**Karen Avila**

**December 4, 2024**

**PPD 631**

# The Digital Divide in Los Angeles County:

# A Limited Snapshot of Compton, Downey, and Beverly Hills

## Defining the Question

My project aimed to analyze if there could be a digital divide in Los Angeles County, California. I wanted to explore if Los Angeles County households have similar internet access or if they have the adequate technology necessary to access web-based services. This paper delves into the technical details behind this project and concludes that there is a digital divide within Los Angeles County, as noted in three of its cities. This conclusion is limited, and further research is warranted to support the validity of this hypothesis. Lastly, please note that the person who read my first draft was my brother, Jesus Avila.

## Digital Divide Background

The term “digital divide” can have a different meaning to various audiences. Van Dijk and Hacker noted that it is usually defined as a gap between those who do and do not have access to forms of communication technology (Van Dijk & Hacker, 2003). Such examples can include access to the internet, broadband speed, internet exchange points, hotspots, and technological devices such as smartphones and computers. This project looked closely at internet exchange points, internet access, and specific households with no computers. Earlier this year, Forbes reported that roughly 94.6% of Americans have access to the internet (Pelchen, 2024). I wondered how this company calculated this number as some people may have access to the internet when using public resources (such as going to a school, coffee shops with wifi, or

libraries). However, they might not necessarily have reliable access to the internet. In 2021, the Affordable Connectivity Program was signed into law, helping more people access the internet for free or at a heavily discounted price. Unfortunately, this program expired in April 2024 and Congress has yet to agree on re-funding the program (Vives & Castillo, 2024). Over 23 million Americans used the program (Johnson, 2024), including in Los Angeles County. Although the data used for this project is not necessarily up to date, I assume that the digital divide will worsen if resources are not adequately distributed to communities in need.

**Using GIS to Approach the Question**

GIS is an appropriate tool for approaching this question because it can reveal trends in populated areas that may be underserved. Using geographical data and census information can help make informed decisions or recommendations on where to increase equity through funding or expansion of internet-based services. People can use different data sources (from Excel files to shapefiles to CSVs) to update and tinker with data to shed more focus on one issue. Overall, it can help us understand the digital divide in this region as society becomes more dependent on technology and internet usage. My project focused on demographics like race, age, median household income, and data points like internet and computer access. GIS is a visually appealing way to analyze if any patterns can shed light on resources needed in Los Angeles County. In addition, it includes various online resources for creating maps and different export features that help analyze data that is not easily accessible on other software such as Microsoft Excel, Tableau, or Canva.

At my current job, I have participated in a few PowerBI training sessions, and many of my colleagues were impressed that the software allows the creation of maps through zip codes. However, the trainer warned that too many zip codes could cause the software to crash, which I

thought was ironic because working in government services includes numerous zip codes. I thought about how GIS can analyze countless pieces of data in a shorter time frame and with more features, such as combining data into different layers and gathering information from various existing databases, unlike PowerBI, where I was informed that using more than three different data sources can slow down the software. It may be due to a virtual desktop infrastructure (VDI) issue but it still made me skeptical. Although I experienced some technical issues when creating this map on GIS, I found other ways to combine the data as necessary.

**Data Used: What, Which, Where, and Why**

After working with the California Counties shapefile for a few weeks I knew I wanted to focus on a county, specifically my hometown of LA County. I used data gathered from the County of Los Angeles Open Data website and the City of Los Angeles Geohub website. My initial thought process was to do a combination of both. I also planned to use census data for demographic information. Overall, I kept six different sources of data. Although I cannot recall which shapefiles I downloaded first, I will describe them in no particular order.

The first data point I looked into was the number of Internet Exchange Points (also known as IXPs) in LA County. I was pleasantly surprised that I could download the data in four different forms: CSV, Shapefile, GeoJSON, and KML. I do not have any experience with the latter forms, so I downloaded a CSV and Shapefile. Ultimately, I decided to use the shapefile because of my experience dealing with error codes when using the cloudapps.usc.edu virtual desktop throughout the semester. I was fortunate to find data that was generally easy to edit. Another reason why I decided to use this data was because it only had 11 records in Los Angeles County as of April 2022. According to the Internet Society, IXPs help improve internet traffic more affordably and seek to improve stability, efficiency, and quality of network connection (n.d.,

2015). I was curious to see the general location of these data points- I had assumed they would be located in different parts of the county but the majority of them were located close to each other in the middle of the county. They appeared to be near government offices so I assumed that they are intended for government employees and not necessarily households. However, I decided to keep it on my map in case of any patterns.

The second shapefile I downloaded was the median household income. This information was last updated in April 2023 as collected by the Los Angeles County Department of Public Health (DPH) for a data initiative project in 2022. I felt that household income was important to include in this digital divide analysis as it can be a contributing factor to internet access and technological use. The third shapefile I used was the public internet access data on households with no internet access and no computer at home. My last two data sources were from the ArcGIS Living Atlas of the World: the ACS Internet Connectivity Variables focused on computer ownership (among other data points) and the Internet Access by Age and Race variables (among other data points) based on the 5-year American Community Survey Census (last updated December 2023). I had data ranging from the federal, state, and county levels. For example, the latter two data sources were divided by tract, county, and state boundaries. I removed state data to focus on the tract desired and narrowed it into LA County. In the next section, I will delve into more detail on my changes.

**General Data Formatting and Analysis**

After importing all the data layers into ArcGIS, I felt overwhelmed with all the data presented. My data was based on zip codes, county name, csa (city), supervisorial district (5), census tract, GEOID, or a longer Geographic Identifier FIPS Code, and not all layers had the exact data fields. Especially with the ACS census tracts, I would get confused after reading a

message that said the group layer could not participate in a symbol layer drawing, realizing I had to focus on one level such as the tract. Other times, it would automatically default to a data point I was not interested in. I had to process that data by using Symbology to remove what I did not need and highlight wanted data elements. However, LA County felt bigger than I expected for my analysis as it showed 88 cities within it. Despite running an outline of the LA County boundary and removing water areas using the pair clip, I felt it contained too much data. At first, I decided to trim the Median Income Census Tract that showed all of LA County to only show the City of Los Angeles. I exported only needed information by selecting attributes and creating an expression that would only focus on the csa (city) if it equaled the City of LA. I was surprised to see that the city of LA was still too big for the in-depth analysis I sought. With the other layers overlapping and pending combinations, I decided to focus on smaller cities instead. Thus, my question became more focused: is there a digital divide within these three cities in Los Angeles County?

**Brief Descriptions of Downey, Compton, and Beverly Hills**

Given their similar size, I chose the cities of Downey, Compton, and Beverly Hills. However, they ranged in population sizes. According to the U.S Census QuickFacts data population estimates as of July 1, 2023, Beverly Hills had 30,974 residents, Downey City had 108,816 residents, and Compton had 90,986 (U.S. Census Bureau, 2023). I was surprised that Downey had over triple the number of residents compared to Beverly Hills. If I had more resources, I would have focused on the number of housing units and building permits to observe any relevant trends explaining the population differences. I also chose them because of the different median household incomes and ethnic makeup. For example, Latinos comprise the following percentages in Beverly Hills, Compton, and Downey, at 7%, 71%, and 74%,

respectively (U.S. Census Bureau, 2023). Given its geographical location, I was surprised to see less than 10% of Latinos live in Beverly Hills. Lastly, Beverly Hill's median household income (2018-2022) was $116,771, compared to Compton's median household income (2018-2022) of $69,728 and Downey's median household income (2018-2022) of $84,236 (U.S. Census Bureau, 2023). These three cities are good examples of different socioeconomic factors that can affect access to digital services.

**Original Format, Processing, and Combining Data**

After removing the layout of the City of Los Angeles, I ran a new selection type of the csas of the three cities to only show median income data there. Next, I changed the primary symbology to graduated colors to showcase the difference of median household income. I clicked on the reverse symbol order to show that darker shades of blue reflected higher income. Next, I combined this layer with the internet and computer access census tract. I changed its primary symbology to dot density to focus on households without internet access and combined it with the first (income) layer. Then, I browsed through the available options under the internet access by education variable and chose the Latino population in households with no computer field under dot density. I wanted to focus on the Latino population given my racial background as a Latina and because Latinos are the fastest-growing population in California (McGhee, 2022). Most of the time, selecting a layer by attribute or location worked for me. Unfortunately, this data did not easily combine with my other layer because it did not include the same field. After some lengthy unsuccessful attempts trying to select by Geographic Identifier FIPS Code manually, I successfully used the geoprocessing pairwise clip to combine these layers. It was important to me to use different symbology options to reduce confusion to the reader of this paper and the map. For the data layer focused on age and race, I decided to focus on percent of

population 25 years and over in households that have no computer. I used proportional symbols to show breakdowns of 1 percent to 100 percent of households that do not have a computer. Lastly, even though I had chosen a specific symbol to represent the internet exchange points, they are not visible in the final map as none were located in these cities.

**Data and Time Limitations and Accuracy Impacts**

In this class, we did not study or practice running chi square regressions, so our conclusions should be considered limited. I led this project with curiosity, educated guesses, and noting patterns but my conclusions should be taken with a caution of further research needed to validate such claims. These observations are set to inform or spark interest in the reader. One limitation with this data is that it will not be real-time and it may contain outdated data. For example, assessing internet access by household can change on a daily basis or by the hour and the data available online only included information before 2023. It is possible that these cities look different in 2024 especially with the passage of new broadband access programs and the removal of others. Second, the census data is updated every ten years and thus, a lot of my data relied on the projected census estimates. Third, because all my data was not taken on a single day or time in a specific year, it is possible that the data is skewed, especially if under-reported. Another limitation is the data quality as further information may be needed to understand how an organization collected the information and what biases or errors may be included. While it was important to me to use a variety of sources, I did question how the LA County Department of Public Health collected household income information and the possibility of any errors. For example, I think it is possible that a resident makes more than the highest income reported here. After reading the revisions of my first paper, I would have preferred if the data had highlighted

areas with zero residents as separate categories to reduce assumptions such as households with no income.

**If I could do this project with no resource limitations, what would I do differently?**

Any ArcGIS maps created should be presented carefully as correlation does not mean causation. If I had resources such as unlimited time, ArcGIS credits, and access to its more premier services, I would focus on more data accuracy checks, partnering with important agencies such as the City of Long Beach's Digital Inclusion Initiative, and embed my ArcGIS maps to be hosted online so anyone can view and filter them as needed. Especially as this is my first-time using GIS, I would appreciate corrections and recommendations. These resources would be important to me to spread knowledge, training opportunities for myself and colleagues that might have access to data needed but do not have the tools to analyze it, and improve the reliability of the maps for online use. Although I did not focus on the City of Long Beach, I would have loved to work on an initiative to reduce the digital divide in the cities I analyzed by creating a similar initiative focused on improving connectivity and providing much needed hardware. I believe these activities would improve the accuracy of the data and inform important stakeholders (like residents, elected officials, and relevant departments that oversee digital service for the public) more efficiently.

Specifically, for this project, I wish I had more time to focus on an education data point. I initially had a graduation cap symbol representing areas with households with a bachelor's degree but I removed it as I felt the maps already contained too much information. I also wanted to focus on K-12 schools to see if there were any notable patterns but it felt like a completely different project. I chose not to highlight the population under 18 years old in households that have a computer without an internet subscription because, during these years, some K-12 school

districts were providing computers to students, and I did not want my data to be affected based on a local policy. However, I am aware that other confounding variables may influence the data presented in my map. Another thing I would have done differently is take more time to review the various data fields and consider running a similar categorical map from another time period to compare and contrast any changes. On the bright side, if I still have an active ArcGIS pro subscription, I can run a similar analysis using a newer data set once available. Lastly, I wish we could have incorporated aerial footage into a digital divide analysis focused on infrastructure durability and development to assess for any interesting trends.

**Conclusion**

Yes, there is a digital divide in Los Angeles County as shown by the maps focused on Compton, Downey, and Beverly Hills detailing on access to the internet, computers, age, and median household income and race. Prior to starting this project, I expected to see trends showing low internet access in lower-income households. My second expectation was to see Latino households across these cities since Latinos represent about 49% of the county's population. My third expectation was a higher (if not highest) income level in Beverly Hills.

Most of my expectations were on the right track. For example, in Compton and Downey, lower income households had more representation of those without internet and computer access than other households. However, we saw some lack of internet access in more affluent areas across all three cities, a contrast to my fourth expectation of higher income areas having little to no issues accessing the internet. On the other hand, Beverly Hills had fewer dots representing no internet, but I was surprised to see more than what I would have imagined. I wonder if other confounding variables can be at play, such as preferences. According to an article on Beverly Hills residents at a community center, the coordinator noted how out of the 80 residents, most

weren't tech savvy and only nine had their own technology device; those residents still did not access the internet (Levin, 2011). It is important to note that this is a low representative sample and it can be subjective. However, it could help us explain some dots in high-income Beverly Hills. Lastly, I was not surprised to see that Beverly Hills had no low-income household regions covered and had the highest income compared to the other cities.

Generally speaking, Downey was an interesting webbing of both cities. It did have low income sections like Compton, but it also had high-income sections like Beverly Hills. In the low-income parts of the city we saw the mass gray clusters as seen in Compton, and in the 97,000 to 127,000 median household income sections, we saw low grey dot clusters. Downey did not have a richer section like Beverly Hills. Compton had at least six 60% or more red circles, Downey had four of the 50% of red circles, and Beverly Hills had two of the 50% - unsurprisingly, in the lowest income bracket of the city.

There was a moment in my ArcGIS processing where I assumed the model was not working or I had run an analysis incorrectly. Specifically, the Latino population with no computer access was not showing up in my Beverly Hills layer, despite changing the order. Zooming out the map, I noticed that it did show dots in Downey and Compton. Based on this observation, I cannot assume there are no Latinos in Beverly Hills. Instead, I think that it means there is a strong likelihood of no Latinos without computers in Beverly Hills. According to the census estimates as of July 2023, Latinos make up about 7% of their population.

**Conclusion Implications**

At first, I assumed that the more affluent the area, the more computer and internet access we would see. However, this map shows that it is not just about the median income. Some of the digital divide can be attributed to preferences or cultural practices. Another implication is to

always question the data presented. For example, in the upper left section of Downey, there is one red dot in a lighter blue section. I was curious about this as it appears to be an anomaly, and looking at another street layer, I realized that section was a rehabilitation center. In other words, looks can be deceiving. A second in-depth analysis should be recommended to any new user such as myself. Overall, despite possible outliers, it is safe to assume that low-income communities struggle with internet and computer access.

**References:**

Internet Society (2015) *Policy Brief: Internet Exchange Points (IXPs)*. https://www.internetsociety.org/policybriefs/ixps/

Johnson, K. (2024). Millions of Californians weigh options after losing an affordable internet subsidy. *CalMatters.* https://calmatters.org/economy/technology/2024/04/digital-divide/

Levin, S. (2011). Seniors Get Online in Beverly Hills Sunrise Assisted Living is one place in town where you can find some tech-savvy older folks. *Patch*. https://patch.com/california/beverlyhills/seniors-get-online-in-beverly-hills

McGhee, E. (2022). California's Hispanic Community. *Public Policy Institute of California.* https://www.ppic.org/blog/californias-hispanic-community/

Pelchen, L. (2024). Internet Usage Statistics In 2024. *Forbes.* https://www.forbes.com/home-improvement/internet/internet-statistics/#:~:text=As%20of%202024%2C%2094.6%25%20of,have%20access%20to%20the%20internet.&text=As%20of%202024%2C%20the%20internet,according%20to%20a%20U.S.%20Census

Van Dijk, J., & Hacker, K. (2003). The Digital Divide as a Complex and Dynamic Phenomenon. *The Information Society*, 19(4), 315–326. https://doi.org/10.1080/01972240309487

Vives, R., & Castillo, A. (2024). Millions could lose internet access in California if this program expires. *Los Angeles Times*. https://www.latimes.com/politics/story/2024-04-22/federal-internet-subsidy-expires-this-month-leaving-2-8-million-californians-at-risk-of-losing-access

U.S. Census Bureau. (2023). American Community Survey 1-year estimates, 2023 “QuickFacts Beverly Hills City, California”. U.S. Department of Commerce https://www.census.gov/quickfacts/fact/table/beverlyhillscitycalifornia/PST045222

U.S. Census Bureau. (2023). American Community Survey 1-year estimates, 2023 “QuickFacts Compton, California”. U.S. Department of Commerce https://www.census.gov/quickfacts/fact/table/downeycitycalifornia/PST045223

U.S. Census Bureau. (2023). American Community Survey 1-year estimates, 2023 “QuickFacts Downey, California”. U.S. Department of Commerce https://www.census.gov/quickfacts/comptoncitycalifornia