

**University of Washington**

Information School

**Study – US Immigrants 1980-2013**

Shipra Gupta

04/23/2017

# The Data

The dataset used in this visualization is taken from the website of United Nations’ Department of Economic and Social Affairs under population division,

<http://www.un.org/en/development/desa/population/migration/data/empirical2/migrationflows.shtml>

The dataset contains annual data, version published in 2015, on the flow of international migrants as recorded by the United States of America from 1980 to 2013. The data presents the inflow according to the place of birth, citizenship or place of previous / next residence both for foreigners and nationals. Immigration data are based on the number of foreigners granted permanent residence status and are reported by fiscal year. The United States does not collect data on immigrating citizens or on emigration.

This compilation of data on international migration flows is useful for analytical purposes, but also serves to raise awareness about the problems of comparability among available statistics.

The data was originally in excel format, but compiled to CSV for the purpose of this project. This CSV file contains 220 rows and represents data of 220 countries on the following 40 variables:

## Variables

* COUNTRY – the name of the country
* AREA – the code of the country
* AreaName – the continent or sub-continent that the country belongs to
* RegName – the region of the country
* DEV – development identifier code
* DevName – Less developed area or More developed area
* 1980 – Year 1980
* 1981 – Year 1981
* 1982 – Year 1982
* 1983 – Year 1983
* 1984 – Year 1984
* 1985 – Year 1985
* 1986 – Year 1986
* 1987 – Year 1987
* 1988 – Year 1988
* 1989 – Year 1989
* 1990 – Year 1990
* 1991 – Year 1991
* 1992 – Year 1992
* 1993 – Year 1993
* 1994 – Year 1994
* 1995 – Year 1995
* 1996 – Year 1996
* 1997 – Year 1997
* 1998 – Year 1998
* 1999 – Year 1999
* 2000 – Year 2000
* 2001 – Year 2001
* 2002 – Year 2002
* 2003 – Year 2003
* 2004 – Year 2004
* 2005 – Year 2005
* 2006 – Year 2006
* 2007 – Year 2007
* 2008 – Year 2008
* 2009 – Year 2009
* 2010 – Year 2010
* 2011 – Year 2011
* 2012 – Year 2012
* 2013 – Year 2013

## Sample Data

Country,AREA,AreaName,RegName,DEV,DevName,1980,1981,1982,1983,1984,1985,1986,1987,1988,1989,1990,1991,1992,1993,1994,1995,1996,1997,1998,1999,2000,2001,2002,2003,2004,2005,2006,2007,2008,2009,2010,2011,2012,2013

Afghanistan,935,Asia,Southern Asia,902,Less developed regions,722,1881,1569,2566,3222,2794,2831,2424,2873,3232,3187,2879,2685,2964,2344,1424,1263,1127,830,877,1011,1202,1759,1252,2137,4749,3417,1753,2813,3165,2017,1648,1617,2196

Albania,908,Europe,Southern Europe,901,More developed regions,30,11,23,22,32,45,53,62,82,71,78,142,682,1400,1489,1420,4007,4374,4220,3695,4755,4358,3765,3362,3840,5947,7914,5737,5754,5137,4711,3612,3364,3186

Algeria,903,Africa,Northern Africa,902,Less developed regions,175,184,190,201,197,202,183,172,199,230,302,269,407,360,364,650,1059,717,804,789,906,875,1030,759,805,1115,1300,1036,1037,1485,1305,1364,1369,1241

# Audiences

The Analysis of this dataset and visualizations generated can be helpful for the following audiences:

1. **U.S. Government**

* **Trend Analysis**

The U.S. Government is one of the most important audiences of this visualization as they would use this in order to perform a trend analysis on the migration of people from different countries to the U.S. in the period of 33 years.

* **Comparison of inflow of population from different parts of the World**

The U.S. Government would be interested in answering the question of the proportion of the incoming people from different parts of the world, including the countries, continents and the type of country (less developed/more developed).

* **Number of Permanent Residencies Allotted During these 33 years**

The U.S. Government would be highly interested in knowing exactly how many permanent residencies have been allotted to people during the period of 1980-2003.

1. **United Nations**

* **To Better Understand international co-operation among US and other countries**

The United Nations is an intergovernmental organization to promote international cooperation. The visualization can be highly useful to the U.N. in understanding the migration patterns of people from different parts of the world to the U.S. and also understanding the collaboration among the different countries.

1. **All other countries’ governments that experience outflow of their population**

* **Understand outflow patterns of people to the U.S.**

Another audience for this visualization would be the governments of all the other countries whose people’s migration to the U.S. is being analyzed here. This may be helpful in knowing how many people have been considering permanent residencies in the U.S. during these years, as opposed to staying in their country of original residence.

# The Motivation/Questions

The visualizations and the analysis of this dataset can help answer the following questions for the following audiences:

1. **U.S. Government**

* Which countries have highest inflow of people to the U.S.?
* From which continent does the U.S. get the most number of immigrants?
* Whether U.S. is a preferred destination of permanent residence among developed nations’ population or under-developed nations’ population?
* Which year saw the highest number of migrants to the U.S.?

1. **United Nations**

* Are there any countries/regions to which the U.S. government do not provide any/very few permanent residencies in the U.S.?
* Impact of the total population in the U.S. by the inflow of people from different regions of the world.

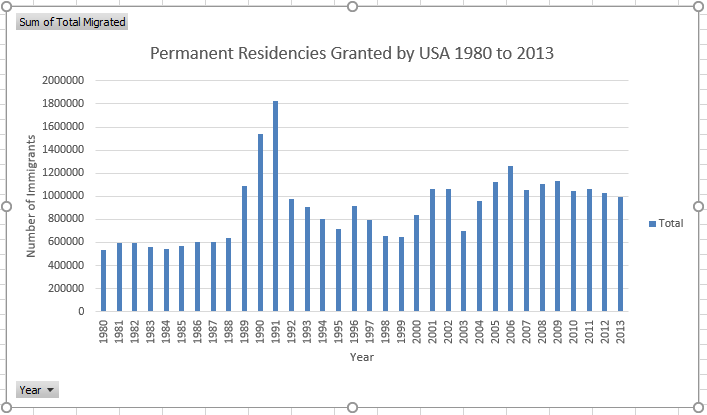
1. **Other Countries**

* How many people have been moving out of a particular country to the U.S. for their place of permanent residence?
* In what years was the outflow of population from their country to the U.S. was the highest.
* They can utilize this information to perform a greater analysis on why people decided to move to the U.S. during these years, i.e., the probable underlying causes.

# The Sketches

## Sketch 1 – Number of Permanent Residencies granted by USA from 1980 to 2013

### Rough Sketch (Created in Excel)

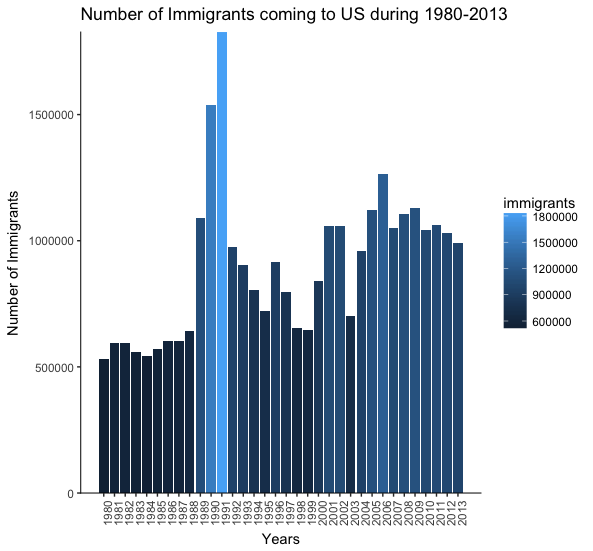


### R Code

install.packages("ggplot2")  
library(ggplot2)  
immigrants <- read.csv(file = "Immigrants-USA.csv", stringsAsFactors = FALSE)  
sum1980 <- sum(immigrants$X1980)  
sum1981 <- sum(immigrants$X1981)  
sum1982 <- sum(immigrants$X1982)  
sum1983 <- sum(immigrants$X1983)  
sum1984 <- sum(immigrants$X1984)  
sum1985 <- sum(immigrants$X1985)  
sum1986 <- sum(immigrants$X1986)  
sum1987 <- sum(immigrants$X1987)  
sum1988 <- sum(immigrants$X1988)  
sum1989 <- sum(immigrants$X1989)  
sum1990 <- sum(immigrants$X1990)  
sum1991 <- sum(immigrants$X1991)  
sum1992 <- sum(immigrants$X1992)  
sum1993 <- sum(immigrants$X1993)  
sum1994 <- sum(immigrants$X1994)  
sum1995 <- sum(immigrants$X1995)  
sum1996 <- sum(immigrants$X1996)  
sum1997 <- sum(immigrants$X1997)  
sum1998 <- sum(immigrants$X1998)  
sum1999 <- sum(immigrants$X1999)  
sum2000 <- sum(immigrants$X2000)  
sum2001 <- sum(immigrants$X2001)  
sum2002 <- sum(immigrants$X2002)  
sum2003 <- sum(immigrants$X2003)  
sum2004 <- sum(immigrants$X2004)  
sum2005 <- sum(immigrants$X2005)  
sum2006 <- sum(immigrants$X2006)  
sum2007 <- sum(immigrants$X2007)  
sum2008 <- sum(immigrants$X2008)  
sum2009 <- sum(immigrants$X2009)  
sum2010 <- sum(immigrants$X2010)  
sum2011 <- sum(immigrants$X2011)  
sum2012 <- sum(immigrants$X2012)  
sum2013 <- sum(immigrants$X2013)  
year <- c(1980, 1981, 1982, 1983, 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013)  
immigrants <- c(sum1980, sum1981, sum1982, sum1983, sum1984, sum1985, sum1986, sum1987, sum1988, sum1989, sum1990, sum1991, sum1992, sum1993, sum1994, sum1995, sum1996, sum1997, sum1998, sum1999, sum2000, sum2001, sum2002, sum2003, sum2004, sum2005, sum2006, sum2007, sum2008, sum2009, sum2010, sum2011, sum2012, sum2013)  
numberOfImmigrants <- data.frame(year, immigrants)  
ggplot(numberOfImmigrants, aes(year, immigrants)) +  scale\_y\_continuous(expand = c(0,0)) +  
  geom\_bar( aes(fill = immigrants),position = "dodge", stat="identity") +  
  scale\_x\_continuous(breaks=year) +  theme(axis.text.x=element\_text(angle=90,hjust=1),   
        panel.background = element\_blank()) + theme(axis.line.x = element\_line(color="black"),  
        axis.line.y = element\_line(color="black"))+  xlab("Years") + ylab("Number of Immigrants") +   
  ggtitle("Number of Immigrants coming to US during 1980-2013")

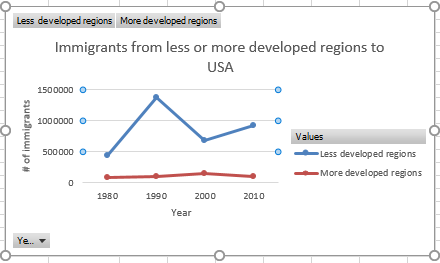
The following visualization illustrates the number of immigrants to the U.S. in each year from 1980-2013. The barplot is able to showcase a trend of the inflow of population to the U.S. for becoming permanent residents. The visualization shows that while the number of permanent residencies given to the immigrants from the years 1980-1988 were almost constant, they saw a huge increase from 1989 and in the following two years. From 2000 onwards, the number pf permanent residencies granted have remained high constantly.

### R – Visualization



## Sketch 2 – Number of immigrants to USA – by the country’s development status

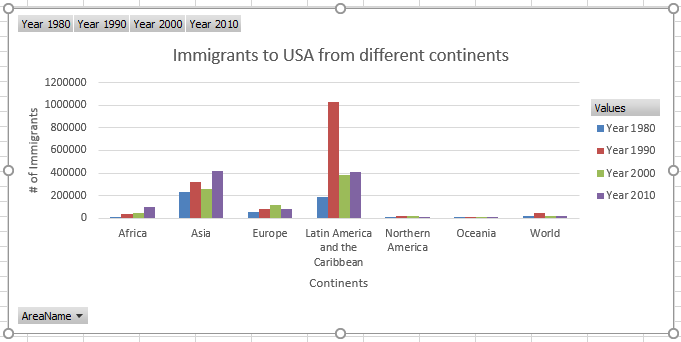
### Rough Sketch



The above line graph illustrates how popular has been U.S. as a destination for permanent residency, when compared among the less developed and more developed regions of the world. The graph is built by comparing the statistics of population migration after every 10 years and is clearly able to depict that the number of immigrants to the U.S. from the more developed regions of the world have been very less consistently. On the other hand, when we look at the migration from the less developed regions of the world, it saw a sharp rise from the year 1980 to 1990 but then gradually declined over the course of next 10 years. To conclude, many more immigrants come to the U.S. from the less developed regions of the world as compared to the more developed regions.

## Sketch 3 – Number of immigrants to USA by continents

### Rough Sketch



### R Code

install.packages("ggplot2")

library(ggplot2)

library(dplyr)

library(reshape2)

immigrants <- read.csv(file = "Immigrants-USA.csv", stringsAsFactors = FALSE)

immigrantsSummary<- immigrants %>%

group\_by(AreaName) %>%

summarise(X1980 = sum(X1980), X1990 = sum(X1990), X2000 = sum(X2000),

X2010 = sum(X2010))

immigrantsSummary

names(immigrantsSummary) <- c('AreaName','1980','1990', '2000', '2010')

meltedSummary <- melt(immigrantsSummary[,c('AreaName','1980','1990', '2000', '2010')],id.vars = 1)

names(meltedSummary) <- c('AreaName','Year','Immigrants')

ggplot(meltedSummary,aes(x = AreaName,y = Immigrants)) +

geom\_bar(aes(fill = Year),stat = "identity",position = "dodge")+

theme(axis.text.x=element\_text(angle=90,hjust=1),

panel.background = element\_blank()) +

theme(axis.line.x = element\_line(color="black"),

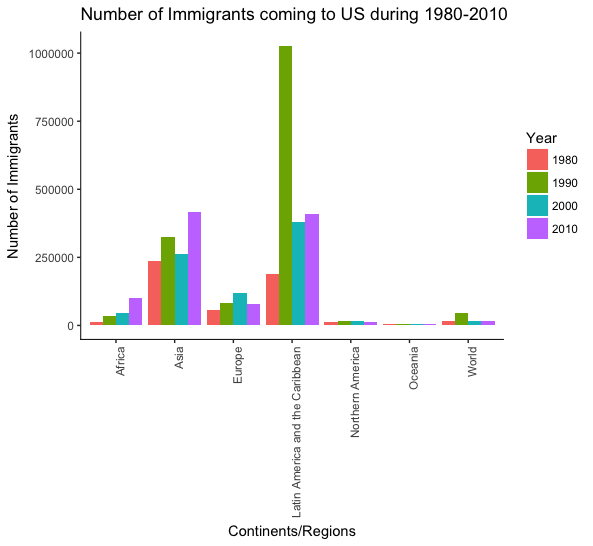
axis.line.y = element\_line(color="black"))+

xlab("Continents/Regions") + ylab("Number of Immigrants") +

ggtitle("Number of Immigrants coming to US during 1980-2010")

The following visualization is a grouped bar chart that depicts the number of immigrants coming to the U.S. from different continents in the world in the years 1980, 1990, 2000 and 2010. This visualization can clearly show which continent had the highest number of people coming to the U.S. in a particular year. Different colors have been used for each year to make easier comparisons. In this visualization, “World” represents all other regions of the world, except the other continents listed in the dataset.

### R – Visualization



# The Visualization

The sketch chosen to generate a refined visualization based on it using R and Illustrator is the “Grouped Bar Plot” for the following reasons:

* The information provided in the grouped barplot is very helpful in making comparisons among different continents/regions of the world while understanding the migration pattern of population to the US.
* This visualization is able to provide a lot of important information and answer several questions through the single visualization.
* This visualization will help answer some of the questions of all our audiences including the U.S. Government, the United Nations as well as other countries in the world whose people have migrated to the U.S.

