Predicting Immigration Into The U.S. From Mexico Using Multi-Dimensional Regression

By Alexander Bidas, Jonathon Dowdy, Sarah Yake

CMSC 455: Numerical Computations

University of Maryland, Baltimore County (UMBC)

Baltimore, MD 21250

{abidas1, jdowdy1, syake1}@umbc.edu

December 5, 2015

## 

## 

## Motivation

Our motivation for choosing to model immigration trends from Mexico into the United States stems from concern over the harsh political climate towards Mexican immigrants. It is our hope that in creating this model, we will allow Americans to understand why this pattern of immigration is occurring, allow the U.S. to plan for future immigration, and possibly help Americans become more open to the idea of having immigrants from Mexico. We believe that in helping the United States correctly estimate the flow of immigration from a given region, the country will be better equipped to accommodate immigrants as well as satisfy the public needs of the American people.

## Techniques

### Implementation

Initially, we wanted to implement our project in Python because we could use ArcGis (Python compatible) to create a visual model of immigration over time, as well as use of the NumPy library to assist in the statistical analysis of our data. After submitting the Project Proposal and as we collected more data, we found that creating the visual model would not be of very practical use. Instead, we decided to implement our model using Matlab and R because we discovered that both have many tools that would help us implement multiple regression faster, more easily, and more efficiently. In Matlab, we used the regress function, as well as the regression learner app and curve fitting app. The regress function in particular was especially useful because it gave us the ability to easily retrieve the coefficient estimates for a multilinear regression. In R, we were able to calculate the residual standard error and the multiple R-squared with ease.

### Design Choices

Initially, we thought that a multiple linear regression may be a good fit for the data first dataset we found, mostly because we did not notice a distinct nonlinear trend in our first set of data at first glance. Upon deeper analysis, we found that a polynomial regression provided a better fit for the data.

For our first model, we collected data on Mexico’s population, life expectancy, birth rate, GDP per capita, United States minimum wage adjusted for inflation, and immigration from Mexico to the United States, from 1997-2015 (as this was the time period in which we could most reliably find data). After our first model failed, we realized that we did not account for several other factors that could lead to changes in immigration patterns. For instance, we mainly focussed on the push-factors within Mexico that could spur immigration, while completely overlooking most of the social and economic conditions that would draw immigrants to the host country (the United States). For this reason, we significantly expanded our data collection to account for changes in demographics, changes in the economy, changes in criminal activity, and historical events in both countries.

## Discussion

### Difficulties Encountered

One of the most difficult obstacles we encountered was simply finding the data, which (in a few cases) had to be pieced together from different datasets. For example, some of the data obtained from the UN database were incomplete, so it was necessary to hunt down data in the UN archives to complete the dataset for our selected time period (a time period that reliably had data). Even still, we were not able to find all of the data for every field. So, in our data table, we just entered “NA” to represent a missing value. To fix this issue in Matlab, we simply replaced “NA” values with NaN. As for R, we used a polynomial regression to interpolate the missing values.

Another set of problems we encountered involved selecting and learning programming languages in order to generate our models. As previously mentioned, we shifted from using Python, to using Matlab, to using both Matlab and R to generate our models. This, not surprisingly, created a somewhat of a learning curve for each of our team members. We finally settled on using Matlab and R because both Matlab and R had exactly the tools for data analysis and plot generation that we needed, minimizing the error that could have occurred due to implementing regression functions from-scratch.

Our last major difficulty was attempting to figure out why our initial model was bad and how to move forward from that to create a model that made sense with our data. After doing more background research, we found that there were major historical events that we believe may have affected our data: the birth of the Department of Homeland Security in late 2002 in response the 9/11 attacks in 2001, and the housing crisis that began in late 2008.

### Limitations

There were many limitations to our models in our implementation. The main limitation is that it is difficult to anticipate the effect of government policy and certain historical events on immigration patterns. Both may have a major impact on the flow of people in and out of a country, but are not easily quantifiable and thus difficult to accurately represent in our model.

Then there are the limitations to using the regression models themselves, in that the models are noticeably very sensitive to outliers in the data we used. We also found that multiple regression did not perform well with incomplete data, especially within our limited data set. We also observed that as the polynomial regression extrapolated data further outside of the initial data range, the regression model became less reliable for prediction.

The programming languages we used also had their own limitations. For example, we found that R can be very slow when computing larger equations. Matlab was not perfect either. Plotting the data using Matlab proved to be a bit of a pain at times, but this limitation pales in comparison to the biggest limitation of using Matlab: the very unreliable data produced by the built in functions. Whenever we used the Regression Learner application, we would get different results each time, whereas R consistently gave us the same outputs. Also we found out that the regress function could only take in a limited amount of variables.

### Benefits

For the models, we found that multiple regression allowed us to easily see which predictor variable had the most influence on our response variable and quickly pinpoint outliers in the data. We later found that our immigration data did not have a distinct linear shape, but the polynomial regression fit the data very well in representing the curvatures in the data.

There have been some great benefits to using our selected programming languages as well. Although Matlab may not produce as reliable outputs as we would have hoped, Matlab was overall fairly easy to use, with a GUI as an added bonus. Matlab also offered many built in functions and applications that were designed specifically for the type of analysis we had been doing. We especially appreciated the built in Regress function’s ability to ignore NaNs in a data set while still performing fairly well (as opposed to R). We were also able to use Matlab to perform interpolation with polynomial regression to fill in data within our data to use in our analysis with R. Even though R may have been finicky about missing data, the built in statistical tools still offered many options for us to generate different models while consistently providing us with the best output.

## Results

### Summary of Data

For our first multiple regression model, we used Mexico’s population, life expectancy, birth rate, GDP per capita, and the United States minimum wage adjusted for inflation as our independent variables to explain immigration from Mexico to the United States in 1998-2015. The results of the first regression model had a Residual standard error: 26050 on 12 degrees of freedom and Multiple R-squared: 0.3356. For clarification, residual standard error shows of how much the observed dependent variable (immigration from Mexico) differs from the fitted values, while the Multiple R-squared value is used to determine the percentage of the data that can be explained by the regression model. With a high residual standard error and a low Multiple R-squared, we have determined that our first model poorly represented our data.

We noticed a drastic decline from 2002-2003 on the graph of immigration from Mexico to the United States. We believe that this event in our data set could be due to the creation of the Department of Homeland Security (DHS) in response to the September 11 attacks in 2001. To account for this event in our regression, we implemented a binary indicator variable (as previously mentioned) that is only true for 2003 and false for all other years. The presence of the indicator variable significantly improved our model, decreasing the residual standard error to 16980 on 11 degrees of freedom increasing the Multiple R-squared to 0.7411.

At this point, we thought we had a decent regression model for our data until we noticed a multicollinearity between three of our independent variables, so we decided to use add new variables to make a better regression model. The new model’s independent variables became thus: Mexico Total Intentional Homicides, Mexico Total Unemployment (% of total labor force), United States minimum wage adjusted for inflation, DHS (indicator variable).

We later discovered that we overlooked another major event, marked by a previously unexplained drastic increase in the unemployment rate in Mexico between the years 2007 and 2008, as well as a steep increase in immigration during this same period. Since this drastic changes in the data coincided with the 2008 financial crisis in the United States, we determined that Mexico’s economy may have been greatly affected by this event, and perhaps this was heavily influencing the spike in immigration. This theory was confirmed plausible after another indicator variable was added to account for this event when our final regression model yielded a Residual standard error of 9562 on 8 degrees of freedom with a Multiple R-squared of 0.9403, our best results yet.

### Initial Model

Using the multiple regression approach in our initial model, we have produced the following model using the data sets of legal immigration from Mexico, Mexico’s Population, Life Expectancy at Birth, Number of Births in Mexico, GDP per capita in Mexico, and US minimum wage adjusted for inflation:

f1(x) = -1.942e+07 + -1.401e-02xi+2.564e+05xi+7.130e+04xi+1.168e+01xi+2.475e+04xi

Where x represents years between 1998-2015 and f(x) represents the function of total immigration from Mexico to the United States.

Residuals:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Min | 1Q | Median | 3Q | Max |
| -60081 | -12527 | -1621 | 16728 | 33883 |

Coefficients:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate Std | Error | T value | Pr(>|t|) |
| Intercept | -1.942e+07 | 1.139e+07 | -1.705 | 0.114 |
| Population | -1.401e-02 | 8.013e-03 | -1.748 | 0.106 |
| Life expectancy | 2.564e+05 | 1.450e+05 | 1.768 | 0.102 |
| Birth Rates | 7.130e+04 | 4.950e+04 | 1.441 | 0.175 |
| GDP per capita | 1.168e+01 | 1.113e+01 | 1.049 | 0.315 |
| US Min Wage | 2.475e+04 | 2.155e+04 | 1.149 | 0.273 |

|  |  |  |
| --- | --- | --- |
| Residual standard error | Multiple R-squared | p-value |
| 26050 on 12 degrees of freedom | 0.3356 | 0.3611 |

Based on the output, we found that our initial model failed to accurately represent the data. For instance, the R-squared value is very low, meaning that a very small percentage of the data that can be explained by the regression model of the data. To try and achieve a better fitting output we implemented a dummy variable in the year 2003 in an attempt to account for the change in immigration policy due to the creation of the Department of Homeland Security.

These are the final outputs after the use of the first dummy variable:

f2(x) = -2.906e+07 + -2.045e-02xi + 3.861e+05xi + 9.934e+04xi + 3.977e+00xi + -8.272e+04xi + 2.937e+04xi

|  |  |  |
| --- | --- | --- |
| Residual standard error | Multiple R-squared | p-value |
| 16980 on 11 degrees of freedom | 0.7411 | 0.008812 |

After using the dummy variable to account for the a policy change in year of 2003, we found that our model has a much better fit. From simply accounting for a policy change in one year, our R-squared rose from .3356 to .7411, signifying that the new model had become significantly better.

### Final Model

We were not satisfied with the multiple linear regression model, so we decided to implement to a polynomial regression in hopes that we would achieve a better fit for our data. However, because we know the limitations of the polynomial regression, decided to only to only make a prediction for the year 2016, as predictions later than that may prove to be unreliable with our model. We also wish to note that for this model, we added another change to our model: our data. We decided to expand our data sets in order to determine which factors may correlate with the observed immigration patterns through trial and error.

After a series of trials, we determined that the data sets Mexico total intentional homicides, US minimum wage adjusted for inflation, and total unemployment (as a percent of total labor force) produced the best model. However, both the intentional homicides and total unemployment had missing data. Thus, we needed to use Matlab's curve fitting app to perform polynomial regression to interpolate those missing data points.

Once done, we used our polynomial regression model to determine the best fit. It is worth reiterating that we kept the dummy variable to account for 2003 from the previous model and added another one to account for 2008. Below is the final model of our polynomial regression.

g(x) = 1.262e+06 + -1.841e+02xi +2.358e-02xi2 + -8.952e-07xi3 + -1.561e+05xi + 2.013e+04xi2 + -7.884e+04xi +

3.143e+03xi2 + -8.936e+04xi + 6.347e+04xi

Where x represents years between 1998-2015 and g(x) represents the function of total immigration from Mexico to the United States.

Coefficients:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate Std | Error | T value | Pr(>|t|) |
| Intercept | 1.262e+06 | 5.129e+05 | 2.460 | 0.039302 |
| Total Homicides | -1.841e+02 | 3.386e+01 | -5.436 | 0.000619 |
| I(Total homicides )^2 | 2.358e-02 | 4.056e-03 | 5.813 | 0.000399 |
| I((Total homicides )^3 | -8.952e-07 | 1.479e-07 | -6.051 | 0.000305 |
| Total unemploy | -1.561e+05 | 5.136e+04 | -3.039 | 0.016097 |
| I(Total unemploy^2) | 2.013e+04 | 7.410e+03 | 2.717 | 0.026379 |
| US min wage | -7.884e+04 | 1.648e+05 | -0.478 | 0.645188 |
| I(US min wage^2) | 3.143e+03 | 1.235e+04 | 0.254 | 0.805549 |
| DHS | -8.936e+04 | 1.495e+04 | -5.976 | 0.000332 |
| FIN | 6.347e+04 | 1.518e+04 | 4.182 | 0.003070 |

|  |  |  |
| --- | --- | --- |
| Residual standard error | Multiple R-squared | p-value |
| 9562 on 8 degrees of freedom | 0.9403 | 0.0005382 |

In 2016, our model predicted that approximately 203670 people will immigrate from Mexico into the US, give or take about 32160 people.

|  |  |  |
| --- | --- | --- |
| FIT | LOWER BOUND | UPPER BOUND |
| 203669.9 | 171509.8 | 235830.1 |

Conclusion

After a lot of trial and error, we believe that we have generated a fairly viable predictive model of Mexican immigration into the United States for the year 2016 using multiple polynomial regression. When immigration data for 2016 becomes available to the public, we hope to confirm the validity of this prediction. For the future, we hope that this model may be improved upon, perhaps using other methods of approximation.

Outputs

Call:

lm(formula = Immigration ~ Population + expectancy + Birth +

GDP\_percapita\_Mex + USmin\_wage, data = Data)

Residuals:

Min 1Q Median 3Q Max

-60081 -12527 -1621 16728 33883

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -1.942e+07 1.139e+07 -1.705 0.114

Population -1.401e-02 8.013e-03 -1.748 0.106

expectancy 2.564e+05 1.450e+05 1.768 0.102

Birth 7.130e+04 4.950e+04 1.441 0.175

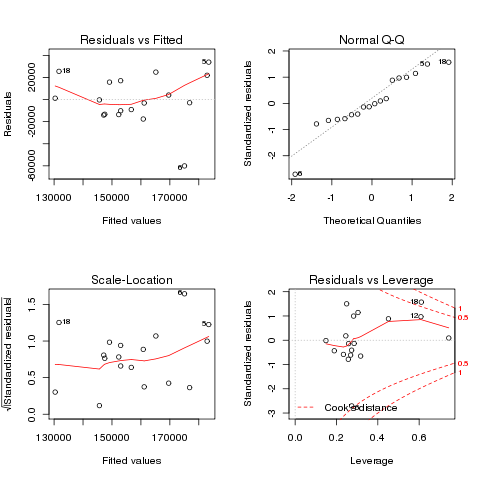
GDP\_percapita\_Mex 1.168e+01 1.113e+01 1.049 0.315

USmin\_wage 2.475e+04 2.155e+04 1.149 0.273

Residual standard error: 26050 on 12 degrees of freedom

Multiple R-squared: 0.3356, Adjusted R-squared: 0.0587

F-statistic: 1.212 on 5 and 12 DF, p-value: 0.3611



Call:

lm(formula = Immigration ~ Population + expectancy + Birth +

GDP\_percapita\_Mex + DHS + USmin\_wage, data = Data)

Residuals:

Min 1Q Median 3Q Max

-15725 -11565 -2350 9870 29511

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -2.906e+07 7.779e+06 -3.735 0.00329 \*\*

Population -2.045e-02 5.450e-03 -3.752 0.00320 \*\*

expectancy 3.861e+05 9.956e+04 3.878 0.00257 \*\*

Birth 9.934e+04 3.297e+04 3.013 0.01180 \*

GDP\_percapita\_Mex 3.977e+00 7.492e+00 0.531 0.60606

DHS -8.272e+04 1.993e+04 -4.151 0.00161 \*\*

USmin\_wage 2.937e+04 1.409e+04 2.084 0.06122 .

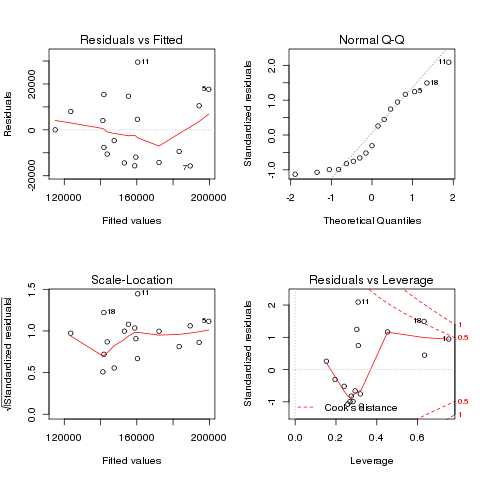
---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 16980 on 11 degrees of freedom

Multiple R-squared: 0.7411, Adjusted R-squared: 0.5999

F-statistic: 5.248 on 6 and 11 DF, p-value: 0.008812



Call:

lm(formula = Immigration ~ C + I(C^2) + I(C^3) + MEX\_EMPLOY +

I(MEX\_EMPLOY^2) + USmin\_wage + I(USmin\_wage^2) + DHS + FIN,

data = Data)

Residuals:

Min 1Q Median 3Q Max

-16571 -2841 0 3965 14857

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1.262e+06 5.129e+05 2.460 0.039302 \*

C -1.841e+02 3.386e+01 -5.436 0.000619 \*\*\*

I(C^2) 2.358e-02 4.056e-03 5.813 0.000399 \*\*\*

I(C^3) -8.952e-07 1.479e-07 -6.051 0.000305 \*\*\*

MEX\_EMPLOY -1.561e+05 5.136e+04 -3.039 0.016097 \*

I(MEX\_EMPLOY^2) 2.013e+04 7.410e+03 2.717 0.026379 \*

USmin\_wage -7.884e+04 1.648e+05 -0.478 0.645188

I(USmin\_wage^2) 3.143e+03 1.235e+04 0.254 0.805549

DHS -8.936e+04 1.495e+04 -5.976 0.000332 \*\*\*

FIN 6.347e+04 1.518e+04 4.182 0.003070 \*\*

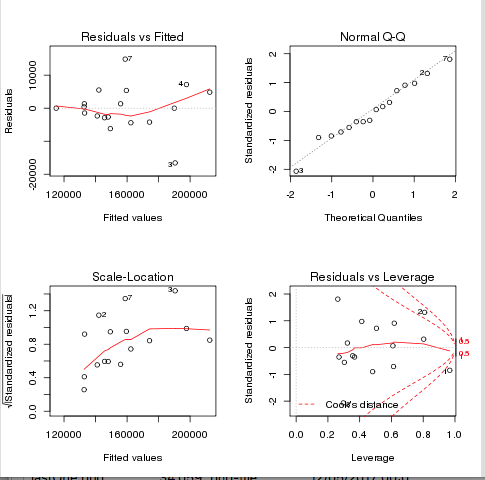
---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

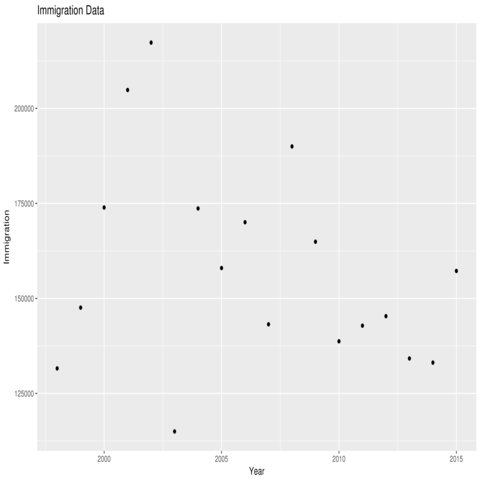
Residual standard error: 9562 on 8 degrees of freedom

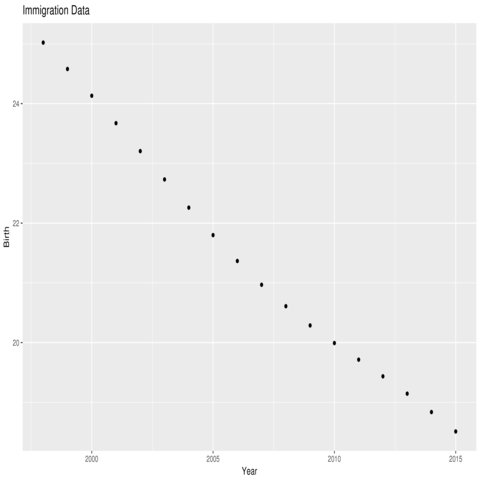
Multiple R-squared: 0.9403, Adjusted R-squared: 0.8731

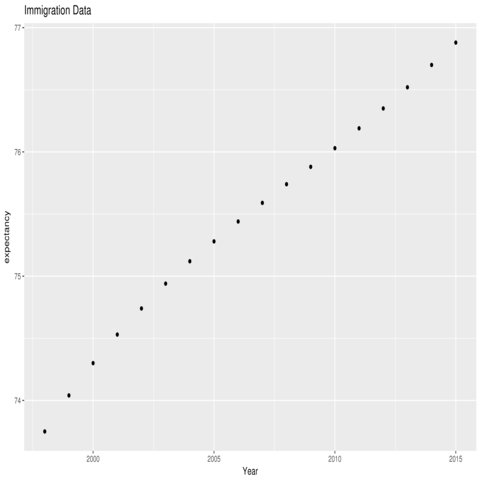
F-statistic: 14 on 9 and 8 DF, p-value: 0.0005382

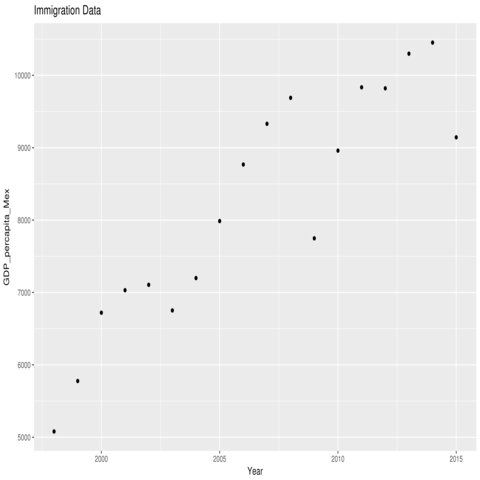


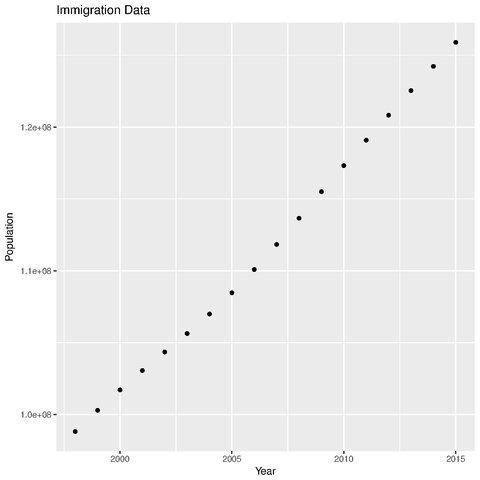
Graphs

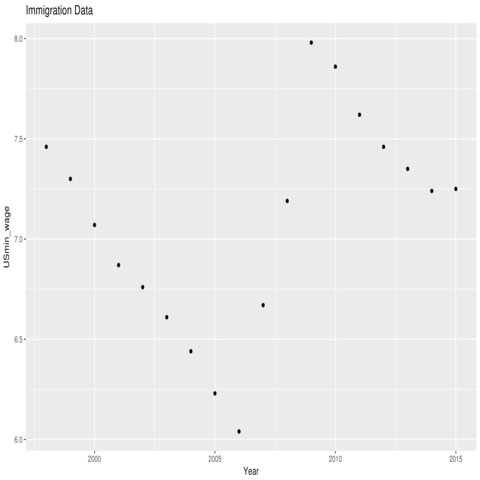


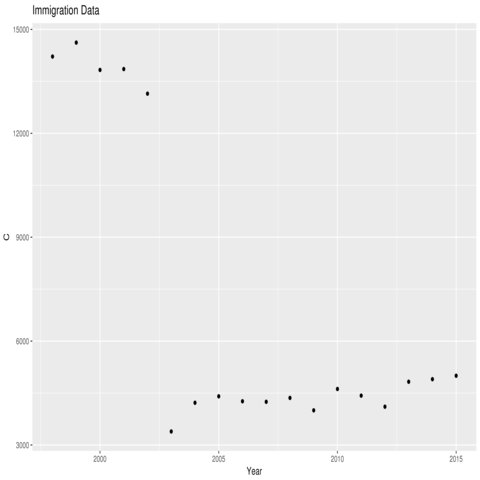


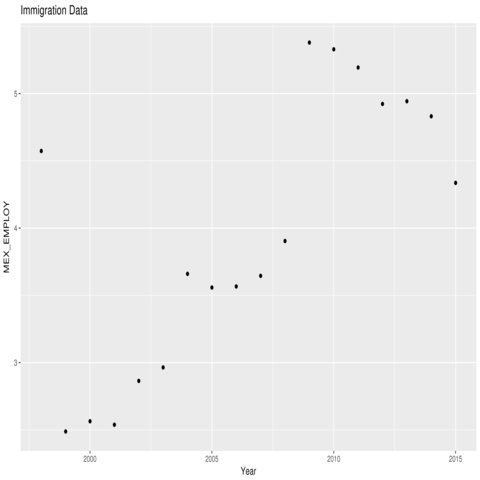












## References

Migration Policy Institute. (2013, January 23). U.S. Immigration Trends. Retrieved from https://www.migrationpolicy.org/programs/data-hub/us-immigration-trends#history

Pew Research Center. (2016, November 3). Unauthorized immigrant population trends for states, birth countries and regions. Retrieved from http://www.pewhispanic.org/interactives/unauthorized-trends/

U.S. Department of Health and Human Services. (n.d.). CDC WONDER. Retrieved November 30, 2017, from https://wonder.cdc.gov/

U.S. Department of Justice. (n.d.). Crime in the U.S. Retrieved November 30, 2017, from https://ucr.fbi.gov/crime-in-the-u.s/

United Nations. (1998). Sixth United Nations Survey of Crime Trends and Operations of Criminal Justice Systems. Retrieved from http://www.unodc.org/pdf/crime/sixthsurvey/publication\_by\_country\_screen.pdf

United Nations. (2001). Seventh United Nations Survey of Crime Trends and Operations of Criminal Justice Systems. Retrieved from http://www.unodc.org/pdf/crime/seventh\_survey/7sc.pdf

United Nations. (2003). Eighth United Nations Survey of Crime Trends and Operations of Criminal Justice Systems. Retrieved from http://www.unodc.org/pdf/crime/eighthsurvey/8sc.pdf

United Nations. (2017). UNODC Statistics Online. Retrieved November 28, 2017, from https://data.unodc.org/

United States Census Bureau. (August). International Programs - Information Gateway - U.S. Census Bureau. Retrieved November 15, 2017, from https://www.census.gov/population/international/data/idb/informationGateway.php

World Bank Group. (2017). Mexico | Data. Retrieved November 15, 2017, from https://data.worldbank.org/country/mexico

World Bank Group. (2017). Population, total | Data. Retrieved November 28, 2017, from https://data.worldbank.org/indicator/SP.POP.TOTL?locations=US

## 

## 

## Source Code

### R Source Code

#!/usr/bin/Rscript  
  
data <- read.csv("data.csv")  
  
  
library(ggplot2)  
  
  
base\_plot <- ggplot(data, aes(x=Year, y=C))  
  
base\_plot + geom\_point()+ ggtitle("Immigration Data")  
  
  
  
ggsave("C.png",dpi=900,width=10)

#!/usr/bin/Rscript

Data <- read.csv("data.csv")

model <- lm(Immigration ~Population + expectancy + Birth + GDP\_percapita\_Mex + USmin\_wage, data \

= Data)

summary(model)

setwd("~/")

home <- getwd()

png(filename="firstOne.png")

par(mfrow=c(2,2))

plot(model)

dev.off()

#!/usr/bin/Rscript

Data <- read.csv("data.csv")

model <- lm(Immigration ~ C + I(C^2) + I(C^3) + MEX\_EMPLOY + I(MEX\_EMPLOY^2) + USmin\_wage + I(U\

Smin\_wage^2) + DHS, data = Data)

summary(model)

newdata = data.frame(C=11145,MEX\_EMPLOY=3.618,USmin\_wage=6.93,DHS=0,FIN=0)

predict(model, newdata, interval="predict")

setwd("~/")

home <- getwd()

png(filename="lastOne.png")

par(mfrow=c(2,2))

plot(model)

dev.off()

### 

### 

### Matlab Source Code

%Matlab code to compute the predicted y

years=[1997

1998

1999

2000

2001

2002

2003

2004

2005

2006

2007

2008

2009

2010

2011

2012

2013

2014

2015];

x1 = MexicoMidyearpopulation;

x2 = MexicoBirthsPerthousandpeople;

x3 = MexicoLifeExpectancy;

x4 = MexicoInfantmortalityrateper1000births;

x5 = MexicoDeathsperthousand;

x6 = MexicoGDPpercapita;

x7 = MexicoInflationconsumerprice;

x8 = MexicoCPI2010100;

x9 = MexicoTotalUnemploymentoftotallaborforce;

x10 = MexicoTotalLaborforceParticipationtotalpop15;

x11 = MexicoAssaultsperthousand;

x12 = MexicoRobberiesperthousand;

x13 = MexicoTheftsperthousand;

x14 = MexicoIntentionalHomicidesperthousand;

x15 = USMidyearpopulation;

x16 = USBirthsPerthousandpeople;

x17 = USLifeexpectancyatbirthyears;

x18 = USInfantmortalityrateper1000births;

x19 = USDeathsperthousand;

x20 = USGDPpercapita;

x21 = USInflationconsumerprice;

x22 = USCPI2010100;

x23 = USTotalUnemploymentoftotallaborforce;

x24 = USTotalLaborforceParticipationtotalpop15;

x25 = USIntentionalHomicidesperthousand;

x26 = USAssaultsperthousand;

x27 = USRobberiesperthousand;

x28 = USTheftsperthousand;

x29 = TotalLegalMigrantsFromMexico;

x30 = TotalMigrantsGrantedCitizenship;

x31 = TotalMigrantsGrantedCitizenshipFromMexico;

x32 = UnauthorizedimmigrantsFromMexico;

x33 = MexicanTotalUnemploymentPercent;

y = TotalLegalMigrantsFromMexico;

%%

N = length(y);

% Compute the regression coefficients for a linear model with an

% interaction term.

X = [ones(size(x1)) x9 x3];

[b] = regress(y,X);

prediction=zeros(1,N);

for i=1:N

prediction(i)= b(1) + b(2)\*x9(i) + b(3)\*x3(i);

end

plot(years,prediction);

## Data

### immigration.csv

Year,Population of Mexico,Total Immigration, Birth Rates in Mexico(per 1000 people),Life expectancy,US Min. Wage Adjusted for inflation(2015 dollars),GDP per capita(Mexico)

1997,97281739,146865,25.461,73.44,7.58,4939.82

1998,98821456,131575,25.022,73.75,7.46,5079.97

1999,100300579,147573,24.58,74.04,7.30,5777.23

2000,101719673,173919,24.132,74.30,7.07,6720.90

2001,103067068,204844,23.673,74.53,6.87,7031.38

2002,104355608,217318,23.205,74.74,6.76,7106.08

2003,105640453,114984,22.731,74.94,6.61,6752.00

2004,106995583,173664,22.259,75.12,6.44,7199.06

2005,108472228,157992,21.799,75.28,6.23,7986.80

2006,110092378,170042,21.366,75.44,6.04,8767.92

2007,111836346,143180,20.968,75.59,6.67,9330.34

2008,113661809,189989,20.609,75.74,7.19,9689.05

2009,115505228,164920,20.287,75.88,7.98,7748.12

2010,117318941,138717,19.992,76.03,7.86,8959.58

2011,119090017,142823,19.714,76.19,7.62,9834.47

2012,120828307,145326,19.435,76.35,7.46,9820.53

2013,122535969,134198,19.145,76.52,7.35,10298.87

2014,124221600,133107,18.837,76.70,7.24,10452.28

2015,125890949,157227,18.512,76.88,7.25,9143.13

### data2.csv

Year,Population,Immigration,Birth,expectancy,USmin\_wage,GDP\_percapita\_Mex,DHS,C,MEX\_EMPLOY,FIN

1998,98821456,131575,25.022,73.75,7.46,5079.97,0,14216,4.573,0

1999,100300579,147573,24.58,74.04,7.30,5777.23,0,14619,2.488,0

2000,101719673,173919,24.132,74.30,7.07,6720.90,0,13829,2.564,0

2001,103067068,204844,23.673,74.53,6.87,7031.38,0,13855,2.538,0

2002,104355608,217318,23.205,74.74,6.76,7106.08,0,13144,2.864,0

2003,105640453,114984,22.731,74.94,6.61,6752.00,1,3391,2.964,0

2004,106995583,173664,22.259,75.12,6.44,7199.06,0,4220,3.66,0

2005,108472228,157992,21.799,75.28,6.23,7986.80,0,4406,3.558,0

2006,110092378,170042,21.366,75.44,6.04,8767.92,0,4263,3.566,0

2007,111836346,143180,20.968,75.59,6.67,9330.34,0,4248,3.645,0

2008,113661809,189989,20.609,75.74,7.19,9689.05,0,4360,3.904,1

2009,115505228,164920,20.287,75.88,7.98,7748.12,0,4001,5.379,0

2010,117318941,138717,19.992,76.03,7.86,8959.58,0,4618,5.329,0

2011,119090017,142823,19.714,76.19,7.62,9834.47,0,4426,5.194,0

2012,120828307,145326,19.435,76.35,7.46,9820.53,0,4106,4.923,0

2013,122535969,134198,19.145,76.52,7.35,10298.87,0,4827,4.943,0

2014,124221600,133107,18.837,76.70,7.24,10452.28,0,7073,4.831,0

2015,125890949,157227,18.512,76.88,7.25,9143.13,0,8903,4.336,0

### FinalDataSpreadsheet3.csv

Year,Mexico Midyear population,Mexico Births,Mexico Births Per thousand people,Mexico Life Expectancy,"Mexico Infant mortality rate (per 1,000 births)",Mexico Deaths,Mexico Deaths per thousand,Mexico GDP (Billion USD),Mexico GDP per capita,Mexico Inflation (consumer price %),Mexico CPI (2010 = 100),Mexico Total Unemployment (% of total labor force),Mexico Total Laborforce Participation (% total pop 15+),Mexico Total Intentional Homicides,Mexico Total Assaults,Mexico Total Robberies,Mexico Total Thefts,Mexico Intentional Homicides per thousand,Mexico Assaults per thousand,Mexico Robberies per thousand,Mexico Thefts per thousand,US Midyear population,US Births,US Births Per thousand people,US Life expectancy at birth (years),"US Infant mortality rate (per 1,000 births)",US Deaths,US Deaths per thousand,US GDP (Billion USD),US GDP per capita,US Inflation (consumer price %),US CPI (2010 = 100),US Total Unemployment (% of total labor force),US Total Laborforce Participation (% total pop 15+),US Total Intentional Homicides,US Total Assaults,US Total Robberies,US Total Thefts,US Intentional Homicides per thousand,US Assaults per thousand,US Robberies per thousand,US Thefts per thousand,Total Legal Migrants From Mexico,Total Migrants Granted Citizenship,Total Migrants Granted Citizenship From Mexico,US minwage,DHSEvent,HousingCrashEvent

1997,95659000,2410000,25.19365663,72,28,466000,4.871470536,480.555,5023.625587,20.626,42.784,4.055,60.84,NA,NA,NA,NA,NA,NA,NA,NA,272657000,3880894,14.23361219,76.429,7.5,2314245,8.487752011,8609,31574.46902,2.338,73.613,4.9,67.1,18210,2419000,498530,7743800,0.05175201375,8.39349801,1.828414455,28.40125139,NA,NA,NA,7.58,0,0

1998,97133000,2397000,24.67750404,73,27,470000,4.838726283,502.01,5168.274428,15.928,49.598,3.573,60.95,"14,216",242144,301512,141225,0.1463560273,2.492911781,3.104114976,1.453934296,275854000,3941553,14.28854757,76.58,7.3,2337256,8.472800829,9089,32948.58875,1.552,74.755,4.5,67.09,14276,2315380,447190,7376300,0.05175201375,8.39349801,1.621111168,26.73986964,NA,NA,NA,7.46,0,0

1999,98469000,2379000,24.15988788,73,25,475000,4.823853192,579.46,5884.694675,16.586,57.825,2.488,60.04,"14,619",248643,265250,104413,0.148462968,2.525089114,2.693741177,1.060364176,279040000,3959417,14.18942446,76.583,7.2,2391399,8.570093893,9661,34622.27638,2.188,76.391,4.2,67.08,12658,2238480,409670,6957400,0.04536267202,8.022075688,1.468140768,24.93334289,147402,837418,207072,7.30,0,0

2000,99775000,2357000,23.62315209,73,24,480000,4.810824355,683.648,6851.896768,9.495,63.315,2.564,59.83,"13,829",255179,215120,98179,0.1386018542,2.557544475,2.156051115,0.984004009,282162400,4058814,14.38467351,76.637,7.1,2403351,8.517616096,10285,36450.63977,3.377,78.971,4,67.07,15517,910744,407842,6965957,0.05499315288,3.227729846,1.445415831,24.68775783,173493,886026,189051,7.07,0,0

2001,101125000,2347000,23.20889988,73,23,489000,4.835599506,724.704,7166.4178,6.363,67.343,2.538,58.82,13855,258783,157170,116013,0.1370086527,2.559040791,1.55421508,1.147223733,284969000,4025933,14.12761739,76.837,7,2416425,8.479606554,10622,37274.22983,2.826,81.203,4.7,66.82,16037,909023,423557,7092267,0.05627629672,3.189901358,1.486326583,24.88785447,205560,606259,102736,6.87,0,0

2002,102470000,2332000,22.75788036,73,22,499000,4.869717966,741.56,7236.84981,5.031,70.732,2.864,58.89,13144,253972,147775,113396,0.1282716893,2.478501025,1.442129404,1.10662633,287625200,4021726,13.98252309,76.937,6.9,2443387,8.495037987,10978,38167.72661,1.586,82.49,5.8,66.58,16204,894348,420637,7052922,0.0563372055,3.109421567,1.462448353,24.52122415,218822,572646,76310,6.76,0,0

2003,103843000,2323000,22.37030902,74,21,510000,4.911260268,713.284,6868.869351,4.548,73.99,2.964,58.43,3391,NA,NA,NA,0.03265506582,NA,NA,NA,290107900,4089950,14.09803042,77.037,6.9,2448288,8.439232437,11511,39678.34037,2.27,84.363,6,66.24,16528,859030,414235,7026802,0.05697190597,8.39349801,1.427865287,24.22133972,115585,462435,55946,6.61,1,0

2004,105231000,2314000,21.98971786,74,19,520000,4.941509631,770.268,7319.782193,4.688,77.415,3.66,59.91,4220,246338,514922,83143,0.04010225124,2.340926153,4.893253889,0.7900998755,292805300,4112052,14.04363924,77.488,6.9,2397615,8.1884276,12275,41922.05537,2.677,86.622,5.5,65.99,16148,847381,401470,6937089,0.05514927496,2.894008408,1.371115892,23.691815,175411,537151,63840,6.44,0,0

2005,106576000,2302000,21.59960967,74,18,531000,4.982360006,866.346,8128.903318,3.988,80.503,3.558,58.87,4406,239166,515916,84827,0.04134139018,2.244088725,4.840827203,0.7959296652,295516600,4138349,14.00377847,77.488,6.8,2448017,8.283856135,13094,44308.84762,3.393,89.561,5.1,66.05,16740,862220,417438,6783447,0.05664656402,2.917670276,1.412570394,22.95453792,161445,604280,77089,6.23,0,0

2006,107916000,2295000,21.26654064,74,17,543000,5.031691315,965.281,8944.74406,3.629,83.425,3.566,59.85,4263,245507,545232,85775,0.03950294674,2.274982394,5.052374069,0.794831165,298379900,4265555,14.29571831,77.688,6.7,2426264,8.131459257,13856,46437.44435,3.226,92.45,4.6,66.18,17309,874096,449246,6626363,0.05800993968,2.929473467,1.505617503,22.20780622,173749,702589,83979,6.04,0,0

2007,109339000,2291000,20.9531823,74,17,554000,5.066810562,1043,9539.13974,3.967,86.734,3.645,60.02,4248,261295,610730,99536,0.03885164488,2.389769433,5.585655622,0.9103430615,301237200,4316233,14.3283532,77.988,6.6,2423712,8.045858878,14478,48061.79316,2.853,95.087,4.6,66.04,17128,866358,447324,6591542,0.05685884745,2.875999379,1.484956041,21.88156708,148640,660477,122258,6.67,0,0

2008,110845000,2287000,20.63241463,74,16,566000,5.106229419,1101,9932.78903,5.125,91.179,3.904,60.42,4360,250932,656877,105398,0.03933420542,2.263809825,5.926085976,0.950859308,304094000,4247694,13.96835847,78.039,6.5,2471984,8.129012739,14719,48402.7965,3.839,98.737,5.8,65.99,16465,843683,443563,6586206,0.05414444218,2.774415148,1.45863779,21.65845429,189989,1046539,231815,7.19,0,1

2009,112426000,2283000,20.30669062,74,15,580000,5.158948998,894.949,7960.338356,5.279,96.009,5.379,60,4001,244623,680566,132345,0.03558785334,2.175857898,6.053457385,1.177174319,306771500,4130665,13.46495682,78.39,6.4,2437163,7.944554823,14419,47002.41059,-0.356,98.386,9.3,65.37,15399,812514,408742,6338095,0.05019697071,2.648596757,1.33239887,20.66063829,164920,743715,111630,7.98,0,0

2010,114061000,2279000,19.98053673,75,14,593000,5.19897248,1051,9214.367751,4.157,100,5.329,59.71,4618,230687,738138,132068,0.04048710778,2.022487967,6.471431953,1.157871665,309348200,3999386,12.92842822,78.541,6.2,2468435,7.97947103,14964,48372.67519,1.64,100,9.6,64.71,14722,781844,369089,6204601,0.0475903852,2.527391464,1.193118305,20.05701342,139120,619913,67062,7.86,0,0

2011,115701000,2272000,19.6368225,75,14,602000,5.203066525,1171,10120.91512,3.407,103.407,5.194,59.62,4426,212141,751750,135171,0.03825377482,1.8335278,6.49735093,1.16827858,311663400,395359,1.268544847,78.641,6.1,2515458,8.07107283,15518,49790.8962,3.157,103.157,8.9,64.11,14661,752423,354746,6151095,0.04704113476,2.414216748,1.13823439,19.73634055,143446,694193,94783,7.62,0,0

2012,117299000,2285000,19.48013197,75,13,610000,5.20038534,1187,10119.43836,4.112,107.659,4.923,60.87,4106,211921,746894,92659,0.03500456099,1.806673544,6.367437063,0.7899385332,313998400,3952841,12.58872975,78.741,6,2543279,8.099655922,16155,51449.30675,2.069,105.292,8.1,63.7,14856,762009,355051,6168874,0.04731234299,2.426792621,1.130741431,19.64619565,146406,757434,102181,7.46,0,0

2013,118818000,2287000,19.2479254,75,13,620000,5.218064603,1262,10621.28634,3.806,111.757,4.943,60.51,4827,211714,728762,90262,0.0406251578,1.781834402,6.133430962,0.7596660439,316204900,3932181,12.43554733,78.741,5.9,2596993,8.213006819,16692,52788.55578,1.465,106.834,7.4,63.25,14319,726777,345093,6019465,0.04528392824,2.298436868,1.091358799,19.03659621,135028,779929,99385,7.35,0,0

2014,120287000,2288000,19.02117436,75,13,630000,5.237473709,1298,10790.85853,4.019,116.248,4.831,57.77,NA,39937,177990,175490,NA,0.3320142659,1.479711024,1.458927399,318563500,3988076,12.51893579,78.741,5.8,2626418,8.244566625,17393,54598.21982,1.622,108.567,6.2,62.89,14164,731089,322905,5809054,0.05175201375,2.294955323,1.013628366,18.23515249,134052,653416,94889,7.24,0,0

2015,121737000,2286000,18.77818576,76,12,640000,5.257234859,1151,9454.808316,2.721,119.411,4.336,59.81,NA,45452,164063,144999,NA,0.3733622481,1.347683942,1.191084058,320896600,3978497,12.3980653,78.741,5.7,2712630,8.453283706,18037,56208.13683,0.119,108.696,5.3,62.65,15696,764449,327374,5706346,0.05175201375,2.382228419,1.020185318,17.78250689,158619,730259,105958,7.25,0,0