



Location Data

ADVAN
Research through Technology





Research through Technology

ADVAN RESEARCH CORPORATION
380 LEXINGTON AVE, 17TH FLOOR
NEW YORK, NY 10168
WWW.ADVAN.US
PHONE: 646 880 6656

Advan Location Data

Advan provides data that powers decision making for investors and traders. In particular, Advan focuses on what is known as “alternative data sets”, which collectively comprise any type of data that is different from what investors traditionally have had access to, that is anything outside stock prices, credit, ratings, SEC filings, and such. Alternative data sets in contrast encompass things such as measuring traffic in stores, estimating how successful cellphone applications are, measuring transactions and transaction receipts, measuring web traffic, and other such non-traditionally available information.



Seeking to become **the de facto go-to company for alternative data sets**, Advan started its business by focusing on measuring foot traffic. The reasons are two: first, physical stores represent the majority of where the average consumer spends their disposable income in; despite the advent of the Internet, 94% of sales are still occurring in physical stores^{1,2}; and second, the retail sector, i.e., consumer spending, represents 27% of the GDP³ and is the main driver behind the financial markets as a whole⁴. Therefore, **measuring consumer traffic in physical stores provides the single largest economic indicator**, both on a macro level, as well as for individual companies.

Traffic Data

One of the most important metrics of a company is the amount of sales that it generates, also known as “top line revenue”. In turn, for companies that operate physical stores, which represent 94% of consumer spending, their sales are a function of the traffic in their stores. More foot traffic translates to more sales. Of course every company’s metrics are different, with the average person entering an Apple store spending quite a bit more than the average walk-in at Dollar General, but when comparing foot traffic within a specific company then the relationship between it and sales is clear.

There are different ways of measuring sales at retail stores: by measuring foot traffic, by measuring transactions, by conducting surveys filled by the retailers themselves, etc. Of those, the one which has the potential for (a) widest coverage (capture all the retailers), (b) unbiased measuring, and (c) the largest data sample, is measuring foot traffic. Let's examine all of these in turn.





Research through Technology

ADVAN RESEARCH CORPORATION
380 LEXINGTON AVE, 17TH FLOOR
NEW YORK, NY 10168
WWW.ADVAN.US
PHONE: 646 880 6656

1. Coverage.

Transaction receipts are a great measure of sales, when they are available; but how do you measure transactions for retailers that do not accept credit cards (Bright Horizons, a preschool company; Service Corporation, an undertaker), or the majority of their customers do not have or use credit or debit cards (dollar stores, like Dollar General), or the majority of their transactions are not occurring with credit or debit cards (hospitals; Brookdale Senior Living, operating senior living communities)? Foot traffic in contrast is universal. This has enabled Advan to offer coverage in more companies than any other research company in the sector, currently covering more than twice as many publicly traded companies worldwide as the closest competitor and set to expand to twice that number in the near future.



2. Unbiased.

Just like an American Express Black Card holder has a completely different spending pattern than most of us, so does the average credit or debit card holder differ from the unbanked, which represent 20% of the US population⁵ and clearly are over-represented in certain retailers: if you are living hand-to-mouth you are probably not buying a Coach bag any time soon. However, foot traffic is universal; as long as a retailer's profile does not change drastically within a month or a fiscal quarter, the breakdown of their high vs low paying customers within their stores remains mostly unchanged.

3. Data sample.

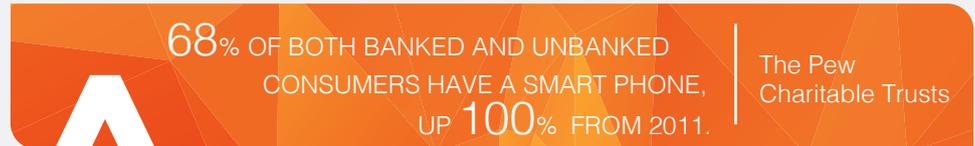
It is self-evident that everybody that shops in a physical store needs to be present at the location. But they may not carry a credit card, and the store manager filling a survey may not count them properly. This is not the case if the foot traffic can be measured in a way that scales and can encompass a large percentage of the population.

For all the above reasons, a scalable means to measure foot traffic is the best way to estimate the performance of most companies in the most accurate and unbiased manner. Which brings us to the next question that guided Advan's data collection, what is the most scalable way to measure foot traffic?



Cellphone Location Data

The majority of the US population, 68%, up 100% in the last 4 years⁶ has a smart- phone on their person, and that percentage is consistent between the banked and the unbanked⁷, further establishing its wide coverage and lack of bias. Clearly, if one wants to be able to measure foot traffic without actually having to install a person in every storefront, the best way is to not measure people, but their cell phones.



The advantage of measuring foot traffic using cellphones is that the method is scalable, potentially being deployed to every cellphone in the world, and the data collection can be automated, resulting in (a) more consistent results, (b) faster delivery of those results, and (c) much less cost on a per-volume basis.

There are however a few different ways to geo-locate a cellphone, and they are not all created equal. In this respect, Advan chose again to focus on the most scalable, accurate and cost efficient way of data collection. Let's review the available data sources and their pros and cons:

Cell phone providers.

Verizon, AT&T, Sprint, and their counterparts, clearly have a very wide user coverage. However, cellphone providers can only triangulate a cellphone using the cell tower signal, which results in accuracy of a few hundred meters. This may be ok if you want to get a measure of the number of people in the vicinity of an airport, but not particularly helpful if you want to measure the traffic in your local Starbucks. In addition, blanketly using and reselling the location data from a cellphone without explicitly notifying the end user that their location is being broadcast to third parties, for whatever usage, could be considered a violation of privacy. For these reasons, accuracy and privacy, cell phone providers' data is at best iffy for our purposes, and one of the main reasons why not a lot of companies use them for financial research.

Advertising exchanges.

Many advertisements popping up on cellphones these days are powered by exchanges; the cellphone application that wants to show an advertisement, logs into an online programmatic exchange, asks "who pays the most for displaying an ad at this time, at this location?"



gets an immediate response and selects the highest paying ad to display. In the process, the ad exchange collects, saves and resells the cellphone's location. It results in a huge data set, providing very wide coverage, however it has a few serious downsides.

First, accuracy is an issue:

The location of the cellphone is not and cannot be verified without extra software checks which would significantly diminish the collection sample (not to mention, the main business of the exchange!); in fact, the ad exchanges do not collect an "accuracy" measure for their data, and a large portion of their cellphones appear to be located in ... Kentucky, because all they can discern is "the phone is in the US" so they place it in the middle of the country.

Second, fraud is prolific.

The application gets paid every time they send an ad request, which results in a lot of fraudulent requests (there is no ad displayed, there is no user that sees it, but the application developer gets paid); how many is anyone's guess, but online ad fraud of this kind is especially prolific and by some measures accounts for 33% of all ads and \$18.5 billion in fraud⁸!



What fraudulent (i.e., fake) data points essentially mean is that you have an active adversary trying to insert random errors into your data sample. It would be a big waste of everyone's time to work on a dataset that is bound to constantly be peppered with new and innovative errors, on purpose, by third parties who are being paid to insert those errors!

Third, bias.

At best, the location of the cellphone is only provided when a user is actively looking at an advertisement. This clearly generates a bias. We do not know how much of it, but we are not particularly in favor of biased data.





Research through Technology

ADVAN RESEARCH CORPORATION
380 LEXINGTON AVE, 17TH FLOOR
NEW YORK, NY 10168
WWW.ADVAN.US
PHONE: 646 880 6656

Fourth, privacy.

Ad exchanges claim that they collect data from tens of thousands of applications. Has every single one of these apps notified their users that their location is being sent to a third party and may be resold? Extremely doubtful. Has every user whose location is being utilized explicitly allowed the application to collect it? Maybe, maybe not. We do not really wish to find out the hard way that user privacy has not been safeguarded properly.

Cellphone Application Location Data.

After researching the above options, Advan settled on using cellphone applications to collect the location data. The way it works is an application that operates on a cellphone installs a small piece of code which collects the user's location, with the explicit permission of each user. The benefits are clear:

Accuracy.

The code that collects the location data uses all available sources to geolocate the phone: cell tower signal; GPS signal; WiFi; and Bluetooth. Note that the phone does not need to be connected to either WiFi or Bluetooth in order for it to be triangulated using those methods; the only thing necessary is for the WiFi and Bluetooth sensors, respectively, of the phone to be turned "on". Because of the extended triangulation methodology, location data collected in this manner are several times more accurate than the methods above.

Scalability.

Today the number of devices (a.k.a. the "panel") which utilize one of the applications that collect the location data used by Advan is over 40 million, of which 80% is in the US. That means over 15% of the US population that has a cellphone! In addition, the number has grown by 100% in the last 6 months alone, and we expect the growth rate to continue. The data sample is not only large, but it has the potential to become larger than the other methods above, even compared to a cellphone service provider's data!

Privacy. Every user explicitly approves their location data collection.

This is extremely important for our purposes, as Advan serves the financial industry which has been under increasingly higher scrutiny, and needs to demonstrate that each one of their data sources is "above board".





Research through Technology

ADVAN RESEARCH CORPORATION
380 LEXINGTON AVE, 17TH FLOOR
NEW YORK, NY 10168
WWW.ADVAN.US
PHONE: 646 880 6656

Summary

At Advan we set out to answer a few simple questions:

What is the most important sector of the economy to measure?

Consumer Spending.

Where does most of the consumer spending occur?

In physical stores.

What is the best way to measure foot traffic in stores?

Cell phone location data.

What is the best and most privacy-respecting way to collect cellphone location data?

Cellphone application location data.

Guided by those answers, Advan's first and currently the largest dataset is cell phone geo location data, which is used as a proxy for top-line revenue of a large set of publicly traded companies worldwide.



Advan's vision is to empower the investors using these types of data sets to make educated decisions that benefit both the investors and the companies deserving of the investments. Transparency of sales reporting is what Advan advocates and strives for.

Citations

¹ Retail Data: 100 Stats About Retail, eCommerce & Digital Marketing, nChannel, July 9, 2015, <https://www.nchannel.com/blog/retail-data-e-commerce-statistics/>

² In-Store vs Online, ICSC, 2013 data, http://www.icsc.org/uploads/press/WSJ_AD.pdf

³ Total US Retail Sales Top \$4.5 Trillion in 2013, Outpace GDP Growth, eMarketer, <http://www.emarketer.com/Article/Total-US-Retail-Sales-Top-3645-Trillion-2013-Outpace-GDP-Growth/1010756#sthash.IHxPhJ4c.dpuf>

⁴ The Relationship between the Retail Industry and the US Economy, Market Realist, September 1, 2015, <http://marketrealist.com/2015/09/relationship-retail-industry-us-economy/>

⁵ 2013 FDIC National Survey of Unbanked and Underbanked Households, <https://www.fdic.gov/householdsurvey/>

⁶ Pew survey shows 68 percent of US adults now own a smartphone, PC World, November 1, 2015, <http://www.pcworld.com/article/2999631/phones/pew-survey-shows-68-percent-of-americans-now-own-a-smartphone.html>

⁷ What Do Consumers Without Bank Accounts Think About Mobile Payments? The Pew Charitable Trusts, June 22, 2016, <http://www.pewtrusts.org/en/research-and-analysis/issue-briefs/2016/06/what-do-consumers-without-bank-accounts-think-about-mobile-payments>

⁸ Report: For Every \$3 Spent on Digital Ads, Fraud Takes \$1, Advertising Age, October 22, 2015, <http://adage.com/article/digital/ad-fraud-eating-digital-advertising-revenue/301017/>

