Healthcare Coverage Analysis in Mexico Using R

Brien Navarro

2025-03-16

#Description

In this project, we use R to analyze a dataset extracted from INEGI (Instituto Nacional de Estadística y Geografía) that contains data on the percentage of the population with access to a healthcare provider through their employment. The goal is to explore trends, visualize key insights, and assess healthcare coverage distribution across different regions in two different periods of time, 2010 and 2015, respectively.

#Key Objectives

[] Perform exploratory data analysis (EDA) to understand healthcare access patterns. [] Use data visualization (ggplot2) to display trends in healthcare coverage. [] Clean and preprocess data for accurate analysis. [] Compare regional variations in access to healthcare services.

#Data source

1. Dataset: Extracted from INEGI (Instituto Nacional de Estadística y Geografía).
2. Metric: Percentage of people with healthcare benefits through employment.
3. Coverage: Data segmented by state and year.

#Fisrt Analysis

#Percentage of the population with access to a healthcare provider through their employment in 2010 per state in Mexico.

#Rcode to achive the objective

Bar plot for 2010 Image

#Conclusion 1

#Based on the provided data and descriptive statistics, we can draw the following insights:

Central Tendency:

The mean percentage of the population with access to a work-related healthcare provider is 68.14%, while the median is 69.4%. This suggests that the distribution is fairly symmetric, with values centered around 69.4%.

Distribution Shape:

The data follows a normal distribution, indicating a balanced spread of healthcare coverage across different Mexican states. There is no significant skewness, as the mean and median are close.
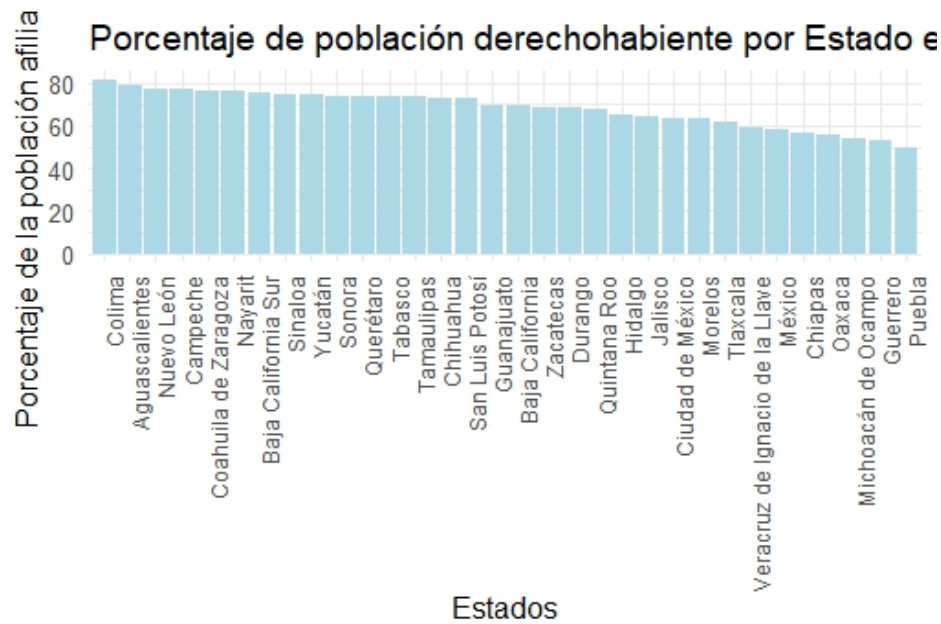
Figure 1: image

State-Level Analysis:

On average, 69.4% of the population in each Mexican state has access to a healthcare provider through employment. This highlights a consistent level of healthcare access across states.

#Outlier Detection: Notably, no outliers were detected in the dataset. This suggests no extreme irregularities, reinforcing the stability of healthcare coverage distribution across the states. #Key takeaway The data indicates a uniform distribution of healthcare access across Mexican states, with no extreme variations. This suggests that work-related healthcare coverage is relatively consistent nationwide.

#Second Analysis

#Percentage of the population with access to a healthcare provider through their employment in 2015 per state in Mexico.

#Rcode to achive all the objectives

```r
#Installing required packages and loading them

install.packages("rio", repos = "https://cran.rstudio.com/")
```

```
## Warning: package 'rio' is in use and will not be installed
```

```r
library(ggplot2) #Load the library
library("psych") #Load the package Psych
install.packages("statip", repos = "https://cran.rstudio.com/")
```

```
## Warning: package 'statip' is in use and will not be installed
```

```r
library(statip)
install.packages("DescTools", repos = "https://cran.rstudio.com/")
```

```
## Warning: package 'DescTools' is in use and will not be installed
```

```
library(DescTools)
install.packages("patchwork", repos = "https://cran.rstudio.com/")
```

Warning: package 'patchwork' is in use and will not be installed

```
library(patchwork)
library(dplyr)


#Exporting and filtering the downloaded data base provided by INEGI

#Exporting process
data_set_1 <- import("C:/Users/DELL/Downloads/morbilidad2023/sap.csv", encoding = 'Latin-1')

#Filtering process
data_set_1_2010 <- data_set_1[data_set_1$Periodo == 2010, ]
data_set_1_2015 <- data_set_1[data_set_1$ Periodo==2015,]
data_set_1_2015 <- data_set_1_2015 %>%
  filter(`Entidad Federativa` != "Estados Unidos Mexicanos")
data_set_1_2010 <- data_set_1_2010 %>%
  filter(`Entidad Federativa` != "Estados Unidos Mexicanos")

#Bar plot of the percentage of access to a work healthcare service per state in Mexico in 2015
plot2 <-ggplot(data_set_1_2015, aes(x = reorder(`Entidad Federativa`,
                                            -`Porcentaje de la población afiliada o derechohabiente a algún
                          y = `Porcentaje de la población afiliada o derechohabiente a algún servicio
  geom_col(fill = "purple", color = "black") +
  labs(
    title = "Porcentaje de población derechohabiente por Estado en 2015",
    x = "Estados",
    y = "Porcentaje de la población afiliada en %"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, size = 8))

describe(data_set_1_2015$`Porcentaje de la población afiliada o derechohabiente a algún servicio de salu
```

```
##    vars  n  mean   sd median trimmed  mad min  max range  skew kurtosis  se
## X1    1 32 83.76 3.39   84.7      84 3.26  74 89.5  15.5 -0.76     0.23 0.6
```

Bar plot for 2015 Image

#Conclusion 1

#Based on the 2015 data and descriptive statistics, we can infer the following insights:

#Normal Distribution: The mean percentage of the population with access to a healthcare service is 83.7%, while the median is 84.7%. Since the mean and median are close, this suggests that the data follows a normal distribution, with values symmetrically distributed around the central tendency.

#State-Level Healthcare Access:

The average percentage of people with access to healthcare services across all states is 83.7%. This means that, on average, approximately 83.7% of the population in each Mexican state is affiliated with a healthcare provider.
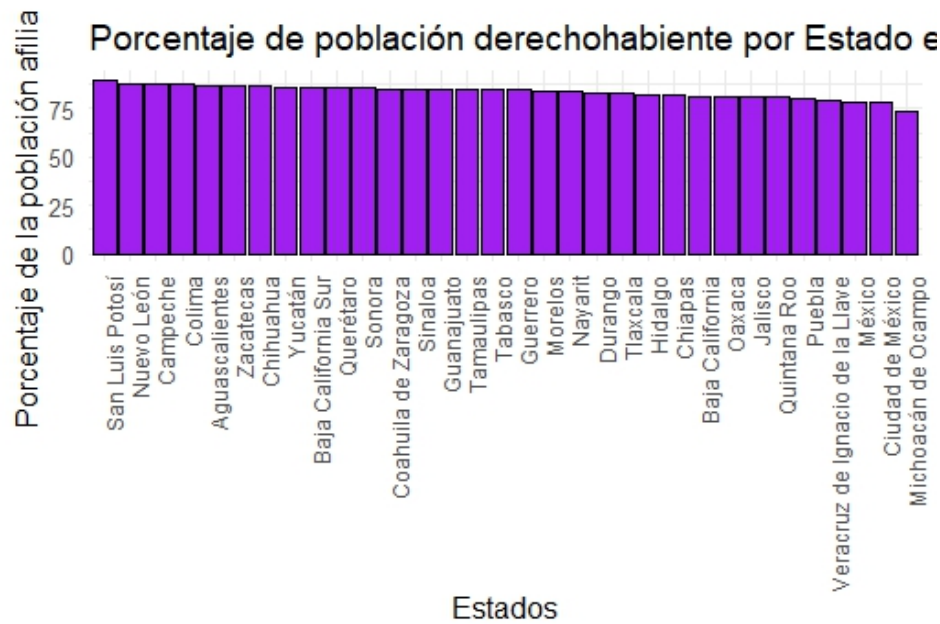
Figure 2: image

#Key Takeaway The small difference between the mean and median supports the assumption of a normal distribution, indicating a relatively consistent level of healthcare access across the states in 2015.

#Comparison Between Two Time Periods (2010 vs. 2015)

Bar plot for 2010 vs 2015 Image

#Visualizing the Change To effectively compare healthcare access trends between 2010 and 2015, we combine both plots into a single visualization (See the last R code): # Key Findings & Conclusion

#Significant Increase in Healthcare Access:

The average percentage of the population with access to work-related healthcare providers increased from 68.4% in 2010 to 83.7% in 2015. This represents a notable upward trend in healthcare coverage across Mexican states.

#Normal Data Distribution:

Both datasets (2010 & 2015) follow a normal distribution, as indicated by the close mean and median values. There were no outliers or irregularities in the data, ensuring reliability.

#State-Level Growth: The visualization clearly shows a consistent increase in healthcare coverage across all states. This trend can be observed in the bar charts, where each state's percentage of population with healthcare access has risen from 2010 to 2015.

#Key Takeaway The data suggests a substantial improvement in work-related healthcare coverage over the five-year period. The absence of outliers and the normal distribution further support the accuracy and consistency of this positive trend.
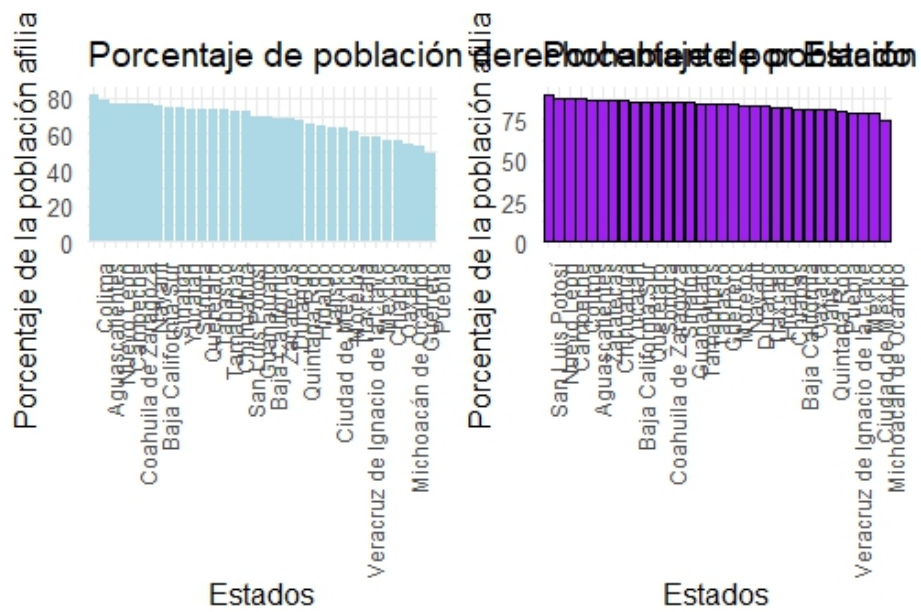
Figure 3: image