

Traffic Flow Demographics – an Innovative Tool for Enhancing Accuracy in Business Site Evaluation and Customer Visit Projections

Gregory Brodski, Hanna Hranich, Arthur Stepanyan

October 3rd, 2024

Abstract

This study introduces a novel methodology for identifying the customers demographics that offers improved target audience definition and more reliable customers visit projections for convenience stores and other entities in the retail sector. Traditional methods for determining target audience rely on trade area analysis, neglecting the demographics of passing traffic flow. That is why traditional methods often fail to capture the full scope of potential customers, as businesses such as quick-service restaurants, convenience stores and gas stations depend more on passing traffic than on local population for their customer base. The findings offer a practical tool for real estate professionals, retailers and convenience store operators, empowering them to make more informed choices in site selection, enabling refined financial forecasts, supporting operational insights and broader strategic planning.

Introduction

In the context of retail site selection, a location with the highest Annual Average Daily Traffic (AADT) may not inherently guarantee the best outcome for a store. Adhering to the "the more, the better" principle often leads to mistakes such as choosing the site with the highest volumes of transit traffic, where drivers have no intention of stopping; locations with high traffic fluctuations, peaking only during holidays, where revenue for other active months barely compensates for the off-season; or locations with a small target audience.

Therefore, a thorough evaluation of a site's potential profitability requires considering a broader range of factors, like drivers' behavior that characterize their willingness and ability to stop, traffic congestion, road density, AADT on adjacent roads, and more. In this article we'll concentrate on the target audience factor, since, first of all, it is crucial to determine whether a sufficient number of individuals who are likely to engage with the products or services offered, is present. Sufficient size of target audience ensures that the convenience store, gas station or other business can achieve a robust return on investment (ROI) and meet projected profitability expectations, of course if traffic patterns are favorable and sufficiently stable to bring that audience to the store.

Failure to employ accurate data in target audience estimation can cause severe consequences. It may lead to misidentifying potential customers, resulting in poor site selection, pitfalls in assortment planning, ineffective marketing strategies, and more. The changing demographic and behavioral patterns of Americans further complicate the identification of target customers in different locations.

Existing solutions for estimating target audiences are often inadequate due to their reliance on over-averaged or incomplete data and a lack of demographic specificity.

Firstly, many solutions draw from overly generalized data sources, which fail to capture the nuanced behaviors and preferences of specific audience segments. This results in inaccurate estimates that do not reflect the true size or makeup of the target market.

Secondly, these methods overlook key data provided by traffic flow demographic analysis, which is essential for understanding the demographic parameters of drivers passing by a given location. For meaningful estimation of these parameters it is important to understand the origination point of travel with high resolution. Without this insight, businesses miss critical details about potential customers, leading to lost opportunities for optimization and growth. Accurate audience assessment requires a more tailored approach, integrating detailed local and traffic flow demographic data to provide a clearer picture of market potential.

Studies indicate that approximately 83% of convenience store customer visits are generated by vehicle traffic [4]. With the average commute distance between 8.6 and often much more due to specific location factors (such as being near a highway, etc.), which surpasses the standard 5-mile trade area, the exclusive focus on local demographics deprives retail operators of key data on a large segment of their potential client base. That is why demographic data on drivers passing by specific locations is critical for convenience stores (c-stores), quick-service restaurants (QSRs), and other businesses that depend not only on local residents but also on passing traffic flows.

In such cases, the target audience must be assessed by considering two key data categories: the demographics of the trade area and the demographics of the traffic flow.

Methodology

The methodology of traffic flow demographic estimation leverages traffic engineering principles, Ticon's traffic analysis algorithms, as well as *Target Audience Index* (TAI), as developed and detailed in the study [3]. The practical application of this methodology focuses on improving the accuracy of visit projections for business locations. Applying the methodology developed in this study, which includes the percentage distribution of various age, gender, and income-based social groups among convenience store's potential customers, we aim to estimate the target audience for traffic flow, just as we do for primary and secondary trade areas. The proportional relationship between these three fractions should be further established with the use of statistical regularities available from the article [2].

The target audience of the traffic flow assesses the likelihood of drivers passing by the store with specific demographic characteristics, such as age, gender, and income, to be interested in visiting a store, based on statistical data on the frequency of visits to similar entities.

Traffic flow demographics, by its inherent nature, represents a probability-based approach. It is important to note that no personal data on individual drivers is collected or accessible for traffic flow demographic estimation. The Ticon algorithm operates on the assumption that

the driver is a representative of the average demographic profile of individuals residing within a 1-mile geo-cell where the driver's point of origin is located.

To effectively identify the demographics of the traffic flow passing by a convenience store, a multi-step process is employed, ensuring both accuracy and the capture of long-term patterns. The process involves two primary tasks: those related to processing and analysis of traffic data, and those related to processing and analysis of demographic data.

1. Tracking Passing Vehicles

The tracking of passing vehicles is conducted using Ticon's traffic analysis algorithms, which involve collection, analysis, and processing of data from various sources, including geospatial, navigation, mobile/cellular, IoT, traffic monitoring, and more. To avoid biased data, long-term tracking (over the course of a year) is essential for capturing seasonal trends, ensuring that the data accurately reflects stable traffic patterns and minimizes short-term anomalies.

To enhance the accuracy of customers visit projections, Ticon actively employs the advantages of highly granular data in traffic data analysis [1]. This includes intraday traffic volume estimations, as well as analysis of driver's behavior.

Joint analysis (Fig.1) of intraday speed/volume traffic patterns (left) that should be estimated for each LOI and visit probability rate (right) that is known from industry statistics [5] indicates that the percentage of visits varies throughout the day depending on the location (Fig.1). There are two different metrics to be considered: traffic volume and traffic flow speed variation. The latter is less known but very important since it reflects the driver's willingness and ability to stop. Calculating this metric requires simultaneous consideration of the distribution of vehicle speeds together with the geospatial and geometrical characteristics of the road graph.

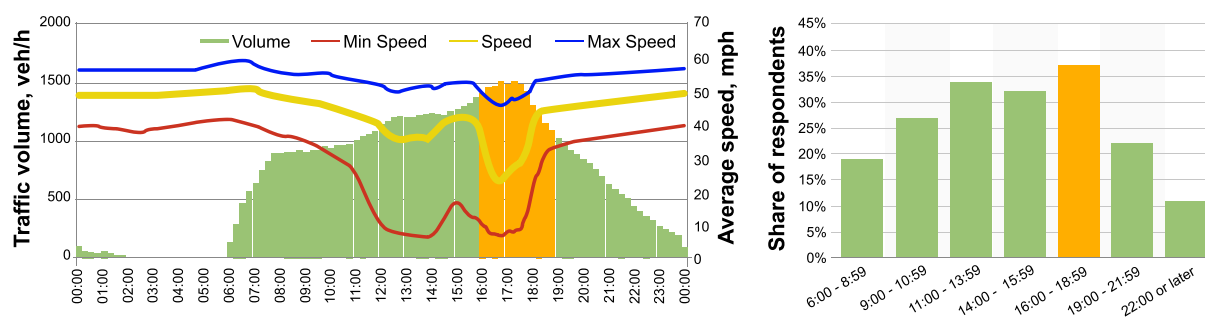


Fig. 1. The Correlation Between Intraday Traffic Volumes, Drivers Behavior and Convenience Store Visits

2. Defining the Point of Origin

To assess the demographic of the passing audience it is necessary to have the information not only on traffic flow, but also on individual moving vehicles. Once the vehicle is tracked, Ticon algorithms discover its points of origin. For convenience store site evaluation tasks the analyzed area corresponds to a up to 2 hour trip. Hence, the maximal distance between

a point of origin and the location of interest may easily exceed 50 miles. Certainly, these temporal and spatial parameters should be adjusted according to the specific nature of the business under evaluation and the geospatial context.

3. Collecting Demographic Data for Each Point of Origin

For each point of origin, the demographic data, including age, gender, race, education, and income, is collected based on the geo-cell corresponding to the specific location. For adequate evaluation, the availability of highly granular data is crucial.

Since the analysis for trade areas smaller than one mile is generally impractical, it is fair to assign one-mile level of demographic data granularity as the desirable one. Unfortunately, conventional sources offer this level of granularity only for major metropolitan areas, or other densely-populated places (Fig. 2). That is why it is important to use sources that enable the extraction and application of demographic data for just one-mile geographic cells [3].

Fig. 2. depicts the typical evaluation of four competitive sites (locations of interest, aka LOI) for rural C-Store (top picture). On the bottom left, the presented demographic data is identical for all four locations, neglecting the unique characteristics of each LOI. This lack of granularity results in the loss of essential insights regarding demographic differences.



Fig. 2. The importance of Data Granularity

In contrast, the right image provides granular data that captures the specific demographic details relevant to each LOI. This level of detail allows for the identification of significant differences among the areas, and facilitates better decision making.

The significance of granular data also lies in its ability to accurately determine the origin of an individual vehicle, as demographic characteristics can vary considerably depending on the point of departure. These variations in demographic data substantially affect the accuracy of consumer behavior forecasts and, consequently, the effectiveness of marketing and operational decisions in the retail sector.

4. Aggregating Demographic Data and Building a Demographic Distribution

The next step is an aggregation of the demographic characteristics of all vehicles registered near a certain road segment. This combined dataset represents a distribution for each demographic parameter (Fig. 3), which enables the identification of patterns and trends within the traffic flow.

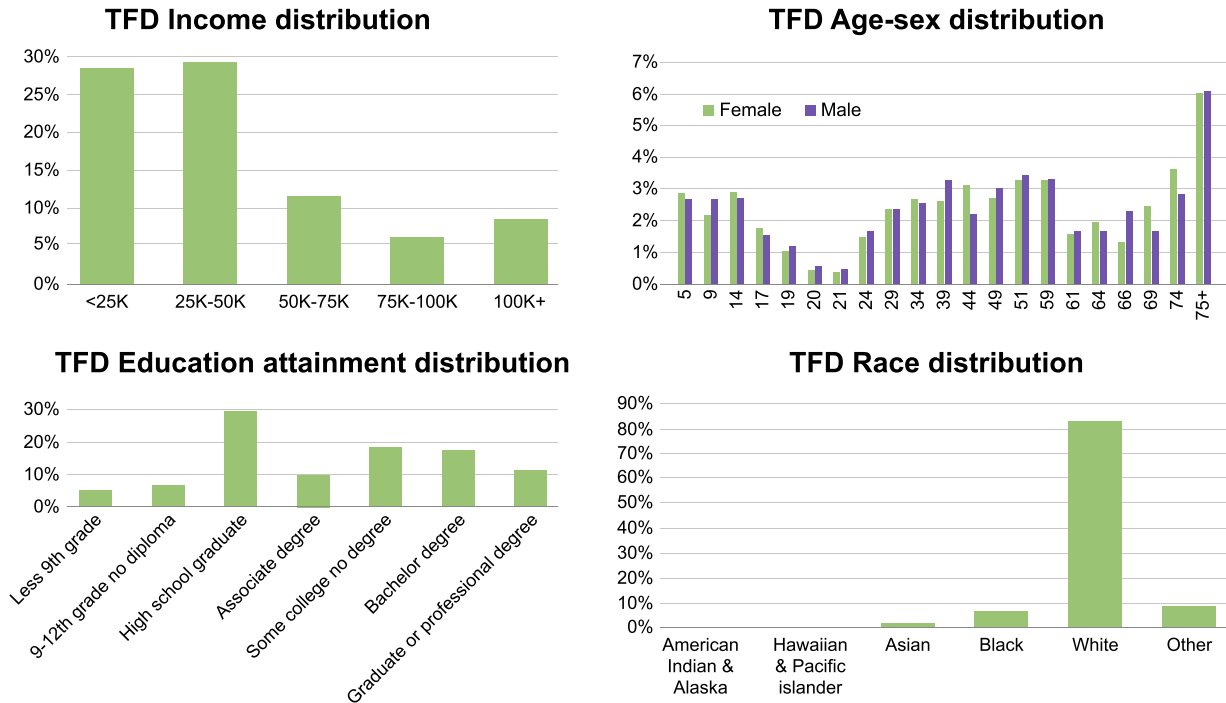


Fig. 3. Traffic flow Demographics Insights

These distributions can and should be directly used for calculation of the target audience. Also, this analysis serves as a foundation for a more accurate assessment of potential customers and their needs, which is crucial for fine-tuning of the offerings, strategic planning, and marketing efforts.

5. Defining the Proportion of the Target Audience

Based on the analysis described above, it is possible to provide estimations for calculating the percentage of the target audience within the traffic flow, utilizing the TAI.

As mentioned before, traffic flow demographics must be considered in conjunction with the demographics of the trade area. However, there is no universally accepted definition of what constitutes a "local trade area" for any certain business entity. Using convenience stores as an example, we can note that the size of the trade area varies based on a range of factors, including the type of a location (urban/suburban/rural), road infrastructure, traffic patterns, including speed/congestion factors, competition, and more. Ultimately, the size of a c-store's trade area may cover a radius from 1 mile to 10 - 20 miles.

The study [2] investigates the relationship between the number of customer visits to convenience stores and influencing factors such as traffic data, demographics, competition, and others, and discovered that number of visits depend on the said parameters in up to 20

miles radius from the location of interest. Using the data from field studies on trip generation [6] together with the results of traffic flow analytics [2], we distinguish 2 zones inside the trade area (Fig. 4): zone of immediate access about 1 mile – in blue, and zone of close access - in red, is from 3 to 20 miles depending on road density and traffic patterns.

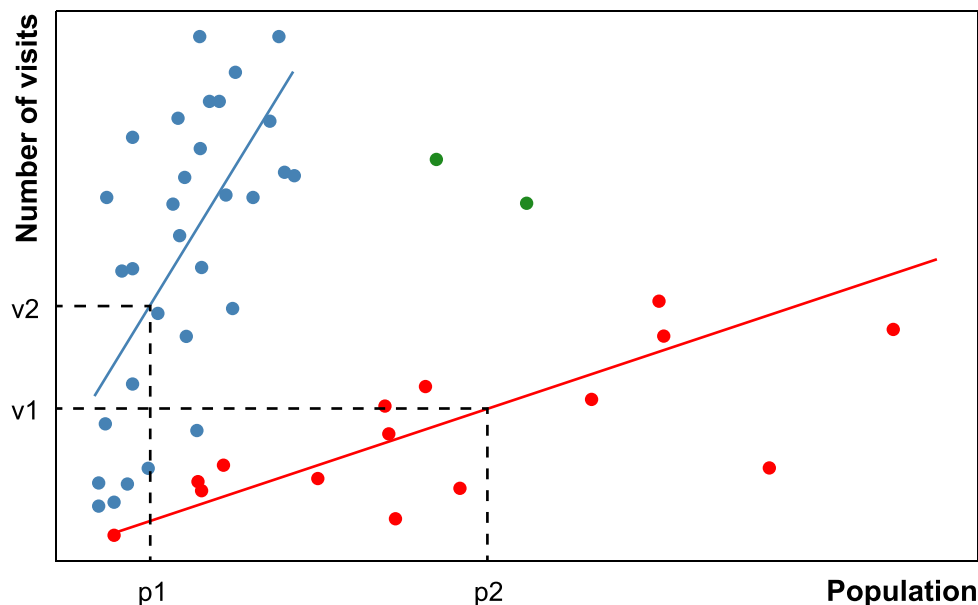


Fig. 4. Convenience Store Visits for the Distinguished Zones

To better understand the influence of primary and secondary trade areas on the number of visits, consider the following example. Assume that a certain population resides within a one-mile radius of the point of interest, while the population within a five-mile radius is significantly larger. Based on this, we can anticipate a specific number of visits from the first zone, denoted by the red line (v_1 on the graph), and a different number from the second zone, indicated by the blue line (v_2 on the graph). By applying a coefficient for the one-mile area, we can derive weights for both the one-mile area and the five-mile area. With these weights, we can incorporate the demographic data of each zone.

The Significance of Distinguishing Target Audience for Trade Areas and Traffic Flow

In order to identify the significance of the proposed solution, we selected 29 sites across 10 states of the U.S. to compare two target audience identification methods: the conventional one, limited to trade area analysis only; and the proposed one, which combines trade area analysis with traffic flow analysis. The selected sites represent a diverse range of locations across the United States, encompassing both business areas and residential areas; the sites with predominantly foot traffic vs vehicular traffic prevalent; sites with minor seasonal fluctuations vs sites with highly noticeable fluctuations. The same applies to the demographics, varying from 12,000 in population to 225,000 population per 5mi radius.

Through a comparison of traffic flow Target Audience and Trade Area Target Audience across 29 locations (Fig. 5), it was found that the difference in the size of target audience members in some cases exceeds 15%, which number in most cases results in thousands or even tens thousands of people.

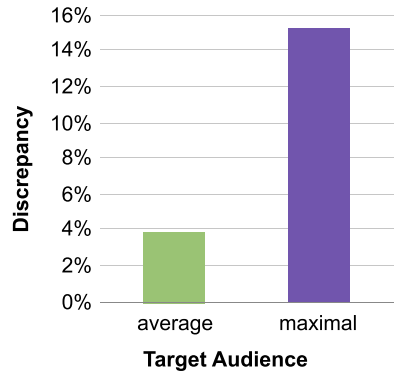


Fig. 5. The Discrepancy Between Traffic Flow and Trade Area Demographics for Target Audience Proportion

The demographic indicators themselves vary even more significantly - up to 30% (Fig. 6,7).

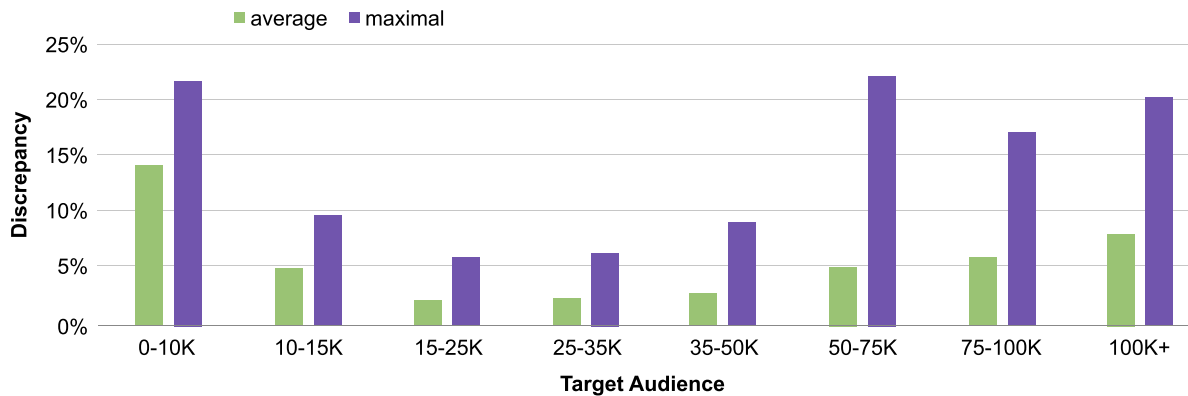


Fig. 6. The Discrepancy Between Traffic Flow and Trade Area Demographics for Median Household Income

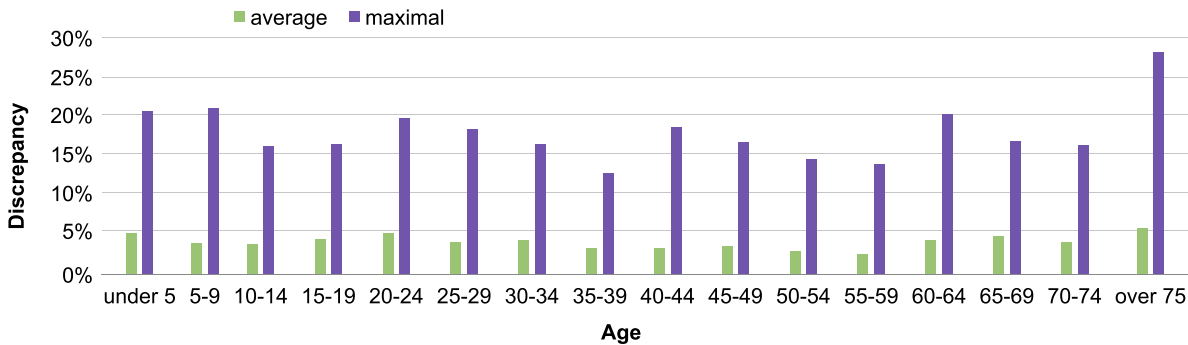


Fig. 7. The Discrepancy Between Traffic Flow and Trade Area Demographics for Age Distributions

It is obvious from the above conclude that the demographic data for trade areas and traffic flow are fundamentally different. Hence, relying solely on the target audience from either group may lead convenience store operators to make misguided decisions in site selection and business forecasting.

Conclusions

1. Highly granular demographic data play a critically important role in the process of defining the target audience and estimating visits, precise analysis of consumer behavior, identification of traffic patterns over time, and understanding of demographic differences.

2. Utilizing detailed traffic data that include intraday volumes and traffic flow speed variations as an indicator of drivers behavior, provides deeper insights into when and how potential customers interact with a business. These insights allow convenience store operators to effectively conduct their site selection, financial projections.
3. Employing a comprehensive approach to the description of a trade area, together with the consideration of traffic flow demographics, allows for more accurate assessment of the potential customer base, enhances strategic planning, and optimizes marketing efforts. Ultimately, this approach increases the likelihood of successful site selection, business strategy implementation and improve the profitability of the entity.

References

1. Brodski G., Chaihorsky A. (2018). AADT Estimation by Various Methods: Accuracy and Reliability https://www.researchgate.net/publication/365806682_AADT_Estimation_by_various_methods_accuracy_and_reliability
2. Brodski G., Kozakevich T., Vyazinko I., Otarov Y., Stepanyan A., Granich A. (2023). Exploring the Visitor Rate in the US Convenience Store & Gas Station Industry https://www.researchgate.net/publication/373144449_Exploring_the_Visitor_Rate_in_the_US_Convenience_Store_Gas_Station_Industry
3. Brodski G., Stepanyan A., Kozakevich T. (2023). Determining the Target Audience in the US Retail Market https://www.researchgate.net/publication/371991213_Determining_the_Target_Audience_in_the_US_Retail_Market
4. Gitnux. 2020 U.S. Convenience Store Sales: \$654 Billion, Key Statistics Revealed <https://gitnux.org/convenience-store-sales-statistics/> (2024)
5. Distribution of times of day for shopping at convenience stores among consumers in the United States as of January 2022. Statista. <https://www.statista.com/statistics/1367227/times-of-day-shopping-convenience-stores/> (2024)
6. Wisconsin Department of Transportation. (2022). Convenience Store/Gas Station Trip Generation Study, Wisconsin – Statewide. <https://wisconsin.gov/dtsdManuals/traffic-ops/programs/analysis/tripgenstudy-gas.pdf> (2023)