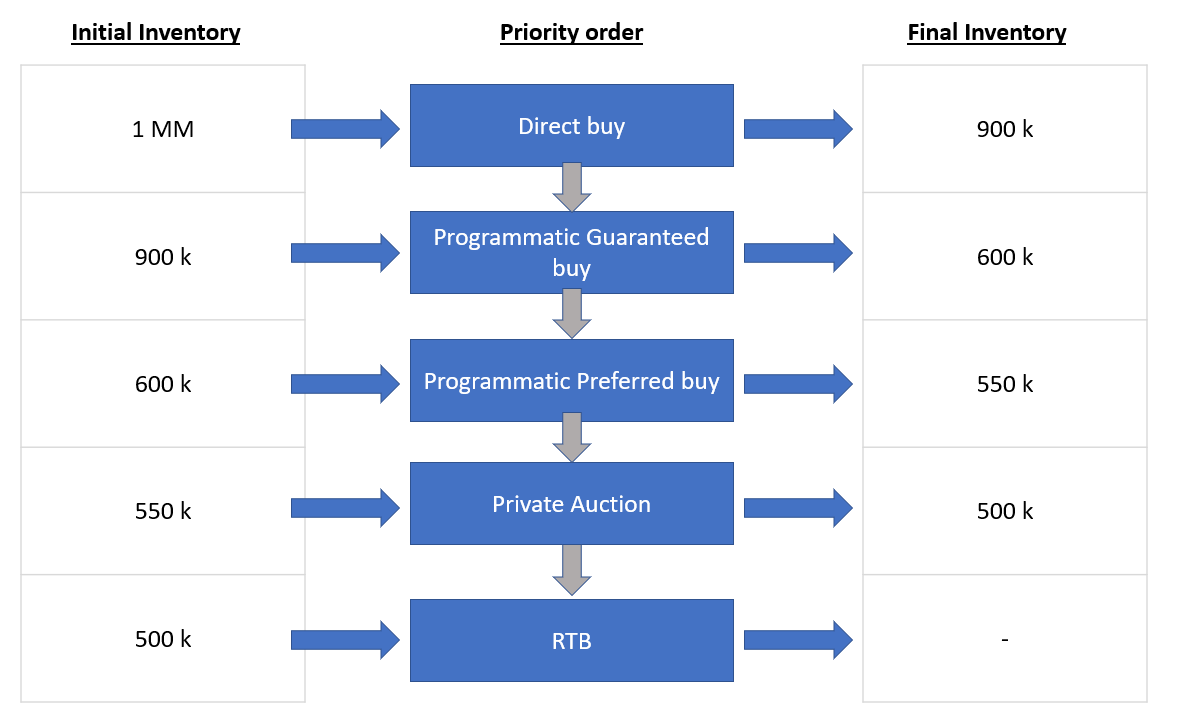
I must confess this is the most confusing and difficult aspect of digital marketing. Feel free to skip this section if it becomes too complex. I will try my best to make these concepts as simple as possible. All these terms are related to technologies used primarily in display media. So before getting into these technologies, let’s first try to understand the display media ecosystem.

Search media is always bought through programmatic buying, however, display media can be bought directly or programmatically. Programmatic buying is basically technology assisted buying of media. Consider the following scenario:

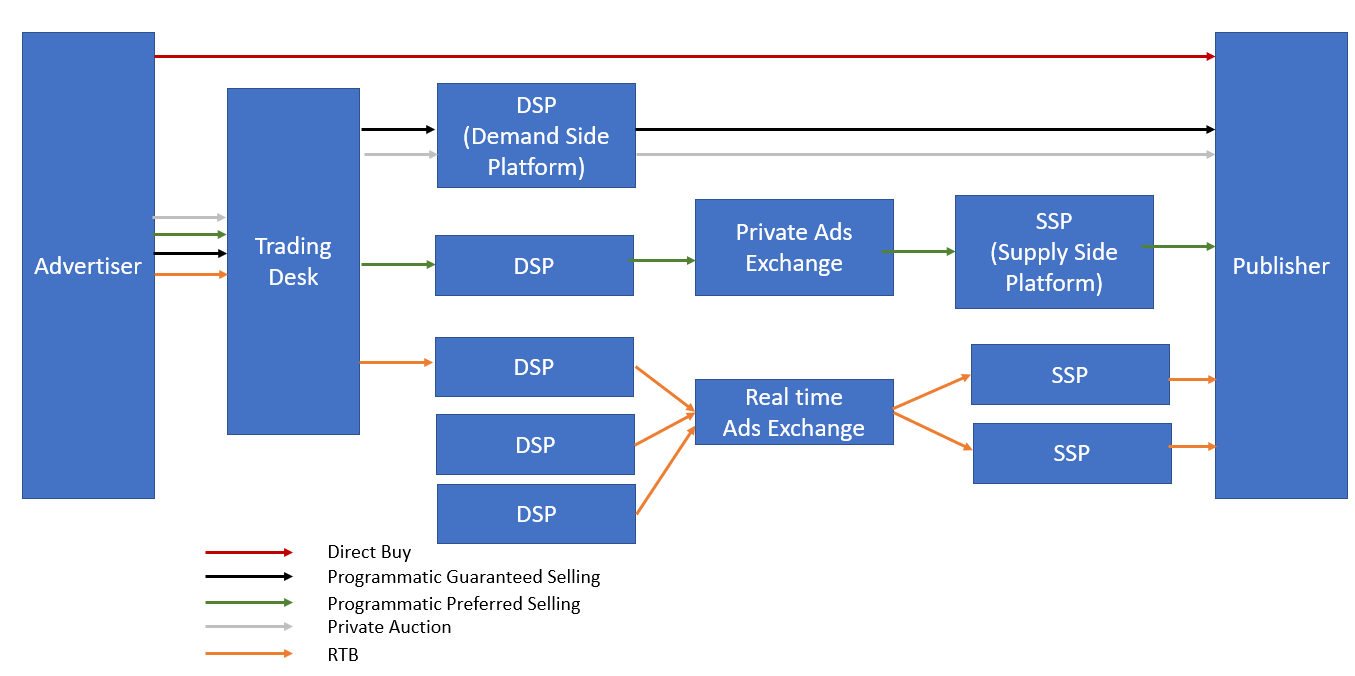
Victor runs a very popular travel blog. He gets about a million visits a month. He now wants to monetize his blog by putting an ad banner at the bottom of the article in the square placement. Out of the million visits, 500k visits will be on blogs that are related to hotel stays and the remaining 500k visits on blogs related to other topics. Victor wants the hotel blogs to specifically have Hotel related ads. He leverages various ad selling options to sell all his ad spaces.

1. **Direct Selling to Advertiser:** Victor finds a travel advisory website interested in buying ads directly from Victor in the hotel related blogs. This advertiser needs just 100k impressions at a fixed price regardless of the timing during the year
2. **Programmatic Guaranteed Selling:** Victor finds an international hotel chain who are interested in buying direct ad inventory from Victor but only from May-July every year. They at least need 300k impressions in this duration, so Victor is able to convince them to pay a fixed rate that is higher than the market price. Generally, this is what happens – guarantee comes at an additional cost
3. **Programmatic Preferred Selling:** Victor also finds a local hotel that wants to purchase ad inventory but wants to target only people on the US East Coast.They also cannot afford a premium rate so are not putting any threshold on the number of impressions. They will pay a fixed cost for these ads. Victor expects about 50% of his visitors to be from the East Coast, so expected impressions sold will be 50k
4. **Private Auction:** Victor expects 50k ad inventory to still be left after the above 3 deals. He collects a list of 50 hotels that are interested in participating in a private auction to buy this ad inventory. No other advertisers can participate in the auction
5. **Open Exchange/Real Time Bidding (RTB):**Finally, Victor plans to sell off the remaining 500k ad inventory in the open exchange to get the best rate available in the market. Any advertiser can bid for this ad inventory and potentially win

The above list is in the priority order of how ad inventory is distributed. The below diagram provides more clarity on the hierarchy:



To make this entire process possible, we use technologies like DSP, SSP and Ads Exchange. Here is a schematic of how it generally works:



* **DSP (Demand Side Platform):** Software Advertiser used to make the ad purchase
* **SSP (Supply Side Platform):** Software Advertiser used to make ad inventory available for purchase
* **Ads Exchange:** Real-time trading of ads between DSP and SSP

As you can clearly see from the diagram, Direct Buy is done without any trading involved. Programmatic Guaranteed Selling is through DSP to the Publisher. If the inventory is unsold in these two options, it will be made available as Preferred Ads buyer, before it goes into the pool of Ads Exchange. Finally, we go through the private auction and RTB stages respectively.

*Google DoubleClick Bid Manager is a Demand Side platform. However, Adwords is not truly a Demand Side platform because of multiple reasons – it restricts the ad inventory to only the Google Display Network. There is no option for programmatic Direct Buy or private auctions.*

## Data Management Platform

Data Management Platform is at the heart of analytics for Display Marketing. Let’s try to understand it with an example. First, a quick concept – **Remarketing** is a way to win back a prospect who has visited your website but has not made an immediate purchase or enquiry. You might have witnessed Amazon ads following you on the web – that’s remarketing.

XYZ co. is an e-commerce company that wants to create a niche audience for display marketing. They only want to use the display channel for customers:

* Who left the cart after adding a Laptop A
* Should have made at least $5000 purchase in last 1 year
* Checked out the Laptop A at least 3 times
* Should be a data science enthusiast

The remarketing ad will be an offer of 5% off on this Laptop A.  XYZ wants to publish these ads by placing bids on Ad Exchanges. How can XYZ execute such complex targeting?

***Short Answer – through a Data Management Platform (DMP).***

If you want to know the technical details of how this can be done, continue reading this section. Otherwise, all you need to know is that DMP can make specific targeting possible. DMP also provides third-party data, like preferences, interests, etc. You can skip the rest of this section if you want to ignore the technical details.

Here is what actually goes on behind the scenes:

1. Visitor comes to the XYZ site
2. DMP fires a pixel and places a cookie (if not already present) in the visitor’s browser. This cookie has a certain ID (say – DMP1)
3. DSP fires a pixel and places a cookie (if not already present) in visitor’s browser. This cookie has a certain ID (say DSP1)
4. DMP server tries to match this visitor to third party data and gathers all the information it can. It also brings in first party information from XYZ
5. DMP decides if this visitor falls under the segment defined by the advertiser
6. DMP server syncs with DSP server to tell DSP that (DSP1 = DMP1) and if DMP1 falls in the target segment found in step #5
7. Now, the visitor goes to read a blog. This blog is connected to the same DMP and a SSP
8. DMP again fires a pixel and finds the cookie already present in the visitor’s browser (DMP1)
9. SSP fires a pixel and places a cookie (if not already present) in the visitor’s browser. This cookie has a certain ID (say SSP1)
10. DMP server syncs with SSP server to map DSP1 = SSP1
11. DMP now informs DSP that DSP1 = SSP1 and asks DSP to bid aggressively in the auction

The above 11 steps happen in less than a mini-second! DMP is used for many different use cases, such as:

1. Enable model building for targeting by giving additional third-party information to both advertisers and publishers
2. Provide solution to find lookalike audience to advertisers
3. Suppression of selected audience
4. Aggregation of first-party information that can be used for on-site personalization
5. Remarketing

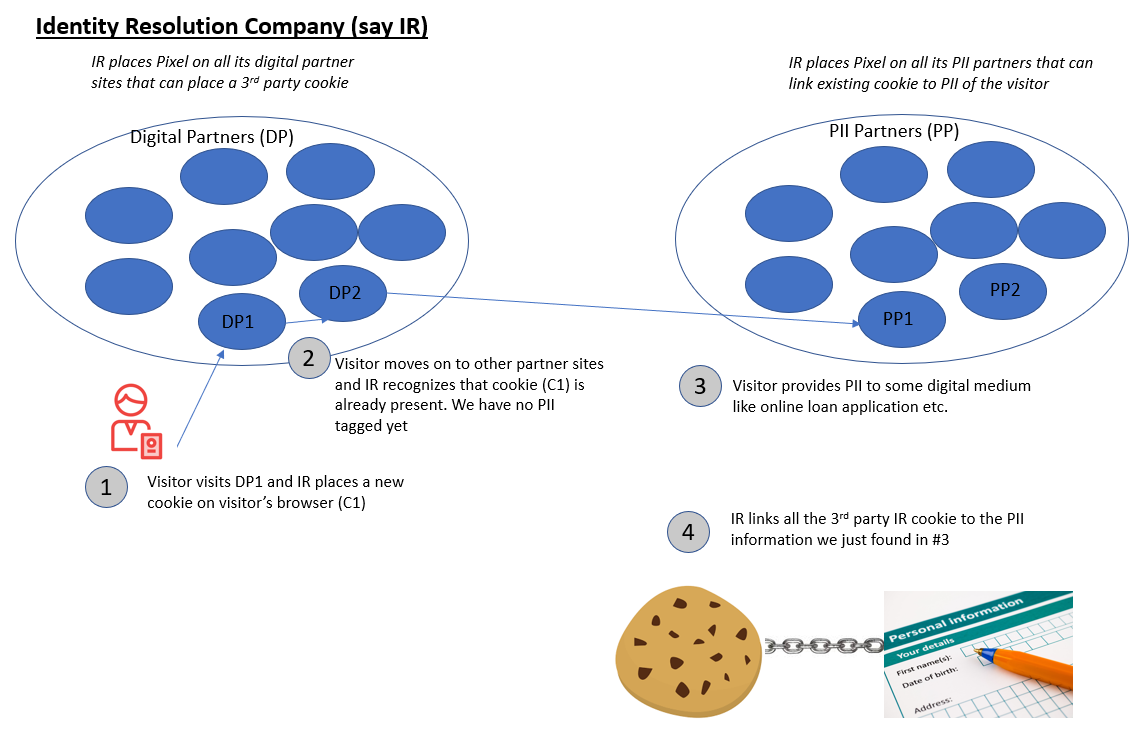
## Identity Resolution

This is the concept that fascinated and scared me the most. We already know that the entire web world is tracking visitors through cookies. Cookies have been around for a long time. Using cookie DoubleClick can really stitch a persona – for instance, a visitor that reads AV blogs clicks on an online analytics course, but does not convert. The same visitor comes back to another university analytics course and as he/she was tagged as an analytics enthusiast, this new university gave him/her a discount of 20% on the course through on-site optimization. The visitor finally converts and complets the course in 60 days. And so on….

So, essentially we know everything about the visitor. What’s left? A very key information, which is, “This visitor is John Bell”.

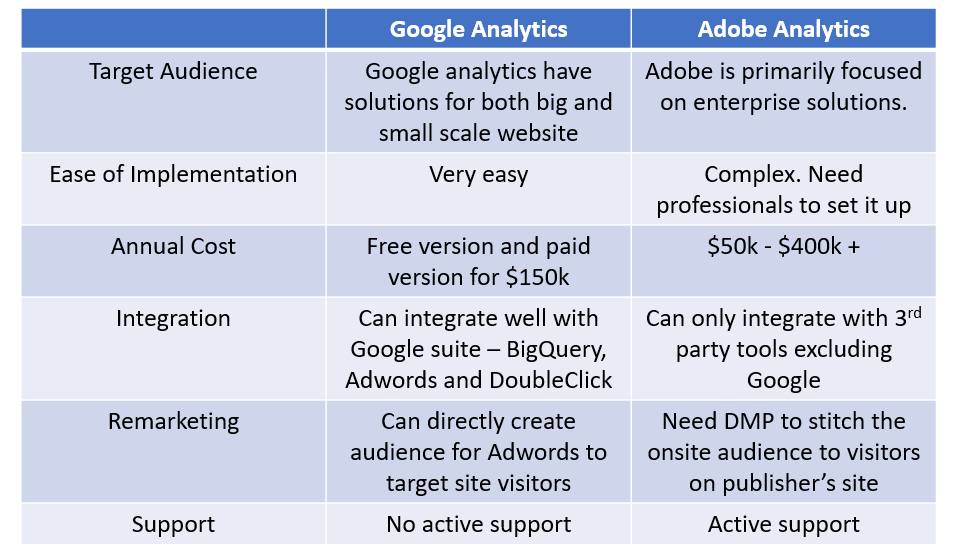
What is so important about this new information? This is a Personal Identification Information (PII) and will never change over time. Cookies get deleted all the time. But if we can link these cookies to PII, we not only have a persona that can describe a small subset of the population, but we directly have an individual that only describes one person on the planet.

This is exactly what an Identity Resolution does. Let’s try to get a broad idea of how Identity Resolution works in the industry currently:



## Digital Analytics Tools

Given all the background information, let’s now talk about the various tools you have at your disposal that can process huge amounts of data coming from a number of channels. These help you analyze the data and implement your strategy in real-time. Even though there are many solutions in the market, most of the companies (big or small) are using one of the two – Google Analytics or Adobe Analytics. Let’s review them briefly by comparing a few key attributes:



Both the tools have their pros and cons. If you are a small scale company, choosing Google Analytics is a no-brainer. Even if you are a large corporate, the choice is tricky because Adobe on one hand gives active support, but Google integrates seamlessly with well-managed ad inventory. Note that Adobe can also integrate third-party tools for ad targeting, but not all of them are well-managed and might be prone to bot attacks, thus wasting your ad spends.

## End Notes

You might have realized that each concept we covered in this guide is closely linked with each other. A comprehensive view is very important to appreciate the entire digital ecosystem. Trust me, it was impossible to find all this information in one place. As a data scientist, we do the hard part of learning all these concepts on the job. So I decided to put everything in one place with the aim of helping any future analyst get up to speed really quickly and keep up with the pace of this dynamically evolving industry.

With the knowledge provided in this article, you will not only understand how to build a successful digital analytics driven strategy, but will also start appreciating how your strategy fits into the broader world of digital marketing. This combination of knowledge is extremely rare in the industry because most professionals focus on only one of these aspects. What it takes to create a successful marketing campaign on digital media lies at the intersection of the digital ecosystem, marketing and analytics.