Nationality-Based Discrimination in Mexico's Refugee Determination Process

Juan C. Venancio

Introduction

The 1951 United Nations Refugees Convention establishes the fundamental framework of international asylum law for member states of the United Nations. Since 1951, member states have met in other conventions to re-contextualize and re-interpret this legal framework in new correspondence to their national interests and needs (OAU Convention 1969 & 1984 Cartagena Declaration), but the legal principles establishing the definition of a refugee remain identical since its conception. According to the 1951 Convention, a refugee is any person that based on a well-founded fear of persecution on the basis of their race, religion, nationality, membership to a particular social group, or political opinion, merits the international protection of a sovereign entity outside of their country of origin. As it stands, in 2022, approximately 108.4 million people were forcibly displaced, and of those forcibly displaced, 35.3 million were categorized as refugees. According to the UNHCR, 76% of the world's refugees are hosted in low and middle-income countries across the world.

Since its inception, Mexico has not historically been a destination country for migrants and refugees. At most, the country became a transit country for Central and South American migrants that headed to the United States during periods of increased generalized drug violence, civil wars, and economic destitution between the late 1980s and early 2000s. Over the last decade alone, the number of refugee applications submitted to and processed by the Mexican government has exponentially increased, and this data is indicative of Mexico's evolving role as a destination country for many immigrants. However, given the exponential increase in applications and Mexico's limited institutional capacity as a middle-income country, questions have emerged whether COMAR, Mexico's refugee processing institution, selective prioritizes applications of certain nationalities and therefore discriminates against others. I would like to explore this phenomenon, and compile/analyze data that uncover potential discriminatory institutional practices on the basis of Hispanic and non-Hispanic categorizations, as well as suggest potential policy recommendations. I am planning to use data from the Mexican Commission for Refugee Assistance.

Project Abstract & Methodology (Data Section)

Research Question: To What Extent Does Hispanic Nationality Impact Refugee Status Determination in Mexico?

Hypothesis: I hypothesize, that migrants who indicate a Hispanic nationality tend to have higher possibilities to successfully obtain refugee status in Mexico. As a result, I expect that migrants from Haiti, the non-Spanish Caribbean, and other parts of the world tend to have lower possibilities at refugee resettlement. Additionally, though the Mexican Government has made advances to expand their refugee granting criteria, because the 1951 UN Convention of Refugees provides a large number of successful reugee claims to migrants with "well-founded fear of persecution", then I also predict that Mexican institutions accept Spanish-speaking Latin-Americans at a significant larger rate. For this project, I will analyze the Mexican Government's Commission on Refugee Assistance open data set that I will use to assess the several reasons that migrants tend to leave their countries of origin. If Venezuelans, Hondurans, and other spanish speaking Latin Americans have a higher rate of refugee success, then that will make my hypothesis correct. If other types of nationalities have a higher rate of success, then I will be proven wrong. In any case, I am also interested in observing the variability of application success and the type of application submitted.

Methodology: My unit of measurement is whether an immigrant was granted refugee status or not. If the immigrant was provided with a refugee status, implying that the migrant was considered to have had a well-founded fear of persecution, then I will indicate that with 1 for successful application. If a migrant was rejected, then I will indicate that with 0. The independent variable will be the migrant's nationality. The explanatory variable of interest is whether the nationality impacts the likelihood that a migrant gets refugee status in Mexico.

To test this hypothesis, I used the Mexican Government's Commission on Refugee Assistance (COMAR) open data set on refugee applications to conduct a statistical data analysis and produce a linear regression model assessing refugee status determination in Mexico between 2013 and 2023. I downloaded the dataset and coded my visualizations, results, and subsequent models using R.

For this study, the independent variable is whether a refugee applicant's nationality was of Hispanic origin. The dependent variable was the outcome of a migrant's refugee application. The dataset provided by COMAR is composed of 172,053 refugee applications submitted since 2013. The data is separated into the 6 following categories: Year, Nationality, Sex, Age, Migration Motive, and Application Outcome. I created a seventh column in the dataset called "Spanish Speaking", indicating whether Spanish is the official language in the applicant's country of origin. To specifically signal this in R code, if the refugee application received a positive result, then I indicated that with "1" for successful application. If a refugee application had a negative result, then that was indicated with "0". The explanatory variable of interest is whether a nationality of Hispanic origin impacts the likelihood that a migrant gets refugee status in Mexico. If my hypothesis stands true, Venezuelans, Hondurans, and other Hispanic migrants will have a higher rate of refugee success. Conversely, migrants from non-Hispanic origin such as, but not limited to, Haitians, Brazilians, and other nationalities will subsequently have a lower rate of refugee application success.

Using this setup, I first coded a table calculating the percentage of positive resolutions for refugees in the top 6 most-prevalent Hispanic nationalities and top 5 most prevalent non-Hispanic nationalities present in the dataset. I chose to use 6 entries in the Hispanic nationality type and 5 non-Hispanic nationalities because there were missing variables in other entries. Then, I assess statistical differences of the top 6 Hispanic nationalities and the top non-Hispanic nationalities applying for refugee status in Mexico. In this statistical comparison I calculated the percentage of positive resolutions of all applications per nationality between 2013 and 2023. Next, I conducted a deep-dive comparison between Haitian and Venezuelan migrants applying for refugee status. I specifically chose to isolate these Hispanic and non-Hispanic cases specifically because of their similar political and economic status contexts. Additionally, because migrants from both countries submitted the largest volume of applications of both variable groups, it made comparing the dependent variable easier in this deep-dive analysis. Then, I conducted a linear regression analysis to study the relationship between Hispanic nationality and refugee application outcome of all 172,053 applications. In this linear regression, I used an applicant's age and sex as control group variables because of the data randomization in this natural experiment. Lastly, I coded a graphical model depicting the statistical outcomes of the linear regression.

```r
#set up all libraries needed for this project
library(tidyverse, ggplot2)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.3     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.3     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```r
#import data needed
submitted_applications <- read.csv("//Users//jvenancio//Downloads//COMAR_Thesis_Data_Solicitudes.csv")
application_results <- read.csv("//Users//jvenancio//Downloads//COMAR_Thesis_Data_Resoluciones.csv")
```

Results

Results Summary

After thoroughly assessing the results of the statistical analysis and linear regression models, there is statistically significant evidence that demonstrates that refugee applicants of Hispanic origin are on average 31% more likely to be granted refugee status in Mexico. Out of any nationality, Venezuelans have overwhelmingly been granted refugee status in Mexico over the last decade. More specifically, nearly 95% of all refugee applicants from Venezuela have been approved. Out of 17,246 processed Venezuelan applications, 16,366 received positive outcomes. Conversely, Haitian applicants have the highest rejection total out of all non-Hispanic immigrants that apply for refugee status. Out of 12,676 Haitian applicants that submitted refugee applications throughout the last decade, only 2190 have ever been approved. Furthermore, whereas Venezuelans, Hondurans, and Salvadorans had the highest approval rates among the top Hispanic applicant nationalities, Indians, Brazilians, and Haitians had the lowest rates. It is important to note that among the top Hispanic and Non-Hispanic nationalities evaluated, there is variability in the percentage of positive application outcomes for each nationality. For example, whereas Venezuela applicants experience a 95% refugee admissions rate, Central American countries like Guatemala and Nicaragua have nearly a 50% refugee admissions rate. On the contrary, the highest approval rating of any non-Hispanic nationality was 50%, and the majority had an approval rating below 40%. Despite the variability per individual nationality, on average, the percentage of positive resolutions among the Hispanic applicant pool is considerably higher than the non-Hispanic demographic, ultimately proving a strong correlation between Hispanic nationality and application outcome.

```r
#count how many applications were submitted per nationality + create data table
apps_sub_per_nationality <- submitted_applications |>
  count(NACIONALIDAD) |>
  rename(applications = n, Nationality = NACIONALIDAD)
```

```r
#count how many applications were submitted per year + create data table
apps_sub_per_year <- submitted_applications |>
  count(AÑO) |>
  rename(applications = n, year = AÑO)

# Assuming 'submitted_applications' is your data frame
```

```r
#count how many applications were processed per year + create data table
apps_proc_per_year <- application_results |>
  count(AÑO_RESOLUCION) |>
  rename(applications = n, year = AÑO_RESOLUCION)
```

```r
#count the end result of the applications that were processed in total from 2013-23
apps_proc_result <- application_results |>
  count(SENTIDO_RESOLUCIÓN)

# Print the data table
print(apps_sub_per_year)
```

```
##    year applications
```

```
## 1  2013          911
## 2  2014         1370
## 3  2015         2105
## 4  2016         5002
## 5  2017         9594
## 6  2018        18325
## 7  2019        41063
## 8  2020        27977
## 9  2021        74489
## 10 2022        82326
## 11 2023        68205
```

```r
#count the end result of the applications that were processed (positive) in total from 2013-23 per nati
apps_proc_result_pos <- application_results |>
  filter(SENTIDO_RESOLUCIÓN == "Resolución Positiva")

table(apps_proc_result_pos$NACIONALIDAD)
```

```
## 
##                      Afganistán                          Angola 
##                               6                               6 
##                         Argelia                       Argentina 
##                               1                               6 
##                      Bangladesh                          Belice 
##                               7                               4 
##                           Benín                          Brasil 
##                               6                               4 
##                    Burkina Faso                         Camerún 
##                              14                              44 
##                            Chad                           Chile 
##                               1                               6 
##                           China                        Colombia 
##                               1                             257 
##                           Congo                 Costa de Marfil 
##                              16                               2 
##                      Costa Rica                            Cuba 
##                               5                            4001 
##                         Ecuador                          Egipto 
##                              59                               4 
##                     El Salvador                     EL SALVADOR 
##                            8470                              19 
##                         Eritrea                          España 
##                               1                               3 
##       Estados Unidos de América                         Etiopia 
##                               1                               2 
##             Federación de Rusia                         Georgia 
##                              48                               1 
##                           Ghana                         Granada 
##                              38                               1 
##                       Guatemala                       GUATEMALA 
##                            2211                              10 
##                          Guinea               Guinea Ecuatorial 
##                              38                               1 
##                   Guinea-Bissau                           Haití
```

```
##                              1                                2190
##                       Honduras                           Hong Kong
##                          27115                                   1
##                        Hungría                               India
##                              1                                   4
##                           Irak                                Irán
##                             12                                  10
##                         Israel                             Jamaica
##                              2                                   7
##                     Kirguistán                                Mali
##                              1                                   2
##                      Marruecos                          Mauritania
##                              5                                   6
##                          Nepal                           Nicaragua
##                              1                                1291
##                      NICARAGUA                             Nígeria
##                              3                                  34
##                  Países Bajos                            Pakistán
##                              1                                   8
##                     Palestina                              Panamá
##                              5                                   3
##                           Perú            República de Kazajstán
##                             14                                   1
##          Republica de Turquia República Democrática del Congo
##                              5                                  13
##          República Dominicana      República Islámica de Irán
##                             25                                   1
##                          Rusia                              Rwanda
##                              4                                   1
##                        Senegal                        Sierra Leona
##                              3                                   9
##                          Siria                             Somalia
##                             32                                   2
##                      Sri Lanka                            Tanzania
##                              3                                   1
##                     Tayikistán                                Togo
##                              1                                  18
##                        Ucrania                              Uganda
##                             49                                   2
##                        Uruguay                          Uzbekistán
##                              4                                   1
##                      Venezuela                           VENEZUELA
##                          16366                                 198
##                          Yemen
##                             29
```

```r
apps_proc_result_neg <- application_results |>
  filter(SENTIDO_RESOLUCIÓN == "Resolución Negativa")

table(apps_proc_result_neg$NACIONALIDAD)
```

```
##
##                     Afganistán                             Albania
##                              1                                   4
```

```
##                     Alemania                        Angola
##                            5                             5
##                    Argentina                       Austria
##                          120                             1
##                   Bangladesh                       Bélgica
##                            2                             1
##                       Belice                         Benín
##                            8                             2
##                      Bolivia                        Brasil
##                           10                            27
##                 Burkina Faso                       Camerún
##                           15                            24
##                       Canadá                         Chile
##                            2                            22
##      China, República Popular                     Colombia
##                           14                          1013
##                        Congo                 Corea del Sur
##                            7                             1
##              Costa de Marfil                    Costa Rica
##                            9                            22
##                         Cuba                          CUBA
##                         4614                             7
##                     Dominica                       Ecuador
##                            2                            88
##                  El Salvador                   EL SALVADOR
##                         3431                            35
##                       España   Estados Unidos de América
##                           14                            23
##                      Etiopia           Federación de Rusia
##                            1                            42
##                      Francia                        Gambia
##                            3                             2
##                        Ghana                     Guatemala
##                           64                          2242
##                    GUATEMALA                        Guinea
##                            9                             8
##            Guinea Ecuatorial                 Guinea-Bissau
##                            1                             1
##                       Guyana                         Haití
##                            2                         10486
##                        HaitÍ                      Honduras
##                            3                          8111
##                      Hungría                         India
##                            2                            33
##                         Irak                          Irán
##                            1                             1
##                       Israel                        Italia
##                            2                             6
##                      Jamaica                         Japón
##                            4                             1
##                       Líbano                         Libia
##                            2                             1
##                    Macedonia                          Mali
##                            1                             6
```

```
##                          Marruecos                            Mauritania
##                                  1                                     2
##                          Nicaragua                             NICARAGUA
##                               1324                                     8
##                            Nigeria                               Nígeria
##                                  2                                    41
##              Países Bajos (Holanda)                              Pakistán
##                                  1                                     8
##                             Panamá                              Paraguay
##                                 11                                     2
##                               Perú                               Polonia
##                                 70                                     3
##                           Portugal                           Reino Unido
##                                  6                                     1
##                    República Checa      República de Guinea-Bissau
##                                  1                                     1
##              República de Suriname              Republica de Turquia
##                                  2                                     4
## República Democrática del Congo              República Dominicana
##                                  4                                    83
##                            Rumania                                 Rusia
##                                  2                                     2
##                            Senegal                          Sierra Leona
##                                  5                                     6
##                           Singapur                               Somalia
##                                  1                                     8
##                          Sri Lanka                                 Suiza
##                                  1                                     1
##                          Tailandia                                  Togo
##                                  1                                     8
##                  Trinidad y Tobago                               Ucrania
##                                  1                                     5
##                            Uruguay                             Venezuela
##                                  9                                   880
##                          VENEZUELA
##                                  2
```

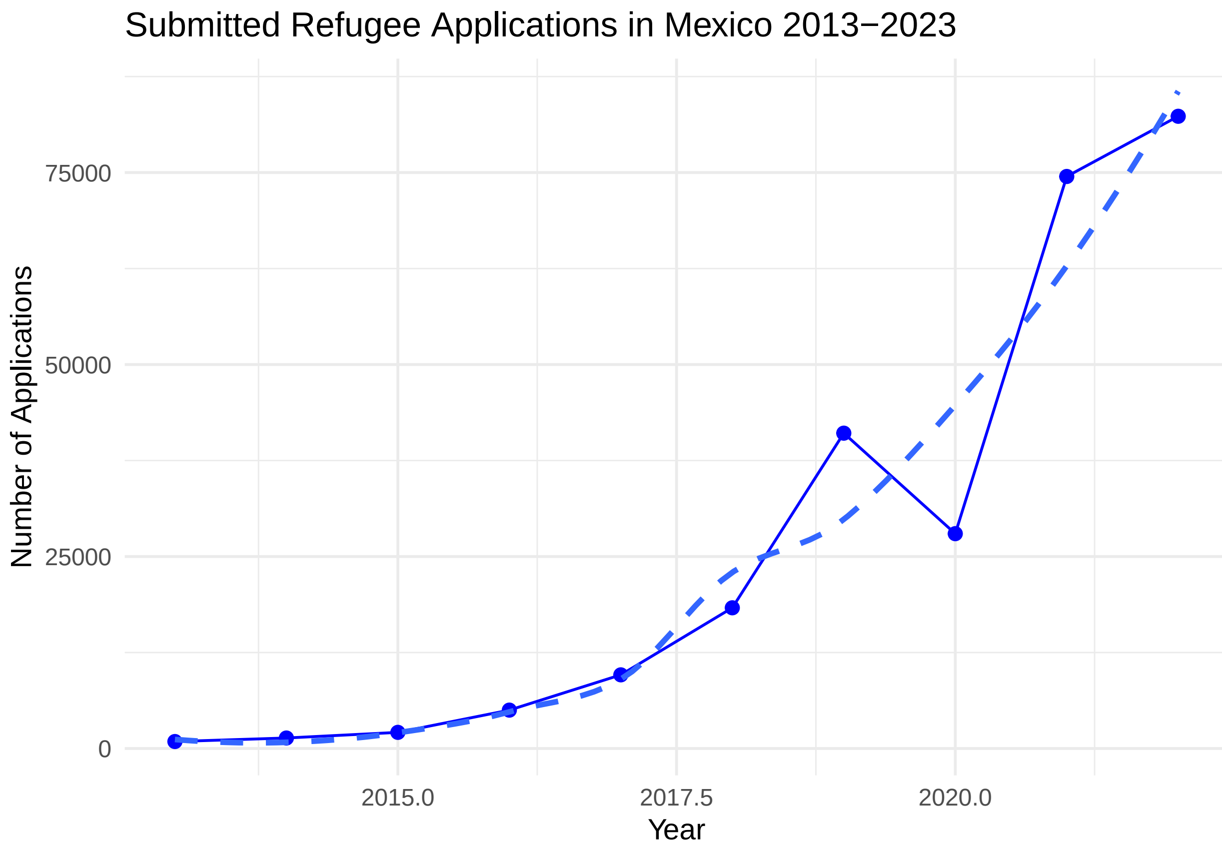
These three tables demonstrate the total number of positive resolutoins, and the total number of negative resolutions for all refugee applicants in Mexico between 2013-2023. I proceed with more interpretations of the data.

These three tables are plotted summaries of my dependent variable.

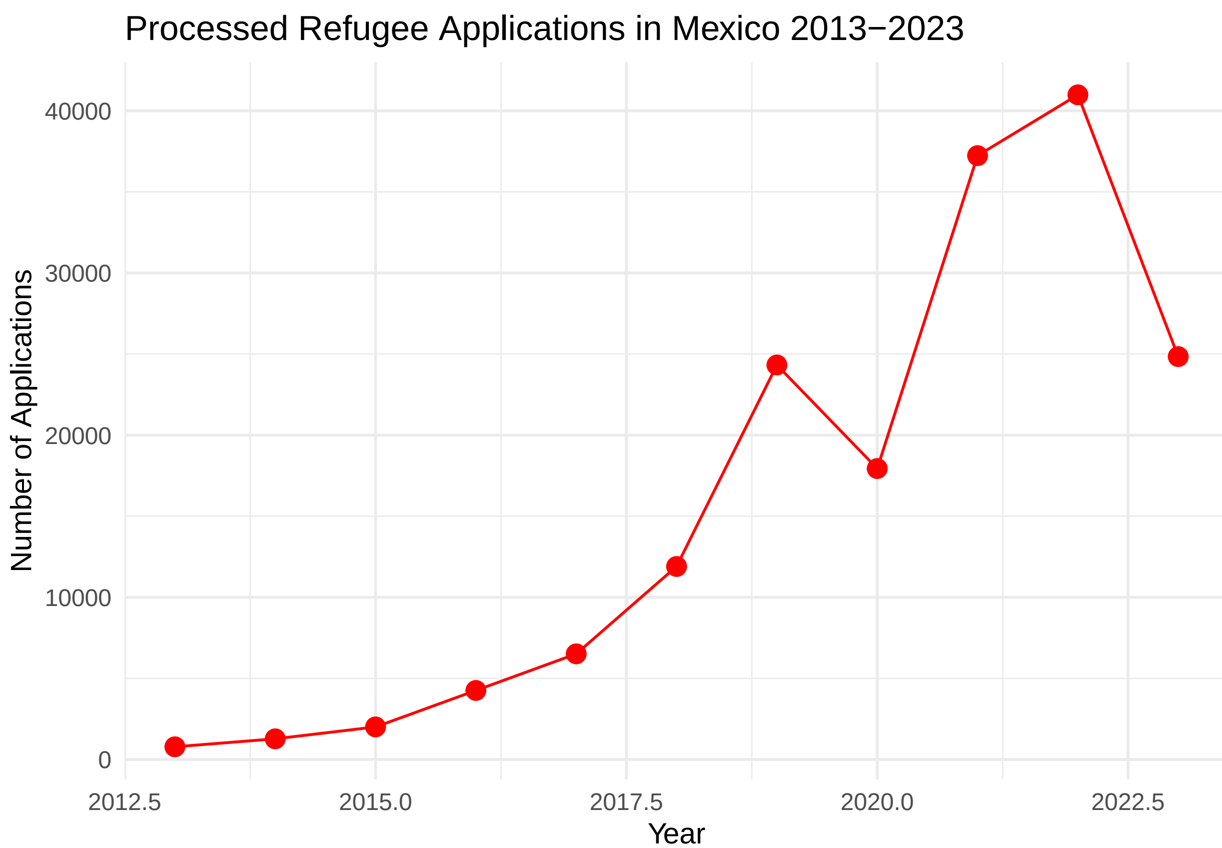
```r
library(ggplot2)

# data frames: apps_sub_per_year, apps_proc_per_year, apps_proc_result

# Creating the line graph for submitted applications per year
ggplot(data = apps_sub_per_year %>% filter(year != 2023),
       aes(x = year, y = applications, group = 1)) +
  geom_line(color = "blue") +
  geom_point(color = "blue", size = 2) +
  geom_smooth(linetype = 2, se = FALSE) +
  labs(title = "Submitted Refugee Applications in Mexico 2013-2023",
```

```r
      x = "Year",
      y = "Number of Applications") +
  theme_minimal()
```

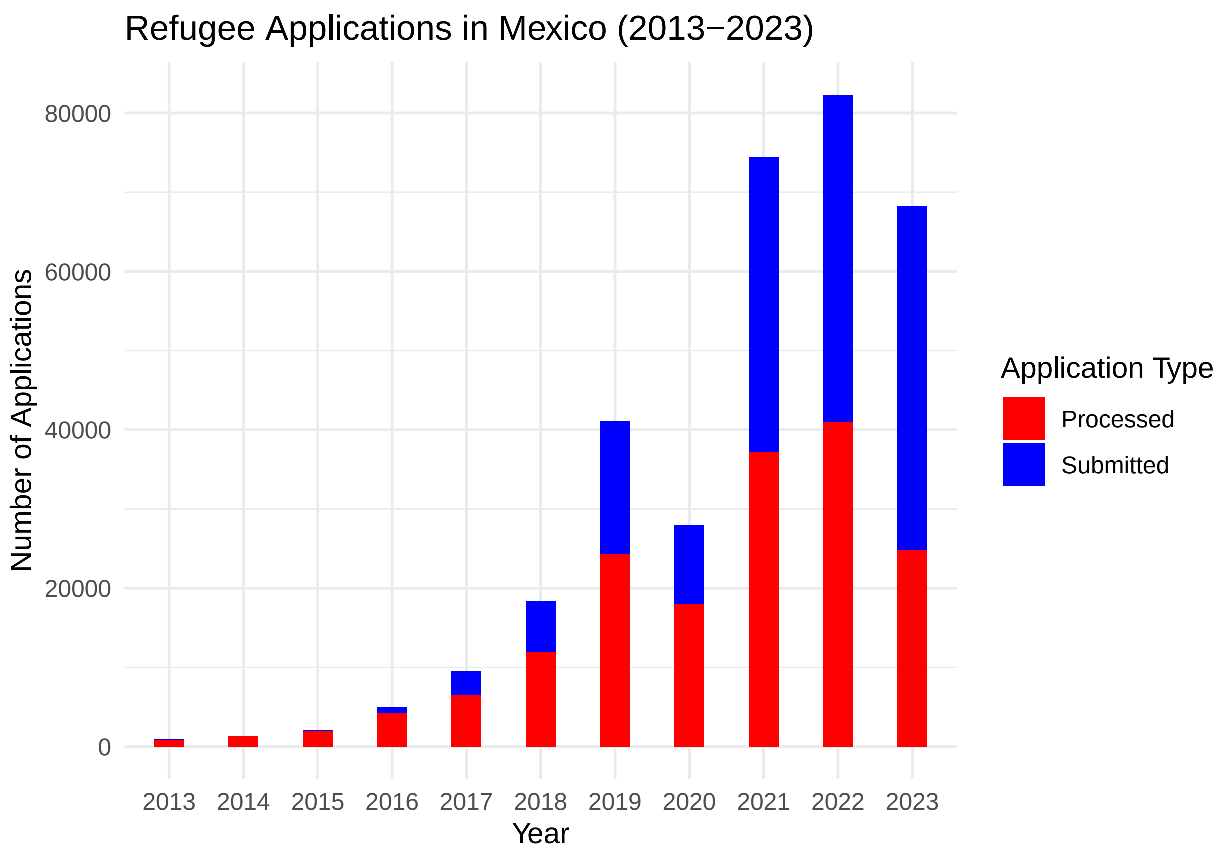
```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```




```r
# Creating the line graph for processed applications per year
ggplot(data = apps_proc_per_year, aes(x = year, y = applications, group = 1)) +
  geom_line(color = "red") +
  geom_point(color = "red", size = 3) +
  labs(title = "Processed Refugee Applications in Mexico 2013-2023",
      x = "Year",
      y = "Number of Applications") +
  theme_minimal()
```




```
library(ggplot2)


merged_data <- merge(apps_sub_per_year, apps_proc_per_year, by = "year", all = TRUE)


ggplot(data = merged_data, aes(x = as.factor(year))) +
  geom_bar(aes(y = applications.x, fill = "Submitted"), stat = "identity", position = "dodge", width =
  geom_bar(aes(y = applications.y, fill = "Processed"), stat = "identity", position = "dodge", width =
  labs(title = "Refugee Applications in Mexico (2013-2023)",
       x = "Year",
       y = "Number of Applications") +
  scale_fill_manual(name = "Application Type", values = c("Submitted" = "blue", "Processed" = "red")) +
  theme_minimal()
```



##Description: Mexico as a Destination Country for Immigrants

Since 2013, the number of refugee applications in Mexico have exponentially increased. Simultaneously, this bar graph indicates how Mexico has also processed, though not completely, refugee applications over the years. This graph indicates that Mexico has become a destination country, and every year there is an increasing number of migrants that attempt to permanently resettle in the country. These visuals are designed to give a visual description of the migration context in Mexico.

```
# 'application_results' is my data frame
library(dplyr)
library(knitr)

# Filter for positive resolutions
apps_proc_result_pos <- application_results %>%
  filter(SENTIDO_RESOLUCIÓN == "Resolución Positiva")

# Count applications per nationality for positive resolutions
table_result <- table(apps_proc_result_pos$NACIONALIDAD)


# Count applications per positive resolution type
apps_proc_result_pos <- application_results %>%
  filter(SENTIDO_RESOLUCIÓN == "Resolución Positiva") %>%
  count(NACIONALIDAD)

# Count applications per negative resolution type
apps_proc_result_neg <- application_results %>%
```

```r
  filter(SENTIDO_RESOLUCIÓN == "Resolución Negativa") %>%
  count(NACIONALIDAD)

# Merge the two tables
merged_table <- merge(apps_proc_result_pos, apps_proc_result_neg, by = "NACIONALIDAD", all = TRUE)

# Rename columns for clarity
colnames(merged_table) <- c("Nationality", "Positive_Resolutions", "Negative_Resolutions")


# List of countries to filter
selected_countries <- c("Honduras", "El Salvador", "Guatemala", "Haití", "Brasil", "Federación de Rusia

# Filter the merged table for the selected countries
filtered_merged_table <- merged_table %>%
  filter(Nationality %in% selected_countries)


# Count applications per positive resolution type
apps_proc_result_pos <- application_results %>%
  filter(SENTIDO_RESOLUCIÓN == "Resolución Positiva") %>%
  count(NACIONALIDAD)

# Count applications per negative resolution type
apps_proc_result_neg <- application_results %>%
  filter(SENTIDO_RESOLUCIÓN == "Resolución Negativa") %>%
  count(NACIONALIDAD)

# Merge the two tables
merged_table <- merge(apps_proc_result_pos, apps_proc_result_neg, by = "NACIONALIDAD", all = TRUE)

# Rename columns for clarity
colnames(merged_table) <- c("Nationality", "Positive_Resolutions", "Negative_Resolutions")

# Add a percentage column
merged_table$Percentage_Positive <- paste0(round((merged_table$Positive_Resolutions /
                                                   (merged_table$Positive_Resolutions + merged_table$Ne

# Filter out the total row
filtered_merged_table <- merged_table %>%
  filter(!is.na(Nationality))

# List of countries to filter
selected_countries <- c("Honduras", "El Salvador", "Guatemala", "Haití", "Brasil", "Federación de Rusia

# Filter the merged table for the selected countries
filtered_merged_table <- filtered_merged_table %>%
  filter(Nationality %in% selected_countries)


# Count applications per positive resolution type
apps_proc_result_pos <- application_results %>%
  filter(SENTIDO_RESOLUCIÓN == "Resolución Positiva") %>%
```

```r
  count(NACIONALIDAD)

# Count applications per negative resolution type
apps_proc_result_neg <- application_results %>%
  filter(SENTIDO_RESOLUCIÓN == "Resolución Negativa") %>%
  count(NACIONALIDAD)

# Merge the two tables
merged_table <- merge(apps_proc_result_pos, apps_proc_result_neg, by = "NACIONALIDAD", all = TRUE)

# Rename columns for clarity
colnames(merged_table) <- c("Nationality", "Positive_Resolutions", "Negative_Resolutions")

# Add a percentage column (rounded to the nearest whole number)
merged_table$Percentage_Positive <- paste0(round((merged_table$Positive_Resolutions /
                                                    (merged_table$Positive_Resolutions + merged_table$Neg

# Filter out the total row
filtered_merged_table <- merged_table %>%
  filter(!is.na(Nationality))

# List of countries to filter
selected_countries <- c("Honduras", "El Salvador", "Guatemala", "Haití", "Brasil", "Federación de Rusia

# Filter the merged table for the selected countries
filtered_merged_table <- filtered_merged_table %>%
  filter(Nationality %in% selected_countries)
filtered_merged_table <- arrange(filtered_merged_table, desc(Percentage_Positive))


# Create a nice-looking table
kable(
  filtered_merged_table,
  format = "html",
  col.names = c("Nationality", "Positive Resolutions", "Negative Resolutions", "Percentage of Positive R
  caption = "Refugee Application Outcomes for Top Hispanic and Non-Hispanic Nationalities (2013-2023)",
   align = "c", # optional, adjust alignment as needed
  width = "100%", # set the table width
  table.attr = 'style="width:100%;"' # set the width attribute in the HTML table tag
)
```

Refugee Application Outcomes for Top Hispanic and Non-Hispanic Nationalities (2013-2023)

Nationality

Positive Resolutions

Negative Resolutions

Percentage of Positive Resolutions

Venezuela

16366

880

95%

Honduras

27115

8111

77%

El Salvador

8470

3431

71%

Federación de Rusia

48

42

53%

Guatemala

2211

2242

50%

Nicaragua

1291

1324

49%

Nígeria

34

41

45%

Ghana

38

64

37%

Haití

2190

10486

17%

Brasil

4

27

13%

India

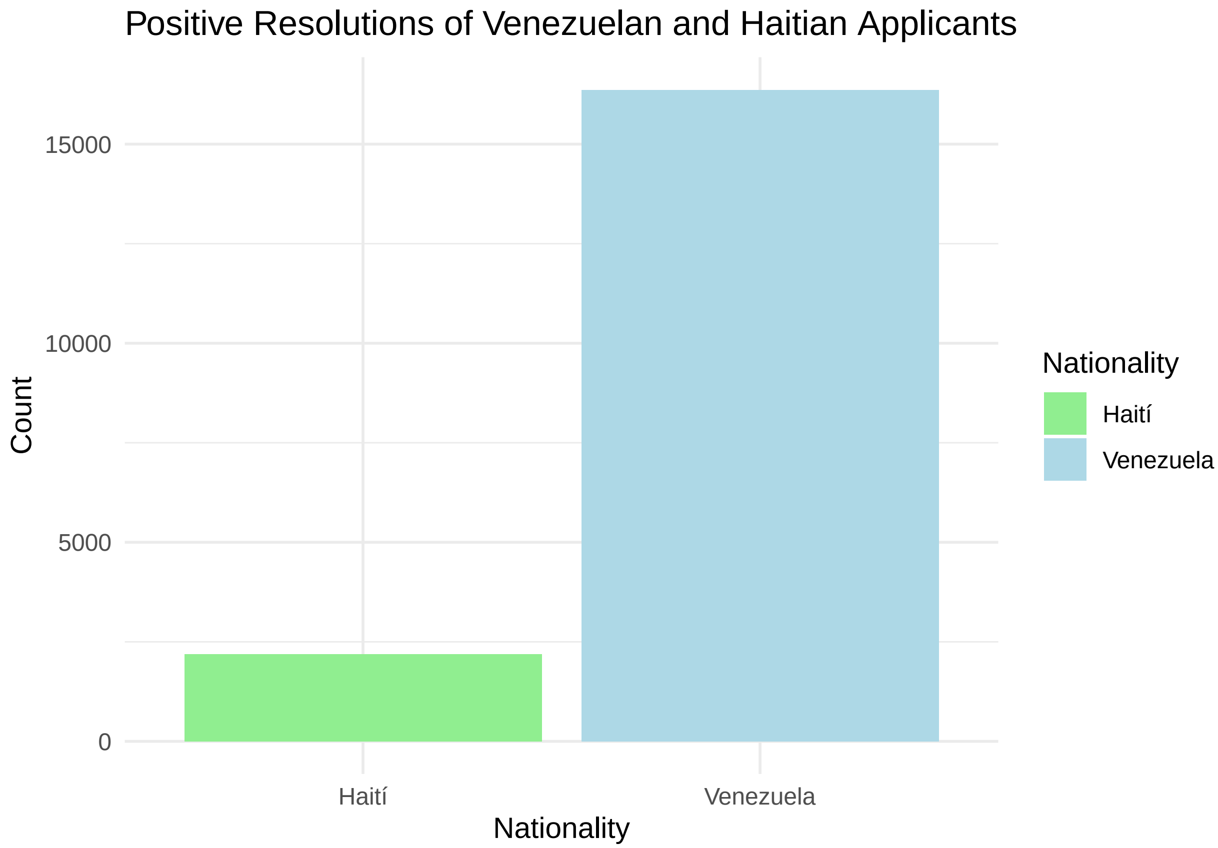
4

33

11%

This chart is the first piece of data that begins to prove my hypothesis. I analyzed the top 6 Hispanic nationalities applying for refugee status in Mexico, and the top 5 non-Hispanic nationalities applying for refugee status. The most telling analysis is that among the top nationalities that obtained refugee status in Mexico were of Hispanic origin. Venezuelan migrants overwhelmingly received refugee status in this decade, following by applicants of Salvadoran, Honduran, Russian, Guatemalan and Nicaraguan origins. The next top nationalities were of non-Hispanic origin, and their application admissions rate was significantly lower. Nigerians and Ghanaians had around a 40% application success rate. Haitians were overwhelmingly rejected, whose admissions rate was nearly 17%. Brazilians and Indians had the lowest acceptance rates among the top non-Hispanic applicants. At first glance, this data supports my hypothesis, and I took a more nuanced approach to compare the Haitian and Venezuelan cases more specifically.

```r
# 'filtered_merged_table' is my data frame
library(dplyr)
library(ggplot2)

# Filter the data for Venezuela and Haiti
venezuela_data <- filtered_merged_table %>%
  filter(Nationality == "Venezuela")

haiti_data <- filtered_merged_table %>%
  filter(Nationality == "Haití")

# Check if the filtered data frames are not empty
if (nrow(venezuela_data) > 0 && nrow(haiti_data) > 0) {
  # Combine the data for Venezuela and Haiti
  bar_data <- rbind(venezuela_data, haiti_data)

  # Create a side-by-side bar plot
  ggplot(bar_data, aes(x = Nationality, y = Positive_Resolutions, fill = Nationality)) +
    geom_bar(stat = "identity", position = "dodge") +
    labs(title = "Positive Resolutions of Venezuelan and Haitian Applicants",
         x = "Nationality",
         y = "Count") +
    scale_fill_manual(values = c("Venezuela" = "lightblue", "Haití" = "lightgreen")) +
    theme_minimal()
} else {
  cat("No data found for Venezuela or Haiti.\n")
}
```



Paragraph Summary:

These two data pieces are also essential to answer my research question. As a reminder, I am investigating whether the nationality of an applicant impacts the likelihood that they receive refugee status in Mexico. My hypothesis was that if an applicant is from a Spanish-speaking country, then they are more likely to acquire refugee status. If an applicant is from a non-Spanish-speaking country, then they are more likely to get rejected.

Based on these data visuals, this deep-dive case analysis further corroborates my initial findings. First, the nationality that is the most negatively impacted is Haiti. According to my data set, while 2190 Haitians managed to acquire refugee status in Mexico, 10486 Haitian applicants have been rejected over the last decade. Haiti is a country in the Caribbean, whose official language is Haitian Creole.

On the contrary, migrants from Venezuela have been overwhelmingly granted refugee status over the last decade. Only 882 Venezuelans were denied refugee status, and 16,366 Venezuelans were granted refugee status. Venezuela is a Spanish-speaking country in South-America.

The reason why I specifically analyzed both cases is because arguably both countries are experiencing the similar political contexts in contemporary Latin-American history, and are often impacted by economic and environmental destitution in the same way. I would have expected that if nationality and the corresponding racial implications were not contributing determinants to refugee acquisition, then applicants from these two countries would have been admitted at statistically similar rates. However, despite their similar political and economic contexts, there is a drastically different outcome in refugee status determination in Mexico. As a result, the incredibly large disparity among admissions rate is telling of a correlation that benefits applicants of Hispanic origin at large.

```r
library(dplyr)
library(modelsummary)
```

```
## `modelsummary` has built-in support to draw text-only (markdown) tables.
##   To generate tables in other formats, you must install one or more of
##   these libraries:
##
## install.packages(c(
##     "kableExtra",
##     "gt",
##     "flextable",
##
##   "huxtable",
##     "DT"
## ))
##
##   Alternatively, you can set markdown as the default table format to
##   silence this alert:
##
## config_modelsummary(factory_default = "markdown")
```

```r
spanish_speaking_countries <- c("Argentina","Chile","Colombia","Costa Rica","Cuba","Ecuador","El Salvado

data <- application_results |>
  mutate(spanish_speaking = if_else(NACIONALIDAD %in% spanish_speaking_countries, 1, 0),
         approved = if_else(SENTIDO_RESOLUCIÓN == "Resolución Positiva", 1, 0))

fit1 <- lm(approved ~ spanish_speaking + SEXO + EDAD, data = data)

summary(fit1)
```

```
##
## Call:
## lm(formula = approved ~ spanish_speaking + SEXO + EDAD, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6321 -0.3826 -0.1856  0.5144  0.9828
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       0.1157192  0.0049220   23.51   <2e-16 ***
## spanish_speaking  0.3134609  0.0029868  104.95   <2e-16 ***
## SEXOM            -0.1255475  0.0023638  -53.11   <2e-16 ***
## EDAD              0.0022549  0.0001109   20.34   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4612 on 172049 degrees of freedom
## Multiple R-squared:  0.08218,    Adjusted R-squared:  0.08217
## F-statistic:  5135 on 3 and 172049 DF,  p-value: < 2.2e-16
```

```r
modelsummary(fit1)
```

	(1)
(Intercept)	0.116
	(0.005)
spanish_speaking	0.313
	(0.003)
SEXOM	-0.126
	(0.002)
EDAD	0.002
	(0.000)
Num.Obs.	172053
R2	0.082
R2 Adj.	0.082
AIC	221957.1
BIC	222007.4
Log.Lik.	-110973.543
RMSE	0.46

In this research study, I ran a linear regression analysis assessing the impact that hispanic nationality has on refugee status determination in Mexico. For this linear regression, I assessed the relationship between the independent variable on the application outcome for all 172,053 processed applications. The linear analysis additionally evaluates the result of the regression in 0.001, 0.01, 0.05, 0.1, and 1 test values.

Null Hypothesis (H0): There is no significant impact of Hispanic nationality on the likelihood of a migrant being granted refugee status in Mexico.

Alternative Hypothesis (H1): Hispanic Nationality Significantly influences the likelihood of a migrant being granted refugee status in Mexico.

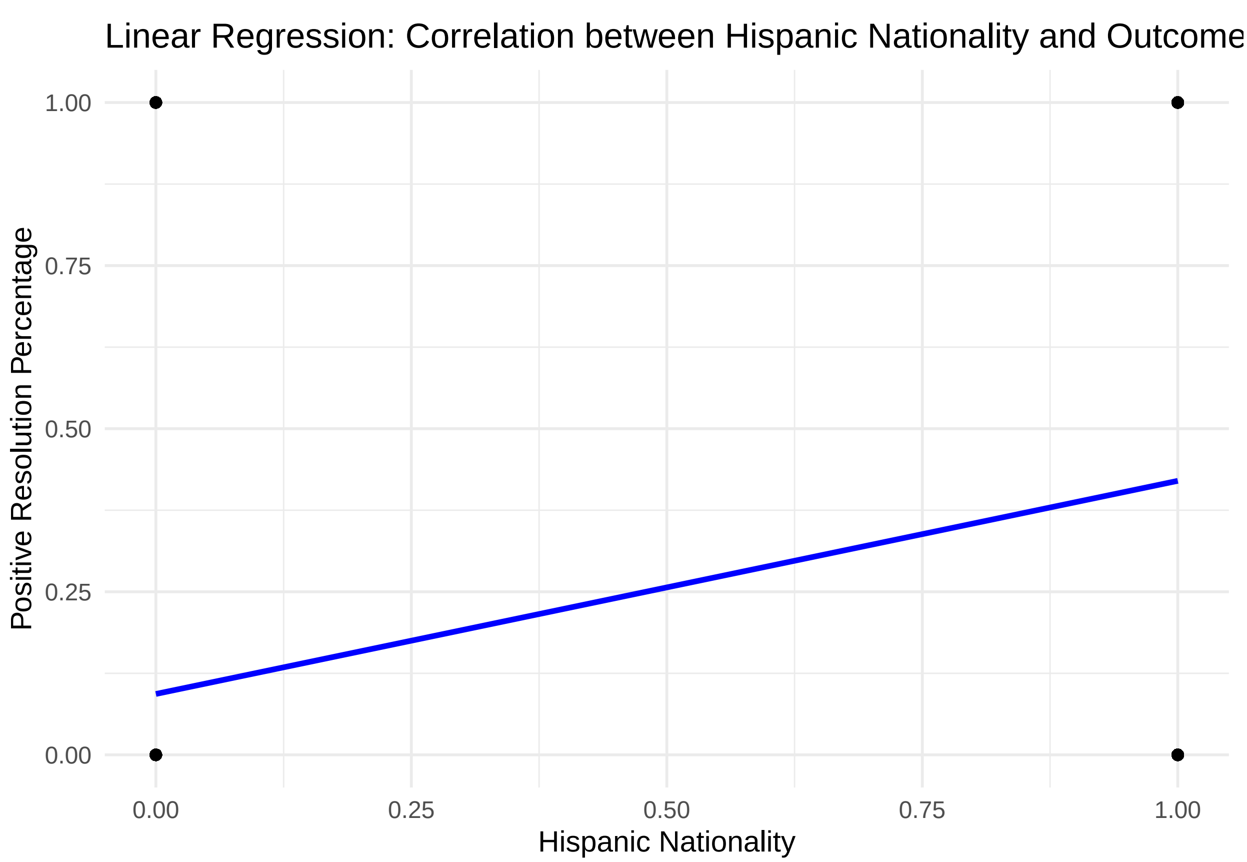
The results of my study indicated a p-value of <2.2e-16. In written form this p-value is 2.2 times 10 to the power of -16, indicating an exceedingly small value that passes is less than all alpha values indicated in my regression.

As a result of this incredibly small p-value, I am able to reject the null hypothesis with all given test values, and my linear regression model indicates that there is statistically significant data to indicate that Hispanic Nationality influences Refugee Status Determination in Mexico. In the context of my research question, the coefficient for the spanish-speaking variable was .313 percentage points, indicating that there is statistically significant data that indicates that an applicant of Hispanic nationality has a 31% higher probability to gain refugee admissions in Mexico, over applicants that indicated a non-Hispanic nationality. Interpreted further, non-Hispanic nationalities only have about a 11-12% chance of gaining refugee status admissions, as demonstrated by the y-intercept of the regression. Similarly, the control variables were an applicants sex and age. In comparison to the spanish-speaking variables, though their p-values were also small, the impact magintude of both variables was not as strong as the spanish-speaking variable.

```r
# Install and load necessary packages if not already installed
if (!requireNamespace("ggplot2", quietly = TRUE)) {
  install.packages("ggplot2")
}
if (!requireNamespace("dplyr", quietly = TRUE)) {
  install.packages("dplyr")
}
```

```r
# Load libraries
library(ggplot2)
library(dplyr)

# Your linear regression model
model <- lm(formula = approved ~ spanish_speaking, data = data)

# Create a dataframe for predictions
predictions <- data.frame(data, predicted = predict(model))

# Plotting the data and the regression line
ggplot(predictions, aes(x = spanish_speaking, y = approved)) +
  geom_point() +  # Scatter plot of the data points
  geom_line(aes(x = spanish_speaking, y = predicted), color = "blue", size = 1) +  # Regression line
  labs(title = "Linear Regression: Correlation between Hispanic Nationality and Outcome",
       x = "Hispanic Nationality",
       y = "Positive Resolution Percentage") +
  theme_minimal()
```

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```



This graphical model visually demonstrates the positive correlation between Hispanic Nationality and

Positive Resolution Percentage. According to the linear regression, the y-intercept value is .116, which indicates that non-Hispanic refugee applicants on average have a 11-12% probability of obtaining refugee status in Mexico. In contrast, the spanish-speaking coefficient variable is calculated to be 0.32, which indicates that through the application of the "Hispanic Nationality" variable, refugee applicants from Hispanic origin on average have a 43% probability of gaining refugee status in Mexico. The data strongly supports my original hypothesis and demonstrates that these refugee status outcomes are strongly associated with Hispanic nationality.

"'

Interpretation and Policy Recommendations

Conclusion