Capstone Case Study: City Population, Median, and Poverty Over the Years (2013 - 2022)

Voltaire

December 8, 2023

### Title: U.S. District of Columbia Poverty Rate, 2013-2022

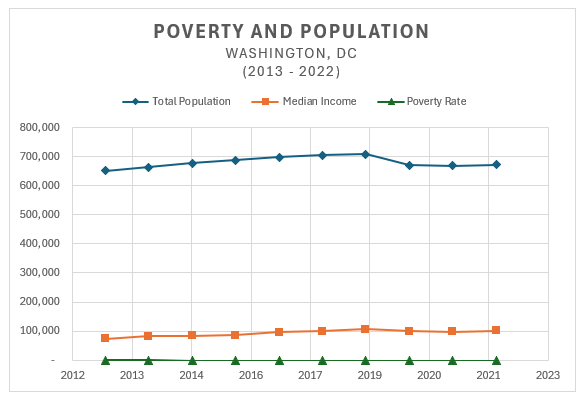
### Case Study: A Quick Summary of the city’s Population vs Poverty

In 2022, about 13.3 percent of the District of Columbia’s population lived below the poverty line. This accounts for persons or families whose collective income in the proceeding 12 months was below the national poverty level of the United States. (cite/source: Statista.com, Nov. 3, 2023).

### Data Analysis Process: an individual theory using data-driven real case scenario

### Summary: Poverty vs Population

# Create ggplot graph using Microsoft Excel, then exported to RMarkdown  
  
knitr::include\_graphics("C:/Users/HP/OneDrive/Desktop/Data Analytics/CAPSTONE Project/PopMedPovXL.png")

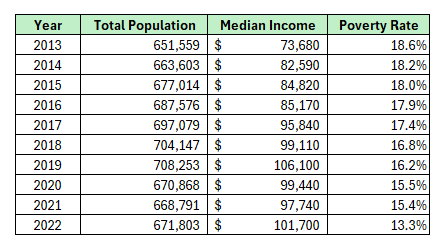


### Population Growth: Overview Summary

The city’s population has undergone a notable change over the specified period. The total population in 2013 was 651,559 compare to 671,803 last year (2022). The calculation below shows we have gained an additional 20,244 for the past decade (despite a slight downturn from 2020 during the pandemic).

### Population Growth: 2013 through 2022

# Import version of analysis from R to Excel - then link back to R Markdown as supplement visualization  
knitr::include\_graphics("C:/Users/HP/OneDrive/Desktop/Data Analytics/CAPSTONE Project/PopulationAndPoverty.png")



### Conclusion:

The good news, my research and analysis shows that the city of Washington (District of Columbia) has significantly reduced it’s poverty rate from over 18% down to 13.3% in 10 years. The city is thriving quite well in eliminating the poverty rate.

Meanwhile, the median income remains about the same for the past decade. But what really caught my attention was the gap between the median income and the vast number of population. The ‘white empty space’ between them. That is quite a gap. Perhaps, for my next case study, I could research and analyze Washington DC’s highest income population.

#### Below are the tools I have learned from Google Data Analytic course. These steps has helped me analyzed and visualized my CAPSTONE: real-scenario case study. Using the Six Steps of Data Analytics [**A**sk, **P**repare, **P**rocess, **A**nalyze, **S**hare, and **A**ct. Which I called the **APPASA Method**].

#### Downloaded .csv files for analysis

# Read csv files  
Pct\_of\_Poverty <- read.csv("C:/Users/HP/OneDrive/Desktop/Data Analytics/CAPSTONE Project/Pct\_of\_Poverty.csv")  
DC\_Pop <- read.csv("C:/Users/HP/OneDrive/Desktop/Data Analytics/CAPSTONE Project/DC\_Pop.csv")  
Household\_Income <- read.csv("C:/Users/HP/OneDrive/Desktop/Data Analytics/CAPSTONE Project/Median\_Household\_Income.csv")

#### Perform INNER JOINS

# Perform Inner Joins, by merging the 3 datasets: Population, Poverty, Median Income  
Population\_table <- inner\_join(Pct\_of\_Poverty, DC\_Pop, by = "Unique\_ID") %>%   
 inner\_join(Household\_Income, by = "Unique\_ID")

#### Get the analysis result - then Create dataset manually for further practice and analysis

# Create dataset manually  
  
PopMedPov <- data.frame(  
 Year = c(2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022),  
 Population = c(651559, 663603, 677014, 687576, 697079, 704147, 708253, 670868, 668791, 671803),  
 Median\_Income = as.numeric(gsub(",", "", c("73,680", "82,590", "84,820", "85,170", "95,840", "99,110", "106,100", "99,440", "97,740", "101,700"))),  
 Poverty\_Pct = c(18.6, 18.2, 18, 17.9, 17.4, 16.8, 16.2, 15.5, 15.4, 13.3)  
)

#### Revisit dataset for additional cleaning and organizing

# More cleaning, organizing, eliminating repeated years  
cleaned\_data <- Population\_table %>%   
 select(Unique\_ID, YEAR, Median\_Income, Pct\_of\_Poverty, Total\_Population)

#### A quick view summary of the dataset

# Quick and simple view of the analysis of the dataset  
summary(cleaned\_data)

## Unique\_ID YEAR Median\_Income Pct\_of\_Poverty   
## Length:11 Length:11 Min. : 73680 Min. :13.30   
## Class :character Class :character 1st Qu.: 83705 1st Qu.:15.85   
## Mode :character Mode :character Median : 95840 Median :17.40   
## Mean : 91564 Mean :16.89   
## 3rd Qu.: 99275 3rd Qu.:18.10   
## Max. :106100 Max. :18.60   
## Total\_Population  
## Min. :635.7   
## 1st Qu.:666.2   
## Median :671.8   
## Mean :676.0   
## 3rd Qu.:692.3   
## Max. :708.3

#### Summary Overview

# Summary overview  
PopMedPov %>%  
 summarise(  
 Mean\_Population = mean(Population),  
 Median\_Income = median(Median\_Income),  
 Mean\_Poverty\_Pct = mean(Poverty\_Pct)  
 )

## Mean\_Population Median\_Income Mean\_Poverty\_Pct  
## 1 680069.3 96790 16.73

kable(PopMedPov, format = "markdown")

| Year | Population | Median\_Income | Poverty\_Pct |
| --- | --- | --- | --- |
| 2013 | 651559 | 73680 | 18.6 |
| 2014 | 663603 | 82590 | 18.2 |
| 2015 | 677014 | 84820 | 18.0 |
| 2016 | 687576 | 85170 | 17.9 |
| 2017 | 697079 | 95840 | 17.4 |
| 2018 | 704147 | 99110 | 16.8 |
| 2019 | 708253 | 106100 | 16.2 |
| 2020 | 670868 | 99440 | 15.5 |
| 2021 | 668791 | 97740 | 15.4 |
| 2022 | 671803 | 101700 | 13.3 |

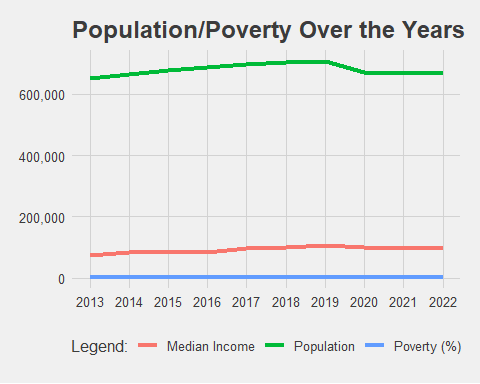
#### Summary: quck view, tabular method

# Summary: quick view, tabular method  
summary(PopMedPov[, c("Population", "Median\_Income", "Poverty\_Pct")])

## Population Median\_Income Poverty\_Pct   
## Min. :651559 Min. : 73680 Min. :13.30   
## 1st Qu.:669310 1st Qu.: 84908 1st Qu.:15.68   
## Median :674409 Median : 96790 Median :17.10   
## Mean :680069 Mean : 92619 Mean :16.73   
## 3rd Qu.:694703 3rd Qu.: 99358 3rd Qu.:17.98   
## Max. :708253 Max. :106100 Max. :18.60

#### R - ggplot Visualization Graph

# Create graph for visualization using ggplot (R Studio)  
  
ggplot(PopMedPov, aes(x = Year))+  
 geom\_line(aes(y = Median\_Income, color = "Median Income"), linewidth = 1.5)+  
 geom\_line(aes(y = Population, color = "Population"), linewidth = 1.5)+  
 geom\_line(aes(y = Poverty\_Pct \* 100, color = "Poverty (%)"), linewidth = 1.5)+  
 labs(title = "Population/Poverty Over the Years",  
 x = "Year",  
 y = "Value",  
 color = "Legend:")+  
 scale\_y\_continuous(labels = scales::comma) +  
 scale\_x\_continuous(breaks = PopMedPov$Year) +  
 theme\_fivethirtyeight()



###### Sources:

**Cite:** *FRED economic data | St. Louis Fed |license:* [*https://fredstlouis.org/legal/#copyright-public-domain*](https://fredstlouis.org/legal/#copyright-public-domain)  
*Dataset Updated: December 8, 2022*  
*Source: U.S. Census Bureau*

**Cite:** *Real Median Household Income in the District of Columbia*  
*Dataset updated: September 2023*  
*Source: fred.stlouisfed.org*

**Cite:** *ACS Demographic Characteristics DC Ward*  
*Dataset updated: February 2023*  
*Source: opendata.dc.gov*

### About Me

Prior to learning Data Analytics, I had a solid background in the hospitality industry as an Administrative Assistant/Coordinator in the Global Sales Department. After my long years of service in the hotel industry, I decided to pursue my dream of becoming a certified bookkeeper due to my analytically mindset. And while being well-versed with Excel spreadsheet and SQL, that’s when I realized that I need to add another set of skills and take the Google Data Analytic which I’m very fond of doing.

When I’m not analyzing data, I indulged in my passion for books, particularly novels and literature. I find joy in the unconventional and strive for a simple, content life surrounded by nature and furry friends. The Fall season, the holidays and snow bring immense delight to my well-being.

### My Portfolio

Now, let’s delve into why employers should be intrigued by my portfolio. My meticulous nature as a Global Sales Coordinator/Bookkeeper translates into a keen attention to detail in every aspect of my work. Studying data analytics has equipped me with a robust analytical mindset, which is reflected in my approach to problem-solving.

The project showcased in my portfolio not only demonstrates technical proficiency but also a profound understanding of the intricacies of data. Employers would find value in my ability to bridge the realms of bookkeeping and data analytics, providing a unique perspective that aligns with my goal of bringing simplicity and efficiency to the professional world of analytics. This combination of skills, coupled with a passion for continuous learning, positions me as a valuable candidate to contribute effectively to any team.