Visitor Mobility in the Park Loop Region

An Interactive Qualifying Project
submitted to the Faculty of
WORCESTER POLYTECHNIC INSTITUTE
in partial fulfilment of the requirements for the
degree of Bachelor of Science

By
Alex Jozitis
Richard Kern
Timothy Lewis
Te Lu
Donovan Shaw

Date:
7/28/2021

Project Center:
Acadia National Park

Report Submitted to:

Abe Miller-Rushing and John Kelly
National Parks Service

Professors Bianchi and Kurlanska
Worcester Polytechnic Institute

Table of Contents

List of Figures

List of Tables

Authorship Page

Section	Primary Author	Editors
Abstract	Donovan Shaw	Richard Kern
Executive Summary	Te Lu, Donovan Shaw	Te Lu, Donovan Shaw
Introduction	Tim Lewis, Alex Jozitis, Richard Kern	Timothy Lewis, Alex Jozitis
Big Data	Alex Jozitis	Richard Kern
Negative Impacts of Congestion	Te Lu, Donovan Shaw	Alex Jozitis, Te Lu
Congestion in the Parks	Richard Kern, Te Lu	Timothy Lewis, Donovan Shaw
Strategies for Addressing Congestion	Alex Jozitis, Donovan Shaw	Richard Kern, Donovan Shaw
Acadia National Park	Timothy Lewis	Donovan Shaw
Congestion in Acadia	Richard Kern	Alex Jozitis, Tim Lewis
Strategies for Addressing Congestion	Richard Kern, Timothy Lewis	Alex Jozitis
Introduction	Timothy Lewis	Te Lu
StreetLight Data	Richard Kern	Te Lu
Direct Observation	Timothy Lewis	Richard Kern, Donovan Shaw
Parking Duration Analysis	Donovan Shaw	Alex Jozitis, Te Lu
Analysis for Duration from StreetLight	Richard Kern	Alex Jozitis, Te lu
Duration from Direct Observation	Richard Kern	Alex Jozitis, Te Lu
Results and Findings	Te Lu	Donovan Shaw
Visitor Travel Patterns	Richard Kern	Donovan Shaw
Average Parking Duration	Richard Kern	Donovan Shaw
Results of Cadillac Mountain	Te Lu	Richard Kern, Donovan Shaw
Visitor Demographics	Donovan Shaw	Alex Jozitis, Richard Kern
Discussion	Te Lu, Donovan Shaw	Timothy Lewis
Recommendations and Conclusions	Timothy Lewis	Te Lu

Abstract

Traffic congestion in National Parks has been an issue the National Park System has been trying to solve for a long time. The goal of this project was to discover how big data can be used to help inform Acadia National Park's future actions. To do this, we used StreetLight, a traffic data aggregation and processing service. We analyzed its effectiveness at finding the travel patterns, parking durations, and demographics of park visitors. These results are then compared to other data collection methods. Additionally, we provided recommendations on how Acadia National Park could utilize big data to improve their current services and inform their future actions.

Visitor Mobility in the Park Loop Region

Worcester Polytechnic Institute
Alex Jozitis, Richard Kern, Timothy Lewis, Te Lu, Donovan Shaw

Executive Summary

Introduction

The United States National Parks were established to preserve the natural beauty for future generations to experience. Many National Parks have roadways, allowing visitors to experience the park in their vehicles. As more vehicles are entering the parks, traffic congestion becomes a more significant issue, which negatively affects both the natural landscape and visitor experience. The National Park Service must ensure that visitors can move around the parks freely while minimizing the impact of visitors on the natural environment.

Acadia National Park, located in Maine, is one of the smallest National Parks in the United States. In recent years, Acadia has become the eighth most visited National Park in the country, serving over 2.7 million visitors (U.S. Department of the Interior, 2020). The high popularity and small size of Acadia has made it especially prone to vehicle congestion.

Nowadays, people carry their cell phone with them wherever they go, and they are constantly sending out location data. Although a few data points alone may be insignificant, larger trends and patterns can be identified when billions of these data points are aggregated. These patterns and trends can be used to determine when and where people arrive at certain locations. More specifically, this information can be used to address vehicle congestion.

Background

Traffic congestion is caused by groups of vehicles occupying a significant amount of space on a given portion of a roadway. Traffic congestion leads to reduced speeds and delays. One factor that can lead to traffic congestion is parking availability. When parking availability is low, people will spend more time “cruising for parking” along a roadway, decreasing the speed of traffic and causing pauses in traffic flow (Zhu, 2020). Congestion can lead to environmental and economic impacts as well. The longer cars are delayed by traffic congestion, the more fuel they use. This means their carbon emissions are higher, and the more money the average driver will have to use on gas. Some studies have shown that even marginally increasing the average speed of cars in traffic congestion can have radical effects, with significantly fewer carbon emissions and an estimated total of 0.1% GDP reduction in the operation costs (Barth, 2008).

Goals/Objectives

The primary goal of this project was to provide an analysis of visitor mobility patterns that would inform the park management about traffic in the Park Loop Road region. These mobility patterns were analyzed using two main methods. First, the team utilized StreetLight Data to gather relevant information about traffic patterns in the region. Second, the team performed direct observation to gather ground truth data in the region that would corroborate data from StreetLight. The efficacy of these data collection methods was also evaluated.

Methods

The first method that the team used was StreetLight Data, a database composed of cell phone and vehicle geolocation data. Through a proprietary algorithm, the platform takes these location data and generates an interactive traffic database. This database can be accessed through an internet browser, and users can investigate how visitors traveled through any region between 2018 and 2020. This is done by drawing areas of interest called zones, which the team placed at entrances to the Park Loop Road, parking lots, key attractions, and popular roadside parking locations. The team was then able to see how many people travelled to, from, and through each zone. The resulting data from StreetLight was also separated into weekdays and weekends, and these data can be further divided by the time of day. We performed these analyses on the peak seasons of 2018, 2019, and 2020.

StreetLight also provides the time and speed distribution for people to travel through these zones. The team was also able to find how many people traveled from one zone to another through a method called origin-destination analysis. Specifically, the team gathered a distribution of visitors going to each zone in the park from the park entrances. StreetLight Data allowed the team to gather information on vehicular traffic patterns throughout the Park Loop Road. By determining where traffic congestion has occurred in the past through StreetLight, we found where we should focus our on-site analyses. This focus was primarily placed on tracking vehicles entering and leaving parking lots. With StreetLight, we analyzed the distribution of vehicle trips entering and leaving individual parking lots. StreetLight can also determine the impacts of park management actions. For example, we investigated the impact of the Cadillac Mountain reservation system that was put in place on May 26th this year. This was done by comparing visitor arrival time distribution across three years during the two-week trial period of the system in October 2020.

The vehicle mobility analysis can also be extended into pedestrians or people-on-foot in the park. For example, by setting up zones at every mountain summit in the park, we gathered a summit usage breakdown. Along with this, by setting up zones on every trail on the mountain, we can investigate the trail popularity on a specific mountain. Moreover, looking at a specific trail, we can get the time distribution of when hikers arrive at those trails.

The other data collection method we used was direct observation, with team members stationed at designated locations to collect traffic data. Using StreetLight, we were able to determine these locations, including the Jordan Pond House parking lots, the Nature Center parking lot, and the Cadillac Summit Road intersection. We recorded information on passing cars, cars entering and cars exiting. Our goal from observing at the Jordan Pond parking lot and the Nature Center parking lot was to calculate the average parking duration. To achieve this, team members would note down how many cars were in the parking lot to start, and record whenever a car enters or leaves a parking spot. The team observed at the Jordan Pond for seven days and at the Nature Center for five days. The team observed approximately 6800 cars in total.

To calculate the parking duration at the two parking lots, we derived an algorithm that could calculate the average duration of stay from both StreetLight data and direct observation data. StreetLight Data can look at all cars that interacted with a specified zone, called a Zone Analysis. These analyses present an hour-by-hour breakdown of when cars entered or left the zone, presented as percentages. We used the product of each percentage and the hour they occurred to make them weighted. We then found the average of those weighted percentages to find the average times cars came into and out of the zone. The difference between the average entrance time and the average exit time expresses the general average time cars stay parked in the zone.

The team performed similar calculations to find the parking duration using data from direct observation. By finding the average difference between the entrance and exit times, we got a rough estimate for the average time a car stays parked. We also had to account for the cars that were already there at the beginning of our testing interval and those that were still there at the end. To accomplish this, we first noted down the number of cars present in a parking lot when we started data collection. When performing the calculation, we subtracted the number of cars noted down previously from the start of the list of exit times. Similarly, we removed the number of cars present at the end of our data collection period from the end of the list of times that people parked. This gave us an accurate count of how long cars stayed in parking spaces on average.

Findings & Discussion

Our research found that StreetLight was accurate when tracking the movements of vehicles along roadways in the park. We tracked popular paths vehicles took on the Park Loop Road, starting with their entrance through one of the marked zones, to their destination from said entrances, to their destination from there. When we could, we used existing data to corroborate our findings. Previous WPI research using car counters showed 58% of cars used the Paradise Hill entrance, which was within 6% of our finding of 52.5% in 2019.

We also used StreetLight to estimate the average times vehicles stayed parked in multiple popular areas along the park loop. We also collected on the ground data in these locations to corroborate StreetLight. There were cases where the data did not line up, which could have been from a number of factors, such as limited sample size and lack of randomization.

The team also analyzed the effects of the Cadillac reservation system by comparing statistics from the reservation period to statistics from the same period in other years. The main model we used was time distributions for when vehicles arrived at the Blue Hill and Cadillac Summit parking lots. These distributions stayed constant across multiple years, suggesting that the reservation system did not impact this aspect of the visitor mobility behavior.

StreetLight also uses census data to provide various demographics of visitors entering and exiting zones. StreetLight's analysis for 2020 showed that 87% of visitors were white, and the next closest was a 5% Asian demographic. It also showed 30% of all visitors came from Maine, 7% came from Massachusetts, and less than 3% came from New Hampshire. Finally, StreetLight determined the family status of visitors, showing 31% of visitors had a child.

Recommendations

The team recommended that the park add big data strategies to their methods for collecting and analyzing data on the Park Loop Road. This project primarily focused on cars, so we believe the park could expand by analyzing pedestrian and bicycle traffic. The information gained from big data can be used to provide visitors with up-to-date information on traffic in the park, which we believe will help to mitigate congestion. Currently on the National Park Service website, there is a section showing the busy times for popular tourist attractions. This could be kept up-to-date using time distribution information from big data. Additionally, we recommended that the park use big-data to evaluate the effectiveness of future management actions, such as implementing further reservation systems.

Citations

Barth, M., & Boriboonsomsin, K. (2008). Real-World Carbon Dioxide Impacts of Traffic Congestion. Transportation Research Record, 2058(1), 163–171. https://doi.org/10.3141/2058-20

U.S. Department of the Interior. (2020). Annual Park Ranking Report. https://irma.nps.gov/STATS/SSRSReports/National%20Reports/Annual%20Park%20Ranking%20Report%20(1979%20-%20Last%20Calendar%20Year)

Zhu, Y., Ye, X., Chen, J., Yan, X., & Wang, T. (2020). Impact of Cruising for Parking on Travel Time of Traffic Flow. Sustainability, 12(8), 3079. https://doi.org/10.3390/su1208307

Introduction

As you move around, your phone is constantly sending and receiving data, whether you realize it or not. While driving to work, riding your bike, walking the dog, and even hiking a mountain your cell phone leaves a digital footprint of your movements and day to day activities. This information may seem insignificant, but when added to the vast pool of data received from billions of people around the world, you start to see a much larger picture. The information collected from these sources is referred to as big data and has the potential to provide incredible insights into how and when people move around the world. On a smaller scale, this data can be used to help address congestion and movement of people and cars in cities, neighborhoods, and even national parks.

The United States National Park Service was established to preserve the natural beauty of America's landscapes for future generations to experience. This purpose, however, presents a conflict in its execution. Visitors must be able to freely experience the park in order to enjoy the natural vistas, but the presence of people can cause damage or alterations to park features. In other words, the people coming to enjoy the park can end up being the biggest threat to it. As a result, when more visitors are traveling to the parks, more effort must go towards managing their impact.

In order to experience the National Parks, many visitors take their personal vehicles on roadways constructed throughout the parks. An increase in vehicles on the roads causes vehicle congestion, which harms both the natural landscape and the visitors' experience of the park. Therefore, one goal of the National Park Service becomes the management of vehicular traffic within the park. Depending on the size of the park issues can also arise when these visitors exit their vehicles, with overcrowding inhibiting the visitor experience.

Acadia National Park spans about 47,000 acres spread across a series of several islands in Maine (U.S. Department of the Interior, 2015). In 2020, Acadia was the eighth most visited United States national park and served 2.7 million visitors (U.S. Department of the Interior, 2020). This was despite the fact that Acadia is one of the smallest National Parks in the U.S. The park's popularity, coupled with its small size, makes Acadia especially prone to vehicle traffic congestion.

This report begins by looking at general issues associated with road congestion, traffic in the context of the national parks, and information on Acadia National Park. The methods chapter details our use of the big data software StreetLight and direct observation to help understand the movement of vehicles and people in the park. Finally, this is followed by our findings and recommendations, which focus on our discoveries in StreetLight and how we believe the park can make use of this program.

Background

Big Data, its Capabilities, and its Limitations

Big data is commonly defined in terms of three Vs: volume, velocity, and variety. These refer to the amount of data that is being gathered, the speed at which data is gathered, and the presence of multiple sources of data, respectively. By 2017, the volume of information collected in big data databases was in units of petabytes and exabytes (Storey & Song, 2017). This high volume of data can be gathered from a variety of sources. Most commonly, databases will use some combination of government-owned public data, business-owned private data, and user-reported data (such as consumer reviews or fitness trackers) (George et al., 2014). To analyze the trillions of data points that can exist in their databases, big data software uses tools to identify patterns in the collected data. In the case of StreetLight, a machine learning algorithm is paired with manual review of each data set (Our Methodology and Data Sources, 2018). To aid in analysis of the data, big data tools also include search and visualization features (Storey & Song, 2017).

Using the large digital footprint left by the average person, big data excels at tracking patterns of activity. For example, big data algorithms on online stores are able to determine which items are ordered together and use this to make recommendations to the user (Karaoulanis, 2019). The patterns that can be analyzed aren't limited to human activity; big data can also help predict the weather (Karaoulanis, 2019). Because of the focus on gathering and processing information rapidly, big data is well-suited to tasks that deal with quickly changing conditions, such as weather prediction.

Big data can tell what people decided to do, but it does not know why those choices were made. Collecting data through qualitative methods, such as questionnaires, is still necessary to gain a more complete understanding of human behavior (Croft, 2014). Additionally, the process by which big data is gathered and analyzed is poorly understood by the public (Croft, 2014). Information that would benefit understanding of big data, such as how the collected data is processed, is often unavailable. Even the users of big data software can be confused by its complexity, which leads them to make mistakes in data processing. Big data struggles to handle outliers, so the process of removing them from the data set often falls to a human (Karaoulanis, 2019). If this is done improperly, the data can be skewed to the point where it is no longer reliable. Overall, big data should be recognized as a useful tool for analysis, but it cannot be relied upon as the sole method to predict behavior.

The Issue of Overcrowding

Negative Impacts of Congestion

There are a large number of vehicles in the United States. Counting only personal vehicles, the Bureau of Transportation Statistics had well over 200 million registered vehicles in 2019 (Bureau of Transportation Statistics, *n.d.*). With so many vehicles, there are often groups traveling at roughly the same time and on the same roads, which becomes vehicle congestion and often leads to reduced speed or delays. This issue is prevalent in many areas across the US, as any extra road space created to try and alleviate the problem was met with a proportional amount of extra vehicles. This phenomenon is dubbed by researchers as “The Fundamental Law of Road Congestion” (Duranton, G, 2011).

Another important consideration that goes hand in hand with driving and congestion is parking availability. Parking availability is typically defined with three factors: occupancy, duration, and turnover. Occupancy is the percentage of total spaces that are filled by vehicles. Optimally, this value should be kept around 85-90%. Duration is the time a vehicle stays in any individual spot and can be averaged for the entire lot. This can be used to find the turnover, which is the time window of the study divided by the average duration (ex. If the average duration is 15 min/vehicle, then the lot has an hourly turnover of 4 vehicles) (Un, 2010). If a parking lot is shown to have high occupancy, high vehicle duration, and a low turnover, it indicates a low parking availability. This directly contributes to traffic congestion as “cruising for parking results in a decrease in the overall speed” of traffic (Zhu, 2020).

Traffic congestion has a negative impact on the environment. Heavy congestion leads to slower speeds and higher speed fluctuation, and in turn raises the fuel consumption. Even a small change in traffic speed can lead to significantly more carbon emissions (Barth, 2008). The higher fuel consumption leads to more carbon emission into the atmosphere, which affects air quality and public health in the congested areas. The noise level also increases with traffic congestion, disturbing the residents in the area and causing more stress (Bull, 2003).

There are also negative impacts associated with the availability of parking. In many cases, when people cannot find a parking space, they circle around the parking lot and go back onto the nearby streets (Ibrahim, 2018). When a large number of cars cannot find parking, traffic congestion is likely to occur. Lack of parking also increases the demand for curbside parking, which occupies one or more lanes on the road, potentially causing more traffic congestion and safety issues (Ibrahim, 2018). Many studies report roadside parking leads to more car collisions, automobile-pedestrian conflict, or even fatalities. (Edquist, 2012) The most common collision occurs when a car parked on the road side exits into oncoming traffic. Another common accident occurs when the parked car opens the door and collides with another car on the main road. On average, more than 3,000 fatalities each year are caused by roadside parking (Ghosh et al., 2017). Cars parked on the roadside limit the visibility of both drivers and pedestrians, thus making accidents more likely to occur (Ghosh et al., 2017).

In general, traffic congestion and parking availability also negatively affects the experience of tourists. Research has also shown that traffic congestion has negatively affected the impressions of visitors towards attractions. Visitors also had to readjust their travel patterns to avoid heavy traffic. The long wait times has also resulted in increased stress and negative attitudes from visitors (Chong, 2015).

National Park Traffic

Congestion in the Parks

The National Park Service (NPS) is a conservation effort officially established by the United States government in 1916. Since then, more than 400 national parks have been established across the United States, collectively receiving 316 million visitors every year (What We Do, 2020). As the popularity of these parks has risen, so has the number of concurrent visitors. While this means the park is serving more of the public, it also means that the human impact on the natural landscapes of the park increases.

Though National Parks are constructed to allow people to enjoy the local natural and cultural history, the overwhelming number of visitors can cause damage to the environment. These include natural resource depletion, pollution, and disturbance to the ecosystem (Finnessey, 2012). The increase in the number of visitors increases the land degradation, since there will be more people trampling the vegetation and parking on natural lands. At the Grand Canyon, park management reported that thousands of visitors wander off the marked trails and tramples on natural vegetation (Briggs, 2010). In Rocky Mountain National Park, there have been numerous instances where visitors park on fragile vegetation and leave plastic wrappers on wild flowers (Kwak-Hefferan, 2020). In Yellowstone National Park, there have been cases where visitors came in close contact with bison. The visitors ventured too close and were then hurt by the bison (Simmonds et al., 2018). Sound pollution is especially significant, due to it disturbing the wildlife and their breeding cycles (Finnessey, 2012). Not only does noise pollution impact the enviornment it also creates a negative user experience. At the Muir Woods in Yellowstone National Park, visitors have reported that they could not listen to the sound of the waterfall and bird calls and experience the peacefulness. This is due to the excessive noise caused by people chattering and excited children giggling (Simmonds et al., 2018).

Both traffic congestion and parking availability also negatively affects the visitor experience in National Parks. At Yellowstone National Park, a news article reported: “Hundreds of parked cars once choked the narrow road leading toward the entrance, threatening the local watershed and wildlife, causing headaches for nearby residents”, further indicating the negative impacts of traffic congestion in National Parks (Simmonds et al., 2018). In another case at Yellowstone National Park, a herd of bison slowly moved across a two-way road, causing a two-mile line of cars. This caused some visitors to sound their horns at rangers directing the traffic (Simmonds et al., 2018). Similarly, the lack of parking has also negatively affected the visitor experience. At Glacier National Park, there have been reported fistfights occurring in parking lots, where people are contending for a parking space. Another instance that occurs frequently is that passengers leave the cars, walk around the parking lot and stand in empty parking spaces to secure them, other cars trying to claim the parking space then bumps them (Simmonds et al., 2018).

Strategies for Addressing Congestion

Congestion can negatively impact visitors' experience and safety, as well as the ability of the NPS to protect natural resources (Managing Congestion, 2020). Because of this, the agency has attempted to find ways to address this issue. The NPS website keeps a regularly updated Congestion Management Toolkit for individual parks (Congestion Management Program, 2021a). This Toolkit compiles and evaluates methods, referred to as "tools," used to address congestion. Each tool's estimated time and monetary cost are listed, as well as their pros and cons. Tools are also broken down by the problem areas they address, such as parking lot congestion or pedestrian safety. Some examples of these tools include implementation of park-operated bike rentals, adding specific lanes for accelerating or decelerating vehicles, and increasing signage at intersections (Managing Congestion, 2020).

Individual parks will also create and test potential congestion solutions. These solutions attempt to cut down on the density of visitors for any given timeframe. For example, in the Great Smoky Mountains National Park, group access reservations and "vehicle-free" days are being tested in sections of the park(Sophie, 2021). Another example is the Gulf Islands National Seashore, where vehicles are barred entry after parking lots are filled (Soehn, 2021). In another instance, a study published in 2009 looked into integrating transportation with user capacity research in Yosemite National Park. The study used visitor surveys and computer simulations, and models to provide an empirical basis for traffic management in the park (Lawson et al., 2009).

Acadia National Park

Located on Mount Desert Island in Maine, Acadia National Park is one of the most popular national parks in the United States. The park currently protects approximately 47,000 acres, encompassing much of Mount Desert Island, surrounding smaller islands, and the Schoodic Peninsula (U.S. Department of the Interior, n.d.). The 27 mile Park Loop Road is one of the park's main points of interest. This road allows visitors to travel along a scenic route that connects to many of the main attractions in the eastern half of the park. Due to the small size of the park and the large number of yearly visitors, the Park Loop Road often experiences severe traffic.

Congestion in Acadia

In 2020, Acadia National Park was visited by more than 2.5 million people, making it the 8th most visited national park in the United States (U.S. Department of the Interior, 2020). During the park's busiest times, roadways become crowded, with large numbers of cars parking on the sides of roads. Due to the one-way nature of the road, many visitors find themselves in traffic jams around points of interest. These crowded roadways have become a common complaint. One study done in the park found that survey respondents considered 40 cars per mile the maximum acceptable level of use (Hallo, 2008). A main focus in managing Acadia National Park has been proposing measures to mitigate the congestion problem. Congestion has become such an issue that there are guidelines posted on Acadia's website on how to avoid traffic in the park. These guidelines include suggestions to use alternative forms of transportation, such as the Island Explorer Bus system, cycling, or even using a taxi service (U.S. Department of the Interior, 2021b). If visitors do decide to bring their own vehicle, the park recommends arriving either before 8am or after 5pm in order to find parking. Even with this recommendation, the park specifies to "visit Acadia with a back-up plan" and "don't expect to find a parking space at popular sites" (U.S. Department of the Interior, 2021b). These suggestions come from years of dealing with traffic congestion in the park. In 2019, the park had 105 road closures caused by traffic congestion (Broom, 2020). While 2020 had significantly fewer visits, the number of visits in October was up by 10% from 2019. This was followed by a 27% increase in the month of November. Coupled with the shut-down of the bus system due to Covid-19, these factors led to a "record amount of [vehicle] traffic," (Acadia on My Mind, 2020).

Strategies for Address Congestion

The management team of Acadia National Park has been attempting to alleviate their congestion problem by conducting years of research (as discussed in Appendix A) and proposing congestion management plans. The 1992 plan not only recognized the growing traffic issue but also proposed a public transportation plan for the entire Mount Desert Island. The initial system included a route between the park and local campgrounds and directly contracted a non-profit bus service called Downeast Transportation, Inc. Since its implementation, the bus service has been continuously expanded to serve Acadia National Park and its neighboring communities. Due to its success, the bussing system was expanded to a locally funded line named the Island Explorer (U.S. Department of the Interior, 2018). This expansion of the bus service continued into the most recent 2019 transportation plan (US Department of the Interior, 2019).

One management action recently implemented is a reservation system for the Cadillac Summit Road. This was the preferred alternative action option in the *Acadia National Park Final Transportation Plan / Environmental Impact Statement*. An entry booth has been constructed at the base of Cadillac Summit Road that checks timed entry reservations from visitors during the peak season. Visitors reserve a half hour time slot to enter the road and are free to stay as long as they like afterwards (US Department of the Interior, 2019). Currently the Cadillac Summit Road system is the only reservation system that has been implemented. There are plans to implement a reservation system in the Ocean Drive corridor and the Jordan Pond House as well (US Department of the Interior, 2019).

Methods

Introduction

The primary goal of this project was to provide an analysis of visitor mobility patterns that would inform park management about visitor movement in the Park Loop Road region. To achieve this goal, we utilized two methods. First, we used Streetlight Data to gather data on visitor movement patterns. Second, we performed direct observation on site to corroborate the data from StreetLight. Using the data from StreetLight and direct observation, we applied an algorithm to calculate the average parking duration in parking lots. Additionally, we investigated the impacts of the Cadillac Mountain reservation system, as well as demographic information from the entire park.

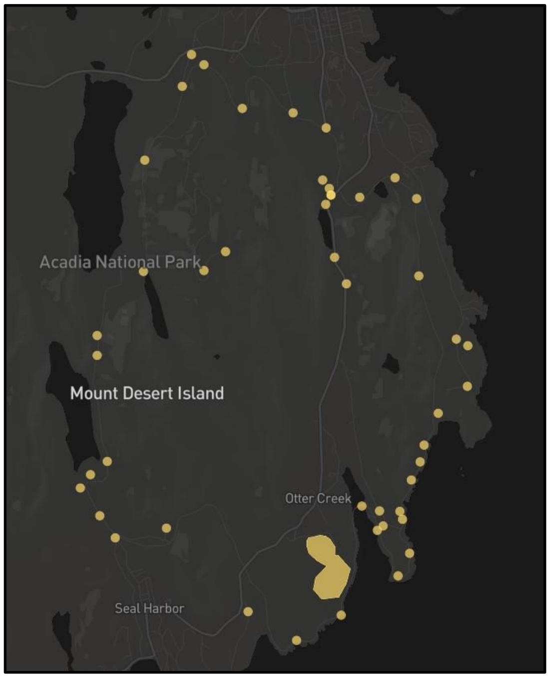
StreetLight Data

The first method of data collection is a traffic database called StreetLight Data. The team utilized this database to investigate the general traffic patterns in the Park Loop Road region. This began before the team arrived on-site, using the data derived from the traffic patterns of previous years.

The StreetLight Data platform is a database composed of cell phone geolocation data and vehicle geolocation data. Through a proprietary algorithm, the platform takes the location data and generates an interactive traffic database. By accessing StreetLight Data through an internet browser, a user can view how visitors traveled through any region between 2018 and 2020. This is done by drawing areas of interest, which StreetLight Data calls zones. As shown in Figure 1, we placed these zones at the areas we wanted to analyze in the Park Loop Road region. This included entrances to the Park Loop Road, key attractions, parking lots, and popular side road parking spots. Once these zones are established, StreetLight Data analyzes how many people traveled to, from, or through the established zone. This can be done for any specified timeframe down to the hour.

Figure 1

StreetLight Zones Used for Vehicles



Note. A screenshot of all of the parking lots and attractions analyzed in StreetLight (vehicles only)

The specific timeframes that we analyzed in StreetLight included the peak seasons of 2018, 2019, and 2020. The first sets of analyses looked into all days from June 1st to November 1st of each year. Once analyses were completed in StreetLight, the resulting data was separated into weekdays (Mon. - Fri.) and weekends (Sat. + Sun.). StreetLight then further separates this data by the time of day. This can be done by the hour or by the time of day (Table 1). This separation of data allows for the overall weekly and daily patterns of visitors throughout the entire summer to be viewed. The next set of analyses separated each summer into weeks to view the visitor patterns of specific weeks and days.

Table 1

Default Time Segments

Time of Day	
Early AM	12am - 6am
Peak AM	6am - 10am
Mid-Day	10am - 3pm
Peak PM	3pm - 7pm
Late PM	7pm - 12am

Note. The default time of day distribution in StreetLight Data

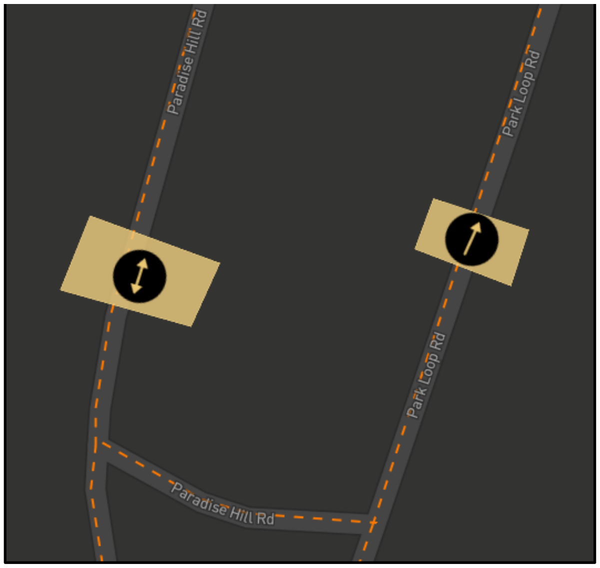
In order to monitor traffic throughout the Park Loop Road region, we split the region into zones within the StreetLight Data program. This allowed us to complete an origin-destination analysis of visitor travel between these zones. More specifically, this means that we used the StreetLight Data program to find the percentage of vehicles that travelled to and from the different zones. There are two types of zones: small "pass-through" zones and large "non-pass-through" zones.

By establishing pass-through zones in StreetLight, we were able to closely analyze the flow of traffic along Park Loop Road. These zones spanned the width of a segment of road, similar to a gate (Figure 2). We placed these zones along lengths of Park Loop Road and ran an origin-destination analysis. This provided us with the time distribution of vehicles passing between the zones. Additionally, this analysis provided distributions of the time it took for vehicles to travel between these zones and the speed at which they were travelling. This allowed us to measure both when and where congestion occurred on the Park Loop Road.

The larger zones cover destinations throughout the park loop region. These zones provided us with the number of visitors that traveled to general regions of the park at any given time. We then completed another origin-destination analysis focusing on visitor travel from park entrances to all parking areas in the region. This showed us the distribution of visitors that went to each zone from every park entrance. Additionally, we analyzed travel between specific parking areas, the distribution of cars taking each exit from parking these parking areas, and travel patterns to the park from Bar Harbor.

Figure 2

Vehicle Pass-through Zones



Note. A screenshot of Two pass-through zones: one placed at the beginning of the one way segment of Park Loop Road, the other placed at a two way entrance to the Park Loop Road.

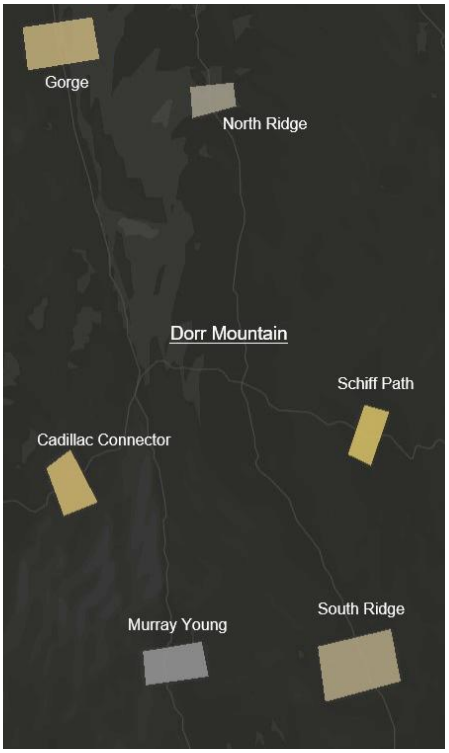
StreetLight Data allowed us to gather information on vehicular traffic patterns throughout the Park Loop Road. By determining where traffic congestion has occurred in the past through StreetLight, we determined where we should focus our on-site analyses. This focus was largely placed on tracking vehicles entering and leaving parking lots. With StreetLight, we were able to analyze the distribution of vehicle trips entering and leaving individual parking lots. This was then corroborated with parking lot turn-over rates derived from direct observation (Section 2.3). We also analyzed congestion to determine how park policy changes have affected traffic congestion. For example, one location we investigated was the Cadillac Mountain Road. By analyzing past traffic volumes on the road, and comparing them to the current volumes, we were able to determine if and how the new reservation system affected traffic congestion.

In addition to gathering the traffic patterns of vehicles, our team also looked into using StreetLight to gather data on the movement of visitors traveling by foot. More specifically, we analyzed visitor travel patterns to the peaks of Acadia's mountains. This was completed with a similar methodology to the way we analyzed vehicle patterns, but with some slight modifications. Just like the vehicle methodology, we set up non-pass-through zones at the locations of interest, which was the peak of every mountain in Acadia National Park. The analysis was set to record only "pedestrian" data, using only location data generated from visitors that stopped in these zones for 5 minutes or more. Cadillac Mountain's summit was excluded due to visitors arriving by car, then walking, having the ability to skew the final distribution. From there, a slightly different analysis, called a zone activity analysis, was run. A zone-activity analysis tracks the number of devices entering a zone. Once a sample is completed for each zone, they are then compared to show the percentage of devices that end at each of the created zones.

After taking a look at the percentage of hikers that went to each summit, we then took a closer look at Dorr Mountain to see the distribution of trail usage at its peak. To do this, we completed another pedestrian zone-analysis, but with pass-through zones. These pass-through zones were placed at the tops of the trails leading to Dorr Mountain's peak (Figure 3). This generated a percentage distribution of the most used trails at the top of Dorr mountain.

Figure 3

Dorr Mountain Pass-through Zones



Note. A screenshot of pass-through zones covering each trail connected to Dorr Mountain's Summit

The data we obtained using StreetLight is difficult to understand using the tools the service provides. Therefore we transferred the data to Microsoft Excel so that we could format the data points into graphs that depict the precise information we want to convey. An example of such a graph is shown below. Using Excel, we had a greater degree of control over how the data is presented and were able to ensure important information would be clear to the reader.

Direct Observation

The other major data collection method used aside from StreetLight was direct observation. Team members were stationed at designated locations to record traffic data. These locations included points of interest such as the Cadillac Mountain intersection, the Nature Center parking lot, and the Jordan Pond House parking lot. The choice to observe these points was informed by the results of StreetLight analysis. We recorded information on passing cars, cars entering and cars exiting.

Our observation at both the Jordan Pond House and the Nature Center focused on the parking lots. Our goal was to determine the average length of stay at both locations. In order to do this we stationed team members in the parking lots and had them record every time a car entered or left a parking spot. We also noted how many cars were parked in the lot when we arrived and when we left. Team members observed at the Jordan Pond House for seven days in total. The first three days we observed from 9 am to 12 pm. The remaining four days we observed from 9 am to 4 pm. We observed at the Nature Center for five days in total. The first three days we observed from 9 am to 12 pm and the remaining two days we observed from 9 am to 4 pm. In total we observed approximately 6800 cars.

Parking Duration Analysis

By looking at the time people stayed parked, we can get an idea of how busy a parking lot gets, which can inform future decisions concerning the parking lot. We were able to determine the average parking duration using data from observation, and via StreetLight analyses.

Analysis for Duration from Streetlight

StreetLight Data is capable of looking at all cars that interacted with a specified zone, called a Zone Analysis. These analyses present an hour-by-hour breakdown of when cars entered or left the zone, presented as percentages. We used the product of each percentage and the hour they occurred during to make them weighted. We then found the average of those weighted percentages to find the average times cars came into and out of the zone. The difference between the average entrance time and the average exit time expresses the general average time cars stay parked in the zone (Appendix B).

We repeated this analysis for each parking lot we have looked at thus far and across all years of data we have available to make as complete of a picture as possible.

Duration from Direct Observation

To complement these analyses, we also performed direct observation, noting down any time a car entered a space or any time a car left a space. By finding the average difference between the entrance and exit times we can get a rough estimate for the average time a car stayed parked. We also have to account for the cars that were already there at the beginning of our testing interval and those that were still there at the end. To accomplish this, we first noted down the number of cars present in a parking lot when we started data collection. When performing the calculation (as referenced in Appendix C), we subtract the number of cars noted down previously from the start of the list of exit times. Similarly, we remove the number of cars present at the end of our data collection period from the end of the list of times that people parked. This gives us an accurate count of how long cars stayed in parking spaces, on average.

Results and Findings

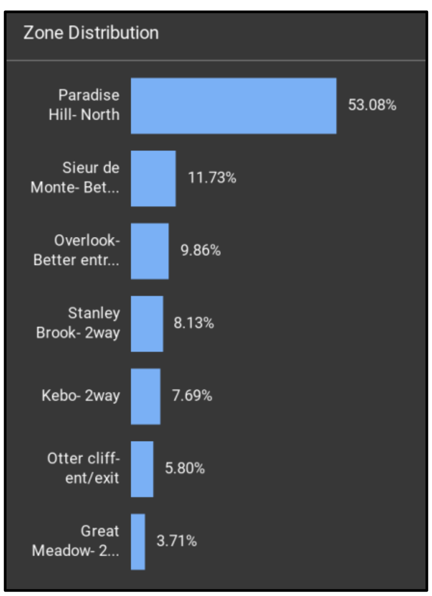
The following results section will lay out the important data collected, identifying a pattern and drawing correlations between these patterns to present findings. There will be four findings presented, each related to one category of study. These include visitor travel patterns, parking duration, Cadillac Mountain reservation system, and visitor demographics.

Visitor Travel Patterns

StreetLight can provide effective and accurate information on the movement of people, vehicles and bikes in the park. Identifying these movement patterns can be advantageous to the park in understanding visitor mobility. By using StreetLight Data, we were able to discern a series of popular travel patterns visitors follow when bringing their personal vehicles into the park. These patterns include where visitors go, when they go to those places and how long they stay there. For example, by setting up pass-through zones at each of the entrances to the Park Loop Road, we determined what percentage of vehicles take each entrance road to the park. The most popular entrance (shown in figure 4) was Paradise Hill Rd. at 53.08% during the peak season (Jun.1-Nov.1) of 2019. This was followed by Overlook Rd. and Sieur de Monte Rd.

Figure 4

Entrance Distribution for the Park Loop Road

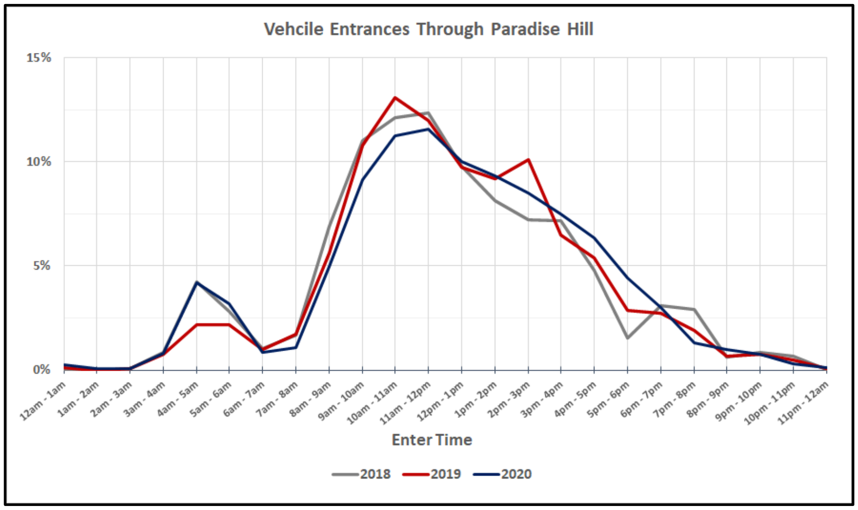


Note. Park Loop Road vehicle entrance distribution from June 1st to November 1st, 2019

StreetLight also gives us the time distribution of when visitors are travelling through these entrances. This included the Paradise Hill entrance, which we decided to take a closer look at. We looked to see how its time distribution had changed over the three year period, keeping the time-frame of June 1st to November 1st consistent between years. The final result is shown in figure 5. As shown in the figure, the time distributions stayed largely consistent between years, with similar peaks around sunrise and mid-day. A higher resolution graph with 15 minute intervals can be seen in Appendix D.

Figure 5

Vehicle Entrance Distribution Through Paradise Hill

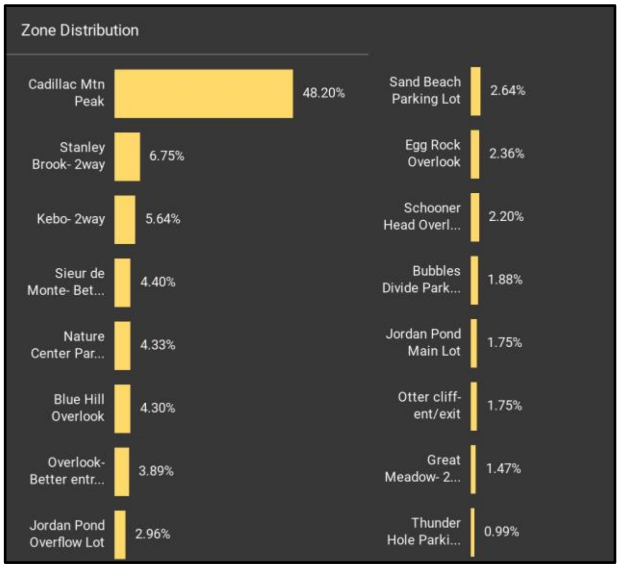


Note. Paradise Hill entrance time distribution between June 1st and November 1st. This time period was compared between the years 2018, 2019, and 2020.

We also found where visitors were traveling after they entered through the Paradise Hill entrance with an origin-destination analysis. This was done by setting the Paradise Hill entrance as an origin point and seeing what destination zones vehicles travelled to. Through this analysis, we found that the first stop most vehicles make is Cadillac Mountain, with 52.50% of vehicles going to the Cadillac Mountain Peak and Blue Hill Overlook (Figure 6). This analysis included all of the exits to the region as well.

Figure 6

Destinations from Paradise Hill

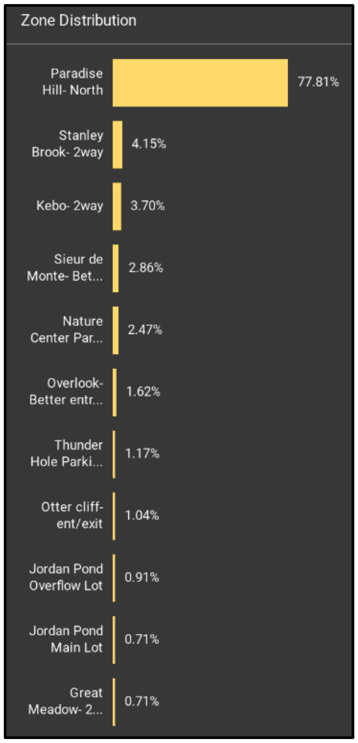


Note. The distribution of vehicle destinations (including parking lots and exits) from the Paradise Hill entrance from June 1st to November 1st, 2019. All destinations shown consisted of 0.99% or more of the trips.

Once we found that Cadillac Mountain was the most popular destination from the Paradise Hill entrance, we decided to look into where all of the visitors to Cadillac Mountain were traveling to once they drove out of the parking lot. This was done with another origin-destination analysis with the Cadillac Mountain Peak and Blue Hill Overlook parking lots as the origin. With all of the lots and exits in the Park Loop Road region as the destination, we found that out of all vehicles leaving the Cadillac Mountain Peak parking lot, 93.06% take an exit from the Park Loop Road (Figure 7).

Figure 7

Destination Distribution from Cadillac Mountain



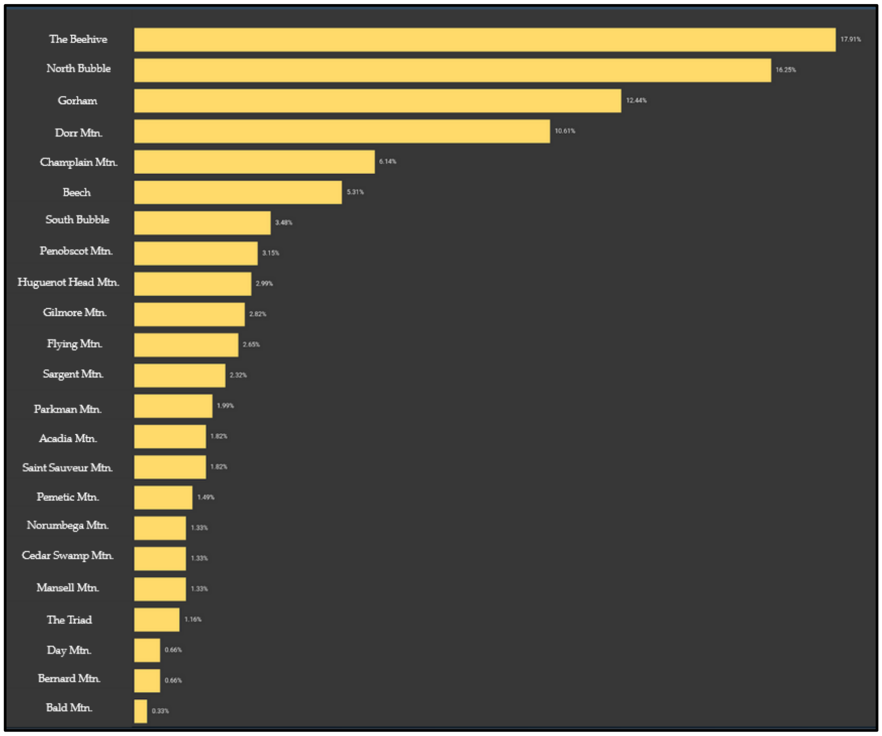
Note. The distribution of the top 11 destinations from Cadillac Mountain from June 1st to November 1st, 2019. All Park Loop Road exits and parking lots act as the destinations

Additionally, we looked at the data generated by our StreetLight analysis on pedestrians moving about the park. We first took a look at the summit analysis, which gave us a percentage distribution of the most popular summits in all of Acadia national park (Figure 8). From this data, we found that the top five most popular summits are within the Park Loop Road region. This included Dorr mountain, which we focused our trail analysis on. The results of the trail analysis can be shown in figure 9, where we found that the two most popular trails to the Dorr summit were Schiff Path and Cadillac Connecter at about 27% each.

Just like the vehicle analyses, time distributions can be generated for pedestrian mobility data. The distribution for Schiff path is shown in figure 10, and shows that the busiest time during the analysis period was between 11am and 12pm.

Figure 8

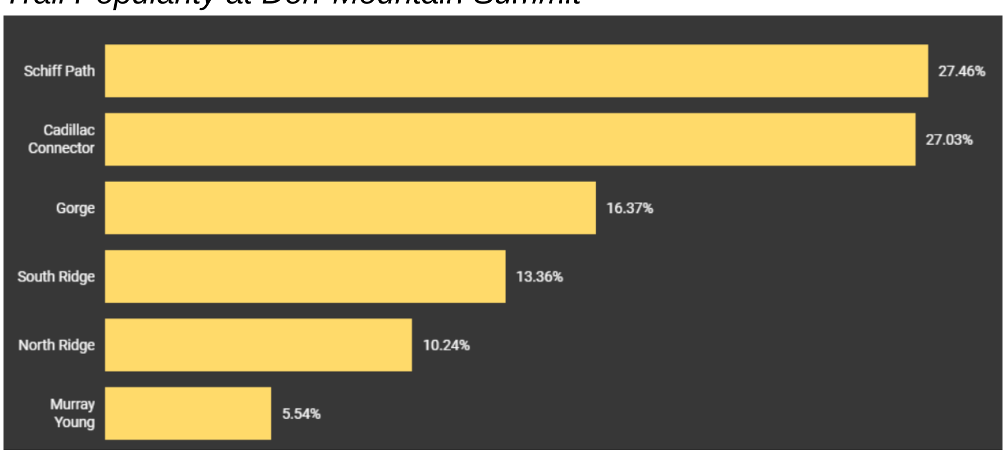
Summit Popularity Distribution



Note. All summits in Acadia National Park ranked by the percentage of visitors that stopped at the summits during all of 2019.

Figure 9

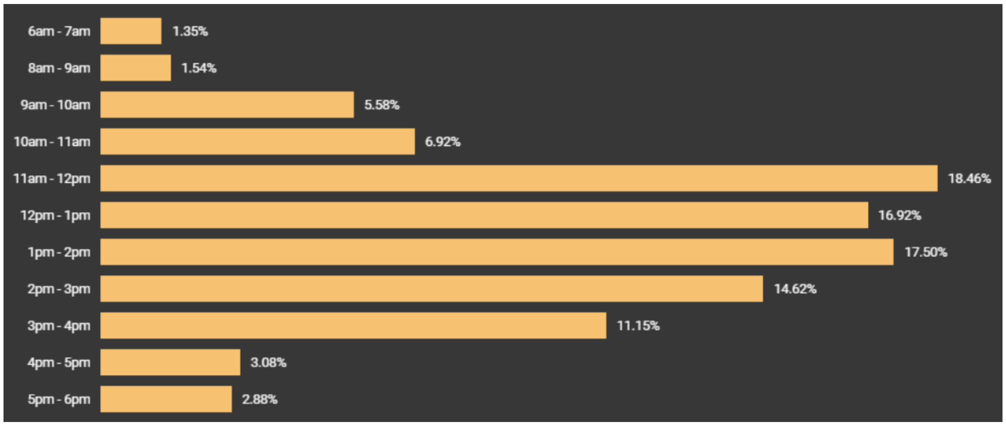
Trail Popularity at Dorr Mountain Summit



Note. All summits in Acadia National Park ranked by the percentage of visitors that stopped at the summits during all of 2019.

Figure 10

Schiff Path Time Distribution



Note. The time distribution of when visitors pass through the top of Schiff Path during all of 2019.

Average Parking Duration Found with StreetLight

The analysis of the big data from StreetLight can generate average parking duration equivalent to that of direct observation, but with a significantly smaller time investment. Using StreetLight Data, we determined the average parking duration in two of the most popular parking lots, which yielded similar results to direct observation. StreetLight, however, has the ability to gather more data on more parking lots than can be covered with direct observation. The two most popular parking lots that we focused on were the Jordan Pond House parking lot and the Nature Center parking lot. With the time distribution of when vehicles enter and leave parking spaces from StreetLight, we calculated the average parking duration during the Summer of 2019, which was defined as between June 1st and August 31st, 2019. The average parking duration for the Jordan Pond House during this period was 86 minutes. For the Nature center, the average parking duration was 64 minutes. This same information was gathered for the Summers of 2018 and 2020. Additionally, the parking duration was calculated for multiple other parking lots throughout the park for the Summers of 2018, 2019, and 2020. The parking duration for the analyzed parking lots can be found in Table 2.

Table 2

Parking Duration from StreetLight and Direct Observation

	Parking Duration		
Parking Lot	StreetLight (Jun. 1st to Aug. 1st, 2019)	StreetLight (Jun. 1st to Aug. 1st, 2020)	Direct Observation (June, 2021)
Cadillac Mtn. Peak	41 minutes	54 minutes	-
Nature Center	64 minutes	84 minutes	84 minutes
Jordan Pond Main	86 minutes	57 minutes	204 minutes
Jordan Pond Overflow	100 minutes	102 minutes	-
Sand Beach	97 minutes (Jun. 1st to Jul. 1st 2019)	103 minutes (Jun. 1st to Jul. 1st 2019)	97 minutes*

Note. Parking duration of every parking lot

*(Crock et al., 2021)

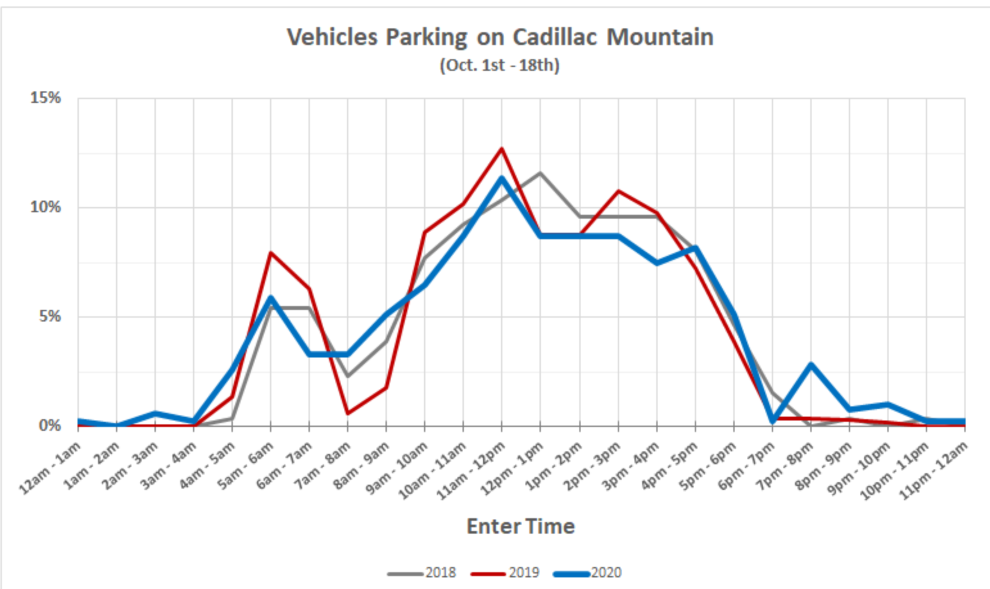
In addition, we calculated the parking duration using direct observation data. This was done for the Jordan Pond Main lot and the Nature Center lot. The resulting parking duration is 209 minutes for the Jordan Pond House and 84 minutes for the Nature Center. The days this data was completed, however, was only 2 days for the Jordan Pond Main Lot and 1 day for the Nature Center. This means that there is significant room for bias in this observation. However, the direct observation data for Sand Beach lot was gathered from a separate study, which found an average parking duration of 97 minutes determined from over a month of camera observation (Crock et al., 2021). The study on the Sand Beach lot was completed during the months of June and July, so the timeframe used in StreetLight was altered to more directly compare the two.

Results of the Cadillac Mountain Reservation System

The reservation system resulted in little to no change to the visitor time distribution on the Cadillac Mountain, shown below in Figure 11 is the time distribution of visitors to the Cadillac Mountain in three consecutive years. These data were collected during the two-week trial period of the reservation system for consistency. Each line in the graph represents data for a different year. It can be seen that the peak hours are around sunrise, after which it decreases until it peaks again at around 11 am, before falling back down for the rest of the day. Comparing the three lines, it is evident that there is little difference between them, meaning that the time distribution of visitors arriving on Cadillac Mountain has remained about the same.

Figure 11

Cadillac Mountain Time Distribution



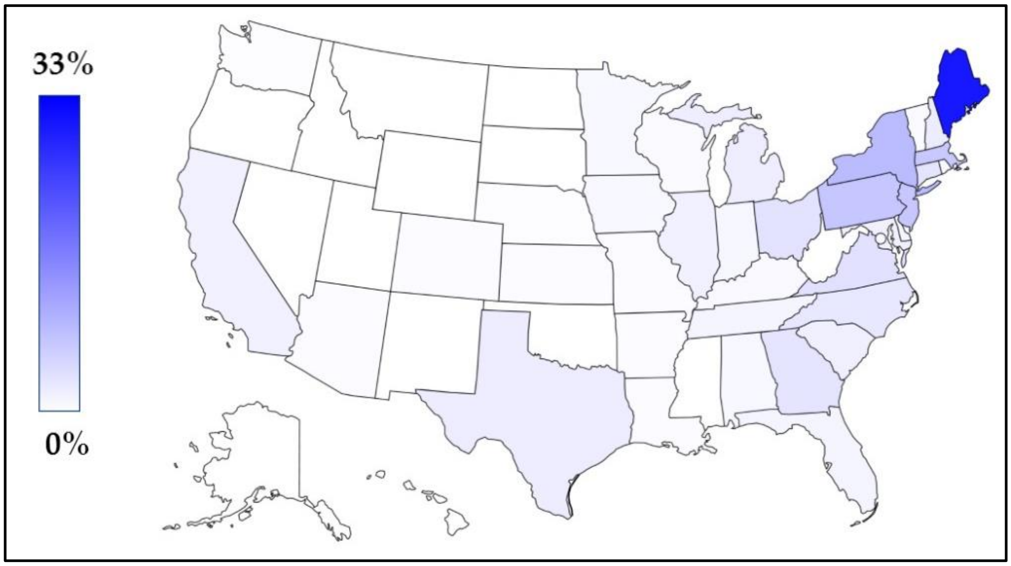
Note. Time distribution of parking at Cadillac Mountain peak across three years

Visitor Demographics

StreetLight Data can also provide demographic breakdowns of visitors, along a number of different aspects. This is done through the Zone Analysis tool, which gives all of these statistics for a specified zone or group of zones. We analyzed where visitors to Cadillac Mountain were from, their ethnicity, and if they had children. Within StreetLight, this is all done by referencing U.S. census data.

Figure 12

Home State Distribution

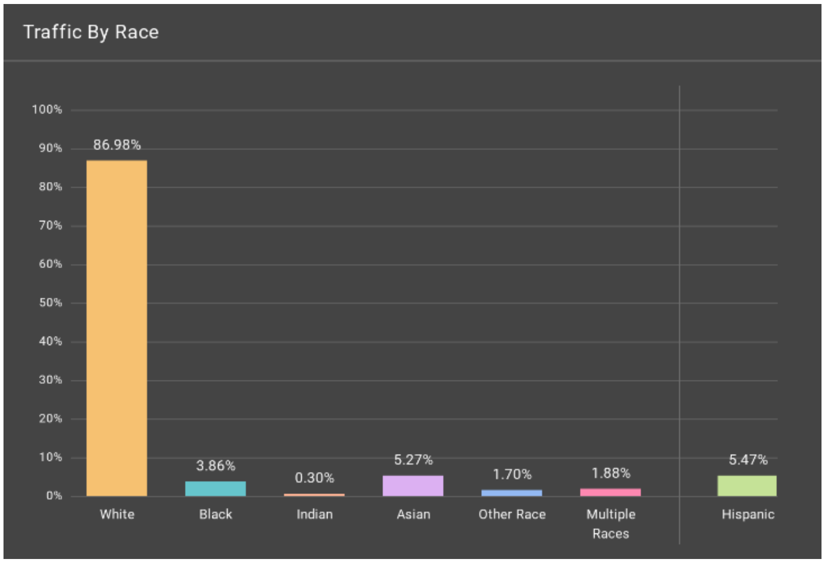


Note. The home-state percentage distribution of visitors traveling to Cadillac Mountain during the peak season (June 1st - Nov. 31st) of 2019 (Specific Values in Appendix E)

From StreetLight's statistics, it can be seen that 30% of visitors live in Maine, and many outside of that percentage coming from other New England states. A heatmap of the states visitors are arriving from can be seen in figure 12, while the total percentage breakdown can be seen in Appendix E.

Figure 13

Racial Background of Visitors

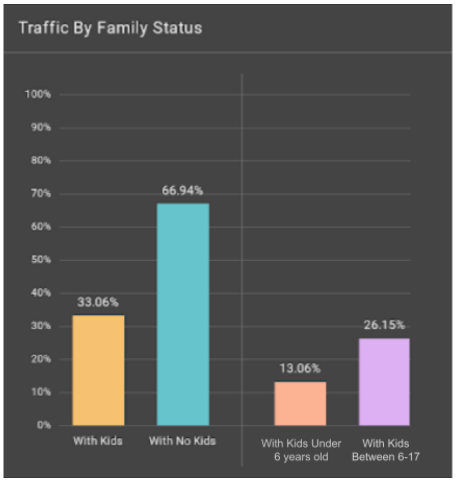


Note. A distribution of the racial background of visitors travelling to Cadillac Mountain Peak during the peak season (June 1st - Nov. 31st) of 2019. Overall, 5.5% of the visitors identified as ethnically hispanic.

Using StreetLight, we are also able to recognize the racial background of all people visiting the park. Shown in figure 13 are the visitors to Cadillac Mountain during the peak season of 2019. As shown, a majority are white, followed by low percentages of Black and Asian. Additionally about 5.5% of the total identified as ethnically hispanic.

Figure 14

Familial Status of Visitors



Note. A family status breakdown of all visitors travelling to Cadillac Mountain Peak during the peak season (June 1st - Nov. 31st) of 2019. The right side is out of a percentage of total families, overlapping due to some families having multiple children.

We can also see the family status of visitors to any location. As shown in figure 14, only 33.06% of visiting families at the peak of Cadillac have children. StreetLight can also show a more specific breakdown, with 13% of visiting families having a child younger than six and 26% having a child in the 6-17 age range.

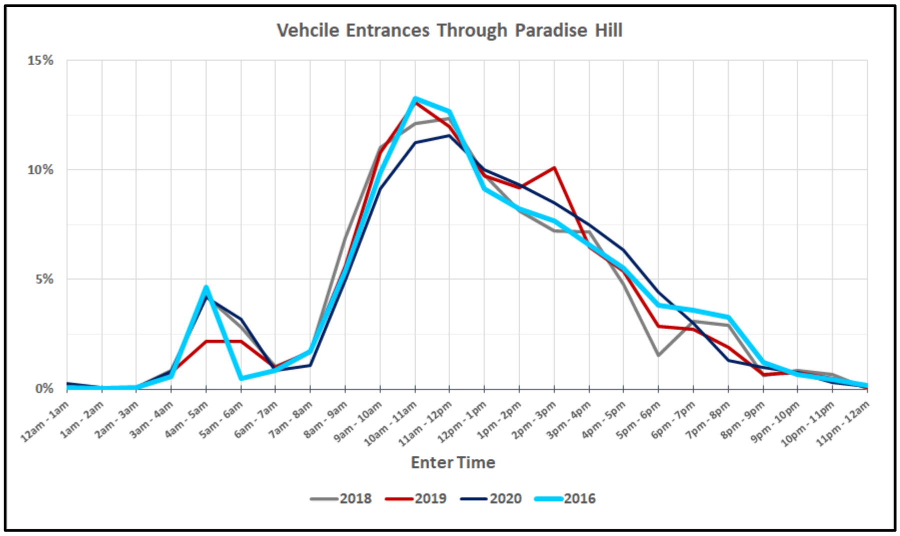
These demographic analyses are easy to obtain with StreetLight and can be found for multiple years in the space of hours. This is opposed to surveying park goers for a number of weeks and looking at the total collected data afterwards.

Discussion

The results from our visitor travel patterns section demonstrate StreetLight's capability of tracking patterns in visitor movement around the park. StreetLight can not only provide a detailed percentage breakdown of visitor destinations, but also can gather the same data across multiple years. We can compare this to existing data to get an idea for how accurate StreetLight is in a given area. A previous research project set up car counters at each entrance onto the Park Loop Road in 2016. Using those collected numbers and comparing them to the data StreetLight provides on the same time frame from a different year, as shown in figure 15, we can see that StreetLight provides an accurate breakdown of when vehicles enter the Park Loop through the Paradise Hill entrance. Unlike other types of data collection methods, such as the traffic counters, which are only operating in the summer months, StreetLight can gather the same type of data all year. In addition, since StreetLight has the capability to provide data from previous years, this means that more data can be accessed and can be used for comparison. Moreover, traditional methods, like the traffic counters, cannot determine where vehicles go. Meanwhile StreetLight can determine the destinations of passing traffic. StreetLight can also obtain a very detailed percentage breakdown of destinations from a set origin. In our analysis, we zoned every possible place that visitors can stop in the Park Loop Road region, these include parking lots, overlooks and attractions.

Figure 15

Vehicle Entrance Distribution Through Paradise Hill (with car counters)



Note. Paradise Hill entrance time distribution between June 1st and November 1st. during the years 2018, 2019, and 2020. Overlaid is the 2016 traffic counter data for the same entrance during the Summer of 2016.

In addition to analyzing vehicles, StreetLight can also analyze people-on-foot in the park. From our results, it has been demonstrated that StreetLight can track hikers going onto mountain summits. Through this, a detailed percentage breakdown of summit usage can be generated. More specifically, zones can be set on every trail on the mountain summits to generate a popularity breakdown, as well as when hikers arrive at each trail. In a similar study done in San Antonio Missions National Historical Park, to monitor the usage of trails, pressure sensors, infrared sensors and human observers were all required (Turner et al., 2013). Compared to traditional on-the-ground methods used to collect this information, which provides less data and requires much more resources, StreetLight is much more efficient. Additionally, using these methods can only generate trail usage data in a certain year, whereas StreetLight has the ability to look at data across different years. This information has the possibility to be crowdsourced and mitigate potential congestion in the park. Furthermore, the same analysis can be done on bikes and e-bikes on the Carriage Roads in the park (Bowles et al., 2021). However, due to our limited research time, we were not able to generate a full analysis on bikes.

To determine the parking duration, the team developed an algorithm that can use data from both StreetLight and direct observation. Using this algorithm, the calculated average parking duration using data from StreetLight is similar to the results calculated using data from direct observations. While they yielded similar results, StreetLight was able to generate significantly more data on all parking lots. Direct observation requires physical presence in the parking lot, which is difficult to maintain at all times. The amount of data generated from a full day of observation is merely a fragment of what StreetLight can provide. Moreover, the bias contained in the direct observation data is crucially higher, caused by factors such as lack of randomization and insufficient data points.(Hammer, 2009). As shown in Figure H, the Jordan Pond House parking duration from StreetLight and direct observation are noticeably different. One potential reason causing this discrepancy could be due to our insufficient data sample from direct observation. The team were only stationed at the Jordan Pond House collecting data for seven days, with two days being the busiest times of the month. However, with a large enough sample size, it is shown that the parking duration calculated from direct observation data corroborates the parking duration calculated from StreetLight data. As shown in Figure H, the Sand Beach parking lot durations were much closer. The direct observation data was provided by another team researching only at the Sand Beach area, where they gathered a much larger sample using camera monitoring. The StreetLight parking duration came out to be 97 minutes in 2019 and 103 minutes in 2020, while the duration from direct observation came out to be 97 minutes. This showed an approximately 5% difference in the result, further proving that with a large enough sample size, the parking duration calculation algorithm can be accurate. In a similar study, cell phone location data was used to determine the visitation to a water recreation area. It was concluded that when using locational data, it is necessary to use on the ground observation data to perform calibration. However, once calibrated, aggregated cell phone location data can provide high-resolution information in both space and time, providing good estimates of visitation in a natural area (Merrill et al., 2020).

All of the results relating to visitor demographics demonstrate StreetLight's usefulness as a substitute for long term on the ground demographic surveys. It uses census data and statistical estimation to determine the demographics of the data it samples from. Sending out demographic surveys can also cause privacy issues, with people feeling uncomfortable answering these types of questions involving sensitive materials (McNeeley, 2012). Additionally, if we would like to find out the visitor demographics information in a certain area in the park, StreetLight has the capability to accomplish that.

Recommendations and Conclusions

Based on our research and what we have learned about the capabilities and reliability of using StreetLight in Acadia National Park we have three key recommendations. These recommendations assume the park has access to big data and provide three areas we suggest it is applied to.

Recommendation 1

We suggest that the park consider utilizing StreetLight or a similar big data software to continue the analysis of visitor movement patterns in the Park Loop Road region. More specifically we believe the park would benefit from conducting more research on vehicle mobility and performing more in depth analysis of pedestrians and cyclists. It may also be valuable incorporating demographic information into these analyses to see if there are any major differences between modes of transportation.

Recommendation 2

We recommend the park make use of these big data tools to maintain the section of the park website pertaining to the busiest times up to date. This will assist in crowdsourcing answers to the traffic issues. The time distributions available on StreetLight will be particularly helpful in this situation.

Recommendation 3

We recommend the park consider using big data software to evaluate future management actions. We have shown that it can be used in relation to the Cadillac Mountain reservation system, so it would be useful in analyzing future tests done on reservation systems at the Jordan Pond House and the Ocean Drive Corridor.

Conclusion

Throughout this report we've investigated the many uses for big data in the context of the National Parks. It is useful in the analysis of visitor mobility, looking at where people and vehicles go, when they arrive, and how long they stay. It can also deliver useful demographic information to provide even more insights into who is moving around the park. The data has been shown to be fairly accurate through the incorporation of our direct observation as well as additional ground truths from other sources. While we did generate recommendations for park management, there is much more that can be done with research like this that we did not have the chance to explore. With time more projects could be done in the park using big data to focus solely on one aspect of mobility, such as just hikers or bikes. Outside of Acadia, data like this could be incredibly useful to other national parks in solving issues of congestion and beyond. At an even higher level someone could even look nationally, at all the national parks, to investigate more of who uses the parks and when. With this vast pool of big data the possibilities are near endless.

Citations

Acadia National Park. (2020, December). Managing Congestion: A Toolkit for Parks. *U.S National Parks Service*. https://www.nps.gov/orgs/1548/upload/Congestion_ Management_2021-508.pdf

Acadia on My Mind. (2020, December 6). Traffic Visits Skyrocket in October at Acadia National Park. *Acadia National Park on My Mind*. https://acadiaonmymind.com/2020/12/traffic- visits-skyrocket-in-october-at-acadia-national-park/

Barth, M., & Boriboonsomsin, K. (2008). Real-World Carbon Dioxide Impacts of Traffic Congestion. *Transportation Research Record*, 2058(1), 163–171. https://doi.org/10.3141/2058-20

Biswas, S., Chandra, S., & Ghosh, I. (2017). Effects of On-Street Parking In Urban Context: A Critical Review. *Transportation in Developing Economies*. 3. 1-14. https://doi.org/10.1007/s40890-017-0040-2

Bowles, P., Lyu, Y., McFann, A., Merchan, J., & Rothstein, M. (2021). *E-bike Usage on the Carriage Roads*. Unpublished Interactive Qualifying Project. Worcester Polytechnic Institute.

Briggs, B. (2010, August 30). National parks feel the effects of human, environmental threats. *NBCNews.com*. https://www.nbcnews.com/id/wbna38883753

Broom, D. (2020, January 2). Gridlock caused 105 road closures in Acadia in 2019. *Mount Desert Islander*. https://www.mdislander.com/maine-news/gridlock-caused-105-road-closures-in-acadia-in-2019.

Bull, A. (2003). Traffic congestion: the problem and how to deal with it. *NU. CEPAL, & German Agency for Technical Cooperation*. https://repositorio.cepal.org/handle/11362/37898

Bureau of Transportation Statistics. (n.d.). Number of U.S. Aircraft, Vehicles, Vessels, and Other Conveyances. *Bureau of Transportation Statistics*. https://www.bts.gov/content/number-us-aircraft-vehicles-vessels-and- other-conveyances.

Calvi, A., Maki, C., Shuai, M., Peters, Jackson. & Wivagg, Daniel (2017) *Intelligent Transportation in Acadia National Park*. Interactive Qualifying Project, Worcester Polytechnic Institute. https://digital.wpi.edu/concern/student_works/ns0646752?locale=en

Caltabiano, J., Charbonneau, J., Fischler, M., & Gillian T. N. (2016) *Visitor Cell Phone Application: An Innovative Design to Monitor Visitor Mobility in Acadia National Park*. Interactive qualifying project, Worcester Polytechnic Institute. https://digital.wpi.edu/concern/student_works/s7526d03d?locale=en

Chong, K. L. (2015). The impact of traffic congestion on tourist behaviour: case study of Chiang Mai. *In Proceedings of the 14th ApacChrie Conference 2016, Bangkok, Thailand, 11-13* May 2016 (pp. 290- 300). Dusit Thani College Thailand.

Croft, C. (2014). The Limits of Big Data. *The SAIS Review of International Affairs, 34*(1), 117-120. https://doi.org/10.1353/sais.2014.0005

Crock, N., Forgione, A., Murguia, E., Van Milligan, J., & Wheelock, R. (2021) *Traffic Mobility Patterns on the Ocean Drive Corridor*. Unpublished Interactive Qualifying Project. Worcester Polytechnic Institute.

Duranton, G., & Turner, M. (2011). The Fundamental Law of Road Congestion: Evidence from US cities. *American Economic Review 101*, 2616–2652. https://doi.org/10.3386/w15376

Dziuban, W., Leahy, A., Sengstaken, J., & Whittle, D. (2016) *Tourist Impact in Acadia National Park*. Interactive Qualifying Project, Worcester Polytechnic Institute. https://irma.nps.gov/DataStore/DownloadFile/580799

Edquist, J., Rudin-Brown, C. M., & Lenné, M. G. (2012). The effects of on-street parking and road environment visual complexity on travel speed and reaction time. *Accident Analysis and Prevention*, 45, 759–765. https://doi.org/10.1016/j.aap.2011.10.001

Finnessey, L. (2012). The Negative Effects of Tourism on National Parks in the United States (Doctoral dissertation, Johnson & Wales University) https://scholarsarchive.jwu.edu/ student_scholarship/4/.

Ganesh, L. (2019, January 16). Can everyday traffic affect our mental health? *White Swan Foundation*. https://www.whiteswanfoundation.org/mental-health-matters/understanding- mental-health/can-everyday-traffic-affect-our-mental-health

George, G., Haas, M., & Pentland, A. (2014). From the Editors: Big Data and Management. *The Academy of Management Journal, 57*(2), 321-326. https://doi.org/10.5465/amj.2014.4002

Hallo, J. C., & Manning, R. E. (2008, December 17), Transportation and recreation: a case study of visitors driving for pleasure at Acadia National Park. *Journal of Transport Geography*.

Hammer, G. P., Prel, J.-B. du, & Blettner, M. (2009). Avoiding Bias in Observational Studies. *Deutsches Ärzteblatt International*. https://doi.org/10.3238/arztebl.2009.0664

Ibrahim, H. E. (November 25, 2017). Car Parking Problem in Urban Areas, Causes and Solutions. 1st International Conference on Towards a Better Quality of Life, 2017.

Karaoulanis, A. (2018). Big Data, What Is It, Its Limits and Implications in Contemporary Life. *Business Ethics and Leadership, 2*(4), 108-114. http://doi.org/10.21272/bel.2(4).108- 114.2018.

Kwak-Hefferan, E. (2020, September 1). 8 ways to ease overcrowding at our national parks. *5280*. https://www.5280.com/2020/09/8-ways-to-ease-overcrowding-at-our-national-parks/

Lawson, S., Newman, P., Choi, J., Pettebone, D., & Meldrum, B. (2009). Integrated Transportation and User Capacity Research in Yosemite National Park: The Numbers Game. *Journal of the Transportation Research Board*, *2119*(1), 83–91.

Manning, Robert E. (2009) Parks and People: Managing Outdoor Recreation at Acadia National Park. *University Press of New England*.

McNeeley, S. (2012) Sensitive Issues in Surveys: Reducing Refusals While Increasing Reliability and Quality of Responses to Sensitive Survey Items. *Handbook of Survey Methodology for the Social Sciences*, 377-396. https://doi.org/10.1007/978-1-4614-3876- 2_22

Merrill, N. H., Atkinson, S. F., Mulvaney, K. K., Mazzotta, M. J., & Bousquin, J. (2020). Using data derived from cellular phone locations to estimate visitation to natural areas: An application to water recreation in New England, USA. *PloS One*, 15(4). https://doi.org/ 10.1371/journal.pone.0231863

Plante, J., & Hogan, J. (2019) *Acadia Visitor Study: A Mobile Tracking Application*. Interactive qualifying project, Worcester Polytechnic Institute. https://digital.wpi.edu/concern/ student_works/dr26z071r?locale=en

Simmonds, C., Whaley, M., Wright, M., Canon, G., Wilkinson, T., Maffly, B., Reilly, P., & McGiveny, A. (2018, November 20). Crisis in our national parks: How tourists are loving nature to death. *The Guardian*. https://www.theguardian.com/environment/2018/nov/20/ national-parks-america-overcrowding-crisis-tourism-visitation-solutions

Soehn, D. (2021, February 10). Park Announces 2021 Plan for Cades Cove Vehicle-Free Days. *National Parks Service.* https://www.nps.gov/grsm/learn/news/park-announces-2021- plan-for-cades-cove-vehicle-free-days.htm

Song, I. & Storey, V. C. (2017). Big data technologies and Management: What conceptual modeling can do. *Data & Knowledge Engineering*, 108, 50-677. https://doi.org/10.1016/ j.datak.2017.01.001

StreetLight Data, (2018). StreetLight InSight Metrics: Our Methodology and Data Sources. *StreetLight Data*. https://www.streetlightdata.com/wp-content/uploads/StreetLight-Data_ Methodology-and-Data-Sources_181008.pdf

Trotter, B. (2017, Jan 18). Acadia Smashes Park Visitation Record in 2016. *Bangor Daily News*, https://bangordailynews.com/2017/01/18/news/acadia-smashes-park-visitation-record-in-2016/

Turner, S., Lasley, P., & Pourteau, C. (2013). MONITORING TRAIL USE: Case Study Applications at San Antonio Missions National Historical Park and Guadalupe Mountains National Park. *Texas A&M Transportation Institute*. https://static.tti.tamu.edu/tti.tamu.edu/ documents/401090-1.pdf

Un, K. (2010, February 8), How to do a Parking Study*. Metropolitan Area Planning Council.* https://www.mapc.org/resource-library/how-to-do-a-parking-study/ Updated in April 2019

U.S. Department of the Interior. (n.d.). Acadia National Park (U.S. National Park Service). *National Parks Service.* https://www.nps.gov/acad/index.htm.

U.S. Department of the interior. (n.d.). *IRMA Portal*. *National Park Service*, https://irma.nps.gov/Portal/.

U.S. Department of the Interior. (2015, February 26). Acadia: Park Statistics. *National Parks Service*. https://www.nps.gov/acad/learn/management/statistics.htm.

U.S. Department of the Interior. (2018, January 24). Acadia's Island Explorer Shuttle (U.S. National Park Service). *National Parks Service*. http://www.nps.gov/articles/island-explorer-shuttle.htm.

U.S. Department of the interior. (2019, March). Acadia National Park Final Transportation Plan / Environmental Impact Statement. *National Park Service*.

U.S. Department of the Interior. (2020). Annual Park Ranking Report for Recreation Visits in: 2020. https://irma.nps.gov/STATS/SSRSReports/National%20Reports/Annual%20Park% 20Ranking%20Report%20(1979%20-%20Last%20Calendar%20Year)

U.S. Department of the Interior. (2021a). Congestion Management Program. *National Parks Service*. https://www.nps.gov/orgs/1548/congestion-management-program.htm.

U.S. Department of the Interior. (2021b, June 8). Avoiding Crowds At Peak Season. *National Parks Service.* https://www.nps.gov/acad/traffic.htm

Wade, S. (2021, April 23). New Traffic Safety Measures at Johnson Beach Begin April 23. *National Parks Service*. https://www.nps.gov/guis/learn/news/2021-04-23-pk-traffic-restrictions.htm.

What We Do (U.S. National Park Service). (2020, September 10). National Park Service. https://www.nps.gov/aboutus/index.htm

Zhu, Y., Ye, X., Chen, J., Yan, X., & Wang, T. (2020). Impact of Cruising for Parking on Travel Time of Traffic Flow. *Sustainability*, *12*(8), 3079. https://doi.org/10.3390/su12083079

Appendix A: Previous Research Details

The management at Acadia National Park spent nearly five years working on the 2019 transportation plan. In order to inform their decision making during this period, the park needed information about the traffic going through the Park Loop Road region. This information came from a variety of sources.

Acadia National Park had been collecting park visitor numbers for about a century. On the National Parks IRMA Portal, an information database for the parks, Recreational Visitor Counts have been recorded since 1919 (U.S. Department of the Interior, *IRMA Portal*). However, the earliest published methodology is from 1960. Since the first publication, Acadia National Park has published nine revisions. Each would add more tally techniques, such as the analysis of public transport, or would focus on editing previous methods. Additionally, each would heavily rely on the use of a few automated traffic counters. In addition to collecting data on specific roads, these traffic counters were used to determine the weight of each tally in the final count. This was an effective estimation technique, and the weights of each tally would be fine-tuned over the years. However, this also means older estimates “used a different methodology for estimating annual visitation” (Trotter, 2017) and that more recent data could not be compared to past estimates. On top of this, there are gaps in the information Acadia’s visitor estimation could provide. Much of the observation techniques included the direct observation of visitors and vehicles within the park. With the small number of traffic counters being the only other method of collecting the location of vehicles, Acadia National Park had no data on where vehicles and visitors had been entering the park (Dziuban et al., 2016). This lack of necessary information meant that Acadia management had to focus on collecting more traffic data.

In order to collect more data for their transportation plan, Acadia National Park employed the help of outside project teams. The first project team began work in 2016, and focused on collecting data on where visitors were entering the park. To accomplish this, the team focused on building inexpensive traffic counters and placing traffic monitoring cameras (Dziuban et al., 2016). While this was an effective data collection solution, the traffic counters were only in place for two months due to their high maintenance requirement.

This same high maintenance issue was what led to the failure of an Intelligent Visitor Tracking (ITS) system in 2003 (Calvi et al., 2017). The aim of the system was to track visitor numbers in key areas such as parking lots so that the information could be displayed to visitors. In 2017, Acadia National park tasked a project team to design a new ITS system by building upon this previous attempt. However the final system was never implemented, likely due to its high cost. This led to Acadia management to task a 2019 project team with building their own visitor monitoring system. The result was the deployment of a series of cameras at key areas in the park. These cameras uploaded images to a website on a set schedule. However, in order to track the location of visitors, images would have to be manually analyzed one at a time. Overall, while project teams were able to provide useful data to the park, nearly none were able to provide visitor travel patterns.

There were two interconnected projects that were the exception in addressing this data gap. They were completed individually in 2018 and 2019, and focused on the collection of visitor cell phone location data. The 2018 project was focused on the feasibility of this method (Caltabiano et al., 2018), while the 2019 project developed an app that visitors could download. The developed app would track the location of visitors through the park while the app was open (Plante and Hogan, 2019). This was noted to be an effective methodology, however the app was never able to be made widely available due to app-store publication issues. This meant that while visitor travel patterns were collected, the data-set was far too small to generate the needed information.

Acadia National Park management has put great effort into gathering visitor data so that they can generate solutions to park congestion. While much of this data has proved useful, there has been a major gap: visitor travel patterns. Two project teams tasked with collecting visitor data found that the tracking of cell phone location data could fill this gap in visitor travel data.

Appendix B: Parking Duration from StreetLight

Calculate Statistical Weight of Each Time

Time x Enter = Weighted Enter

Time x Leave = Weighted Leave

Time	Enter	Leave	Weighted Enter	Weighted Leave
6	0.0041	0	0.0246	0
7	0.0041	0.0082	0.0287	0.0574
8	0.0041	0.0082	0.0328	0.0656
9	0.0288	0.0164	0.2592	0.1476
10	0.0905	0.0328	0.905	0.328
11	0.1646	0.0656	1.8106	0.7216
12	0.2058	0.1066	2.4696	1.2792
13	0.0988	0.1311	1.2844	1.7043
14	0.1276	0.1148	1.7864	1.6072
15	0.0988	0.1311	1.482	1.9665
16	0.0741	0.1721	1.1856	2.7536
17	0.0535	0.1148	0.9095	1.9516
18	0.0247	0.0492	0.4446	0.8856
19	0.0123	0.0246	0.2337	0.4674
20	0.0041	0.0164	0.082	0.328
21	0	0	0	0
22	0.0041	0.0082	0.0902	0.1804

Calculate Average Enter and Leave Times

Then Subtract

Find the summation of each column

Then divide by total number of instances used for the distribution

Weighted Enter	Weighted Leave
0.0246	0
0.0287	0.0574
0.0328	0.0656
0.2592	0.1476
0.905	0.328
1.8106	0.7216
2.4696	1.2792
1.2844	1.7043
1.7864	1.6072
1.482	1.9665
1.1856	2.7536
0.9095	1.9516
0.4446	0.8856
0.2337	0.4674
0.082	0.328
0	0
0.0902	0.1804

$$\frac{\Sigma \text{ Weighted Enter/Leave}}{\text{\# of Instances}}$$

Average Enter Time $= \frac{13.0289}{1.00}$ = **<u>1:02 PM</u>**

Average Leave Time $= \frac{14.444}{1.00}$ = **<u>2:27 PM</u>**

(Because it's out of 100%, it sums to 1)

Average Parking Duration

14.444 - 13.0289 = **<u>1 hr. 24 min.</u>**

Appendix C: Parking Duration from Direct Observation

t = Time a Vehicle Enters a Spot T = Time a Vehicle Leaves a Spot

V = Number of Vehicles that Enter and Leave a Spot

$$\frac{(T_1 - t_1) + (T_2 - t_2)}{V} = \text{Average Duration of Stay Between 2 Vehicles } (V = 2)$$

$$\frac{(T_1 - t_1) + (T_2 - t_2)}{V} = \frac{T_1 + T_2 - t_1 - t_2}{V} = \boxed{\frac{\Sigma(T - t)}{V}}$$

4 cars were observed entering
while 4 cars were observed leaving

t	T
8:50	9:10
9:00	9:30
9:50	10:00
9:55	10:15

$$\frac{\Sigma(T - t)}{V} = \frac{(20 + 30 + 10 + 20)}{4} = \frac{80}{4}$$

$$\frac{\Sigma(T - t)}{V} = \frac{(40 + 10 + 10 + 20)}{4} = \frac{80}{4}$$

Average Duration = **20 minutes**

Observation Period:

9:00 to 10:00

3 cars were observed entering

while 3 cars were observed leaving

t	T
	~~9:10~~
9:00	9:30
9:50	10:00
~~9:55~~	

1 car in lot at the Start

$\boxed{S = 1}$

1 car in lot at the End

$\boxed{E = 1}$

$V =$ Number of Vehicles that Enter **and** Leave a Spot

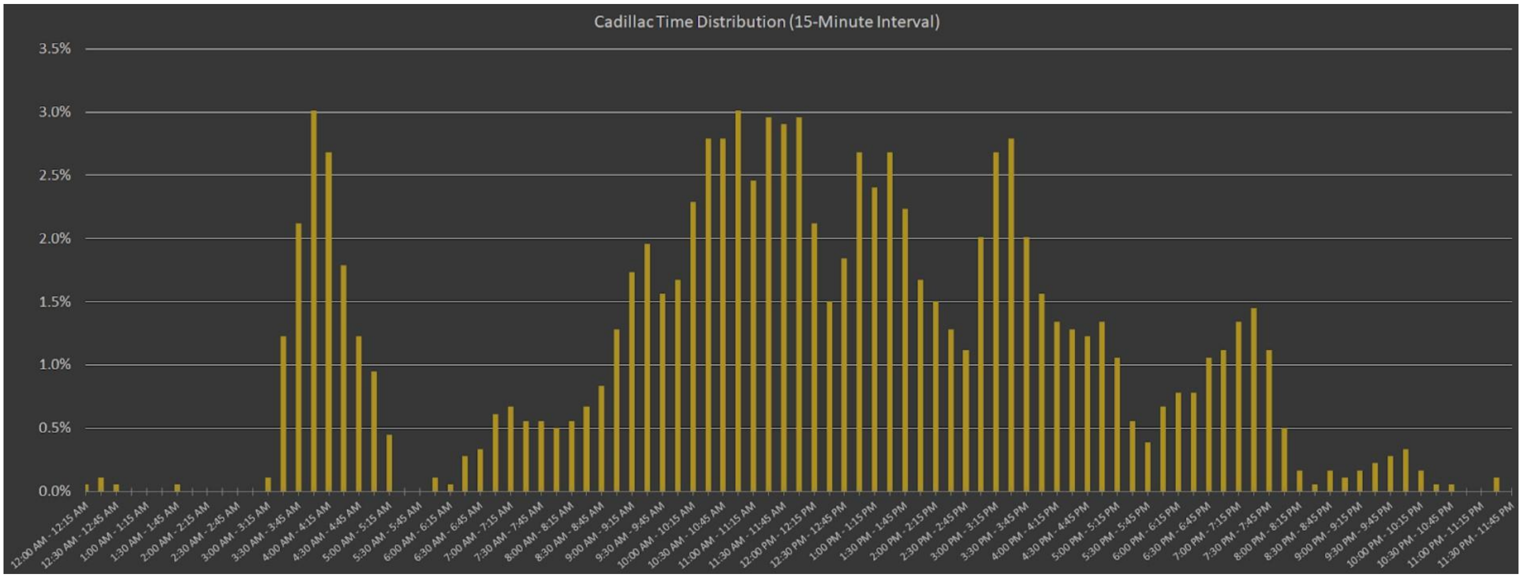
$$\frac{\Sigma(T - t)}{V} = \frac{(30 + 10)}{2} = \frac{20}{2}$$

Average Duration = **20 minutes**

t	T
	~~9:10~~
9:00	9:30
9:50	10:00
~~9:55~~	

Appendix D: Cadillac Mountain Arrival in 15 Minute Intervals

In addition, StreetLight can provide entrance times down to 15-minute intervals, as shown in the bar graph below. The graph shows the times visitors arrive at the top of Cadillac Mountain between June 1st 2019 and November 1st 2019.This allows for more detailed analysis on when visitors are arriving at a certain location.



Appendix E: Home State Distribution

Home State Distribution of all Visitors to Cadillac Mountain (Jun.1-Nov.1, 2019)					
Home State	**Percentage of Visitors From Home State**	**Home State**	**Percentage of Visitors From Home State**	**Home State**	**Percentage of Visitors From Home State**
Alabama	0.72 %	Michigan	2.09 %	Utah	0 %
Alaska	0 %	Minnesota	0.8 %	Vermont	0.42 %
Arizona	0.25 %	Mississippi	0.05 %	Virginia	3.71 %
Arkansas	0.26 %	Missouri	0.25 %	Washington	0.08 %
California	1.75 %	Montana	0 %	West Virginia	0 %
Colorado	0.2 %	Nebraska	0.16 %	Wisconsin	0.38 %
Connecticut	3.22 %	Nevada	0.05 %	Wyoming	0 %
Delaware	1.03 %	New Hampshire	2.26 %		
District of Columbia	0.48 %	New Jersey	6.92 %		
Florida	1.16 %	New Mexico	0.03 %		
Georgia	3.26 %	New York	8.57 %		
Hawaii	0 %	North Carolina	2.9 %		
Idaho	0 %	North Dakota	0.04 %		
Illinois	1.70 %	Ohio	3.53 %		
Indiana	0.95 %	Oklahoma	0.04 %		
Iowa	0.78 %	Oregon	0.01 %		
Kansas	0.24 %	Pennsylvania	7.17 %		
Kentucky	0.68 %	Rhode Island	0.58 %		
Louisiana	0.29 %	South Carolina	1.79 %		
Maine	29.51 %	South Dakota	0.03 %		
Maryland	1.57 %	Tennessee	1.16 %		
Massachusetts	6.76 %	Texas	2.17 %		