

ZIP Code Demographics of In-State Reservation Data for North Carolina Federal Recreation Facilities, 2010 and 2016

Lauren Holmes, Annetta Wilson, Natalie Young
GEOG 574 - Spatial Databases, Spring 2021
University of Wisconsin-Madison



Outer Banks, North Carolina
Source: commons.wikimedia.org

Introduction

The U.S. government provides a centralized reservation system for all federal recreation facilities on Recreation.gov. Users can reserve campsites and lodging, book tickets and tours, secure permits, and purchase activity passes for federal lands, historic sites, museums, and other destinations. There are roughly 3,600 sites & facilities and over 103,000 individual reservable sites across the country, offering experiences such as wildlife viewing, rafting, RVing, horseback riding, boating, and hunting & fishing. Recreation.gov has 12 federal participating partners; among them are the Bureau of Land Management, Bureau of Reclamation, National Park Service, U.S. Forest Service, U.S. Army Corps of Engineers, and the Smithsonian Institution.ⁱ In accordance with the Open Data Policy, the federal government provides recreation reservations data on the Recreation Information Database (RIDB) website - ridb.recreation.gov. From there, users can “access RIDB API endpoints which contain information for federal recreation areas, facilities, campsites, tours, and permits.”

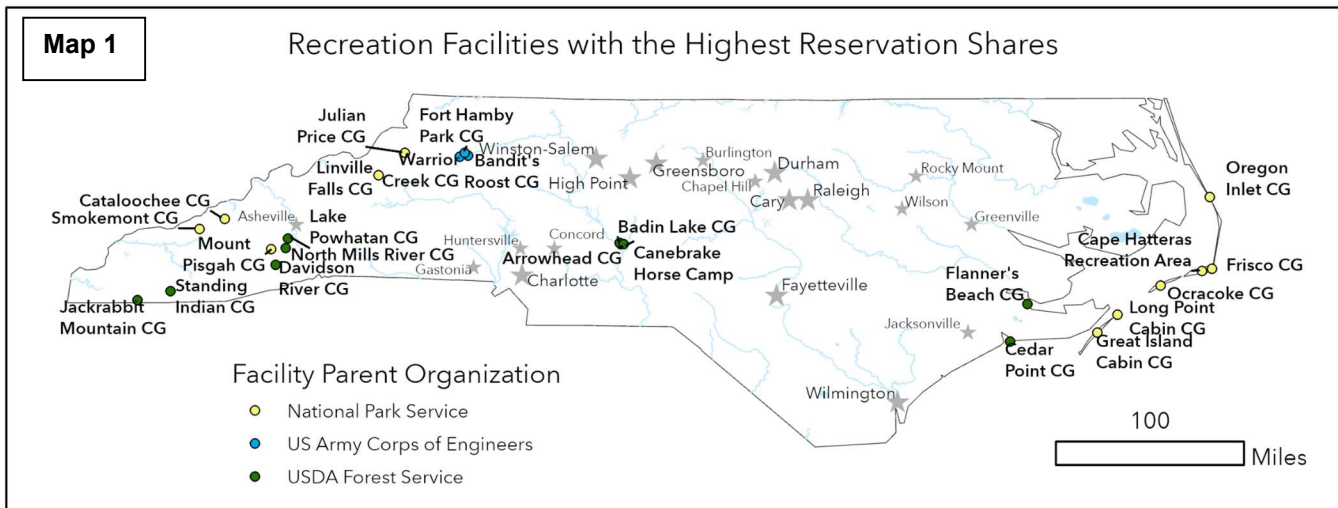
After reading an NPR article about the [National Park Service's diversity problem](#), we decided to explore specific demographics (income, age, and racial-ethnic minority status) of a subset of recreational reservations to identify trends within two years of reservation data - 2010 and 2016 - and differences between these two years.ⁱⁱ However, since the only identifying information given about customers in the reservation data was their ZIP code, we decided to look at the aggregated demographic characteristics of ZIP codes that occurred in the reservation data. We limited our analysis to reservations at North Carolina recreation facilities by in-state ZIP codes. Our findings help describe how well in-state usage of federal recreation facilities reflects and represents North Carolina's statewide demographics.

Data Sources

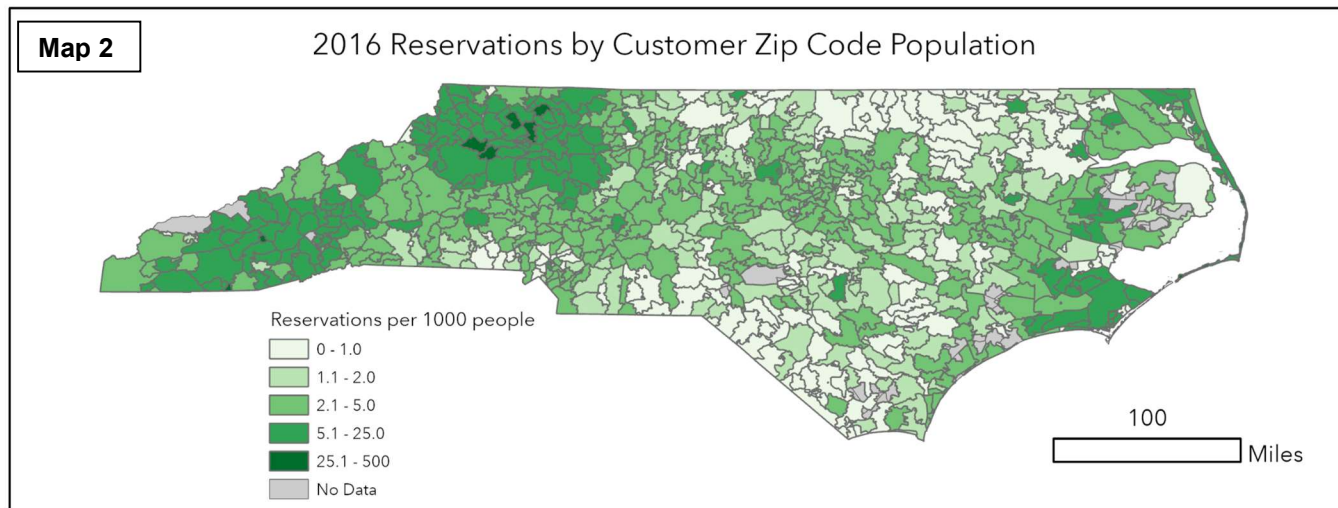
Our first data source is the Recreation Information Database (RIDB). The RIDB is “a part of the Recreation One Stop (R1S) program, which oversees the operation of Recreation.gov -- a user-friendly, web-based resource to citizens, offering a single point of access to information about recreational opportunities nationwide.”ⁱⁱⁱ The site offers two main datasets, a recreation database containing data on facilities, management, activities, tours, etc., and historical reservation data for 2006 through 2020. We downloaded the recreation database in CSV format. We also downloaded two historical reservation data files in compressed CSV format - one for Fiscal Year 2010 and another for Fiscal Year 2016. Our second source of data is the U.S. Census Bureau, where we downloaded TIGER/Line ZIP Code Tabulation (ZCTA) geodatabases that include multiple tables of demographic data from American Community Survey (ACS) 5-year estimates.^{iv} We downloaded the 2008 - 2012 ZCTA ACS geodatabase to use with the FY 2010 RIDB historical reservation data, and downloaded the 2014-2018 ZCTA ACS geodatabase to use with the FY 2016 RIDB historical reservation data.

Each historical reservation CSV file contains the entirety of individual reservations made at recreation facilities for that year. Each row of reservation data includes (among other data): a unique reservation id; the name, unique id, and address of the recreation facility the reservation is associated with; the federal agency that the facility is housed under; the type of reservation; and customer location data, most importantly the customer's ZIP code. The most common types of reservations in the data were for campgrounds, but other reservations could be for tour

tickets or other facilities such as visitor centers, resorts/hotels, museums, or other miscellaneous recreation areas. North Carolina's facilities belong to only three federal agencies: the National Park Service (NPS), U.S. Army Corps of Engineers (USACE), and U.S. Forest Service (USFS). Maps 1-3 below show the North Carolina recreation facilities with the highest shares of reservations and reservations per 1000 people from 2010 and 2016 for North Carolina ZIP codes. ZIP code geographies are represented using the Census Bureau's ZCTA boundaries.

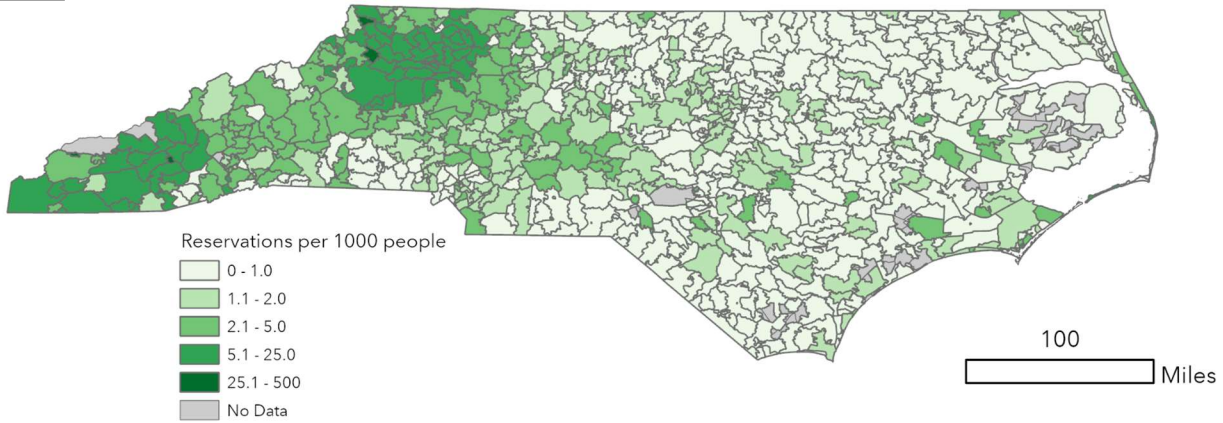


Note: CG = campground. Map 1 displays the 24 facilities with the highest shares of total FY 2016 recreation reservations. Each facility accounts for at least 1% of all reservations for the year. Together, these facilities cover 90.7% of all reservations for 2016. The top five areas with the most reservations were Bandit's Roost campground (USACE), Cape Hatteras National Seashore ORV Permits (NPS), Warrior Creek campground (USACE), Davidson River campground (USFS), and Julian Price campground (NPS).



Map 3

2010 Reservations by Customer Zip Code Population



Model Design

Since the scope of our project involves comparing two different years of data, each entity listed below occurs twice in our database, once for each year of reservation data. For purposes of showing our database design in the simplest terms, we list each entity once in our ER Diagram.

Our database has six entities: a ZIP Code Tabulation Area (ZCTA) Polygons entity (called “ACS Geography Polygon” in ER Diagram), a reservation data entity (called “Reservations table” in ER Diagram), an entity called “ZIP Code Reservation Count Table,” and three entities containing American Community Survey (ACS) demographic information for each ZCTA: one for Age data (“ACS Age table”), one for Income data, and one for Race/Hispanic or Latino Origin data (“ACS Race/Hispanic Origin”).

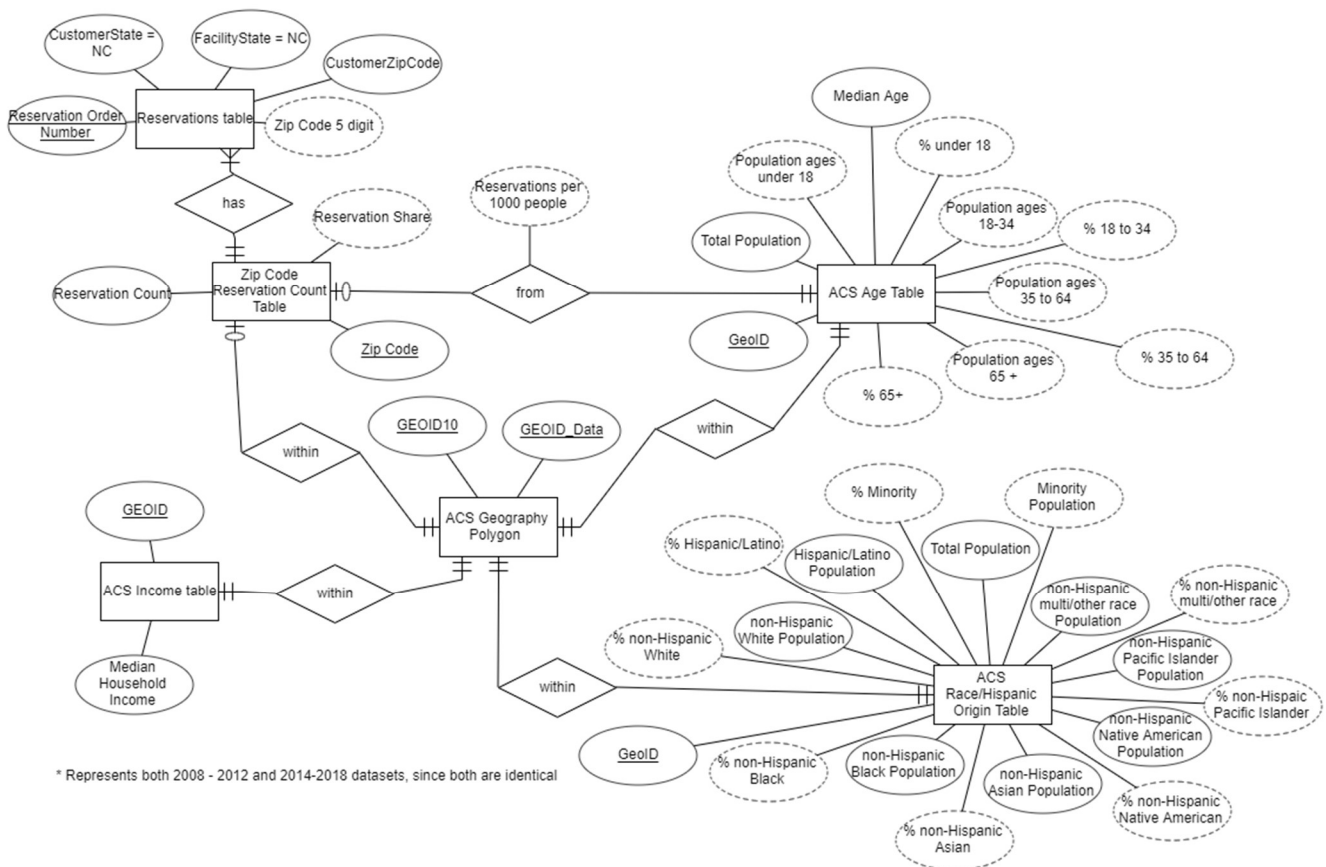
The ZCTA polygon entity has two crucial attributes, the five-digit ZIP code representing the polygon called “GEOID10” and the ZIP code’s census data GEOID (called “GEOID_Data”) that allows the polygons to join to the demographic data tables. Both attributes serve as primary keys and also as foreign keys to point to data in other entities. The reservations data entity has five attributes: (1) reservation order number, which serves as the primary key; (2) the customer’s state of residence, limited to NC; (3) the state the reserved facility is located in, also limited to NC; (4) the customer’s ZIP code (ten digits); and (5) the customer’s ZIP code in five-digit format, which is derived from the previous attribute to only include the first five digits, and serves as a foreign key to connect to other entities. The “ZIP Code Reservation Count Table” entity has three attributes: (1) ZIP code (five-digit) that serves as both the primary key and as the foreign key to the ZCTA Polygon entity and the reservations data entity; (2) the aggregated count of reservations for each ZIP code; and (3) the percentage of total reservations for each ZIP code, which is derived from the aggregated count of reservations divided by the total number of reservations for all ZIP codes.

Each of the ACS Demographic data table entities includes a primary key & foreign key attribute, “GeoID” that is used to join to the ZCTA polygons entity. The Income table only includes one other attribute: Median Household Income. The ACS Age Table entity includes ten other attributes: Total Population, Median Age, Population Under

Age 18, Population Age 18-34, Population Age 35-64, Population Age 65 and older, and then the share of each age group (Population of age group / Total population). The ACS Race/Hispanic or Latino Origin Table includes 18 other attributes: Total Population, and Population totals for several groups: Hispanic/Latino, non-Hispanic White, non-Hispanic Multi/Other Race, non-Hispanic Native Hawaiian or Other Pacific Islander, non-Hispanic Native American or Alaska Native, non-Hispanic Asian, non-Hispanic Black. It also includes shares of each race/ethnicity group (Population of race/ethnicity group / Total Population).

There are relationships between each of the ACS demographic data table entities and the ZCTA Polygon entity - they are all mandatory with one-to-one cardinality on both ends. There is an optional one-to-one relationship between the ZCTA Polygon entity and the ZIP Code Reservation Count Table entity, as there may not be any reservations for a particular ZIP code. Then between the ZIP Code Reservation Count Table entity and the reservations data entity, there is a mandatory one-to-many relationship as there are one or more reservations in the reservations data entity per each ZIP code listed in the ZIP Code Reservation Count Table entity. Finally, between the ZIP Code Reservation Count Table entity and the ACS Demographic Age Table, there is an optional one-to-one relationship as each GeoID (representing the ZIP code) in the ACS Age Table may not have any reservations associated with it. This particular relationship is special as it was used to calculate the number of reservations per 1,000 people for each ZIP code.

Figure 1: Entity Relationship (ER) Diagram



Database Implementation

ArcGIS Pro was used for the database implementation. Our implementation steps and notes are as follows:

1. First, an ESRI file geodatabase was created to contain our data from both timeframes and named recreation_demographics.gdb.
2. The ACS and Reservations data downloads included nationwide data, so we constrained them to North Carolina for speed and ease of data manipulation. The ACS Geography polygons and ZCTA tables didn't contain a state attribute, so a Select By Location was needed to extract the ZIP code polygons within North Carolina. The North Carolina state boundary polygon was sourced from the US Census TIGER database.^v For the Select by Location step, "Have their Center in" the North Carolina polygon was chosen as the relationship, because it selected all the ZCTA polygons that were within the North Carolina polygon without excluding any at the borders or including any from neighboring states that shared a border.
3. The selected features were exported to create new feature classes called NC_2014_2018_ZCTA_Geography and NC_2008_2012_ZCTA_Geography and saved to recreation_demographics.gdb.
4. The following ACS ZCTA demographic data tables were then joined to the associated Geography polygons:
 - X01_AGE_AND_SEX (contains total population and median age by ZIP code)
 - X03_HISPANIC_OR_LATINO_ORIGIN (includes total population and race population by ZIP code)
 - X19_INCOME

The tables were joined to their associative polygons and saved in order to map. An inner join was used, to constrain the demographic data to North Carolina ZIP codes:

- `SELECT * FROM ZTCA_Geography AS Z INNER JOIN X01_AGE_AND_SEX AS X ON Z.GEOID_Data = X.GEOID`
 - `SELECT * FROM ZTCA_Geography AS Z INNER JOIN X03_HISPANIC_OR_LATINO_ORIGIN AS X ON Z.GEOID_Data = X.GEOID`
 - `SELECT * FROM ZTCA_Geography AS Z INNER JOIN X19_INCOME AS X ON Z.GEOID_Data = X.GEOID`
5. The triple- joined feature classes were then exported and saved as NC_2014_2018_ZCTA_ACS and NC_2008_2012_ZCTA_ACS to recreation_demographics.gdb.
 6. The 2010 and 2016 Historic Reservations .csv tables were added to ArcGIS Pro and then exported to GIS table format and saved to recreation_demographics.gdb. In the export window, the following SQL was used to select the data for each year to set data constraints for facilities and customers within North Carolina:
 - `SELECT * FROM ReservationData WHERE FacilityState = 'NC' AND CustomerState = 'NC'`
 7. A visual data quality check revealed that some of the CustomerZip entries were 10-digit ZIP codes. To ensure uniform data, a text field was added to contain the first 5 digits of the ZIP and make it compatible

with the ZCTA Join key field, GEOID10. The following Arcade expression was used to add the left 5 digits to the text field:

- `Left($feature.CustomerZIP,5)`

8. The new ZIP_Code fields were then summarized to get a reservation count for each ZIP code and create a key constraint from the ZIP_Code field to join to the demographic data:

- `SELECT ZIP_CODE, COUNT(ZIP_CODE) FROM NC_Reservations`

These tables were saved as NC_2010_ReservationCount and NC_2016_ReservationCount to recreation_demographics.gdb

9. The Reservation Share attribute was added to the above tables and calculated:

- `ALTER TABLE NC__ReservationCount AS Data
ADD ReservationShr Double
UPDATE Data SET ReservationShr = COUNT_ZIP_Code/SUM(COUNT_ZIP_Code)*100`

10. To normalize the reservation data by ZIP code population and to map it, the NC_ReservationCount tables were joined to their respective polygon feature classes. An outer join was used because some of the ZIP codes don't have reservations associated with them:

- `SELECT * FROM NC_ZTCA_ACS AS Z
FULL OUTER JOIN NC_ReservationCount AS R
ON Z.GEOID10 = R.ZIP_Code`

These feature classes were then exported and saved to the geodatabase as

NC_2014_2018_ZCTA_ACS_ReservationData and NC_2008_2012_ZCTA_ACS_ReservationData

11. Another data constraint is no NULL values. The outer join of the ReservationCount data left NULL values in the ZIP code features that didn't contain reservations. The NULL entries were selected and '0' was added:

- `SELECT * FROM NC_ZCTA_ACS_ReservationData AS Data
WHERE COUNT_ZIP_Code IS NULL AND ReservationShr IS NULL
UPDATE Data SET COUNT_ZIP_Code = 0
UPDATE Data SET ReservationShr = 0`

A quality check was also done on the GEOID entries to look for NULL values from potential incorrect ZIP code entries:

- `SELECT * FROM NC_ZCTA_ACS_ReservationData
WHERE GEOID IS NULL`

Both selections on the associated feature classes didn't yield any results, most likely because the table entries were not associated with spatial data and were left out by the join.

Results, Discussion, and Database Manipulations

Demographic Analysis Background and Limitations

A very important disclaimer to our demographic analysis is that recreation.gov reservations collect an individual's ZIP code when a reservation is made, but not the individual's demographic traits like race, age, income, etc.

Without data representing the actual individuals making reservations, we analyze data for the home ZIP codes of the individuals. Though we assume that this offers some representative information when large amounts of reservations are considered, we recognize that this practice may not directly represent the individuals. For example, a reservation could be made by a 60-year-old person who lives in a ZIP code where the median age is only 27 years. However, the analysis of large numbers of reservations and their source ZIP codes still offers meaningful information and suggests trends worthy of attention.

Our analysis emphasizes distributions of how selected demographic traits are represented in reservation data and in North Carolina's ZIP code areas in general. We create low, medium, and high categories for each trait and compare the shares of all reservations in each category to the shares of all statewide ZIP codes in each category. This means that exact demographic values for age, race, and income are less important to this analysis than the comparison of shares in each category.

The tables in this section are labeled with year 2010 or year 2016. This analysis associates 2010 reservations with 2008-2012 ACS 5-year demographic data. 2016 reservation data are associated with 2014-2018 ACS 5-year demographics. North Carolina ZIP code demographics are also described using the 2008-2012 and 2014-2018 ACS. The ACS data represent the survey responses gathered anywhere within the 5-year period and should be treated as representing the entire period. We note that the 2008-2012 data includes and is influenced by the Great Recession. There are several factors that can play a role in differences across time periods. This analysis describes demographic differences, but doesn't seek to explain them. Additionally, we recognize that as we are not stewards of the recreation reservations data and have only done preliminary analysis of it, there may be data integrity issues or other factors we are unaware of.

Due to space limitations in this report, we were not able to analyze all the demographic variables we had calculated and prepared in our database. Instead we focused on one variable for each of three categories: racial and ethnic minorities, median age, and median household income. Full definitions of these variables can be found in online ACS resources.^{vi}

Racial and Ethnic Minorities

In this analysis, the term "minority" includes all individuals who identify as a race other than white, as multiracial, or as being Hispanic or Latino. Put differently, "minority" includes anyone who does not identify as being non-Hispanic white alone ("alone" indicates that someone selected only one race category). We calculated minority totals and percentages of the population using the race and ethnicity categories used in American Community Survey data. Hispanic or Latino is the only ethnicity defined in the data. Race categories include White, Black, American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, Other, or combinations of these.

Table 1: Demographic Comparison of Recreation Reservation ZIP Codes and All NC ZIP Codes: Minority Population, 2010

		Fiscal Year 2010 Reservations				
		NC Reservations (Observed) Demographics		NC ZIP Code Area (Expected) Demographics		
ZIP Code Area Minority Share of Total Population	Category	Reservations	Share of Total Reservations	Count of NC ZIP Code Areas	Share of Total NC ZIP Code Areas	Difference in Shares
<=22%	Low	9,097	61.6%	376	46.7%	14.8pp
22.1-50%	Medium	4,542	30.7%	280	34.8%	-4.0pp
>50%	High	1,140	7.7%	149	18.5%	-10.8pp
	Sum	14,779	100.0%	805	100.0%	—

To analyze the minority share of the total population, we created three general groups: Low minority share (0-22.0% minority), medium minority share (22.1-50.0% minority), and high minority share (>50.0% minority). The “high” group indicates ZIP code areas where “minorities” account for over half of the population and thus actually make up the majority population of the area.

2010 recreation reservation data show that 61.6% of reservations come from areas with “low” minority populations, 20.7% from “medium” minority areas, and 7.7% from high-minority areas (Table 1). In comparison, North Carolina ZIP codes in general are 46.7% “low”, 34.8% “medium”, and 18.5% “high” minority areas. Comparing these two distributions, the largest difference was that 2010 recreation reservations were much more likely to be from low-minority areas (61.6% vs. 46.7%, a 14.8 percentage point difference). They were less likely to be from medium and high-minority areas (differences of 4.0 and 10.8 percentage points, respectively). Stated in other terms, our analysis shows that 2010 recreation reservations were much more likely to be from ZIP codes with higher shares of white, non-Hispanic or Latino residents than would be expected if reservation data reflected state demographics.

Table 2 shows the results of our analysis of 2016 reservations. As with 2010, we found that 2016 reservations were more likely to come from ZIP codes with a lower minority shares. However, a potentially positive sign is that the gap in these distributions compared to statewide ZIP codes was smaller: about 50% of reservations came from low-minority areas, compared to about 45% of NC ZIPs being classified as low-minority. The resulting difference was less than five percentage points, a much smaller difference than the 14.8pp difference seen in 2010. In contrast to 2010, 2016 reservations were more likely, rather than less likely, to come from medium-minority ZIP codes. However, 2016 reservations remained less likely to be from high-minority areas than would be expected from statewide ZIP code demographics.

Table 2: Demographic Comparison of Recreation Reservation ZIP Codes and All NC ZIP Codes: Minority Population, 2016

		Fiscal Year 2016 Reservations				
		NC Reservations (Observed) Demographics		NC ZIP Code Area (Expected) Demographics		
ZIP Code Area Minority Share of Total Population	Category	Reservations	Share of Total Reservations	Count of NC ZIP Code Areas	Share of Total NC ZIP Code Areas	Difference in Shares
<=22%	Low	15,355	49.5%	320	44.8%	4.7pp
22.1-50%	Medium	12,547	40.4%	271	37.9%	2.55pp
>50%	High	3,148	10.1%	124	17.3%	-7.2pp
	Sum	31,050	100.0%	715	100.0%	—

Table 1 and 2 Notes: “pp” = Percentage Points. Reservation sums only include reservations from ZIP codes that matched a North Carolina ZIP code area. Sums of reservations or NC ZIP code counts exclude ZIP codes areas with zero population or with no demographic data for the variable in question.

Median Age

We divided median age into three groups: low median age (less than or equal to 36.3 years), medium median age (36.4 through 54.9 years), and high median age (55 and older). 2010 recreation reservation data show that 25.1% of reservations come from areas with a low median age, 73.7% from areas with a medium median age, and just 1.2% from areas with a high median age (Table 3). In comparison, North Carolina ZIP codes are fairly similar overall, at 24.2% with a low median age, 71.5% with a medium median age, and 4.4% with a high median age. Our analysis reflects that 2010 recreation reservations are a little more likely to come from low and medium median age areas. However, they were less likely to come from areas with a high median age (1.2% vs. 4.4%, a 3.2 percentage point difference).

Table 4 shows the results of our analysis of 2016 reservations. There was little difference in the share of reservations from low median age areas compared to statewide ZIP codes (0.3pp). In other words, reservation data represented low median age areas well. As with 2010 data, 2016 reservations were more likely to come from medium median age areas, and less likely to come from high median age areas. Though the share of all reservations from high median age areas increased, the overall number of ZIP codes with high median age increased as well. The comparative underrepresentation of areas with high median ages seemed to worsen, moving from a 3.2 to a 4.3 percentage point difference.

Table 3: Demographic Comparison of Recreation Reservation ZIP Codes and All NC ZIP Codes: Median Age, 2010

		Fiscal Year 2010 Reservations				
		NC Reservations (Observed) Demographics		NC ZIP Code Area (Expected) Demographics		
ZIP Code Area Median Age	Category	Reservations	Share of Total Reservations	Count of NC ZIP Code Areas	Share of Total NC ZIP Code Areas	Difference in Shares
<=36.3	Low	3,711	25.1%	194	24.2%	1.0pp
36.4-54.9	Medium	10,893	73.7%	574	71.5%	2.2pp
>=55.0	High	174	1.2%	35	4.4%	-3.2pp
	Sum	14,778	100.0%	803	100.0%	—

Table 4: Demographic Comparison of Recreation Reservation ZIP Codes and All NC ZIP Codes: Median Age, 2016

		Fiscal Year 2016 Reservations				
		NC Reservations (Observed) Demographics		NC ZIP Code Area (Expected) Demographics		
ZIP Code Area Median Age	Category	Reservations	Share of Total Reservations	Count of NC ZIP Code Areas	Share of Total NC ZIP Code Areas	Difference in Shares
<=36.3	Low	5,798	18.7%	152	19.0%	-0.3pp
36.4-54.9	Medium	24,486	78.9%	594	74.3%	4.6pp
>=55.0	High	763	2.5%	54	6.8%	-4.3pp
	Sum	31,047	100.0%	800	100.0%	—

Table 3 and 4 Notes: “pp” = Percentage Points. Reservation sums only include reservations from ZIP codes that matched a North Carolina ZIP code area. Sums of reservations or NC ZIP code counts exclude ZIP codes areas with zero population or with no demographic data for the variable in question.

Median Household Income

ACS median household income data represent median household income in the past 12 months. We divided income data into three groups: low (earned less than or equal to \$50,000), medium (over \$50,000 and up to \$75,000), and high (over \$75,000). These categories are used to create relative groupings for our analysis. The low, medium, and high categories shouldn’t be interpreted as representing strong definitions of “low income” or “high income” areas.

The distributions of 2010 recreation reservations incomes and North Carolina ZIP code area incomes are shown in Table 5. Our analysis reflects that 2010 recreation reservations were less likely to come from the lower and medium income groupings than would be expected based on state ZIP code demographics (by 3.1 and 0.9 percentage points, respectively). Reservations were more likely to come from the highest income group than would be expected from state demographics: 8.8% of reservations are from this group, although only 4.9% of state ZIP codes are in that group. Though lower and middle income areas make up the majority of overall reservations and North Carolina ZIP codes, higher income groups still appear to be overrepresented among reservation data.

Table 5: Demographic Comparison of Recreation Reservation ZIP Codes and All NC ZIP Codes: Median Household Income, 2010

		Fiscal Year 2010 Reservations				
		NC Reservations (Observed) Demographics		NC ZIP Code Area (Expected) Demographics		
ZIP Code Area Median Household Income	Category	Reservations	Share of Total Reservations	Count of NC ZIP Code Areas	Share of Total NC ZIP Code Areas	Difference in Shares
<= \$50,000	Low	10,375	70.3%	580	73.3%	-3.1pp
\$50,001-75,000	Medium	3,085	20.9%	172	21.7%	-0.9pp
\$>75,000	High	1,306	8.8%	39	4.9%	3.9pp
	Sum	14,766	100.0%	791	100.0%	—

Table 6 shows the results of our analysis of 2016 reservations. As with 2010 data, 2016 reservations were less likely to come from the lowest income grouping, and more likely to come from the highest income grouping. These differences are even larger than those observed in 2010 reservation data. While inflation-adjusting the data might adjust the difference in shares somewhat, the differences are large enough that we are confident that they still indicate an actual difference in representation of these groups. Areas in the medium income group were somewhat underrepresented in reservation data in 2010, but now appear slightly overrepresented (36.5% of reservations, vs. 34.7% of state ZIP codes in the same category).

Table 6: Demographic Comparison of Recreation Reservation ZIP Codes and All NC ZIP Codes: Median Household Income, 2016

		Fiscal Year 2016 Reservations				
		NC Reservations (Observed) Demographics		NC ZIP Code Area (Expected) Demographics		
ZIP Code Area Median Household Income	Category	Reservations	Share of Total Reservations	Count of NC ZIP Code Areas	Share of Total NC ZIP Code Areas	Difference in Shares
<= \$50,000	Low	14,837	48.0%	392	56.5%	-8.5pp
\$50,001-75,000	Medium	11,275	36.5%	241	34.7%	1.8pp
>\$75,000	High	4,798	15.5%	61	8.8%	6.7pp
	Sum	30,910	100.0%	694	100.0%	—

Table 5 and 6 Notes: “pp” = Percentage Points. Reservation sums only include reservations from ZIP codes that matched a North Carolina ZIP code area. Sums of reservations or NC ZIP code counts exclude ZIP codes areas with zero population or with no demographic data for the variable in question.

Database Manipulations: Queries and Analysis Details

Queries for 2008-2012 demographic data (associated with 2010 reservations) are shown in this section. 2014-2018 queries (associated with 2016 reservations) for both reservation count data and of NC ZIP code area data are not shown. The queries are identical to the 2008-2012, except with data years updated as necessary: 1. They reference the 2014-2018 NC ZCTA ACS Reservation Data layer and 2. They use the columns minority_2018shr, med_age_2018, and med_hh_inc_2018.

Racial and Ethnic Minorities:

Queries for Selecting Reservation Totals:

2008-2012 Queries: The “COUNT_ZIP_Code” column contains the count of 2010 reservations for the ZIP code. We used the “where” clauses shown in the Select by Attribute tool to limit rows, then viewed statistics for the frequency column of the resulting selection. However, those steps could also be expressed in full queries as follows:

- SELECT SUM(COUNT_ZIP_Code) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE minority_2012shr <= 22
- SELECT SUM(COUNT_ZIP_Code) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE minority_2012shr > 22 AND minority_2012shr <= 50
- SELECT SUM(COUNT_ZIP_Code) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE minority_2012shr > 50

Queries for Selecting Counts of NC ZIP Code Areas:

2008-2012 Queries:

- SELECT COUNT(*) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE minority_2012shr <= 22
- SELECT COUNT(*) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE minority_2012shr > 22 AND minority_2012shr <= 50
- SELECT COUNT(*) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE minority_2012shr > 50

Median Age:

Queries for Selecting Reservation Totals:

2008-2012 Queries: The "COUNT_ZIP_Code" column contains the count of 2010 reservations for the ZIP code. We used the "where" clauses shown in the Select by Attribute tool to limit rows, then viewed statistics for the frequency column of the resulting selection. However, those steps could also be expressed in full queries as follows:

- SELECT SUM(COUNT_ZIP_Code) FROM NC_2008_2012_ZCTA_ACS_ReservationData demog WHERE med_age_2012 <= 36.3
- SELECT SUM(COUNT_ZIP_Code) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE med_age_2012 > 36.3 AND med_age_2012 < 55
- SELECT SUM(COUNT_ZIP_Code) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE med_age_2012 >= 55

Queries for Selecting Counts of NC ZIP Code Areas:

2008-2012 Queries:

- SELECT COUNT(*) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE med_age_2012 <= 36.3
- SELECT COUNT(*) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE med_age_2012 > 36.3 AND med_age_2012 < 55
- SELECT COUNT(*) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE med_age_2012 >= 55

Median Household Income:

Queries for Selecting Reservation Totals:

2008-2012 Queries: The "COUNT_ZIP_Code" column contains the count of 2010 reservations for the ZIP code. We used the "where" clauses shown in the Select by Attribute tool to limit rows, then viewed statistics for the frequency column of the resulting selection. However, those steps could also be expressed in full queries as follows:

- SELECT SUM(COUNT_ZIP_Code) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE med_hh_inc_2012 <= 50000

- SELECT SUM(COUNT_ZIP_Code) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE med_hh_inc_2012 > 50000 AND med_hh_inc_2012 <= 75000
- SELECT SUM(COUNT_ZIP_Code) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE med_hh_inc_2012 > 75000

Queries for Selecting Counts of NC ZIP Code Areas:

2008-2012 Queries:

- SELECT COUNT(*) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE med_hh_inc_2012 <= 50000
- SELECT COUNT(*) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE med_hh_inc_2012 > 50000 AND med_hh_inc_2012 <= 75000
- SELECT COUNT(*) FROM NC_2008_2012_ZCTA_ACS_ReservationData WHERE med_hh_inc_2012 > 75000

Opportunities for Further Analysis

For our project, we analyzed a limited set of demographic variables from a small selection of the tables provided in the American Community Survey (ACS) geodatabase. Demographic analysis of the reservation data could be expanded greatly by investigating more of the variables we created for our database, particularly detailed age and race groups, or by testing for statistical significance in differences in reservation data vs. state demographics and across time periods. There are also opportunities to analyze hundreds of variables from other data tables, such as Educational Attainment, Poverty, Disability, or Industry Occupation.

Additionally, the Recreation Information Database (RIDB) includes recreation area, facility, and site level that did not fit into the scope of our demographics analysis but could provide several other angles and variables for analyzing the recreation reservation data. Most of this data includes a "FacilityID" column that could be used as a foreign key to join to reservation data. For example, a "Campsites" table is available and includes fields such as "CampsiteType," "TypeOfUse," and "CampsiteAccessible" (whether the campsite is accessible by car). Other options for analysis include studying North Carolina's out-of-state reservations and their demographics, the reservations for a broader region, or a higher number of years to better substantiate trends.

Conclusion

Our database creation and analysis led to preliminary findings about demographic characteristics in-state federal recreation reservations in North Carolina for 2010 and 2016 and how they compare to demographics across the state. We found that reservations underrepresented high-minority areas and overrepresented low-minority areas in both 2010 and 2016 compared to statewide demographics. We expected this finding based on background information related to the National Park Service. However, the 2016 reservations seemed to move toward a more representative balance. Though there are many factors we have not considered, this may be a sign of improving racial and ethnic representation in North Carolina's in-state federal recreation. Our analysis also found that areas with a middling median age were overrepresented in recreation data in 2010 and 2016, and higher median age areas were underrepresented (vs. state age distributions). Finally, we found that higher-income areas were overrepresented and lower-income areas were underrepresented in recreation reservations for both years. The

differences intensified from 2010 to 2016, suggesting that recreation use across income categories may need improvement in order to better represent the residents of North Carolina. Our research findings offer interesting considerations and raise many topics for further understanding of federal recreation usage in North Carolina and beyond.

References

ⁱ Full list of federal recreation.gov partners: <https://www.recreation.gov/about-us>

ⁱⁱ Full link: <https://www.npr.org/2016/03/09/463851006/dont-care-about-national-parks-the-park-service-needs-you-to>

ⁱⁱⁱ <https://ridb.recreation.gov/>

^{iv} TIGER/Line with Selected Demographic and Economic Data: <https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-data.html>. We downloaded 2018 and 2012 5-year geodatabases for ZCTA geographies.

^v North Carolina boundary source: the National Nation-Level Geography Geodatabase State feature class: <https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-geodatabase-file.html>.

^{vi} ACS Code Lists, Definitions, and Accuracy: <https://www.census.gov/programs-surveys/acs/technical-documentation/code-lists.html>