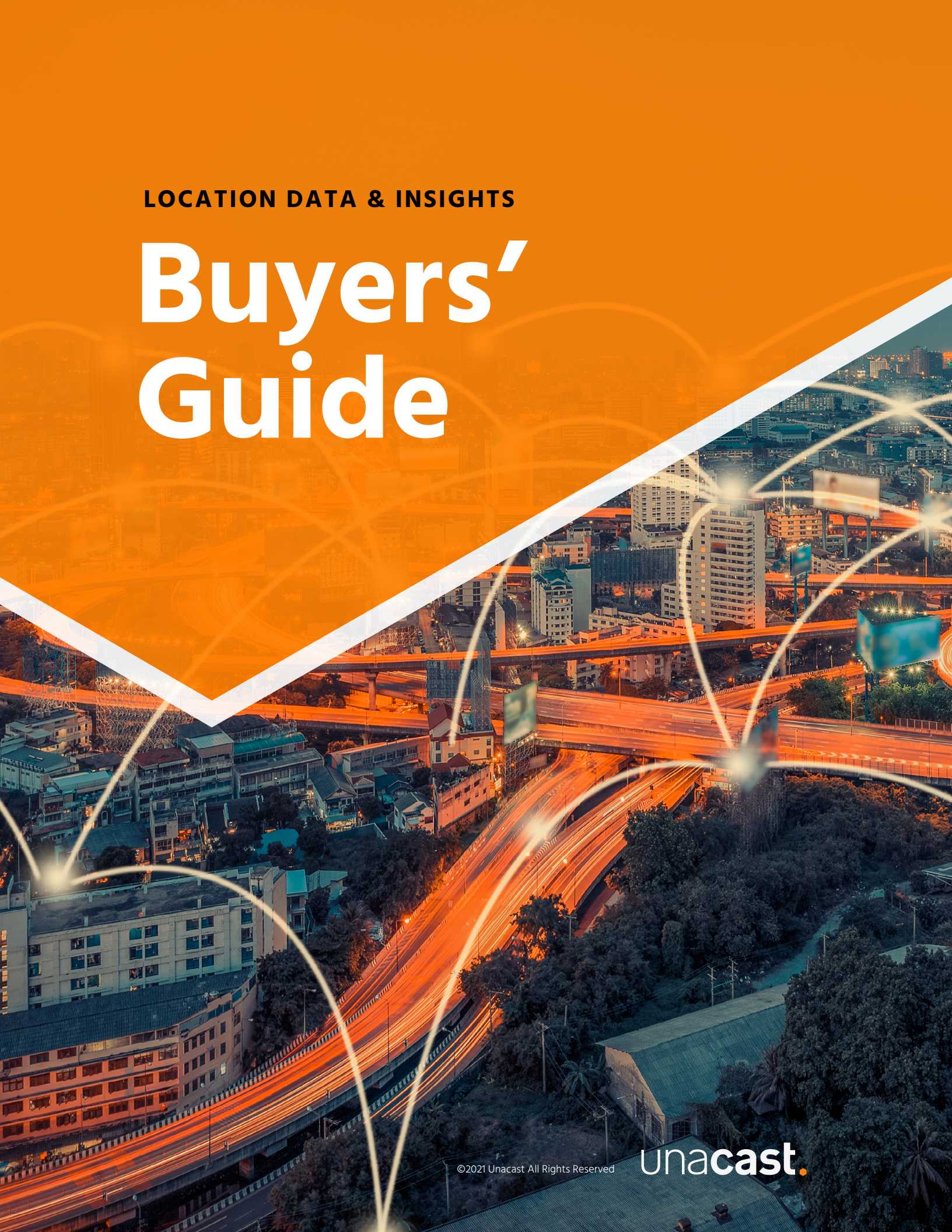


LOCATION DATA & INSIGHTS

Buyers' Guide



Welcome to the 2021 Edition of Unacast's Location Data Buyers' Guide — which we've renamed the Human Mobility Insights Buyers' Guide to better align with where the location data market is headed.

Whether you're just dipping your toes in location data, or already diving deep, this guide can help navigate the complex world of using data to understand trends and measure population movements in the physical world.

Whenever you see an asterisk (*) in the following pages, it means that we're using a technical or industry term that you may not be familiar with. Check out the [Glossary of Terms on pages 13-15](#) for clarification on these terms and on other location data lingo.

- 04** Why an Insights Buyers' Guide?
- 06** What Are Human Mobility Insights?
- 08** Who Are Human Mobility Insights For?
- 10** The Process: How Location Data Becomes Human Mobility Insights
- 11** Six Questions to Ask During the Buying Process
- 13** Glossary of Terms
- 16** Parting Thoughts

Why an Insights Buyers' Guide?

The location data industry is growing in both size and complexity. As a result, the needs of buyers are evolving in three major ways:

1

MORE BUYERS WANT INSIGHTFUL, AGGREGATED DATA PRODUCTS VERSUS DEVICE-LEVEL DATA SETS

The number of companies who want to use location data is growing significantly faster than companies who can use location data. Here's why: data without interpretation is like paint in a jar — it's not until you apply it to a canvas that it begins to mean something. Device-level data products are like that proverbial jar: un-curated, and sometimes un-processed, sets of raw data. Companies without a data science team and/or the proper technological tools and/or dedicated time won't have ways to interpret it, let alone make it actionable. They need aggregated products — those that feature built-in processing, curation, and analysis as a matter of course (such as visualizations, reports, or dashboards). This kind of shortcutting offers clients at all different levels of data sophistication the opportunity to leap directly to interpretation so that they can turn location data into actionable insights for their businesses.

2 SMART BUYERS ARE REALIZING THAT DATA QUALITY IS MORE VALUABLE THAN QUANTITY

Many location data providers count on sheer volume as a selling point. While some quantity is critical to achieving statistical significance, more doesn't always mean better. Some providers even pad their coverage* metrics with invalid or low-accuracy data (which is then immediately removed from the set during processing). A good rule of thumb is "garbage in, garbage out"; insights created from high-quality data not only give you a stronger grounding from which to make decisions, but you also won't spend money on data you can't possibly use.

3 THE PRIVACY ISSUE IS PARAMOUNT — AND A MUCH BIGGER DEAL THAN BEFORE

An occupational hazard of location data is the potential "creepiness" factor — something the general public is growing more aware of with each passing day. Privacy regulations like GDPR* and CCPA* notwithstanding, overly intrusive location data is a PR nightmare waiting to happen, and savvy buyers want as much insulation as possible. Any provider unwilling to provide transparency into its operation — in sourcing, processing, sales, or all of the above — should be viewed with suspicion. Location data should be about tracking aggregated population movements and trends, not creepy individual device tracking.

HOW TO TURN OFF LOCATION DATA ON YOUR MOBILE DEVICE

A location data company telling you how to stop creating location data? Absolutely. These are our Norwegian values at work: while most people have decided that handing over their data for convenient services is a worthwhile trade-off, everyone deserves to make their own determinations on privacy.

All major mobile device operating systems allow the users of mobile devices to control the types of data collected in and through apps.

Apple iOS

Apple instructions on how to control iOS device location settings can be found [here](#).

Apple instructions on how to limit ad tracking on iOS devices can be found [here](#).

For other privacy-related settings on iOS, go to Settings from your mobile device's home screen, and select "Privacy."

Google Android

Android instructions on how to control device location settings can be found [here](#).

Android instructions on how to limit ad tracking on Android devices can be found [here](#).



What are Human Mobility Insights?

There are three critical ingredients — leave out any one of them, and the data won't mean much.

An insight is the meaning we derive from data: what's happening, what's likely to be causing it, and what conclusions you should be drawing from it. And it's the result of some very technologically-complex processes, combined with good old-fashioned human brainpower.



INGREDIENT #1: THE WHO & WHEN

Location Data

There are many kinds of location data out there. It's possible to determine a device's location through the bluetooth devices it connects to, the wifi networks it joins, and the towers it draws a cellular signal from. But the most reliable option is an oldie but a goodie: the Global Positioning System. GPS works by sending signals, or "pings," from your mobile device to a constellation of satellites in low earth orbit, and back again.

GPS uses triangulation to determine where on the planet your device is, and describes that position using **latitude** and **longitude**. And to prevent the satellites from confusing your device with any of the billions of others using the system, your device has a unique **identifier**. Finally, each ping from your device to the satellites has a **timestamp** to record when the device made an appearance. These four (latitude, longitude, identifier, and timestamp) are the critical components, but GPS can also include enhanced data such as horizontal accuracy*.

THE TRUTH ABOUT GPS

The appeal of GPS data is that it reveals the exact location of any device at any time. Except that it doesn't — at least not all the time, and smart users will always be cognizant of the fact that it's fallible.

For example, determining exactly where a device within a city is, is in many cases, impossible. Tall buildings block the signal to the satellites, while small storefronts packed closely together make it difficult to discern which one the person is in. Any company that claims to have 100% accurate GPS data is either lying or doesn't really understand how the technology works.

This doesn't mean that all location data derived from GPS satellites is useless. It just means that those who use it for insights need to calibrate their expectations, know the data they are buying, and factor the limitations into their partnership agreements and business strategies. It means looking under the hood at the source of the data — especially if it's being sold with a promise of pinpoint accuracy.



INGREDIENT #2: THE WHO & WHEN

Map Data

Most humans couldn't begin to guess which latitude and longitude they're at at any given moment. That's why it's important to "translate" these pairs of numbers into a human context: a map. Unacast's Digital Atlas has multiple kinds of maps to use, depending on which point of view is most relevant.

Venues + Brands: Because we work with several companies who need data on footfall trends and patterns, our Digital Atlas includes nearly every location for the top 4,500 (and counting) retail brands.

Municipal Maps: The smallest scale in our Digital Atlas is the census block group (CBG) — divisions of land and population based on the most recent census. While this is not the finest-grain map out there, we want to make sure that private information, such as people's homes, are obfuscated and not discernible to an address. We can add CBGs together to measure neighborhoods, cities, counties, states, and even the entire U.S.

BYOPOI: Bring-Your-Own-Places-of-Interest is our term for a custom location at any scale. In cases where your venue may not be in our Digital Atlas, we work with you to add it so that you're still able to access all of our groundbreaking human mobility insights.



INGREDIENT #3: THE WHY & HOW

Strategic Intelligence

It's tempting to think that once you put GPS pings on maps, you're done. Au contraire! There are literally an uncountable number of ways to make these data sets say different things, some of which are esoteric, and others that are downright conflicting. This is where the human touch becomes most important: when the answers are infinite, you need a robust team of data scientists and business strategists that helps you "ask" the data just the right questions. In other words, a correlation strategy tailored to your unique needs.

Luckily the smart Unacast brains have also spotted a shortcut: any question can be answered using a combination of just six kinds of calculations:

- **Visits:** how many people visit your location and when
- **Dwell Time:** how long they stay
- **Foot Traffic:** how many people are in the area around your location
- **Capture Rate:** how much of that foot traffic turns into actual visitors
- **Catchment Area:** where your visitors come from
- **Cross-visitations:** where else your visitors go

Who Are Human Mobility Insights For?

(Short Answer: Anyone)

When it comes to technology, companies can exist anywhere on the sophistication spectrum. But you shouldn't have to have an in-house data science team to harness the power of human mobility data. Unacast will meet you where you are, providing the right services and tools for your business to access human mobility insights, in whichever industry it resides.

CURATED DATA SETS

What It Is: Filtered and processed data downloadable in JSON files and other import-friendly formats.

Who Benefits Most: Organizations with in-house data science capabilities who don't need extra help ingesting data and drawing insights from it.

UNACAST REPORT STUDIO

What It Is: A brief, one-time analysis and templated set of data visualizations describing the venue, neighborhood, event, or brand of the client's choosing.

Who Benefits Most: Companies who don't need to dig deep, or who want a fast, low-cost way to determine if human mobility insights are a fit for their company.

UNACAST DASHBOARD

What It Is: Conveniently accessible via web browser, Unacast's dashboard instantly generates insights and visualizations for the time and place of the user's choosing.

Who Benefits Most: Self-navigating users who want answers to common location data questions available at their fingertips.

STRATEGIC RECOMMENDATIONS

What It Is: Clients can engage Unacast data scientists and business analysts as consultants-for-hire.

Who Benefits Most: Organizations with highly-specific or uncommon questions that are looking to fold human mobility insights into their overall business strategies.

API*

What It Is: A mechanism that directly connects Unacast's human mobility insights and data feeds to clients' in-house systems.

Who Benefits Most: Companies that have an established technology infrastructure that incorporates many different kinds of data, including location data.

UNACAST TURBINE

What It Is: Our Platform-as-a-Service offering enables clients to utilize Unacast technology to privately create insights from their proprietary data.

Who Benefits Most: Companies that have their own location data, but are lacking the means to process or monetize it without exposing it to privacy regulations.



Retail

USE CASES

Performance Benchmarking:

Measure the impact of increased staff engagement; new products or services; or changes to pricing, advertising, or merchandising.

Location Prioritization: Identify potential store cannibalization or enhance strategies for new locations or store closures.

Competitive Intelligence:

Learn where, when, how often, and which kinds of customers visit your competitors' locations instead of yours (and vice versa).



Real Estate

USE CASES

Site Selection: Identify gaps in the local marketplace and other opportunities to expand your portfolio.

Neighborhood Analysis:

Determine which community characteristics and population demographics help create a lasting visitor base.

Business Development & Sales

Enablement: Craft data-driven messages to attract target brands and tenant types.



Other Industries

USE CASES

Travel & Tourism: Form new partnerships with other brands or venues shared by your visitors.

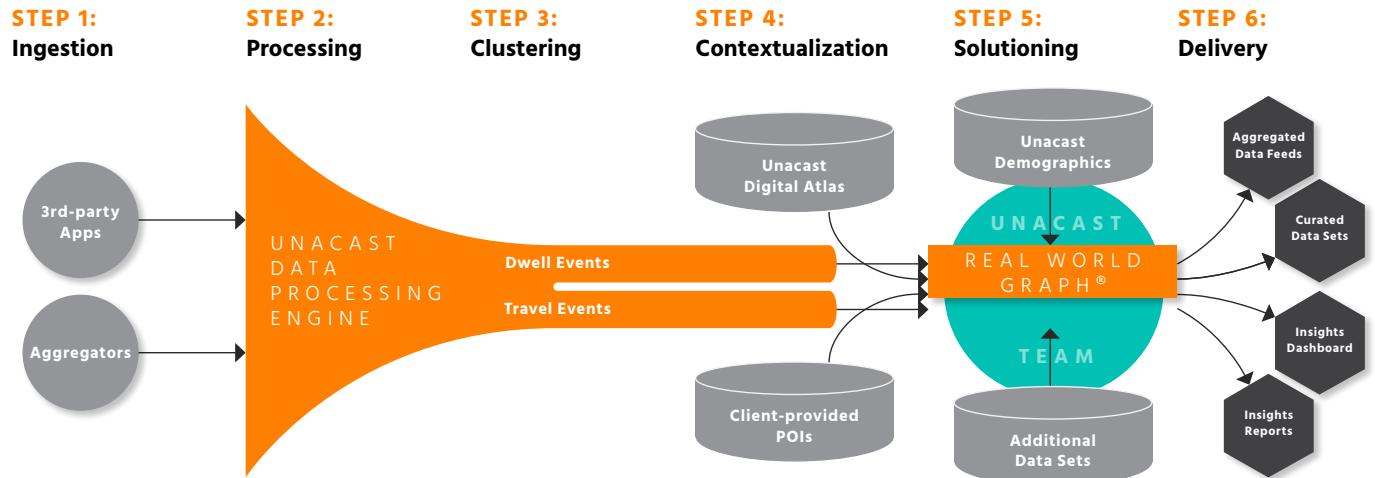
Disaster Planning: Analyze post-event human mobility patterns in order to enhance future evacuation strategies, or to help predict the spread of infectious diseases.

AdTech & Marketing: Enhance clients' bottom lines via segmentation and retargeting, measurement and attribution, and target audience analytics.

The Process

How Location Data Becomes...

...Human Mobility Insights



Unacast sources data from three different kinds of sources:

Third-party Apps:

Mobile apps that collect GPS data for their own purposes are generally high-quality sources.

Aggregators*:

We maintain partnerships with several collectors of GPS data — but only from those that share our commitment to transparency.

Most of the GPS data out there, frankly, isn't very good, and our data is processed to enhance quality:

Filtering: Weeding out invalid data, like duplicates, pings in the middle of the ocean, unrealistic speed or timestamps in the future.

Standardization: Giving all our data a common nomenclature ensures our various sources are apples-to-apples.

Enhancement: Employing several data-science-based techniques, such as bias correction* and extrapolation*.

The Unacast data engine also clusters individual GPS pings into events that indicate human activity:

Dwell Events: GPS pings within 150 meters of each other, in a 1.5 minute - 4 hour time period. A dwell on a certain place (venue or neighborhood) is classified as a visit.

Travel Events: GPS pings less than 10 minutes apart, when the distance between the first and last ping is greater than 100 meters.

Using the Unacast Digital Atlas or client-specified POIs*, proprietary algorithms place events in certain locations to indicate where in the physical world the activity took place:

Maps: Events can be contextualized in the context of venues (such as a store or a brand) or neighborhoods (using municipal maps).

Home & Work: Determined from dwell events that align with typical work or non-work hours, repeat in a 28-day period, and are obfuscated to the census block group to protect privacy.

It's not enough to have great, contextualized data. It still has to be correlated* in a way that gives it meaning and inspires action.

Using tools like the Real World Graph®, our data scientists and business analysts **team up to create a strategy** that accurately answers questions pertaining to your business.

Additionally, we can factor in non-location data (such as demographics, or point-of-sale) to add even more dimension to the data story.

Our insights do more than just inform — they're tailored to help make smart changes and delivered via the method best suited to each client, including:

Aggregated data sets: JSON files and other import-friendly formats.

Unacast API: Easy connections to your in-house systems.

Visualizations: Data formatted into charts, graphs, and tables.

Strategic Recommendations: Bespoke documents that illustrate data and clearly articulate likely interpretations.

6

Questions to Ask During the Buying Process

1

What methods do you use to ensure your location data is representative of the market population?

The challenge of creating high-quality location data is ongoing and multi-faceted — and can be a bit of a moving target.

Your location data provider should be willing to describe the various means and methods they use to create an accurate location data set. At a minimum, these techniques should include supply correction*, bias correction*, and extrapolation*.

Without these, you may end up with phantom fluctuations that skew your location data set and result in inaccurate human mobility insights.

2

What methods do you use to increase the quality of your contextualization?

Assigning visits by placing GPS pings on specific places on a map — otherwise known as contextualization — is another minefield of potential inaccuracies that needs to be carefully navigated in order to arrive at high-quality human mobility insights.

For example, if you're looking to measure visits to a particular store, how does your provider parse out the devices that actually visited the store from the ones that were merely passing by? You'll want to ask if their maps utilize centroids* or polygons*, and how many GPS pings constitute a dwell event in their methodology (hint: it should never be less than two).

3

What correlation* strategies will you use to enable my use case or achieve my goals?

Correlation can get tricky — depending on which method you choose, you can make data say anything you want. But you want to make sure that the data is telling you the truth, not just what you want to hear.

Your provider's team should be deploying data scientists for two main purposes: to make sure that the data is high-quality and reflects real-world events; and to make sure that the deliverable enables your specific use case. A great provider will work hard to understand which problems you're trying to solve, which questions you're trying to answer, and which audiences you're trying to reach before they propose specific solutions.

4

How do your products integrate with other data sets I work with?

The vast majority of location data buyers are also buying other kinds of data in order to flesh out a fuller picture of their target audience. The data scientists and business experts on your location data provider's team should be able to help you discern how their products and delivery mechanisms complement other data sets and existing tech infrastructure to help you easily draw new conclusions.

5

What's your approach to privacy issues?

With privacy being an increasingly important topic, your provider, at a minimum, should be well-versed in global standards and where they apply, particularly GDPR* in the European Union and CCPA* in California. But regulations are constantly evolving in favor of strictness, and your provider should be making an everyday practice of figuring out how to get just as much value out of location data in spite of regulatory headwinds.

6

How do you measure the validity of your data?

Iteration is the best path to quality. That's why a good provider makes a regular practice of comparing its products to third-party or independent sources to identify where improvements can be made. Unacast employs truth sets* such as the number of devices that connected to a venue's wifi as a proxy for visits. We've also been known to stand outside that venue and physically count the number of people who walk in.

Glossary of Terms

Unfamiliar jargon can be intimidating. This glossary takes the mystery out of location data terminology so that you can be a better-informed buyer of human mobility insights.

ACTUALIZATION

Actualization is a catch-all term for data that becomes useful information or action. It can take the form of: using one or more metrics to identify challenges or opportunities; benchmarking to discover patterns or make comparisons; making predictions based on past patterns; and creating strategic recommendations based on cause-effect relationships.

ADAPTIVE SPATIAL INDEX

Insightful location data requires the identification of areas that represent the boundaries of a place. These boundaries are required so that when a device falls inside the boundary, it is tied to that specific area; in other words, it's how we know the difference between a device that was visiting a place, and a device that was just passing by. Instead of sourcing predefined polygons* from maps, an adaptive spatial index uses algorithmically-derived clustering on human mobility data to identify these boundaries.

AGGREGATOR

An aggregator is a company whose primary purpose is to collect and resell data. They typically provide high volumes of data, but quality can be very hit-or-miss. Many aggregators lack transparency and it can be difficult to determine which methods they use to collect and process the data or to what extent they comply with privacy regulations (if at all).

ANALYTICS

Analytics is the discovery, interpretation, and communication of meaningful patterns in data. Using the simultaneous application of statistics, computer programming, and operations research, analytics often take the form of a prescribed series of metrics that quantify performance, and are captured at regular intervals of time.

API

An Application Programming Interface is what allows different technological systems to talk to each other. For example, Unacast clients can use our API to import location data directly into their content management system. This is particularly useful for sophisticated organizations that collect and use several types of data, including location, in order to create their own insights.

BENCHMARKING

In the context of location data, benchmarking is the act of evaluating a metric against a standard and measuring its difference. For example, you can benchmark the number of visitors to a place and then determine how that number changes month by month. Another benchmarking example would be to compare that visitor count number to the visitor count of a different place in order to measure how the two places differ.

BIAS CORRECTION

Depending on regional attributes, our datasets represent between 5-15% of the US population. In order for us to extrapolate* to the larger population, we need to understand the potential biases that exist. In other words, we need to determine in which ways our source data may not align with the population's demographics, and therefore skew our insights in an inaccurate way. For example, data we source from a given app might over-represent people from certain metropolitan areas, or certain demographic behaviors, because those kinds of people that are the app's target audience. Therefore we use both publicly-available and privately-acquired datasets to adjust our data to ensure that our data sample, and therefore our insights, accurately represent the real world.

CCPA

As of January 1, 2020, the California Consumer Privacy Act (CCPA) is the world's strictest privacy regulation. While it is a state statute intended only for Californians, any location data provider that is deeply committed to privacy, like Unacast is, extends CCPA's benefits to all US residents.

CENTROID

On a map, a centroid is a mathematically-derived center of a given venue. A visit is considered any dwell event that falls within a certain radius of the centroid. But since almost no venues are circular in shape, this method of contextualization can easily result in false positives and false negatives. By contrast, high-quality insights are derived from polygons*.

COVERAGE

Coverage is the percentage of the population captured by the location data set, prior to extrapolation*. Typically, good quality coverage hovers around 5%. Anything significantly higher should be viewed as suspicious, since low-quality, invalid, or otherwise unusable data may not have been filtered out.

CORRELATION

Correlation is the overlaying of different data sets in order to derive meaning. A simple example is a bar chart: the X axis is one data set, and the Y axis is another, and correlation determines how tall each bar is, creating a new understanding of the facts reflected in the data.

DEMOGRAPHICS

As the name implies, demographic data categorizes a population based on certain attributes, some of which are collected with location data. This data often includes information on a device user's gender, age, and ethnicity, but can also include other segmentation types, such as household income. Because not all humans behave the same way, segmenting a population by demographics can uncover patterns within a segment that might not have appeared in the population at large.

DENSITY-BASED SCANNING

Also known as density-based spatial clustering of applications with noise, or DBSCAN, this is a widely-used and reliable data clustering algorithm. It recognizes pings that are close to each other in space and in timestamp (high density), and groups them together. Conversely, it also identifies outlying pings whose closest neighbors are too far away to be clustered (low density). Density-based scanning is a critical component for determining dwell and travel events.

DEVICE FIDELITY

Device fidelity refers to the persistence of a device, or how many times a device is seen in a given time period. The higher the device fidelity — multiple hours within a day, or multiple days within a month — the higher the likelihood that the device accurately represents real-world human behavior. Consistent devices contribute to a more stable data set, which can more readily demonstrate patterns and trends.

EXTRAPOLATION

Inevitably, and for a variety of reasons, location data supply does not capture 100% of any given population. Extrapolation is a data-science-based method that uses census data to fill in the gaps and create a demographically-representative reflection of a given population.

GDPR

The General Data Protection Regulation, or GDPR, governs protection and privacy for data generated in the European Union and Economic Area. It was the world's highest privacy standard until CCPA.

GROUND TRUTH

This term refers to information gleaned from direct observation rather than from an interference source. In the case of location data, GPS technology would be the interference source used to count visitors at a given location; the corresponding ground truth would be an observer physically standing at that location and hand-counting visitors. Employing ground truths ensures that data and processing practices are accurate.

HORIZONTAL ACCURACY

GPS pings comprise a latitude and longitude created via triangulation; horizontal accuracy is a measurement of how likely a device is to be where GPS says it is (see "Precision" below). You can see this in action on map apps: the dot in the center is the GPS' best guess for actual location, and the translucent circle around it represent anywhere else the device might be. The bigger the circle, the less accurate the signal — which is why high thresholds for horizontal accuracy are critical to creating high-quality data products.

OVERLAP

No third-party source of location data collects everything on a given device, which is why companies like Unacast often have multiple sources working in unity. However, that presents a challenge: if the same device appears in two different sources, then the overlap must be resolved so that the device is not counted twice, which would reduce the accuracy of the data set.

POI

“Place of Interest” is a common industry term that can mean a given spot in the physical world. POIs are map inputs that tell us where to focus our insights, and can comprise individual venues (like a store) or a larger area (like a neighborhood or city).

POLYGON

On a digital map, a polygon is a customized area that represents a venue, typically a building footprint, parcel boundary, or similar. Using polygons (as opposed to centroids) is currently the most accurate way to determine if a device visited a place. This is also referred to as venue-level assignment.

PRECISION

Refers to the number of decimal places in a latitude or longitude. The more decimal places, the more precise a reading is. Current mobile device technology creates 6-8 decimal places of precision — more than enough precision for accurate human mobility insights.

RAW SIGNALS

In its most unaltered form, location data comprises billions of “pings”. Every ping, at the very least, consists of a device identifier, latitude, longitude, and a timestamp — in other words, what device was where and at what moment. The challenge with raw signals is twofold: first, they can be malformed and may contain non-human or fraudulent events that need to be filtered out; second, with billions of rows ingested daily, raw signals alone are incomprehensible and require a significant amount of data engineering and processing to be transformed into meaningful data sets or insights.

SCHEMA

At its simplest, a schema is the blueprint for a database: what kinds of data are included, what kinds of calculations can be made, and how they all relate to each other. Often location data and human mobility insights providers will share documentation that outlines the schema, in the form of a table or list.

SUPPLY CORRECTION

An occupational hazard faced by every location data provider is that the incoming supply of raw GPS data constantly waxes and wanes, both quantitatively and qualitatively. As a result, it can be difficult to distinguish which visitation fluctuations are due to supply changes and which ones represent true shifts in human mobility behavior. Your provider should be employing mathematically-derived correction techniques that enable them to parse this out, ensuring that your insights are as close to real-world behavior as possible.


TRUTH SETS

A truth set is a data set that is collected in order to confirm the accuracy of a different data set or insight. For example, visits to a store are counted using GPS data, but could be checked against a truth set of the number of users who were logged on to that store’s wifi. We use truth sets to ensure that our algorithms are calibrated to reflect real-world events.

VERTICAL ACCURACY

Vertical accuracy is a measurement of how far above the ground a device is. This is particularly useful for determining human behavior in the context of multi-story buildings. For example: is the device visiting the street-level retail, or the office above? Today’s technology isn’t quite sophisticated enough to capture a lot of this kind of data, which makes creating insights for high-density areas, like cities, very challenging.



 Clothing store ▶ 50
❤️ 742 💬 53

 Museum ◀ 5
❤️ 2879 💬 651

 Cafe ◀ 110
❤️ 1154 💬 231

Parting Thoughts

Trustworthiness is the highest aspiration of location data and human mobility insights.

It's what will give you confidence that the information you get is accurate. It's how you'll know that you have the best possible jumping-off point for strategic decision-making. And it's what will insulate you from third-party privacy negligence.

That's why, whether you select Unacast or another partner as your human mobility insights provider, our hope is that you'll use this guide to get the highest levels of quality, transparency, and value out there — though you'd be hard-pressed to find a better partner than us.

We look forward to hearing from you.

About Unacast

Creating usable
and transformative
insights from unwieldy
amounts of data

Showing the core
data as it is, without
bias, skew, or
preconceived notions

We **harness the power** of location data by building
the gold standard for **depicting human movement:**
accessible, accurate, and trustworthy.

Lowering the
threshold by providing
delivery options for a
variety of data users

Reflecting real-world
events through the
highest-quality
data available

Delivering privacy-
forward, ethical, and
transparent data you
can rely on

