

ADOLESCENT MATERNAL MORTALITY

EMILY SCHOOF
SPRINGBOARD CAPSTONE
WINTER 2019

MEXICO

MEX



WHAT FACTORS INFLUENCE THE PERSISTENCE OF YOUNG MATERNAL MORTALITY?

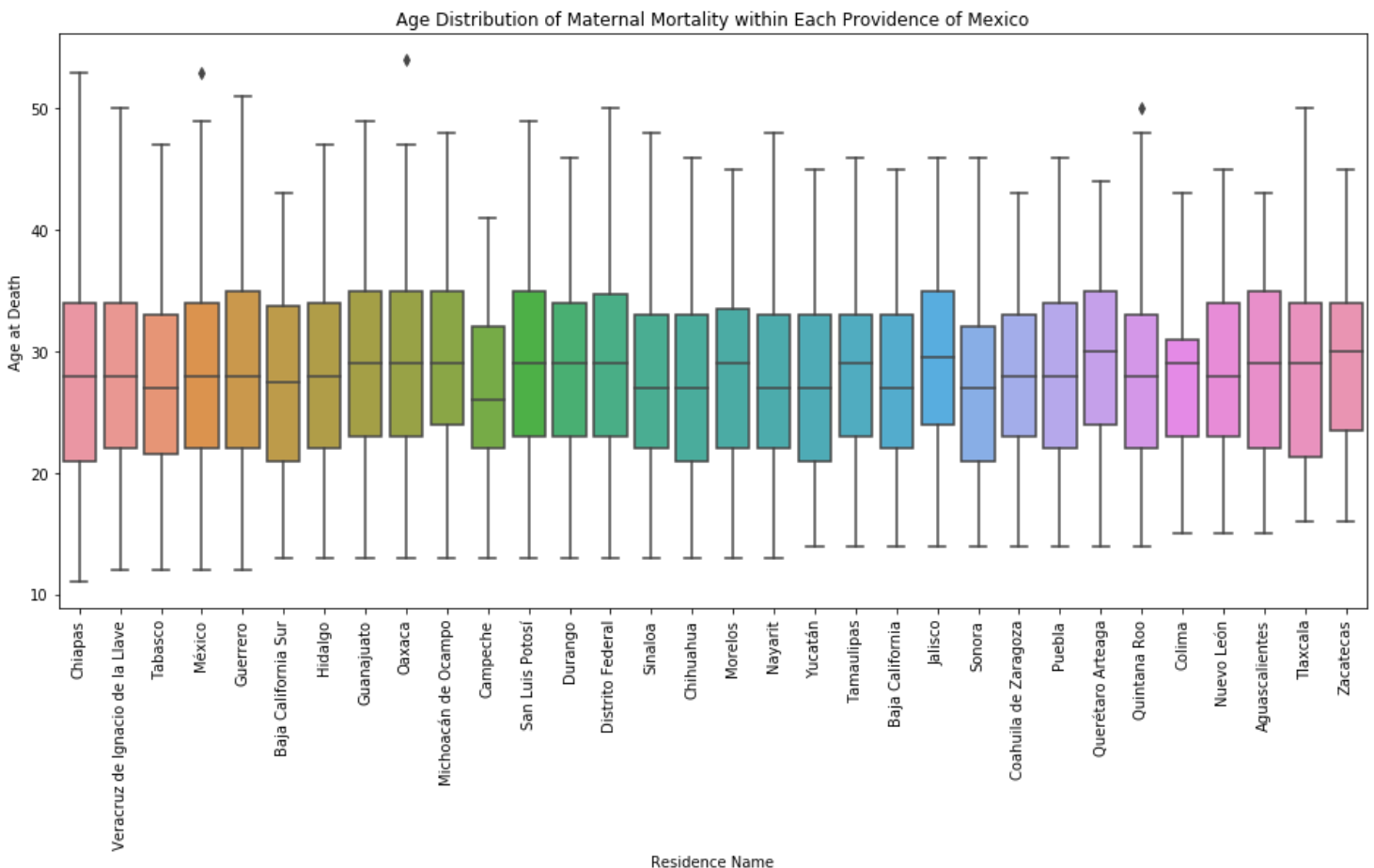
A DATA SCIENCE ANALYSIS BY EMILY SCHOOF

This project addresses a real-world problem that has the potential to directly impact the society of Mexican women and young girls. It is a challenging project, as it entails the collection real-world data various sources, including the World Bank, University of Chicago, and Wikipedia, in order to solve a problem that is less known and not fairly addressed. The objective was not just to investigate the mortality rate, but also uncover the hidden factors that contribute towards its significance. Learning from them, the ultimate pursuit was to REDUCE maternal mortality, especially in adolescents.

PROJECT IN REVIEW

Based of off the factors discussed on the World Health Organization (WHO)'s webpage on Maternal Mortality worldwide, some of the top factors discussed included region and local community population sizes, region GDP, education level, and presence of medical assistance at time of death (ATD). These 5 key features [region population, region GDP, local community size, education level, and presence of medical care] were collected and organized by region to predict if a region of Mexico's likelihood of maternal mortality of adolescents using a Logistic Regression machine learning model.

Prior to analysis, it was statistically proven via ANOVA analysis that the mean age of maternal mortality is statistically different across at least one region.



MODEL SELECTION AND SET-UP

Logistic Regression is a simple, probabilistic algorithm for linear and binary classification that is built on the mathematical logistic sigmoid function, which calculates the probability that a certain sample belongs to a particular class. Since the overarching goal is to identify factors that increase the likelihood of adolescent mortality in Mexico, this model is the natural choice.

TARGET VARIABLE: ABOVE(0) OR BELOW(1) MEX MEAN

To increase the predictive significance of the model, a binary target variable was created based on if a region's mean age of maternal mortality within the dataset was above or below the mean age of maternal mortality in Mexico (dataset mean was verified to be statistically accurate via Bootstrap).

PROJECT CONCLUSION

The scaled regression model predicted the likelihood of maternal mortality of a region to be above or below the national mean. The model outcome was based on five key features namely, region population, region GDP, local community size, the education level, and the medical care availability. The accuracy of the model was significantly high (87.5%), which complimented our methodology and scientific primitives. Additionally, it served as a preliminary step towards the systematic identification of areas within Mexico which could benefit the most from the government resource interventions.

To enhance the accuracy of the model, some notable directions could be the incorporation of level-sex education, nearest hospital proximity, and the cases of child-bride instances within each region. These factors can provide detailed information, specific to the adolescent maternal mortality. Additionally, these features can provide the correlation data regarding the potentially needed resources such as increased access to the healthcare, and socio-economic factors such as child brides. As a result, we can obtain quantitative measurements to further reduce the maternal morality rate.

Overall, the outcome of the project can be leveraged to aid the national and the local governments, as well as, the foreign-aid institutions, in the fair and timely allocation of the intervention funds. Naturally, these will be driven by the intent to reduce the maternal mortality in the regions that are struggling with it. Our analysis, however dense, may still provide the informative and actionable methodology eradicate the maternal mortality across all age groups and regions in Mexico.