Investigating the Relationship Between Transit Planning for Leisure and Household Vehicle Ownership in King County

Grace Anne Young

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Urban Planning

University of Washington

2021

Committee:

Qing Shen

Jan Whittington

Program Authorized to Offer Degree:

Urban Planning

University of Washington

Abstract

Investigating the Relationship Between Transit Planning for Leisure and Household Vehicle Ownership in King County

Grace Anne Young

Chair of the Supervisory Committee:

Qing Shen

Department of Urban Planning

Historically, transportation planning has focused on accommodating peak hour commuting behavior, but in reality, these trips represent just 15% of daily travel (Holden & Linnerud, 2011; Bureau of Transportation Statistics, 2017). Meaningful efforts to curb emissions from the transportation sector must consider a wider range of trip types. Leisure travel makes up an estimated 27% of everyday trip-making behavior in the United States and is typically more spatiotemporally complicated than commuting, increasing the likelihood of driving (Bureau of Transportation Statistics, 2017; Beirão & Cabral, 2010). This research investigates how transit planning for accessing leisure opportunities impacts household-level car ownership in King County through binary logistic modeling built around three key independent variables: average travel time and the transit/auto travel time ratio to nearby leisure destinations, and evening and weekend transit service area. Additionally, k-means clustering of households by leisure style reveals how activity preferences may impact the likelihood of owning one or more vehicles. Results indicate that household car ownership increases with longer average travel times to leisure-specific destinations and, less so, decreases with larger transit access sheds.

Table of Contents

List of Figures

List of Tables

I. Introduction

Perhaps now more than ever, leisure is a crucial ingredient in a happy life. For some, this means watching a football game at a friend's house or rock climbing, while others prefer to enjoy a meal at a restaurant or to lounge on the beach. For everyone, however, this means some degree of travel and a mode choice decision. Cities concerned about harmful emissions associated with personal vehicle travel have a vested interest encouraging users to choose sustainable modes like transit, walking, or biking. In contrast with commuting which follows nearly the same paths at the same times every day, leisure trips are typically more spatially and temporally complicated (Gkiotsalitis & Stathopoulos, 2016). In addition, given the time constraints of a typical 40-hour work week, leisure generally occurs at night or on the weekends when transit service is often reduced. The trip complexity and relative resource scarcity compound in these situations and may provide insight into why mode splits for leisure or discretionary trips tend to be more auto-dominated.

Traditionally, sustainable transportation planning efforts focus on accommodating the peak hour maximum capacity loads associated with commute traffic (Holden & Linnerud, 2011). However, according to the Bureau of Transportation Statistics (2017), commute trips make up about 15% of total trips taken in the United States. By comparison, trips made for social and recreational purposes make up an estimated 27% of total trips while shopping and errands represent another 45% of total trips. Any meaningful efforts to curb emissions from the transportation sector must address trips beyond commuting.

Leisure trips can be thought of as a subsection of discretionary travel, and typically includes trips made for social or recreational purposes like dining in restaurants or going to the park but excludes trips that may be considered essential like grocery shopping, medical appointments, and general errands. To add to the complexity, there are many different leisure styles which creates unpredictable demand. Researchers in Germany went so far as to sort leisure styles into six distinct categories, including FAMOS (family-oriented folks who like to spend their time at home), QUICKFITS (younger folks who live fast lives full of adventure), and CARCULTS (urbanites with a taste for high culture) (Lanzendorf, 2002). All of this is to say that leisure transportation needs are much less clear-cut than commuter needs and cannot be addressed equally for all people.

Seattle has boasted consistent transit ridership growth for the past decade (pre-COVID-19 pandemic) and local residents have repeatedly voted to fund additional transit capacity in the region (Levy, 2017). In 2017, King County Metro and King County Parks partnered to launch Trailhead Direct, a new bus route that links downtown Seattle directly to a popular hiking trailhead nearby. Since its introduction, Trailhead Direct ridership has grown rapidly and expanded service to three additional trailheads, providing round-trip access for an estimated 17,500 hikers during the 2019 season (Constantine, 2019). This service, designed specifically for expanding leisure opportunities, is a great example of targeted transit investments aimed at providing alternatives to the personal vehicle, but effective transit planning for leisure requires more than geographic coverage.

As previously mentioned, nights and weekends are the most popular times for leisure activities but this typically coincides with periods of reduced or suspended transit service. Researchers studied the impacts of extending night service hours on ridership levels and found that the majority of the ridership increase actually occurred during the day where there were no changes to the schedule (Currie & Loader, 2009). They hypothesized that the schedule changes allowed daytime riders that previously would have been stranded at their destination to now have a return transit option. This temporal coverage is the second prong of effective transit planning.

As a car-less individual in Seattle, I personally have experienced late-night long wait times while trying to return home from a friend's house, abandoned bus plans altogether to call a rideshare, or chosen to forego a trip altogether because I missed one bus and the next would not arrive for another 30 minutes. In anecdotal conversations with many of my car-less peers, the main temptation for getting a car is weekend activity access.

Motivated to explore the merit of these discussions, the purpose of my research is to investigate the relationship between transit planning for leisure and car ownership in Seattle. Considerable research has investigated mode choice behavior (Beirão & Cabral, 2010), leisure trip generation (Lanzendorf, 2002), and household vehicle ownership (Clark, Lyons, & Chatterjee, 2016), and interactions between any two of those categories, but there is a gap in knowledge at their intersection. It is well understood that effective transit planning has the potential to reduce vehicle use (Liu & Cirillo, 2015, Kim & Kim, 2004), but this has not been specifically examined in the context of leisure travel. Non-vehicular access to leisure is not only motivated by environmental concerns, but also ensures that all people can equitably access and enjoy a diversity of activities.

Working with data from the 2017/2019 Household Travel Survey, administered by the Puget Sound Regional Council, in conjunction with GTFS data for King County's transit agencies, this research aims to answer the following questions: (1) How does transit access to leisure activities affect car ownership in King County? and (2) How does leisure style impact car ownership at the household level?

In order to answer these questions, I developed three metrics using network analysis tools in ArcGIS Pro to quantify leisure transit access at the household level: transit travel time and the ratio of transit to auto travel time to the nearest 10 leisure destinations, as well as the total area accessible within 20-minutes via transit on weekends and late evenings. Additionally, I performed k-means clustering on households based on their leisure trip-making behavior in order to classify them by leisure style. Finally, I tested these independent variables for significance in a binary logistic regression model predicting whether or not a household owns at least one vehicle.

II. Literature review

Leisure travel is a subset of discretionary trips, which is defined by the Federal Highway Administration as trips with non-fixed locations and times (Federal Highway Administration, 2019). Discretionary trips include those made for the purposes of shopping, general errands, family or personal business, visits with friends or family, or other social and recreational purposes. In the context of the COVID-19 pandemic, leisure trips can be thought of as those made for non-essential purposes. That is, a trip to the grocery store is considered essential travel and thus generally not treated as a leisure trip (CDC, 2020).

Despite representing a greater share of total trips than commuting, leisure travel is historically deprioritized in transportation planning research and practice (Holden & Linnerud, 2011). In fact, researchers found that policies like dense development that are designed specifically to promote multimodal commuting behavior may actually have unintended negative consequences for the environmental impact of leisure trips (Holden & Linnerud, 2011). They contend that while dense development provides shorter vehicle trips, it may create urban conditions that encourage less environmentally friendly leisure activities.

This idea was explored by Næss (2006), examining the leisure travel patterns of people living within the Copenhagen metropolitan area. They found that inner-city residents often lived eco-conscious, transit-friendly lifestyles during the week but compensated by taking longer, more environmentally taxing leisure trips on the weekends. Gim (2018) found similar results when examining weekend household travel survey data in Seoul. Næss theorized that this compensatory travel may be caused, in part, by reduced access to leisure activities in nature during the week.

A systematic review by Czepkiewicz, Heinonen, and Ottelin (2018) of research conducted on the associations between urban form and long-distance leisure travel primarily in western Europe found the story to be a bit more complicated. Considering domestic or regional travel, urbanites actually travel for leisure less frequently than suburban dwellers, regardless of their access to green space. However, when considering international travel as well, urban residents do travel more for leisure than those residing in the suburbs.

Of course, this debate centers around the idea that leisure travel is driven by access to nature or other green spaces. This is a broad generalization, as pointed out by Lanzendorf (2002). In a cluster analysis of the weekend travel patterns and questionnaire responses of households in four neighborhoods in Cologne, Germany, Lanzendorf identified six distinct mobility styles associated with different leisure preferences. While some had a propensity for spending time in nature, others preferred activities like museum-going or fine dining, activities well suited for urban life. The researcher found that mode options like access to a car impacts a person’s leisure style, but that the opposite is also true (a person’s leisure style impacts their car access).

Analysis of responses from the 2003 American Time Use Survey supported the finding of multiple leisure styles. Fan and Khattak (2012) identified five lifestyle orientations, passive leisure, socializing, family, recreation, and community, and arrived at the similar conclusion that different leisure styles have different levels of car dependence. They found that family-oriented leisure preferences are associated with the highest levels of car dependence while recreation-oriented lifestyles exhibit the lowest levels.

The work of Beirão and Cabral (2010) attempts to connect the differences in the nature of leisure and work trips to mode choice decisions that impact car dependence. They explained that, in general, leisure trips are more susceptible to travel stress because of the spatial and temporal variability. This, combined with the finding that a need for control is highly important in leisure travel, helps to explain higher rates of car usage for leisure activities than work commutes.

When studying the leisure mode choice of seniors specifically, Schwanen, Dijst and Dieleman (2001) found that older people with cars will tend to use it frequently for leisure and will rarely or never substitute public transit options for car trips. However, public transportation did present as a substitute for walking or biking trips. These findings, particularly regarding car use among seniors, persisted throughout different urban and rural settings.

In order to inform future sustainable transportation policy strategies, it is crucial to understand the reasons why households buy cars because it may help reveal fundamental flaws in alternative transportation options. Clark, Lyons, and Chatterjee (2016) found that changes in car ownership levels often happen in concurrence with life events like moving, changing jobs, or the birth of a child. In the short term following these changes, households often resist changing car ownership but ultimately, the researchers hypothesize, the changes in roles, relationships, spatial contexts, and lifestyle preferences may lead to stress that creates dissatisfaction with current car ownership levels.

Nolan (2010) found additional evidence supporting this idea of inertia in car ownership, that is that households are likely to stay in their current state of ownership, in longitudinal data from Ireland spanning between 1995 and 2001. She also found that variables like income, specifically permanent income, household characteristics, specifically the presence of young children, and lifestyle effects significantly predict car ownership. Additionally, changes in income influence the purchase of a first car more than subsequent vehicles.

Van Acker, Boussauw, and Witlox (2011) challenged the notion that lifestyle is a direct predictor of car ownership, but rather that it influences where people live which in turn may impact car ownership. More broadly, they investigated subjective variables like mode specific attitudes and their impact on car ownership and mode choice and found that public transit use is not influenced by attitudes towards transit as much as it is influenced by attitudes towards cars. That is, people who choose to ride transit generally feel more negatively towards cars than they feel positively towards transit.

Researchers have attempted to directly model the relationship between transit service provision and car ownership. Liu and Cirillo (2015) used household travel survey data and General Transit Feed Specification (GTFS) information from the Washington D.C. metropolitan area to estimate an econometric model that found a negative relationship between transit service quality and car ownership and use rates. The reductions in car use were particularly strong when the transit improvements focused on bus systems as compared to other transit services like light rail. Kim and Kim (2004) found a similar relationship using an ordered logit model to predict car ownership using the inverse square root of distance to transit as a measure for transit access.

These models investigated transit services as a whole without taking into consideration specific transit features that benefit leisure trips. For example, Currie and Loader (2009) investigated the impact of extending evening services on ridership and found that the greatest growths in ridership occurred during time periods unaffected by the schedule changes. They hypothesized that extending transit services allows for more feasible transit round trips because return service exists where it previously did not.

Simmons and Haas (2016) provided additional evidence for this hypothesis with their findings across nine agencies in the western United States. Additionally, they found that cutting evening hours increases ridership on the last routes of the night but decreases overall ridership and the expanded hours actually adds new riders rather than just shifting existing trips around. Furthermore, changing midday service impacted ridership more than altering evening service.

On the other end of the spectrum, Lu and Reddy (2012) made a business case for cutting Friday services as more people choose to work 4/10s or telecommute more frequently (pre-pandemic). Accepting the framing that providing transit is first and foremost about running a business, transit agencies have an interest in providing maximally efficient services, meaning reducing service in line with reduced demand.

Gkiotsalitis and Stathopoulos (2016) present timetable rescheduling as a possible demand-response solution for leisure transit trips. Using a heuristic method that adjusts timetables to meet leisure demand while preserving quality of service, the researchers demonstrated possible system improvements on two bus lines in Stockholm, Sweden.

In order to quantify transit service quality for modeling purposes, researchers have developed specific tools for measuring transit accessibility. Fayyaz et al. (2017) created an open-source toolbox with an algorithm designed for large-scale GTFS transit accessibility inquiries that can be used to assess the spatiotemporal variation of transit services. Benenson, Martens, and Rofé (2010) presented another open-source option that exists as an ArcGIS extension for measuring high-resolution transit and car accessibility that is particularly well suited for comparing accessibility levels by different transportation modes.

Inspired by these tools, I developed three leisure transit specific metrics to feed into a binary logistic regression model in addition to demographic considerations like household size and income, similar to Nolan (2010), and household leisure styles, inspired by Lanzendorf (2002). This allows me to test for their impact on car ownership, both in directionality and statistical significance, and contribute to understanding how leisure transit planning in King County relates to household vehicle decisions.

III. Data Sources

In order to answer the key research questions, this work will rely on data from two main sources: the combined 2017/2019 Household Travel Survey (HTS), and General Transit Feed Specification (GTFS) data for King County Metro and Sound Transit. Both sources are publicly available online for free download through the respective agency websites.

The 2017/2019 HTS was conducted within the region by the Puget Sound Regional Council and includes data from both the 2017 and 2019 surveys. Another survey collection is planned for 2021, but this data will not be available in time for use in this research. Because of this, all conclusions drawn from this dataset will be for assumed travel behavior prior to the COVID-19 pandemic, which resulted in widespread changes to daily travel activities. The HTS was conducted on a stratified set of households within the Puget Sound Region, including a special focus on residents within the City of Seattle thanks to a partnership with the city. Participants reported their household's trip activity either via mobile app tracking over a period of 7 days, or via website for a single day. Because the survey is designed to capture regular, everyday travel, collection methods focused on weekday (Monday through Thursday) trips made within the region so households reporting just a single day have data about weekday leisure but not weekend behavior. Approximately one-third of households elected to record data for the full week.

The HTS data is reported in three data tables that may be joined on common identifiers. The household table contains information about the entire household like income, location, and number of vehicles available. It can be joined to the person table by a unique household ID. Each person within the household is represented in the person table with individual information like race, age, and gender. The bulk of the dataset is found in the trips table, where each trip made by an individual is represented with information like mode, purpose, origin/destination location, and travel time. In all, there are over 100,000 trips made by about 6,000 individuals representing over 3,000 households captured in the data tables.

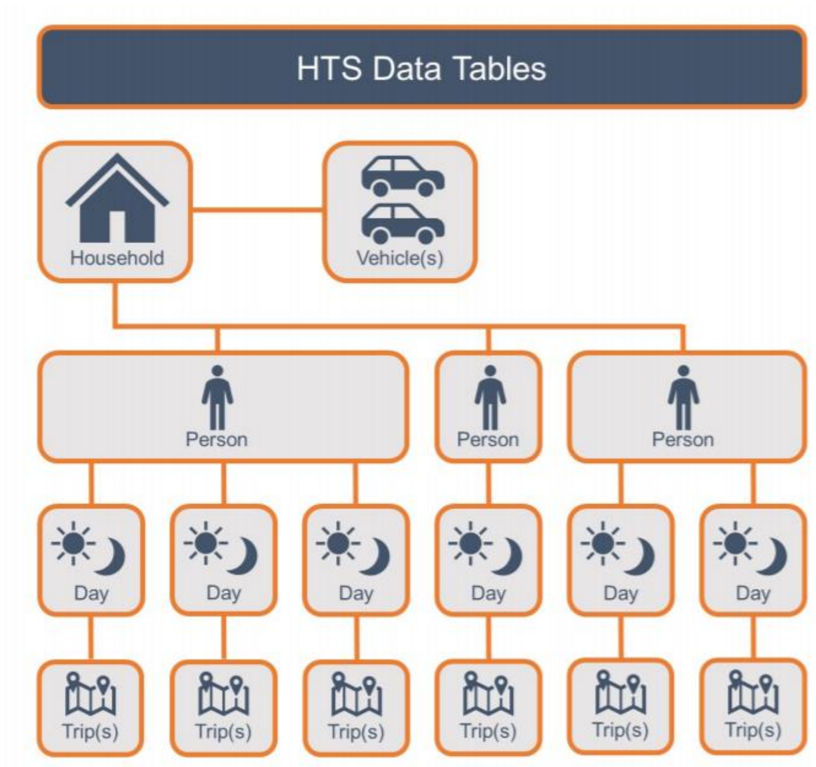


Figure 1 HTS data table relationships, from PSRC and RSG.

The second major data source is GTFS data from King County Metro and Sound Transit, the two public transit agencies operating within the City of Seattle. GTFS data is reported on a bi-weekly basis and contains information about spatial characteristics like stop locations and route connections, as well as temporal characteristics like route frequencies and operating days. Because HTS trip data was all measured pre-pandemic, I used GTFS data reported on January 13th and January 14th, 2020 for King County Metro and Sound Transit, respectively, to capture the appropriate transit conditions that these households would have experienced during their reporting period. Of course, there may have been service changes between 2017 and 2019 that impacted household-level transit access, but for consistency I elected to just use a single reporting period. This data is primarily used for leisure transit metric development via network analysis tools in ArcGIS Pro.

IV. Methods

To answer the main question, "how does transit planning for leisure impact car ownership in Seattle?" I built a binary logistic regression model to predict whether or not a household owns one or more vehicles. This type of statistical model is designed for nominal dependent variables (in this case, if the household owns a car) and can handle both continuous and discrete independent variables. Additionally, this type of model does not require perfectly independent variables which is beneficial in this situation because there may be some correlation between variables like access shed and travel times.

This model determines if the developed leisure transit metrics are statistically significant in predicting whether or not a household owns at least one vehicle. I developed three of these independent variables, using network analysis tools native to ArcGIS Pro in the ESRI suite: total area of a 20-minute transit network access shed using the service area tool, average travel time and auto/transit travel time ratio to the 10 nearest leisure destinations using the nearest facility tool.

Network Analysis in ArcGIS Pro

In order to perform this analysis, I first built a transportation network using the GTFS data for King County Metro and Sound Transit and OpenStreetMap (OSM) road centerline data downloaded from Geofabrik. The network includes all stops on all routes, even those that travel outside of King County. Because the specific geometry of each route is not reported in GTFS data, the Connect Network Dataset Transit Sources to Streets tool snaps stop locations to street centerlines and infers route paths from the shortest path between adjacent stops. The downloaded

OSM data included additional pathways like sidewalks, stairs, and shared use paths that are not suitable for vehicle traffic, so I filtered these network components out of the final street network.

The final logistical step in establishing the complete transit network involves designating which streets allow pedestrians in order to establish walking routes between homes and stops. OSM does not come with an attribute that indicates pedestrian suitability, so I applied my own filter based on the functional class attribute field and restricted pedestrians from segments classified as ‘motorway,’ ‘motorway link,’ ‘track,’ or ‘unclassified.’

Within the ArcGIS Pro online resources, ESRI provides a network template specifically designed for converting GTFS data into a network dataset that implements the appropriate connectivity rules and travel attributes (ArcGIS, 2020). The connectivity is such that a rider can only switch between a street and a transit line at the designated stops and road segments are linked if they share any vertex. There are two key travel cost attributes: walking time and public transit time. Walking time is calculated based on the length of the segment times a constant walking speed of 83.33 meters per minute and transit time is calculated from GTFS data and scheduled stop arrival times. There is also a waiting time component included in the case that a simulated pedestrian arrives at a stop at 10:10 A.M. but the next bus does not arrive until 10:15 A.M. which would add those five minutes of waiting onto the total travel time.

Of course, transit systems are not always on time, nor are travel times between stops necessarily consistent from one trip to another, but for the purposes of this analysis I am ignoring any real-world and random disturbances to the transit network behavior and assuming the system functions precisely as described in the GTFS data to limit the complexity of analysis.

Finally, I implemented a drive mode for transit/auto travel time ratio calculation. Using the same road shapefile from OSM that supports the transit mode network, I built a separate network dataset for driving. The OSM files provided contained speed limits for select road segments so I imputed the rest based on the provided data and typical speed limits by road functional class (e.g. 20 mph on residential roads). The OSM data does not include presence of stop signs, traffic signals, or other traffic control devices that slow down cars from free-flow speed so I implemented a parameter that increases drive time by 10 seconds for every node intersection the route passes through to simulate these missing impediments based on an estimated level of service C and half of intersections with traffic control in place ("Signalized Intersections," 2004). Additionally, these driving time calculations do not account for traffic conditions that may further reduce vehicle speeds.

Calculating Leisure Transit Metrics

Using Network Analysis tools in ArcGIS Pro, I developed three leisure transit specific metrics to test as independent variables in the binary logistic model. The first, total area serviced by a 20-minute transit access shed, was calculated using the Service Area tool with each King County household as an initial facility. This area includes all places accessible within a combined walking, waiting, transit traveling time of 20 minutes or less.

Two metrics depend on travel time to specific leisure destinations. These locations are pulled from the OSM Points of Interest shapefile, included in the full OSM street data download. To specify leisure destinations like parks and cinemas from other types of destinations like hospitals and banks, I selected points with functional classes as listed in Table 1.

Table 1 Selected functional classes of leisure specific destinations

arts center	cafe	florist	museum	restaurant
attraction	cinema	food court	nightclub	sports center
bakery	clothes	garden center	park	stadium
bar	community center	golf course	picnic site	swimming pool
bicycle rental	department store	ice rink	pitch	theater
biergarten	dog park	library	playground	theme park
bookshop	fast food	mall	pub	zoo

The final selection included 13,309 leisure destinations which I inputted as the "facilities" along with King County household locations as the "incidents" in a Closest Facility analysis. The algorithm selects the ten nearest facilities to each incident and calculates the travel time to reach that destination. Using the dissolve tool, I generated an average travel time for each household. One metric includes just the average travel time by transit and/or walking, while the other includes the ratio of transit and/or walking time to the driving time.

I calculated values for each of the three measures based on three different departure times to capture the differences in transit service from day to day and throughout different hours of the day. Additionally, leisure trips typically take place outside of work hours which for most people is evenings and weekends. Because of this, I calculated the metrics at 9 pm on Friday night, and 10 am on both Saturday and Sunday mornings.

K-means Clustering by Leisure Style

In order to capture the differences in households beyond the demographic characteristics reported in HTS, I developed a clustering model inspired by Lanzendorf's (2002) work on grouping German households by leisure style. Designed to explore household behavioral differences beyond the demographic information reported in the HTS, I used trip information reported by each household to produce independent variables similar to those used by Lanzendorf for the k-means clustering procedure. These variables are summarized in Table 2.

In one key departure from Lanzendorf's procedure, I elected to exclude any independent variables that measured mode choice because of the potential of double-modeling these in the binary logistic regression model.

Table 2 Description of independent variables involved in k-means clustering of households by leisure style.

Variable	**Description**
frac_leisure	Percentage of total trips reported by a household designated as leisure trips
med_leisure_distance	Median distance of leisure trips made by a household
food_frac	Percentage of leisure trips made to eat at a restaurant or get takeout
rcv_frac	Percentage of leisure trips made for religious, community, or volunteer activities
rec_frac	Percentage of leisure trips made to attend a recreational event (e.g. movies, sporting event)
social_frac	Percentage of leisure trips made for social events (e.g. visit with friends or family)
exercise_frac	Percentage of leisure trips made for exercise purposes (e.g. went for a run, bike, or walk)

Each trip in HTS has a categorically reported trip purpose. I defined leisure trips as discretionary (non-work or school) activities excluding essential errands like grocery shopping, medical appointments, or trips to the bank. Functionally, in HTS data these are trips with a simple origin or destination purpose listed as “recreation/eat meal.” Following the traditional trip generation definitions, I counted the total number of trips made by a household wherein the leisure activity is considered the trip attraction (“Four-Step Travel Model,” n.d.). By convention, this is any home-based trip to or from a leisure destination, and any non-home-based trip taken to a leisure destination. Trips that originate from a site of leisure but end at non-home locations are not counted as leisure trips. These trip classifications are explained visually in Figure 2. Households with no recorded leisure trips were excluded from the clustering analysis.

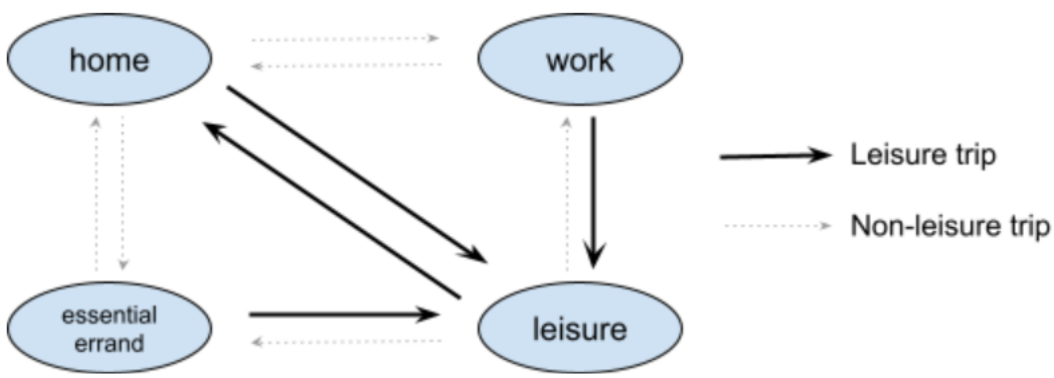


Figure 2 Leisure trip classifications by origin and destination patterns.

Because of the continuous nature of all the independent variables, a k-means clustering technique was most appropriate. K-means clustering requires independent variables with approximately the same scale, otherwise variables with larger scales will disproportionately impact clustering results. Six of the independent variables are reported as percentages, so their ranges naturally fall between 0 and 1, but I had to transform the median distance traveled variable into this range. To preface this procedure, I filtered out extreme outliers using the interquartile range (IQR) test.

With all variables standardized, I used the cluster package from the sklearn library in Python to carry out the unsupervised learning procedure and selected the number of clusters based on the 'elbow method,' as calculated using the KElbowVisualizer from the yellowbrick library in Python (Kodinariya & Makwana, 2013). In this case, the optimal number of clusters is identified as 4, as shown in Figure 3.

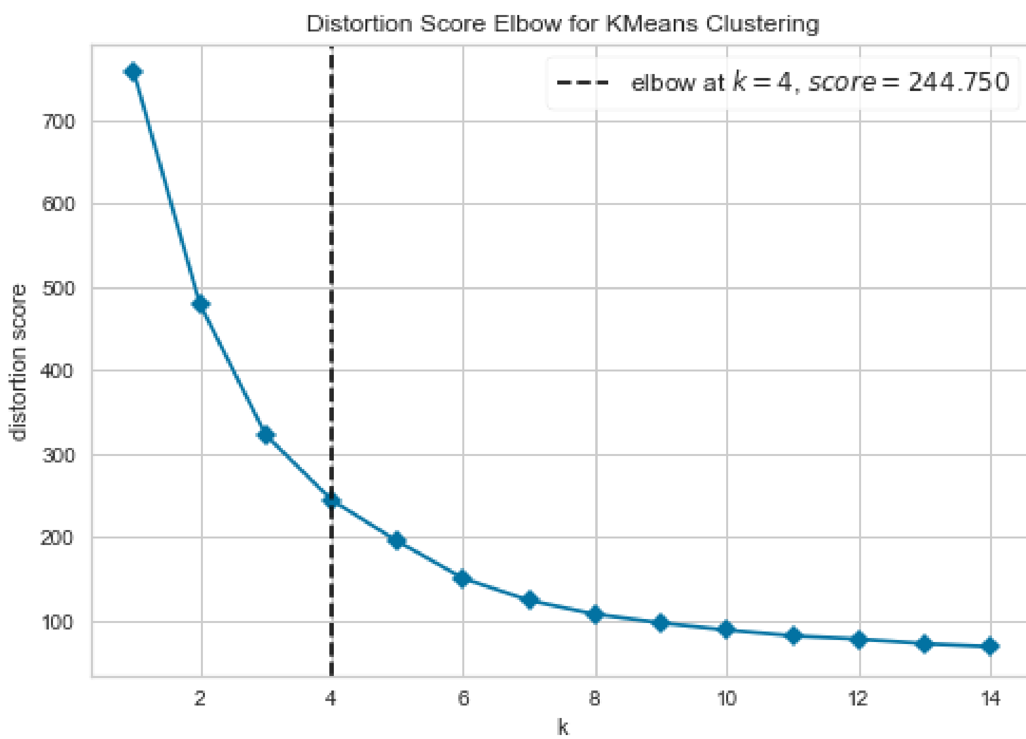


Figure 3 Elbow method using KElbowVisualizer for selecting the optimal number of clusters.

Developing the Binary Logistic Model

Because of the discrete, binary nature of household car ownership which serves as my dependent variable, a binary logistic regression model is most appropriate for analysis purposes. I chose this over a multinomial logistic model which would be suitable for predicting the number of vehicles a household owns because of my focus on baseline car ownership and zero-car lifestyles.

Using the three independent variables I developed in ArcGIS Pro describing leisure transit access for each household, 20-minute transit access shed and average transit travel time and auto/transit travel time to the 10 nearest leisure destinations, I built a baseline binary logistic model using the statsmodels library in Python. Based on previous literature (Clark, 2009; Nolan, 2010), I also included two variables shown to have an impact on household car ownership rates: household size and income which I implemented as a binary variable indicating if the reported household income is greater than $75,000 or not, excluding households that elected not to answer. The survey collected income data as a categorical variable, thus the decision to convert to a dummy variable rather than represent continuously. A conceptual model describing the model inputs and their origins is provided in Figure 4.

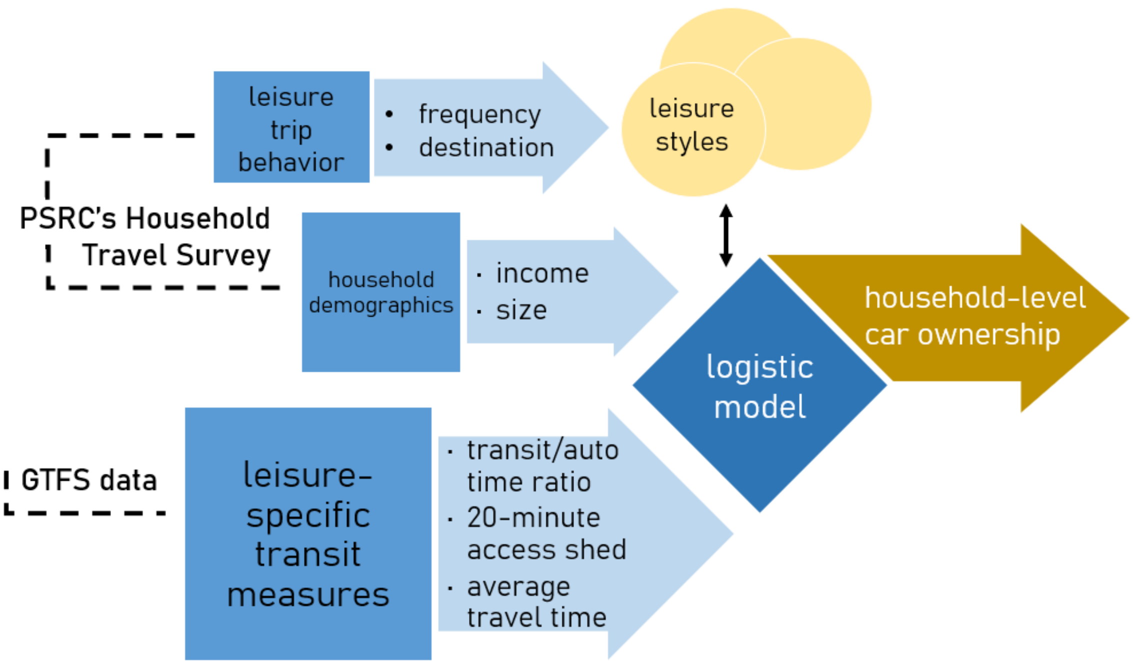


Figure 4 Conceptual model describing the inputs for the binary logistic model predicting household-level car ownership.

Like all regression models, binary logistic models are sensitive to extreme outliers and overall model performance may suffer if outliers are present in the dataset. To address this, I implemented the IQR test for each of these independent variables, removing any records with any attribute value greater than three-times the IQR plus the third-quartile cutoff or less than the first-quartile cutoff minus three-times the IQR. This reduced the total number of household records included in the final model by 4.4%, from 1,711 to 1,635.

Additionally, independent variables in binary logistic regression must be *independent*. That is, variables should not be too highly correlated, or multicollinearity issues may impact the model performance. This is because regression works by measuring the change in dependent variable with changes in each independent variable while holding all other independent variables constant. If two variables are too highly correlated, meaning that they increase or decrease together, the model may struggle to distinguish which independent variable is responsible for a corresponding change in dependent variable.

With the three leisure transit metrics, it is reasonable to assume there may be high levels of correlation between variables because they are, most broadly, measuring the same system in different ways. Additionally, travel times from the same station on Friday and Sunday, for example, are likely highly related because system-wide service reductions on the weekends are probably nearly constant across all lines. I identified and diagnosed multicollinearity problems using a variance inflation factor (VIF) test, removing variables with VIF greater than 5.

Finally, I included household leisure style as determined by the k-means clustering by generating three dummy variables indicating membership in cluster one, three, or four. I used cluster two as the reference variable because it has the greatest number of households.

V. Results and Discussion

K-means Clustering by Leisure Style

The 1,727 households located in King County with at least one leisure trip recorded were clustered into four distinct categories based on seven independent variables derived from trip-making behavior reported in the 2017/2019 combined HTS. The clusters are generally described in Figure 5, with further detail on the distribution of independent variables in Figure 6.

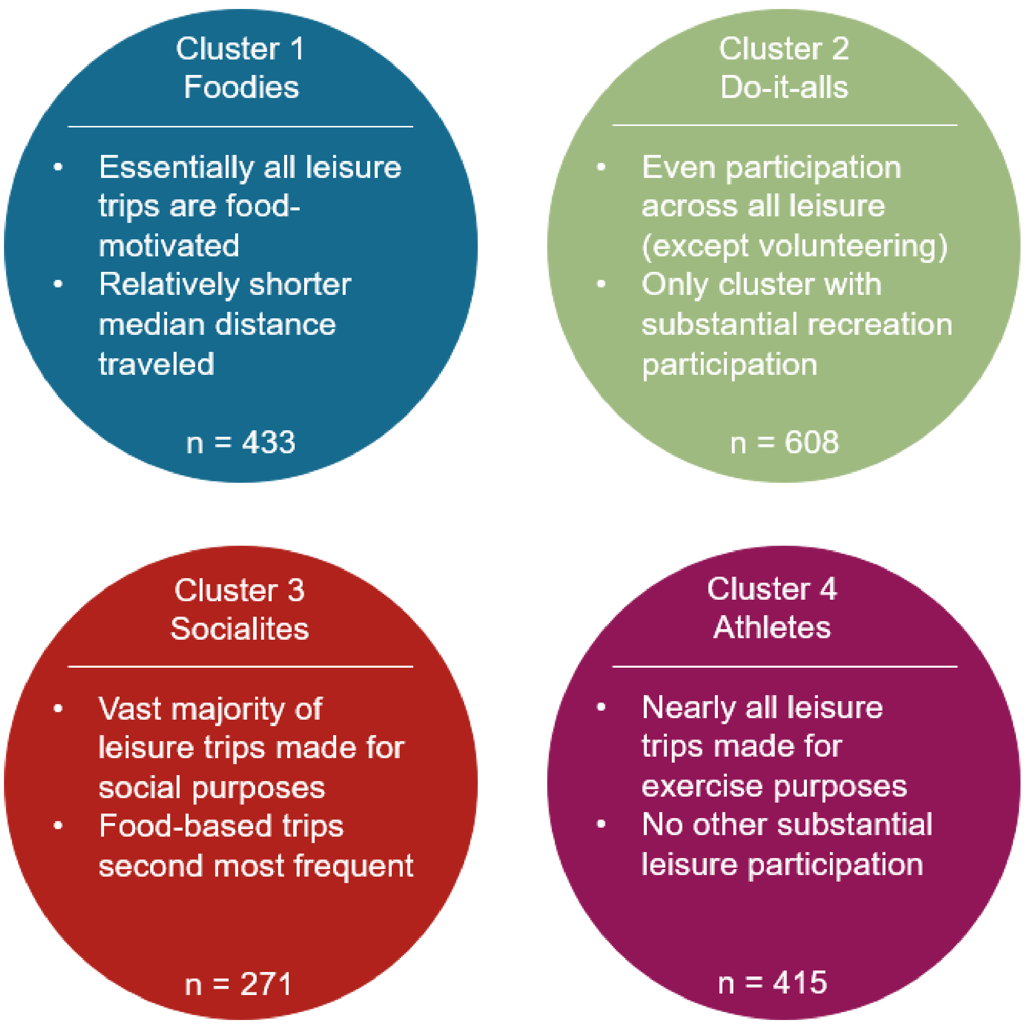
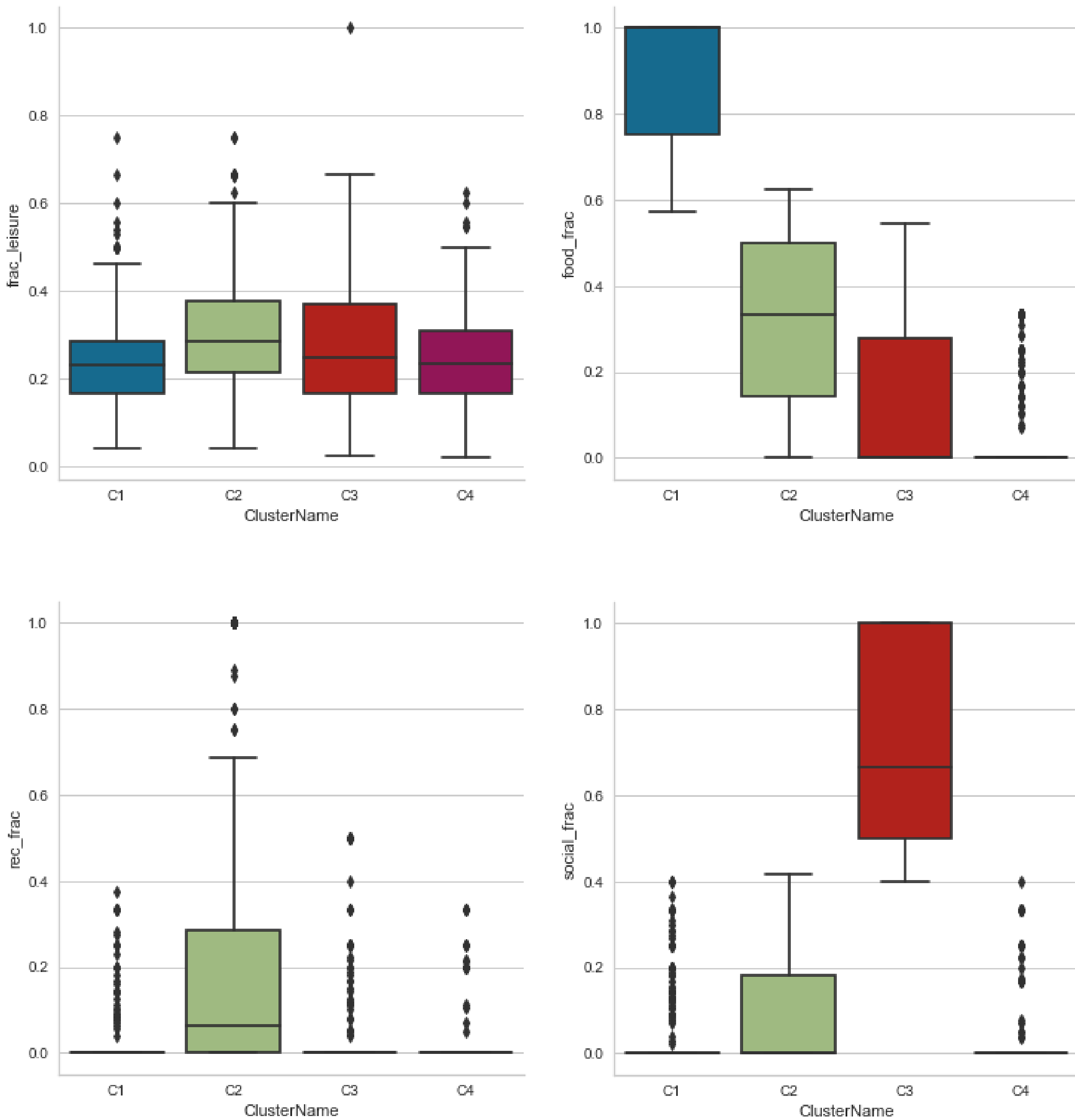


Figure 5 High-level descriptions of four household leisure style clusters.


frac_leisure
food_frac
rec_frac
social_frac
0.0
0.2
0.4
0.6
0.8
1.0
C1
C2
C3
C4
ClusterName

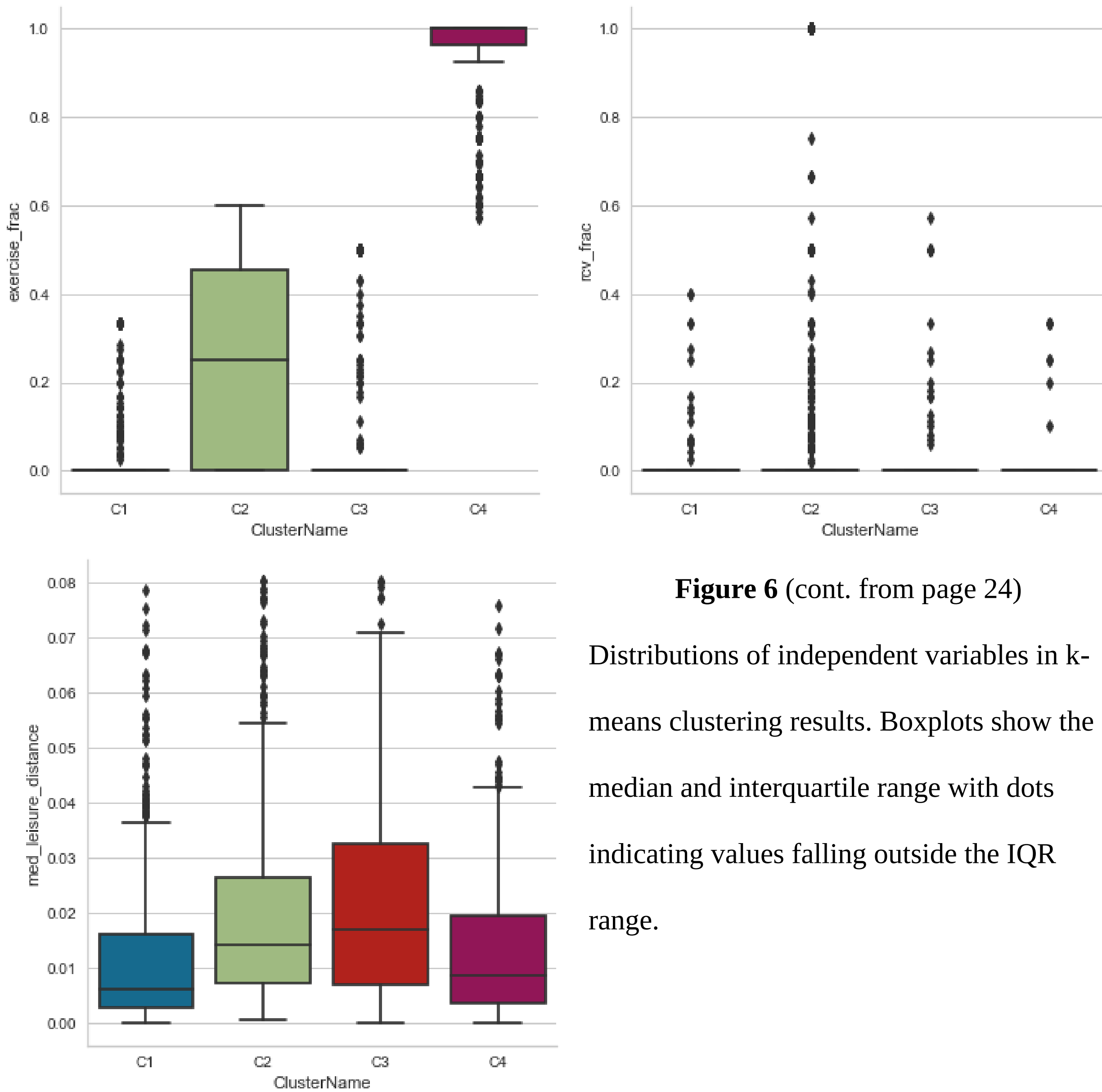


Figure 6 (cont. from page 24)

Distributions of independent variables in k-means clustering results. Boxplots show the median and interquartile range with dots indicating values falling outside the IQR range.

The clusters exhibited modest variation in the frequency of leisure trips as a percentage of the total number of trips taken by a household. The median values fall between about 25% and 30%, consistent with the 2017 Bureau of Transportation Statistics estimate that 27% of trips made in the United States are for leisure purposes. Similarly, the normalized median distance traveled for leisure shows small variations between clusters, with clusters two (do-it-alls) and three (athletes) exhibiting slightly higher medians and larger IQRs.

What primarily differentiates the clusters is the type of leisure activities attended. Other than cluster two (do-it-alls) which exhibits a nearly even distribution of leisure participation across the five categories surveyed, each cluster is marked by significantly high participation in a single type of leisure activity, using IQR as a proxy for confidence interval.

Silhouette analysis is commonly used to assess the quality of clustering results and is built on the intra- and inter-cluster distances of each observation. An overall score of 1 indicates perfect, tight clusters. The overall silhouette score of the leisure style clusters is 0.42 and Figure 7 shows the silhouette scores of each observation.

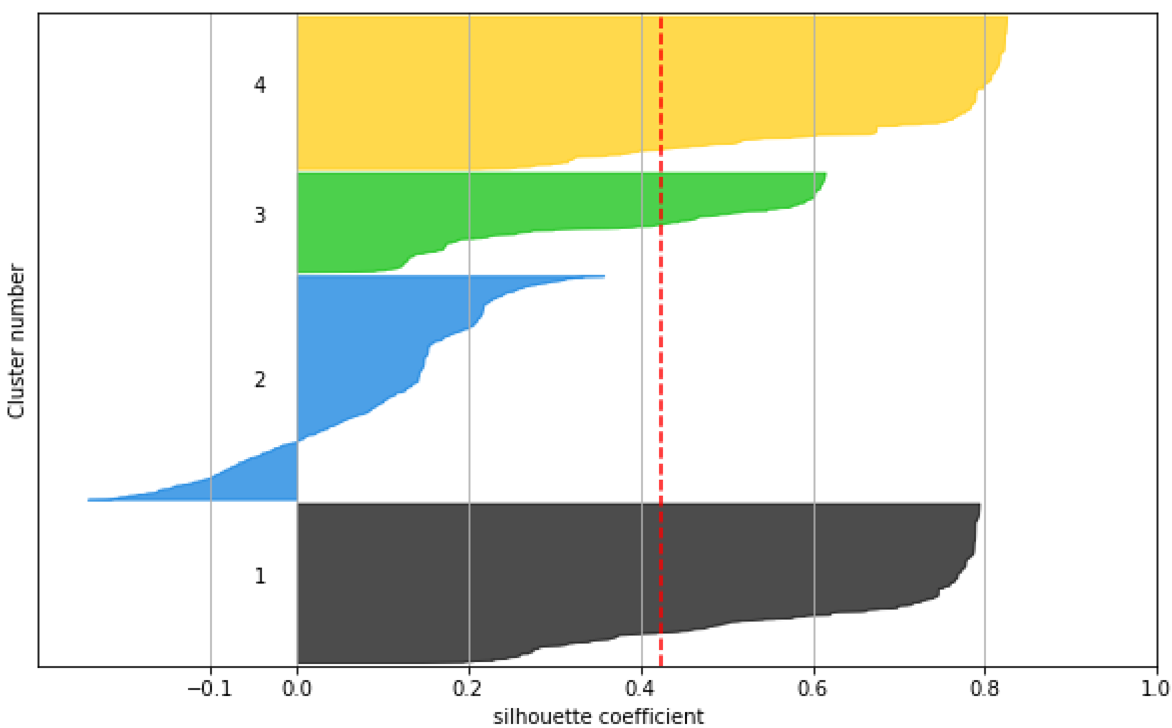


Figure 7 Silhouette analysis of household leisure style clusters.

Clusters one, three, and four have strictly positive scores, indicating good performance, but about one-third of observations in cluster two have negative scores, indicating a possible overlap in clusters. This overlap can be reduced by increasing the number of clusters, but at the risk of overclassifying.

Spatially, there are no obvious groups of household leisure styles as the households in each cluster are quite randomly distributed around the county, per Figure 8.

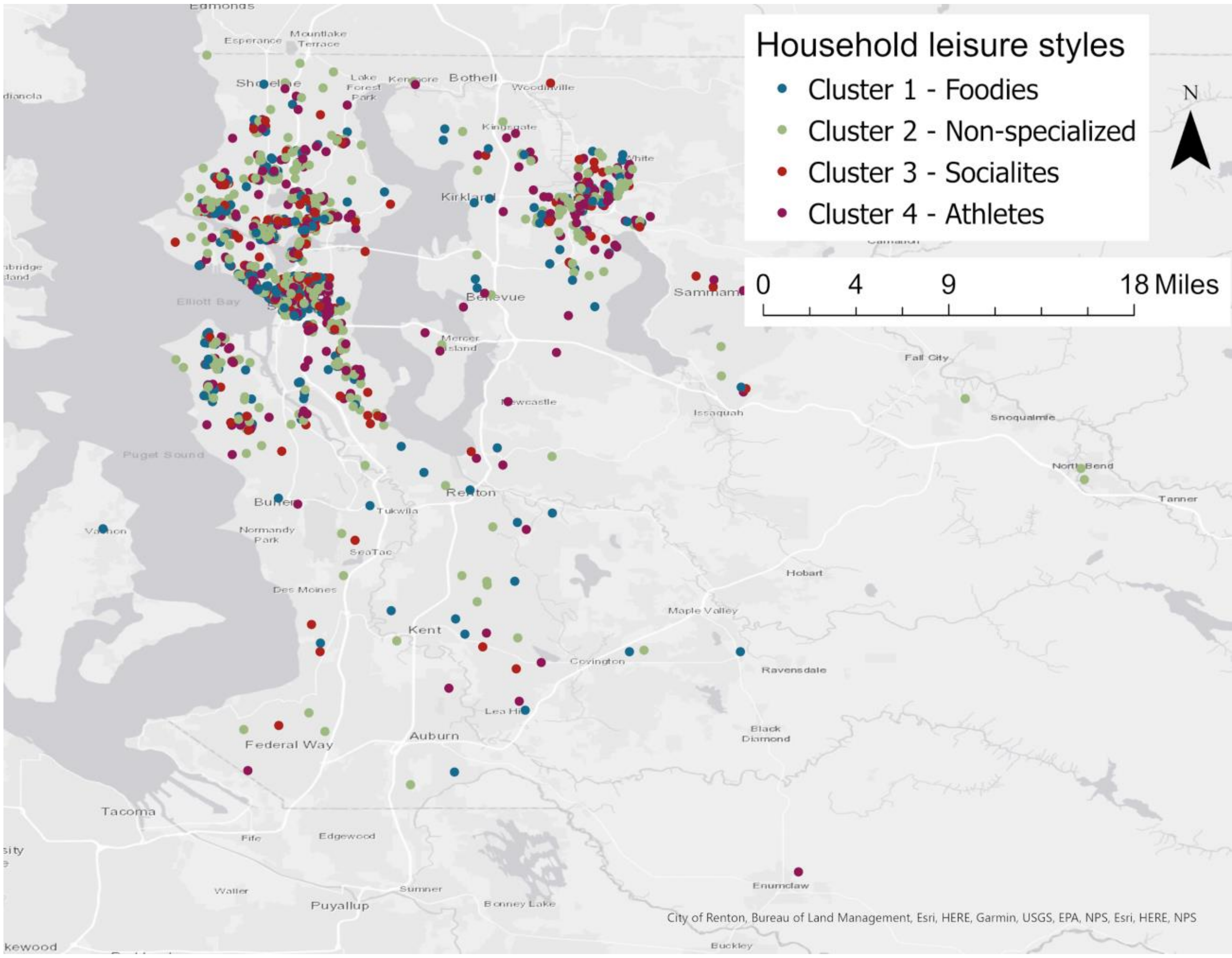


Figure 8 Spatial distribution of King County households by leisure style.

Leisure Transit Metrics

Using network analyst tools in ArcGIS Pro, I calculated three household metrics at three different times of the week to serve as inputs to the binary logistic model predicting car ownership. Figure 9 shows an example calculation of the three metrics for a household located in north Greenlake, Seattle.

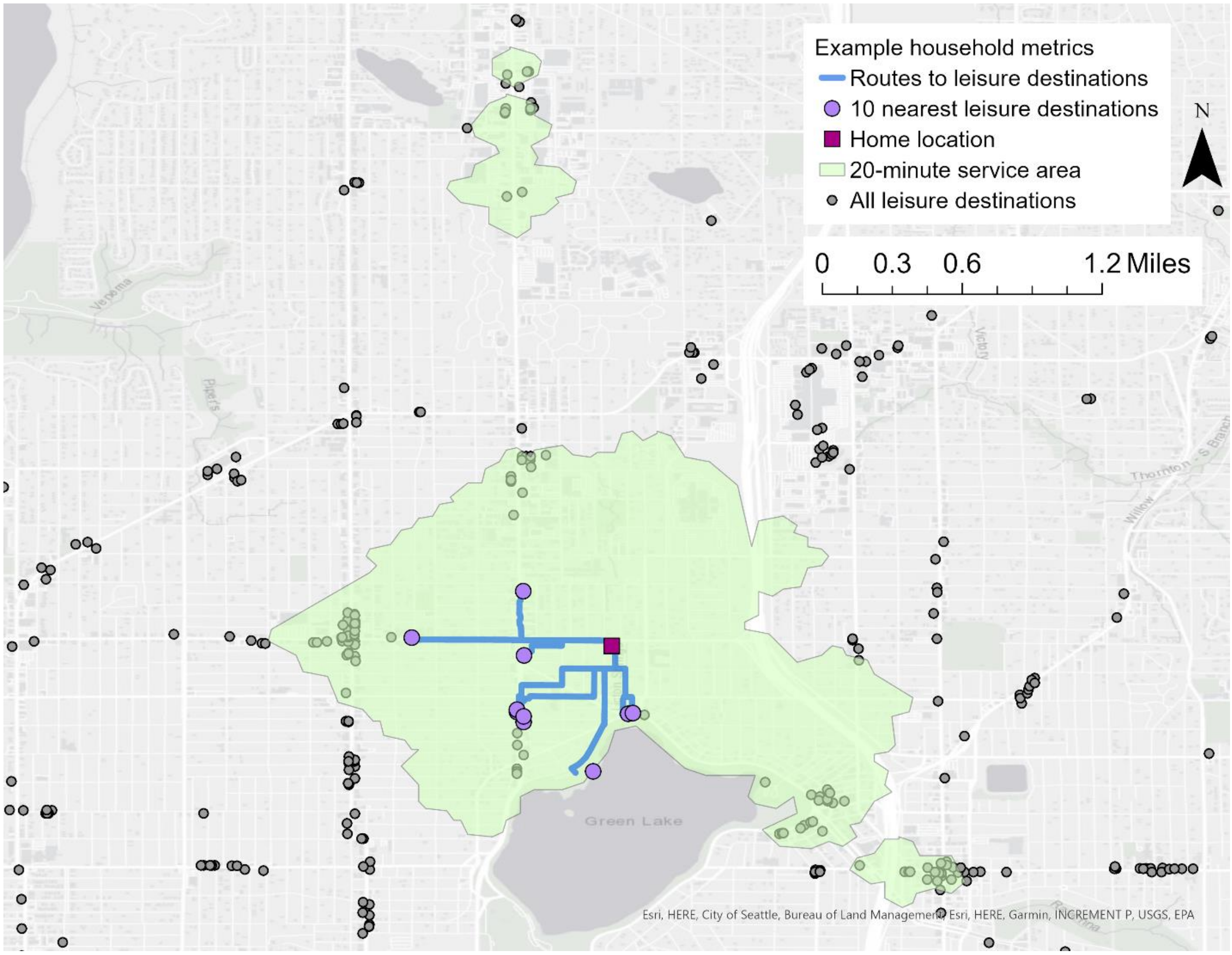


Figure 9 Example household-level calculation of three transit metrics.

Spatially, there are some interesting patterns to the distribution of these variables for households around King County. First, let us examine the average transit travel time to the 10 nearest leisure destinations, calculated assuming a 10 am departure on Sunday morning.

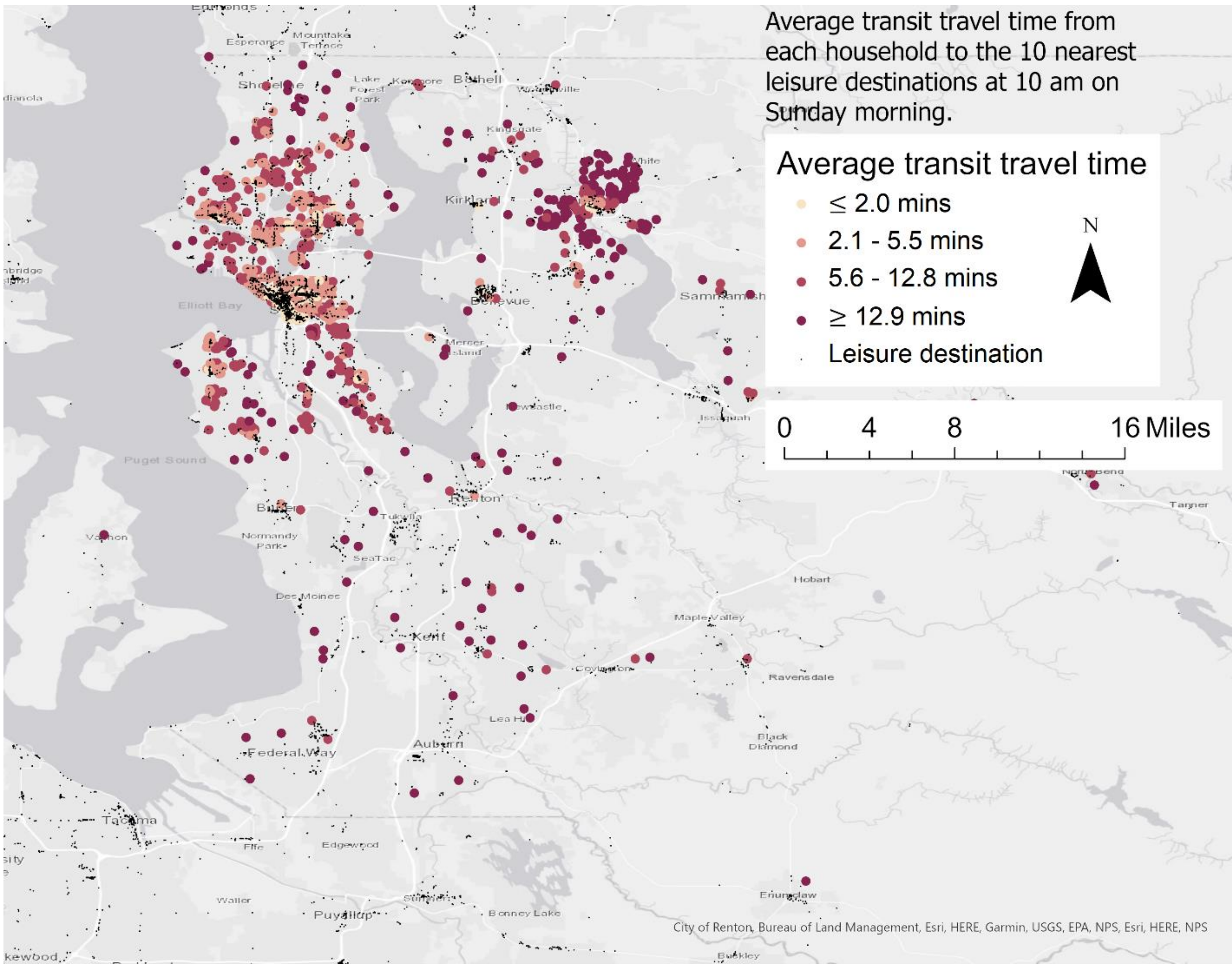


Figure 10 Household-level average transit travel time to the 10 nearest leisure destinations, departing at 10 am on Sunday morning.

Unsurprisingly, households near Seattle's and Redmond's downtown cores and urban neighborhoods like the University District, Greenlake, Fremont, and Ballard have shorter average travel times because of the high concentration of leisure destinations in those neighborhoods. It is interesting to note the dearth of leisure destinations in the areas immediately surrounding Redmond and in some suburban areas like West Seattle and Shoreline, resulting in higher travel times because residents simply must travel further. For residents of south King County, even living geographically near leisure may not be enough to enjoy shorter travel times because of less spatiotemporally comprehensive transit offerings.

In order to tease apart the impacts of these two potential contributions to longer travel times, it is helpful to examine the associated 20-minute transit service area map as well, shown in Figure 11. From this map, it appears that most of the households outside Redmond and in south King County with long average travel times also experience smaller transit access sheds. In general, this trend seems to hold based on visual inspection of the two maps. To highlight a counterexample, however, I would point out the Eastlake neighborhood which has among the smallest transit access sheds but reasonably low travel times to leisure. While a transportation planner may see the accessibility as a failure, a land use planner could identify that leisure needs may be met locally instead.

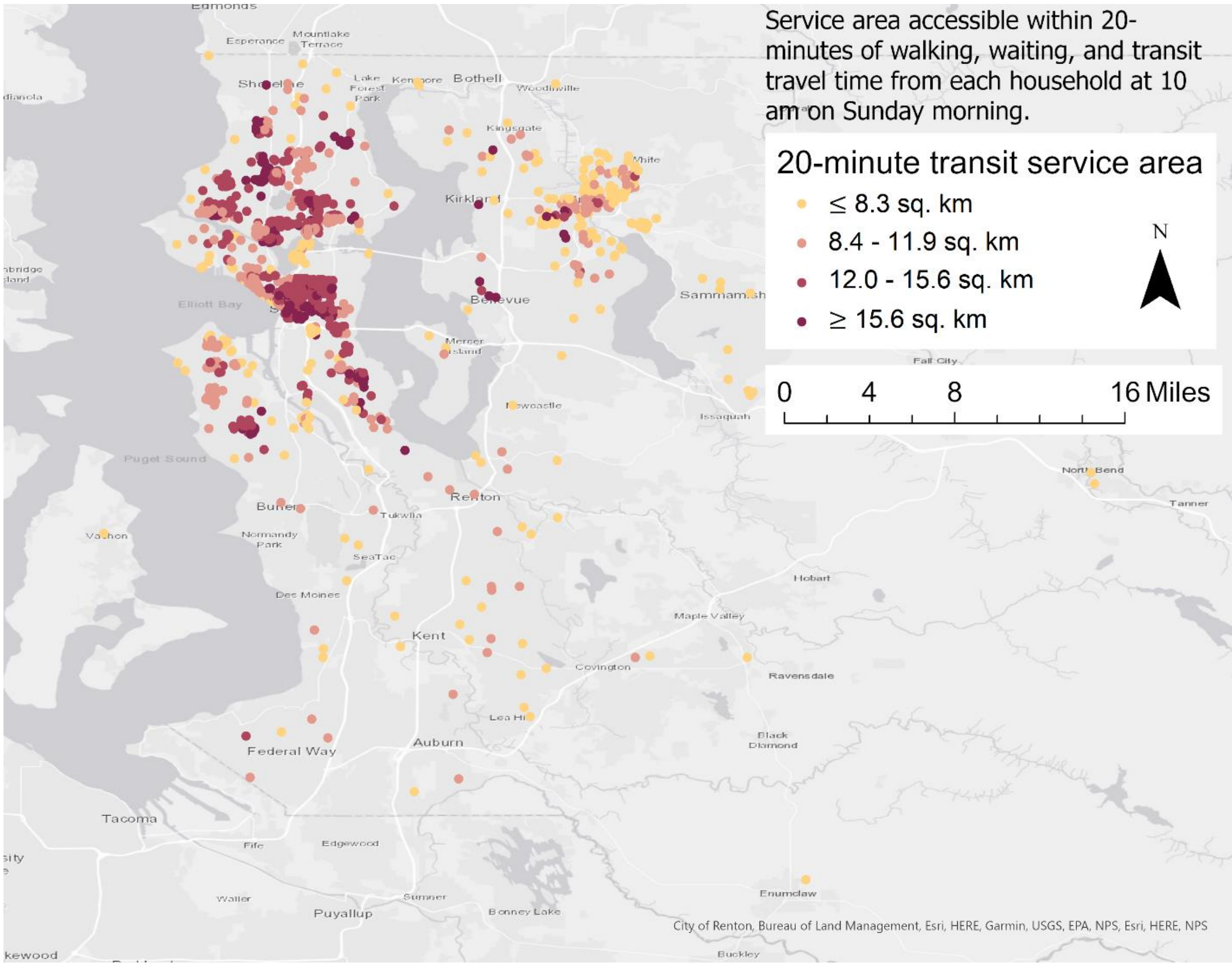


Figure 11 Household-level 20-minute transit access sheds, departing at 10 am on Sunday morning.

The final metric, inspired by Benenson, Martens, and Rofé (2010), computes the ratio of transit travel time to auto travel time to the 10 nearest leisure destinations. In this case, spatial patterns are much less pronounced, except for a strong concentration of low transit-to-auto ratio values in Redmond. This may be due to a more pronounced suburban-urban structure than Seattle where leisure destinations are concentrated in a single central business district and transit lines designed for commuting also function to transport households the co-located leisure sites.

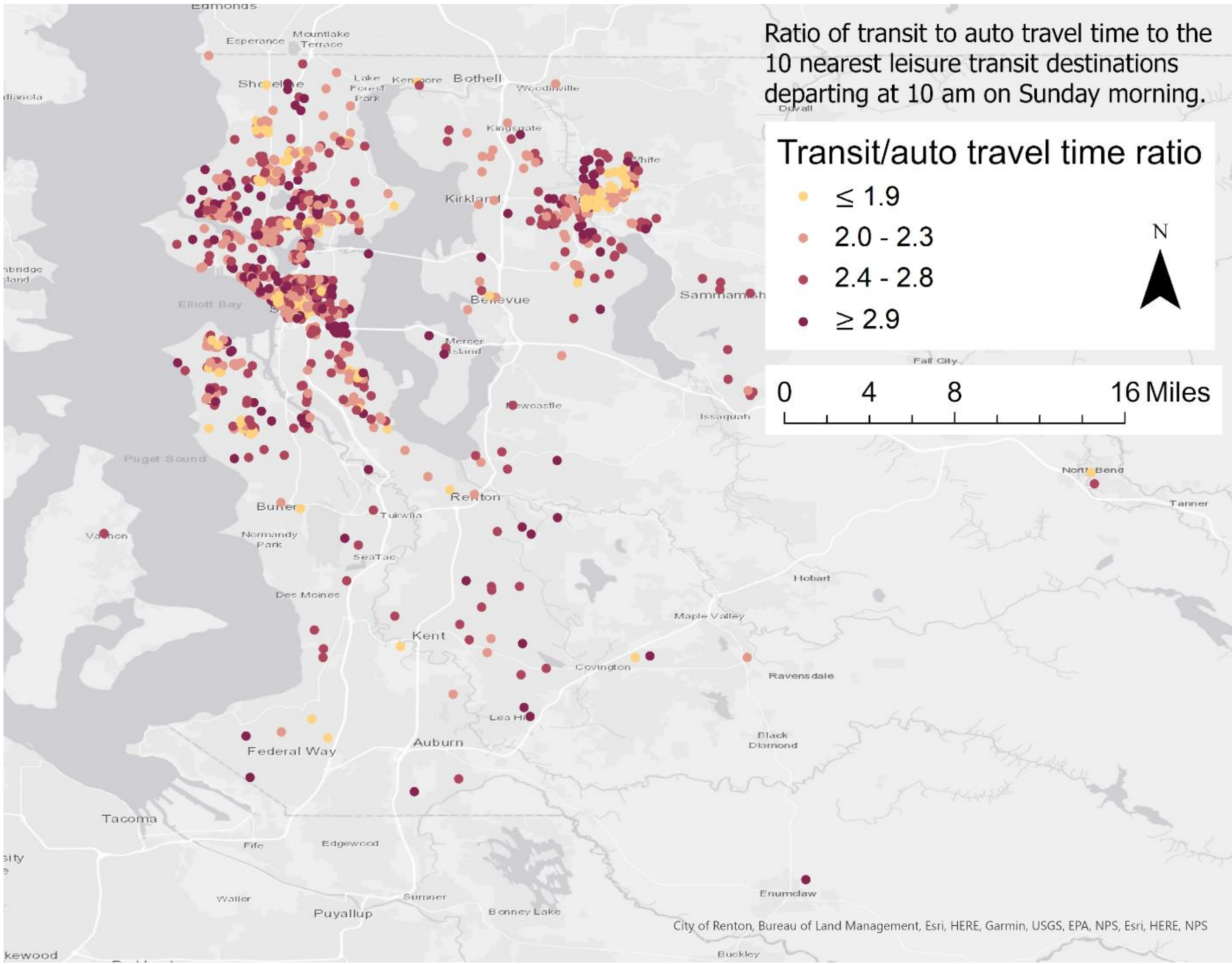


Figure 12 Household-level transit-to-auto travel time ratio for accessing the 10 nearest leisure destinations at 10 am on Sunday morning.

Binary Logistic Regression

Using the leisure style clustering and transit metrics as building blocks, supplemented with household income and size, known significant predictors of household vehicle ownership (Clark, 2009; Nolan, 2010), I developed a binary logistic regression model. As previously mentioned, logistic regression requires low correlation between independent variables which I verified in two ways. First, I computed a correlation coefficient matrix with each of the nine independent variables (three times for each of the three metrics), then further clarified by examining variance inflation factor (VIF) scores. The complete correlation matrix is shown in Table 3.

Table 3 Correlation matrix of the nine transit metric independent variables.

		10 am Sunday			9 pm Friday			10 am Saturday		
		Travel time	Transit/ auto ratio	Service area	Travel time	Transit/ auto ratio	Service area	Travel time	Transit/ auto ratio	Service area
10 am Sunday	Travel time	--	0.30	-0.56	1.00	0.31	-0.51	1.00	0.31	-0.51
	Transit/ auto ratio	0.30	--	-0.15	0.29	0.97	-0.13	0.29	0.96	-0.14
	Service area	-0.56	-0.15	--	-0.56	-0.15	0.86	-0.56	-0.15	0.88
9 pm Friday	Travel time	1.00	0.29	-0.56	--	0.31	-0.52	1.00	0.31	-0.51
	Transit/ auto ratio	0.31	0.97	-0.15	0.31	--	-0.16	0.31	0.96	-0.15
	Service area	-0.51	-0.13	0.86	-0.52	-0.16	--	-0.51	-0.14	0.86
10 am Saturday	Travel time	1.00	0.29	-0.56	1.00	0.31	-0.51	--	0.32	-0.51
	Transit/ auto ratio	0.31	0.96	-0.15	0.31	0.96	-0.14	0.31	--	-0.18
	Service area	-0.51	-0.14	0.88	-0.51	-0.15	0.86	-0.51	-0.18	--

Right off the bat, this indicates that it is infeasible to test the same metric from multiple days in the same model run because the correlation between respective variables across days is at minimum 0.86 and at maximum perfectly correlated after rounding. Isolating metrics from a single day like Sunday reveals weak to moderate correlations between the three metrics, ranging in magnitude from 0.15 to 0.56.

In order to further clarify which variables are too highly correlated to include in the model simultaneously, I conducted a VIF test on the complete model variable set available for each day, then removed variables with VIF greater than 5 one at a time, starting with the highest VIF score and retesting with each new model. The preliminary and final filtered variable results are summarized in Table 4.

Table 4 VIF scores for binary logistic models first including all independent variables with transit metrics split by day, then the final filtered VIF scores after removing one-at-a-time the highest score above 5.

	9 pm Friday		**10 am Saturday**		**10 am Sunday**	
	all	filtered	all	filtered	all	filtered
Travel time	3.76	1.93	3.72	1.91	3.91	1.93
Service area	9.64	3.10	8.75	3.02	9.88	3.14
Transit/auto ratio	17.02	--	16.08	--	17.14	--
Household size	5.68	--	5.64	--	5.66	--
High income	2.40	2.11	2.39	2.12	2.39	2.15
Cluster 1	1.65	1.61	1.64	1.60	1.64	1.61
Cluster 3	1.42	1.38	1.41	1.37	1.42	1.39
Cluster 4	1.65	1.60	1.65	1.60	1.65	1.61

With the final variable selection complete, I ran three total models, one for transit metrics calculated on each day of the week. Because the results are nearly identical for each day, I will show and discuss only 10 am on Sunday here but the complete regression results are available in Appendix I.

Table 5 Binary logistic regression model results using independent transit metrics calculated for Sunday at 10 am.

```
==============================================================================
Dep. Variable:               veh_own   No. Observations:                 1535
Model:                         Logit   Df Residuals:                     1529
Method:                          MLE   Df Model:                            5
Date:               Tue, 18 May 2021   Pseudo R-squ.:                  0.1821
Time:                       09:44:57   Log-Likelihood:                -602.50
converged:                      True   LL-Null:                       -736.62
Covariance Type:           nonrobust   LLR p-value:                 6.714e-56
=================================================================================
                   coef    std err          z      P>|z|      [0.025      0.975]
---------------------------------------------------------------------------------
tt_10a_sun       0.2201      0.022      9.911      0.000       0.177       0.264
sa_10a_sun_km   -0.0208      0.010     -2.000      0.046      -0.041      -0.000
high_inc         1.3501      0.146      9.241      0.000       1.064       1.636
cluster1        -0.2330      0.176     -1.326      0.185      -0.577       0.111
cluster3        -0.1328      0.199     -0.668      0.504      -0.523       0.257
cluster4         0.5269      0.211      2.497      0.013       0.113       0.941
=================================================================================
```

Average travel time, household income, and the cluster four (athletes) dummy variable are all significant at the $p<0.05$ level. Household income is not surprising, as this has been previously reported in the literature (Clark, 2009; Nolan, 2010), but it can serve as a rough litmus test for the model in that it is both significant and with a relatively large positive coefficient. The empirical result suggests that households with total income over $75,000 per year are much more likely to own a vehicle than households making less than that amount.

Examining the cluster variable results, keeping in mind that cluster two (do-it-alls) serves as the reference category, only membership in cluster four (athletes) is significant with a positive sign. This means that, all else being equal, households with the identified “athletic” leisure style are more likely to own a vehicle than households in the reference cluster. Membership in cluster one (foodies) or cluster three (socialites) are not significant in predicting household car ownership when compared to the reference cluster.

It is worth noting that other households are not considered specific leisure destinations in this methodology, so for socialites that may place more importance on visiting other peoples’ homes, the travel time to nearby leisure centers may not accurately represent their true desired destinations.

Finally, the average transit travel time to the 10 nearest leisure destinations and the 20-minute transit service area are both significant predictors of household car ownership, all else being equal. In interpreting the positive sign on the coefficient on travel time, I can conclude that as average transit travel time to leisure increases, so does the likelihood of a household owning a vehicle. Conversely, the negative sign on the service area variable indicates that as the area accessible via transit increases, a household is less likely to own a car. This provides an answer to my central research question, “how does transit planning for leisure impact household car ownership in King County?”

The coefficients on each variable indicate the magnitude of the effect on the outcome caused by a one-unit change in the input variable. The effect of increasing average travel time by one minute is 0.22 whereas the effect of increasing the total service area by one square kilometer is -0.02, meaning that in order to negate the effects of a one-minute increase in travel time on the likelihood of owning a car, the service area would have to increase by 11 square kilometers. This suggests that household vehicle ownership is more sensitive to accessibility to specific types of destinations than the area covered by the general transit access shed. Transit planners may use this result to consider future route planning focusing on bringing people to leisure whereas land use planners can see the importance of providing a variety of leisure destinations within close proximity of residential areas. Reducing travel time for everyone is quite difficult to achieve through transit service alone. Building leisure destinations closer to homes, or vice versa, is one strategy for reducing travel time without requiring improvements to the existing transit network. However, improvements to the transit system may pack a double punch by both reducing travel times and simultaneously increasing the service area accessible.

Compared to these leisure transit metrics, however, the effect that being a high-income household has on car ownership is remarkable. For a household making less than $75,000 per year, their average travel time could be six minutes longer than a comparable high-income household while maintaining the same likelihood of owning a vehicle.

The final significant variable, membership in cluster four (athletes), has a relatively strong effect on the outcome as well at about half that of high-income status. Anecdotally, from my conversations with friends and peers, this makes sense to me. For those with athletic lifestyles in King County, many of the opportunities for these types of activities like skiing, hiking, or mountain biking take place outside of cities which makes transit options highly inconvenient or unavailable altogether and thus would explain the increased likelihood of vehicle ownership.

In order to ensure these model results do not violate any assumptions of logistic regression, I investigated the residuals plot to check for unusual patterns that could indicate poor model suitability. Because of the binary nature of the response variable, a simple residuals plot is somewhat meaningless, so I instead calculated a binned residuals plot, shown in Figure 13. While the residuals are randomly distributed about the y-axis, there is much greater variance for observations with lower predicted probabilities. Known as heteroskedasticity, this is devastating in linear regression, but does not violate any assumptions for logistic regression. Thus, I conclude that the model results are acceptable with no major red flags in the residuals plot.

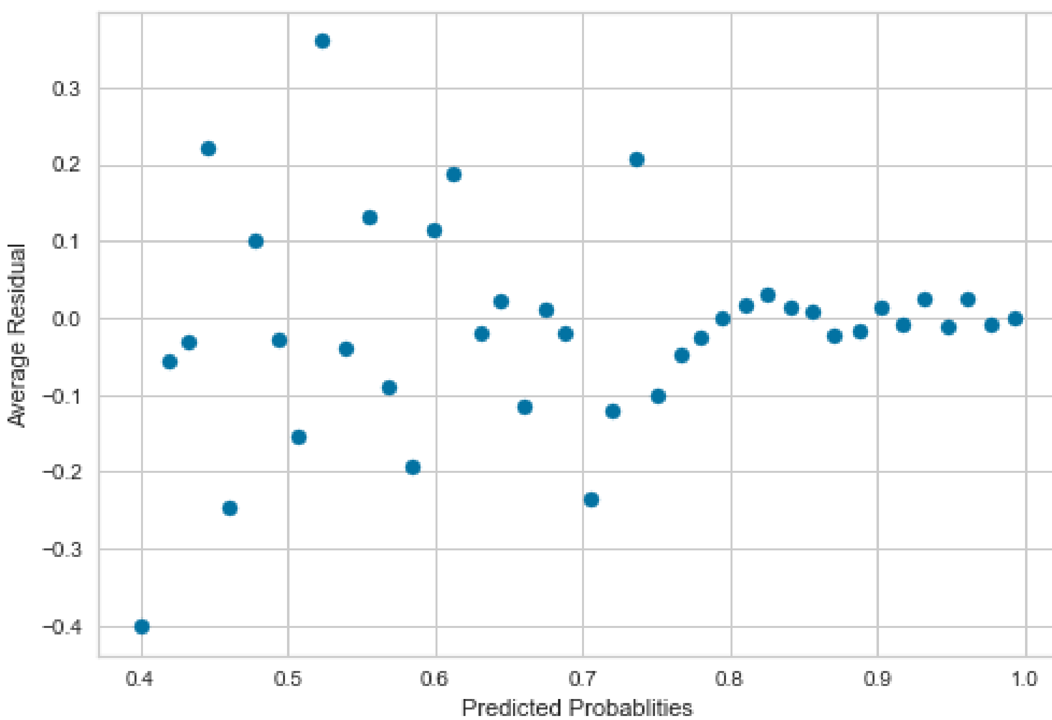


Figure 13 Binned residuals plot of binary logistic regression model results.

There are a variety of urban styles represented within King County that may differ from other urban areas, both American and international. Seattle breaks from the traditional urban-suburban duality with the implementation of an urban village strategy that creates several smaller business districts spread throughout the city, interspersed among residential neighborhoods. By contrast, a city like Redmond follows the much more traditional city pattern of a single central business district surrounded by low density, residential suburbs. Because this analysis considers travel behavior from households in both settings, there may be nuances in the findings that are lost by considering all types together. For example, households in Seattle where transit use is more commonplace may have different attitudes or willingness to use transit than someone from a suburban neighborhood outside of Redmond that, even though they may have nearby transit, does not consider transit as viable because of less daily exposure.

While I would expect leisure style clusters to be relatively consistent from city to city because there are people with diverse interests everywhere, the significance of each leisure style in predicting car ownership may differ. Again, considering the “athletes” cluster, I would suspect that in cities without such a plethora of outdoor recreation options within driving distance of the city, these households may not be at a significantly higher likelihood of owning a car because the athletic activities they attend are concentrated within urban limits. For example, thinking about my past home of Chicago, such few outdoor sports opportunities were available within a day-trip drive that many of those seeking access to nature just turned to local parks like the Lake Michigan waterfront.

Conducting this same analysis on a different city or even at a different geographic scope within King County may reveal how differences in urban structure and broader regional positioning among natural leisure attractions interact with logistic regression results in a way that was not fully captured in the scope of this particular inquiry.

VI. Conclusion

After a year of non-essential travel restrictions, psychological stress from stay-at-home orders (Marroquín et al., 2020), and limited opportunities for leisure outside the home, a return to dining out, visiting friends and family or playing frisbee in the park feels like a breath of fresh air, quite literally. Even in pre-pandemic times, leisure served as a key component to living a happy and fulfilling life (Becchetti et al., 2012). From a personal perspective, I sometimes struggle to access activities as an individual without a car and became curious about the connection between leisure transit planning and car ownership.

Intuitively and scientifically, leisure trips are more complicated both in space and time than commute trips in part because of the wide variety of destinations available to a person (Gkiotsalitis & Stathopoulos, 2016). Additionally, they tend to take place outside of work hours like evenings and weekends when transit services may operate on reduced schedules. These factors make transit use more difficult and thus more people opt for more convenient modes like personal cars, when available (Beirão & Cabral, 2010). From an environmental perspective, transit agencies should investigate ways to better serve leisure travel as it accounts for a larger share of daily travel than commute trips.

This research set out to investigate the relationship between leisure transit planning and household leisure styles and their impacts on household car ownership in King County. To do so, I developed three metrics measured at three different times of the week to quantify access to leisure at the household level. Additionally, I performed k-means clustering on households based on their leisure trip-making behavior and sorted observations into four distinct categories: foodies, socialites, athletes, and do-it-alls.

Considering these metrics and leisure styles as independent variables, I developed a binary logistic regression model to test for significant predictors of household vehicle ownership. After filtering out highly correlated variables, I found that, all else equal, as the average transit travel time to the 10 nearest leisure destinations increases, so does the likelihood of a household owning one or more vehicles. Total service area accessible by 20 minutes on transit is also significant but in the opposite direction and with a much smaller magnitude, indicating that consumers are more sensitive to specific destination accessibility than general spatial coverage.

In order to better accommodate sustainable, car-free lifestyles, transportation planners must consider how transit can better serve a broader range of trip types like leisure. This includes expanding and protecting evening and weekend services as well as designing routes with access to leisure opportunities in mind. Additionally, these conclusions support providing opportunities for diverse leisure activities closer to residential areas.

As the world returns to having fun outside the home, it is my hope that transportation planners will recognize the importance of access to leisure to be as environmentally crucial as we have learned it is emotionally important.

Appendix I - Complete binary logistic model results

Dep. Variable:	veh_own	No. Observations:	1535
Model:	Logit	Df Residuals:	1529
Method:	MLE	Df Model:	5
Date:	Tue, 18 May 2021	Pseudo R-squ.:	0.1801
Time:	10:05:06	Log-Likelihood:	-603.96
converged:	True	LL-Null:	-736.62
Covariance Type:	nonrobust	LLR p-value:	2.845e-55

	coef	std err	z	P>|z|	[0.025	0.975]
tt_9p_fri	0.2155	0.023	9.572	0.000	0.171	0.260
sa_9p_fri_km	-0.0169	0.011	-1.560	0.119	-0.038	0.004
high_inc	1.3276	0.145	9.170	0.000	1.044	1.611
cluster1	-0.2658	0.176	-1.509	0.131	-0.611	0.079
cluster3	-0.1688	0.198	-0.853	0.394	-0.557	0.219
cluster4	0.4948	0.210	2.359	0.018	0.084	0.906

Binary logistic regression results using metrics calculated at 9 pm on Friday night.

Dep. Variable:	veh_own	No. Observations:	1535
Model:	Logit	Df Residuals:	1529
Method:	MLE	Df Model:	5
Date:	Tue, 18 May 2021	Pseudo R-squ.:	0.1815
Time:	10:04:10	Log-Likelihood:	-602.89
converged:	True	LL-Null:	-736.62
Covariance Type:	nonrobust	LLR p-value:	9.875e-56

	coef	std err	z	P>|z|	[0.025	0.975]
tt_10a_sat	0.2204	0.023	9.776	0.000	0.176	0.265
sa_10a_sat_km	-0.0212	0.010	-2.086	0.037	-0.041	-0.001
high_inc	1.3506	0.145	9.295	0.000	1.066	1.635
cluster1	-0.2271	0.174	-1.302	0.193	-0.569	0.115
cluster3	-0.1297	0.196	-0.660	0.509	-0.515	0.255
cluster4	0.5359	0.210	2.552	0.011	0.124	0.947

Binary logistic regression results using metrics calculated at 10 am on Saturday morning.

References

Becchetti, L., Giachin Ricca, E., & Pelloni, A. (2012). The relationship between social leisure and life satisfaction: Causality and policy implications. *Social Indicators Research, 108*(3), 453–490. https://doi.org/10.1007/s11205-011-9887-5

Beirão, G., & Cabral, J. (2010). Travelling to work or just for fun? Exploring differences in perceptions and attitudes by travel modes. Proceedings of the 12th World Conference on Transport Research society. (pp. 470-493).

Benenson, I., Martens, K., & Rofé, Y. (2010). Measuring the Gap between Car and Transit Accessibility: Estimating access using a High-Resolution Transit Network Geographic Information System. Transportation Research Record, 2144(1), 28–35. https://doi.org/10.3141/2144-04

Coronavirus Disease 2019 (COVID-19). (2020) Centers for Disease Control and Prevention.

Clark, B., Lyons, G., & Chatterjee, K. (2016). Understanding the process that gives rise to household car ownership level changes. Journal of Transport Geography, 55, 110–120. https://doi.org/10.1016/j.jtrangeo.2016.07.009

Clark, S. (2009). The determinants of car ownership in England and Wales from anonymous 2001 census data. *Transportation Research Part C: Emerging Technologies*, *17*(5), 526–540. https://doi.org/https://doi.org/10.1016/j.trc.2009.04.004

Currie, G., & Loader, C. (2009). High ridership growth from extended transit service hours: An exploration of the causes. Transportation Research Record: Journal of the Transportation Research Board, 2110(1), 120–127. https://doi.org/10.3141/2110-15

Czepkiewicz, M., Heinonen, J., & Ottelin, J. (2018). Why do urbanites travel more than do others? A review of associations between urban form and long-distance leisure travel. Environmental Research Letters, 13(7), 073001. https://doi.org/10.1088/1748-9326/aac9d2

Fan, Y., & Khattak, A. (2012). Time use patterns, lifestyles, and sustainability of nonwork travel behavior. International Journal of Sustainable Transportation, 6(1), 26–47. https://doi.org/10.1080/15568318.2011.553770

Fayyaz S, S. Kiavash, Liu, Xiaoyue Cathy, & Zhang, Guohui. (2017). An efficient General Transit Feed Specification (GTFS) enabled algorithm for dynamic transit accessibility analysis. PloS One, 12(10), E0185333.

Four-Step Travel Model. (n.d.). Metropolitan Washington Council of Governments. Retrieved May 14, 2021, from https://www.mwcog.org/transportation/data-and-tools/modeling/four-step-model/step-1/

Gim, T.-H. T. (2018). SEM application to the household travel survey on weekends versus weekdays: The case of Seoul, South Korea. European Transport Research Review, 10(1), 11. https://doi.org/10.1007/s12544-018-0285-9

Gkiotsalitis, K., & Stathopoulos, A. (2016). Demand-responsive public transportation re-scheduling for adjusting to the joint leisure activity demand. International Journal of Transportation Science and Technology, 5(2), 68–82. https://doi.org/10.1016/j.ijtst.2016.09.004

Holden, E., & Linnerud, K. (2011). Troublesome Leisure Travel: The Contradictions of Three Sustainable Transport Policies. Urban Studies, 48(14), 3087–3106. https://doi.org/10.1177/0042098010396234

Kim, Hong Sok, & Kim, Eungcheol. (2004). Effects of public transit on automobile ownership and use in households of the USA. Review of Urban and Regional Development Studies, 16(3), 245-262.

Kodinariya, Trupti & Makwana, Prashant. (2013). Review on Determining of Cluster in K-means Clustering. International Journal of Advance Research in Computer Science and Management Studies. 1. 90-95.

Lanzendorf, M. (2002). Mobility styles and travel behavior: Application of a lifestyle approach to leisure travel. Transportation Research Record: Journal of the Transportation Research Board, 1807(1), 163–173. https://doi.org/10.3141/1807-20

Liu, Y., Cirillo, C. Measuring transit service impacts on vehicle ownership and use. Public Transp 7, 203–222 (2015). https://doi.org/10.1007/s12469-014-0098-8

Lu, A., & Reddy, A. (2012). Strategic Look at Friday Exceptions in Weekday Schedules for Urban Transit: Improving Service, Capturing Leisure Markets, and Achieving Cost Savings by Mining Data on Automated Fare Collection Ridership.

Transportation Research Record, 2274(1), 30–51. https://doi.org/10.3141/2274-04

Marroquín, B., Vine, V., & Morgan, R. (2020). Mental health during the COVID-19 pandemic: Effects of stay-at-home policies, social distancing behavior, and social resources. *Psychiatry Research*, *293*, 113419. https://doi.org/10.1016/j.psychres.2020.113419

Næss, P. (2006). Are short daily trips compensated by higher leisure mobility? Environment and Planning B: Planning and Design, 33(2), 197–220. https://doi.org/10.1068/b31151

National household travel survey daily travel quick facts. (2017). Bureau of Transportation Statistics. https://www.bts.gov/statistical-products/surveys/national-household-travel-survey-daily-travel-quick-facts

Network Analyst ArcGIS Pro Tutorial Data. (2020). ArcGIS.

Nolan, A. (2010). A dynamic analysis of household car ownership. Transportation Research Part A: Policy and Practice, 44(6), 446–455. https://doi.org/10.1016/j.tra.2010.03.018

Schwanen, T., Dijst, M., & Dieleman, F. M. (2001). Leisure trips of senior citizens: Determinants of modal choice. Tijdschrift Voor Economische En Sociale Geografie, 92(3), 347–360. https://doi.org/10.1111/1467-9663.00161

Signalized Intersections: Informational Guide (FHWA-HRT-04-091; pp. 145–158). (2004). Federal Highway Administration. https://www.fhwa.dot.gov/publications/research/safety/04091/07.cfm#chp712

Simmons, J., & Haas, P. J. (2016). Impact on bus ridership from changes in a route's span of service. Transportation Research Record: Journal of the Transportation Research Board, 2539(1), 48–54. https://doi.org/10.3141/2539-06

Trends in discretionary travel (2017 National Household Travel Survey). (2019). Federal Highway Administration.

Van Acker, V., Boussauw, K., & Witlox, F. (2011). Understanding modal choices for leisure activities. Is it just objectively determined? Belgeo, 1–2, 105–120. https://doi.org/10.4000/belgeo.6395