unacast.

The 2018 Location Data Buyer's Guide

The ingredients aren't listed on the packaging, you can't tell if location data is fresh by sniffing it, and the person selling it to you likely had no part in harvesting it.

So how are marketers and tech teams supposed to know how to cook with it?

Read this guide. →

unacast.

Table of Contents

Why a Location Data Buyer's Guide? 03

Data Buying for Marketers 05

Data Buying for Technical Teams and Data Scientists 07

Practical Use Cases 08

Industry Specific Use Cases 09

Parting Thoughts 10

unacast.

Why a Location Data Buyer's Guide?

While buying and selling data may seem straightforward - data has been bought and sold for a long time - location data specifically is a complex, young marketplace that all too often lacks transparency for buyers. Why?

Because the true nature of modern location markets is that of a black box. The inputs and outputs may be familiar, but the data's machinations and internal structure are a complete unknown. So, the actual product bought may be entirely unsuitable to the marketer's intended task (aka: ROI is unpredictable and maybe zilch).

Absent the open surfacing of not only the limitations of a given data set, but also the reliability of the data itself (based on its composition and relative quality), the buyer doesn't really know what they're getting. As a result, they don't know how to use - or value - location data for a given use case.

In the case of attribution, for example, where the end objective is about measuring the effectiveness of an ad campaign, the value of location can be quite high, so you want premium quality product. That's a different scenario than in a retargeting use case, where the objective may be moving the user down the path, rather than getting an immediate transaction. In this case, a less expensive type of data may do.

TERMS AND DEFINITIONS

Mobile Advertising ID (MAID)
A Unique Identifier for a mobile device created by Android (GAID) or Apple (IDFA). Not the equivalent of a person (a person can have a device, or many devices, and a device might be used by many people).

Precision
Refers to the number of decimal places in a Latitude or Longitude. The more decimal places, the more precise a reading is: however, technology is not precise enough (on mobile phones) to create more than 6-8 decimal places of precision -- more than enough precision for human mobility.

Horizontal Accuracy
When a phone sends a location ping, it is sending you a circle where it is certain (with a 68% confidence interval) the device is located. The latitude/Longitude point is the center of the circle. These two fields used in conjunction provide you an approximate area of the device's accurate horizontal location.

IP Address
Used to identify the host on a network. The assignment of IP's is done by the network and not the device - therefore the device does not know it's IP. When a device sends data to a server, it is up to the server to collect the IP. Depending on how often a server updates, the IP address used may be recent, or aged.

The point is, location data markets operate absent much in the way of transparency. The net effect is that buyers, who have probably been burned before (by bidstream data perhaps; see below), don't really know what they're buying.

The ingredients aren't listed on the packaging, you can't tell if location data is fresh by sniffing it, and the person selling it to you likely had no part in harvesting it. So how are you supposed to know how to cook with it?

That lack of basic transparency is toxic to the location data industry. It's bad for buyers, it makes it tough for quality sellers to stand out, and it exposes users and everyone else in the market foodchain to risk. To move forward, marketers and tech and data teams need simple tools to help them buy data the right way.

To that end, we offer our 2018 Location Data Buyer's Guide. We hope it will help you - whether you're just dipping your toes in location data, or already diving deep.

Timestamp
The time when a ping occurs. Unless otherwise specified, it is stored in UTC. Converting to Local Time Zones is much trickier than it seems because, within the US, not even states necessarily contain the same time zones.

Daily Active Users (DAU)
The number of Unique MAIDs on a given day. Please note the definition of MAID, above-- Even though the industry standard is "user," it is important to understand the distinction in case it matters for your use case.

Monthly Active User (MAU)
The number of Unique MAIDs seen in a given month. There are only about 330 Million people in the United States. If the number is too high, or even very close to the population of the US, it might signal bidstream data.

Cost Per Mille (CPM)
Cost per thousand MAIDs. Most sellers charge on this metric.

Pings Per MAID Per Day
The number of pings an MAID sends on a given day. Higher does not always mean better, because this number does not factor in duplicate data points.

Bidstream Data
Bidstream data comes from an advertisement served to an app. With no access to the phone's GPS, it will instead use probabilistic methods, such as IP address to attempt to pull the information. This methodology is less accurate, but highly scalable. Therefore, if scale is remarkably high, it is a sign that the data you are receiving is Bidstream, not GPS data.

unacast.

Data Buying for Marketers

Data from apps, Wi-Fi networks, app aggregators, beacons (and other hardware) all help tie users to a specific location. Because we collect so much data from so many different places, there are naturally sources that are better than others. To get a feel for the landscape, start by understanding that there are different categories of location data.

A visits dataset will provide the most value for use cases requiring a high degree of accuracy. Visits datasets are a synthesized pool of a user's known points of interest (POIs), including a timestamp, the latitude-longitude of visited locations, a unique venue-ID, the number and names of nearby venues, and the user's dwell time (or length of visit) in a given location, among several other data points.
In other scenarios, a pure dataset may suffice. These lightly filtered streams of aggregated location data still include a timestamp and the coordinates of a given interaction, but also account for horizontal accuracy, IP address and source-ID. This is a good fit for a sophisticated brand with the capability to attach their own POIs, and analyze and extract value from data in its raw form.

Yet another use case may be for brands seeking to engage users specifically at home or work. In this case AdID, a GeoHash, zip code, work classification and/or relationship score is likely of deeper value to the marketer.

Hotspots/Centroids
A phone can't always figure out where it is. In this case, it widens the likely area to what it can predict. An IP address derived ping typically will select the middle of a zip code, or even a landmark like city hall, as the latitude/longitude. If you don't see any of this in your data, it means it's been processed before it got to you. If you see a lot, it is a sign that there might be bidstream data in your pool.

Exclusivity vs. Uniqueness
Everyone wants unique data, but often it is confused with exclusivity. Even if you have an exclusive source for procuring data, it does not guarantee that the information you are receiving is unique.

Overlap
If source A has MAID X and source B also has MAID X then you would consider them overlapping. No single source collects everything on a device, and if your goal is to build out a picture of a device then you need to look at how multiple sources work in unity.

These different location data sets have different intrinsic value, and they should be sold as such. This is the only way that those who buy location data will be able to do so with confidence and clarity. And that's important, because for marketers, understanding how people and places relate to one another is a gold mine.

Unleashing the power of location data allows the creation of ads that offer more relevant content and are better targeted, and intelligent insights that inform the way brands engage with consumers, mobile and otherwise. This is where our Real World Graph® comes in.

We made The Real World Graph® because mapping out the connection between people and places, and providing insight into how people move and behave in the real world, is crucial for companies, marketers, technical teams, and data scientists, across all kinds of industries.

Did You Know...

The Truth About GPS Data

The appeal of GPS data is that it reveals the exact location of any device at any time. The problem? That isn't true, at least not all of the time. GPS data can be very accurate, but marketers must accept that there is no way of knowing exactly where a device is at all times. For example, using device location to determine exactly where a consumer is within a city is, in most cases, an impossible task. The tall buildings block the signal to the satellites, while small storefronts packed closely together make it impossible to determine if the person is inside one store or another next door. Any brand, agency, or vendor who claims that they always have 100% accurate GPS location data is either lying or doesn't know how the technology behind GPS satellites work.

This doesn't mean that all location data derived from GPS satellites is inaccurate or useless as a marketing tool. It just means that marketers need to better set their expectations, know the data they are buying, and factor the limitations into their partnership agreements and marketing plans. This means looking under the hood at the source of the data, especially if it is being sold with a promise of pinpoint accuracy.

unacast.

Data Buying for Technical Teams and Data Scientists

Although the data itself that comes from apps, location companies, beacon companies and other vendors is inherently omnichannel, the main source of data for consumer location is the mobile phone. After all, smartphones are our constant companions.

To aggregate the mobile user's endless stream of data and make sense of it, technical teams must harmonize fragmented data relating to the mobile native's location, transactions, demographics, and behavioral data, in-app and otherwise. This entails marketing and advertising technology and applications.

For Unacast, the key to making sense of mobile user's Real World Graph® is ensuring that any industry (retail, city planning, events, hospitality), can use our tools to good ends, and with absolutely transparency for all parties.

QUESTIONS TO ASK WHEN DEFINING DATA GOALS, AND THE DATASETS THAT PROVIDE ANSWERS.

1. How much processing do I need?

- None = Pure data
- Some = "Bring Your Own" POI or Pure data clusters
- Lots = Visits data

2. What problems am I trying to solve (or can I solve) with location data?

- Foot traffic to retailers = Visits data
- Finding visits to my POI = "Bring Your Own" or Pure data clusters
- Figuring out where to place a new store = Pure data clusters or Visits

3. How does this integrate or match with other data providers I work with?

- If I work with TV advertising - can I match home to devices?
- If I work with offline sales, how can I tie it to location data?
- How does it complement or not work with other data sets I work with to help me draw new conclusions I wouldn't have been able to before?

QUESTIONS TO ASK DURING THE BUYING PROCESS

1. How do the different data set features enable or block my use case? (coverage of areas, coverage of devices, volume, overlap, etc.)

2. Is the device data collected representative of target population? (goes to quality of data i.e. how much is junk vs useful).

3. How easy is it to take delivery of and ingest data and how often do I need to do that?

unacast.

Practical Use Cases

What follows are a few brief examples of how location data can be applied - both on a high level and through deeper detail across different verticals.

General Use Cases

#1 Use Case: Area-based human movement

- Evacuation planning
- Impact to traffic control and road closings
- Transportation decisions
- Customer journeys between points of interest or between their home and work
- Site Selection based on human densities and profiles

#2 Use Case: Store performance

- Stock predictions based on foot traffic
- Understanding traffic trends at underperforming or over-performing stores
- Understanding macro consumer trends around POIs for economic research

#3 Use Case: Hyperlocal foot traffic insights & cohort analysis

- Market a commercial real estate asset by understanding foot traffic cohorts
- Profile areas and create retail analytics insights based on cohorts (locals vs commuters, loyal shopping trips vs one-time shopping trips)

Industry-Specific Use Cases

City/Transportation Planning
City/Transportation Planners need to understand how and why people move from one area to another. Currently, a combination of survey data and aggregated traffic data is used to build this view, but with location data, planners can build views that are nearly "real-time", with unprecedented granularity and precision.

Retail - CRM Augmentation
Filling in holes in your current CRM (think dining verticals that can't capture customers outside of loyalty card programs or credit card transactions).

Add new rich attributes to your existing CRM (personas, cross-shopping metrics, loyalty metrics, distance traveled metrics, shopper behaviors, etc.)

Retail Performance
Using visits data as a proxy for understanding retail performance, such as: national trends, comparisons across regions, competitive landscape.

Retail - Building a Competitive CRM
Using third party data to determine what your competitor's customer base looks like and how they behave in key moments.

Site Selection
Applying site-level foot traffic data to value potential sites for a new commercial location, thereby offering significant improvement to existing selection models.

Local Trade Area Identification
To understand where customers to a particular store location come from, combine visits data with home/work data and study areas of overlap.

unacast.

Parting Thoughts

Telling you exactly what you're buying is the only way to create a transparent location data market and help companies realize the dream of happy and engaged mobile customers, in-store and otherwise.

At Unacast, we are opening the location black box and investing heavily in transparency. There are all sorts of things that you need to consider to be able to make a determination on the location of a device. Understanding if a device is in an urban area or a rural area, what time of day the visit is taking place, how long the device is spending at the location, how often it is seen there, how many data points make up the visit, what the technical source of the data is (GPS, WiFi, Beacon etc), what other venues are nearby, what is the horizontal accuracy for the data that created the visit. The list goes on.

The data attributes above (and more) are important because they give insight into how confident you can be that a device was at a place, especially when it comes to GPS location data. Any company that tells you they mastered GPS and you should trust them, needs to be questioned. If you question them but they aren't willing to actually give you the ability to inspect and understand the data on an ongoing basis, you need to question your long term relationship with the partner.

As a cure for this, Unacast believes in giving insight into the attributes of every visit to enable you to "create your own truth" based on your business needs. It enables you to see and understand the black box.

Why is this significant? Because sketchy, underperforming datasets with inflated prices waste dollars and hinder otherwise successful marketing and campaigns with less than ideal experiences. That turns both buyers and users sour, and is negative for the location industry as a whole. ***We don't believe that's the best way. We believe in trust through transparency. We believe that helping you become a more informed location data buyer is the better way.***

Ready to learn more? Check out unacast.com or drop us a line: info@unacast.com