DATA606 Final Project

Mael Illien

11/25/2019

Contents

Introduction

Given the extraordinary surge in migration during the recent years, it is interesting to find out if there is a particular change in the demographics of those groups. The data presented here does not account for refugees or population outflows, but it does provide the opportunity to explore some macro trends of migration, which in this case will be gender.

The guiding research questions is: **Is there convincing evidence that the world has seen a change in its gender proportion of migrants between 1990 and 2019?**

In order to address this question, we postulate the following hypothesis:

$$Ho : The\ gender\ proportion\ has\ not\ changed\ (p_{1990} = p_{2019} \rightarrow p_{2019} - p_{1990} = 0 \rightarrow p_{diff} = 0)$$

$$Ha : The\ gender\ proportion\ has\ changed\ (p_{1990} \neq p_{2019} \rightarrow p_{2019} - p_{1990} \neq 0 \rightarrow p_{diff} \neq 0)$$

Environment Setup

Load required libraries

```
library(tidyverse) #readr, dplyr, tidyr, stringr, tibble, ggplot2
library(knitr)
library(kableExtra)
library(readxl)
library(scales)
library(countrycode)
```

A helper function for displaying tables

```
showtable <- function(data, title) {
  kable(data, caption = title) %>%
  kable_styling(bootstrap_options = c("striped", "hover", "condensed"))
}
```

Data

Data Description

We are looking at data collected by the United Nations found here.

In particular, we are looking at Table 1.

The data is collected by the United Nations and the estimates are based on official statistics of foreign-born populations. The data is collected every year, but only 5 year increments as well as the current year (2019) are given.

Each case represents an estimate of a country's international migrant population by gender, for a particular country and year. There are 3216 observations in the data set.

The response variable is the count of migrant population which is discrete and numerical. We will boil it down into the gender proportion of migrants which is continuous and bounded between 0 and 1.

The independent variables are `Year` which is a quantitative discretization of time and `Gender` which is qualitative. We should note that Year can be a different type of variable depending on how it is used. While in this context it makes sense to think of it as numerical time series, it does not make sense to apply numerical operations to it like for example to taking the average of two years or adding two years.

This study is observational because the data collectors had no control over the variables.

Regarding generalizability, the collected data represents the entire population of migrants so there was no random sampling. However, we will be randomly sampling the data to be able to generalize. In this study there is notion of causality since there is no random assignment.

Data Import

The document consists of multiple tabs of data. We will focus on Table 1 which contains International migrant stock at mid-year by sex and by major area, region, country or area, 1990-2019. We will look at the non-aggregated data for each country for both males and females to see what trends we can derive from this dataset.

United Nations
Population Division
Department of Economic and Social Affairs

Workbook: UN_MigrantStockTotal_2019.xlsx
Table 1 - International migrant stock at mid-year by sex and by major area, region, country or area, 1990-2019

Suggested citation: United Nations, Department of Economic and Social Affairs. Population Division (2019). International Migrant Stock 2019 (United Nations database, POP/DB/MIG/Stock/Rev.2(

Sort order	Major area, region, country or area of destination	Notes	Code	Type of data	International migrant stock at mid-year (both sexes)						
					1990	1995	2000	2005	2010	2015	2019
1	WORLD		900		153,011,473	161,316,895	173,588,441	191,615,574	220,781,909	248,861,296	271,642,105
2	UN development groups				..	..	..	..	..	..	..
3	More developed regions	b	901		82,767,216	92,935,095	103,961,989	116,687,616	130,613,460	140,643,317	152,069,261
4	Less developed regions	c	902		70,244,257	68,381,800	69,626,452	74,927,958	90,168,449	108,217,979	119,572,844
5	Least developed countries	d	941		11,060,221	11,681,777	10,063,948	9,833,150	10,432,671	13,631,349	16,289,023
6	Less developed regions, excluding least developed countries		934		59,184,036	56,700,023	59,562,504	65,094,808	79,735,778	94,586,630	103,283,821
7	World Bank income groups				..	..	..	..	..	..	..
8	High-income countries	e	1503		77,802,868	89,346,612	103,029,727	120,543,730	144,282,576	160,366,582	175,811,829
9	Middle-income countries	e	1517		65,124,366	61,040,029	61,158,978	61,623,367	65,980,003	76,123,099	82,237,681
10	Upper-middle-income countries	e	1502		33,285,549	33,494,586	34,642,477	36,657,488	40,238,761	49,653,277	54,279,436
11	Lower-middle-income countries	e	1501		31,838,817	27,545,443	26,516,501	24,965,879	25,741,242	26,469,822	27,958,245
12	Low-income countries	e	1500		9,790,238	10,577,086	8,975,309	8,969,212	10,040,501	11,869,725	13,072,099
13	No income group available	e	1518		294,001	353,168	424,127	479,265	478,829	501,890	520,496
14	Geographic regions				..	..	..	..	..	..	..
15	Africa		903		15,689,666	16,357,077	15,051,677	15,969,835	17,804,198	23,476,251	26,529,334
16	Asia		935		48,209,949	46,418,044	49,394,322	53,439,306	65,938,712	77,231,760	83,559,197
17	Europe		908		49,608,231	53,489,829	56,858,788	63,594,822	70,678,025	75,008,219	82,304,539
18	Latin America and the Caribbean		904		7,161,371	6,688,710	6,570,729	7,224,942	8,262,433	9,441,679	11,673,288
19	Northern America		905		27,610,408	33,340,948	40,351,694	45,363,257	50,970,861	55,633,443	58,647,822
20	Oceania		909		4,731,848	5,022,287	5,361,231	6,023,412	7,127,680	8,069,944	8,927,925
21	Sustainable Development Goal (SDG) regions	f			..	..	..	..	..	..	..
22	SUB-SAHARAN AFRICA		947		13,286,341	14,275,437	13,151,229	14,222,466	15,855,936	21,337,333	23,573,485
23	Eastern Africa		910		5,959,183	4,977,644	4,847,119	4,762,817	4,757,899	6,716,524	7,908,176
24	Burundi		108	B R	333,110	254,853	125,628	172,874	235,259	289,810	321,018
25	Comoros		174	B	14,079	13,939	13,799	13,209	12,618	12,555	12,504
26	Djibouti		262	B R	122,221	99,774	100,507	92,091	101,575	112,351	115,341
27	Eritrea		232	I	11,848	12,400	12,952	14,314	15,676	15,941	16,101
28	Ethiopia		231	B R	1,155,390	806,904	611,384	514,242	568,748	1,161,642	1,253,083
29	Kenya		404	B R	298,089	618,745	707,852	773,354	954,925	1,126,886	1,044,854
30	Madagascar		450	C	23,917	21,177	23,541	26,058	28,905	32,075	34,934
31	Malawi		454	B R	1,127,724	241,624	232,620	221,661	217,722	232,803	247,652
32	Mauritius	1	480	C	3,613	7,493	15,543	19,647	24,836	28,585	28,849
33	Mayotte		175	B	15,229	26,316	45,474	63,176	72,757	73,992	74,643
34	Mozambique		508	B R	122,332	168,256	195,702	264,679	306,471	321,794	334,665

Figure 1: Raw Data

```
untidy <- read_xlsx("UN_MigrantStockTotal_2019.xlsx", sheet = "Table 1")
```

Here is a quick look at the data. There are rows for every country as well as for aggregations by region. We can only find data by year and gender for the range 1990-2019.

Data Transormation/Preparation

We start by selecting the rows and columns of interest. We will ignore the aggregated data and select the country rows and the and the migrant stock data for the individual sexes

```
# Select the rows and columns of interest.
clean <- untidy %>%
  select(2,`United Nations`,13:26) %>%
  slice(33:294) %>%
  filter(!is.na(`United Nations`)) %>%
  select(-`United Nations`)
```

Table 1: Long Format

Country	Year	MigrantCount	Sex
Samoa	2015	2104	Female
Estonia	1990	210306	Female
Canada	2015	3895493	Female
Guadeloupe	2015	43858	Male
Cuba	1990	17278	Male
Slovakia	2010	79381	Female
Syrian Arab Republic	2015	445050	Male
Mali	2010	167309	Female
Tokelau	2019	262	Female
Nepal	2019	342315	Female

```
years <- untidy %>%
  select(2, 13:26) %>%
  slice(11)
```

```
# rename columns
names(clean) <- as.character(years)
names(clean)[1] <- "Country"
```

```
# transform the data into long format
cleaner_m <- clean[1:8]
cleaner_f <- clean[c(1,9:15)]
tidy_m <- cleaner_m %>%
  gather("1990":"2019", key = Year, value = MigrantCount) %>%
  mutate(Sex = "Male")
tidy_f <- cleaner_f %>%
  gather("1990":"2019", key = Year, value = MigrantCount) %>%
  mutate(Sex = "Female")
migrants <- rbind(tidy_m, tidy_f)
migrants["MigrantCount"] <- apply(migrants["MigrantCount"], 2, function(x) as.numeric(x))
migrants <- drop_na(migrants)
```

Let's take a look at 15 random oberservations from this tidy data table.

```
showtable(sample_n(migrants, 10),"Long Format")
```

```
migrants_wide <- migrants %>% spread(Year, MigrantCount)
showtable(sample_n(migrants_wide, 10),"Wide Format")
```

We perform a series of transformations for downstream analysis.

```
grouped_total <- migrants %>%
  group_by(Year, Sex)  %>%
  summarize(Total=sum(MigrantCount)) %>%
  mutate("Total_m" = round(Total/1000000,1))
```

Table 2: Wide Format

Country	Sex	1990	1995	2000	2005	2010	2015	2019
Ireland	Female	115868	118808	177593	279136	370102	382918	419112
Philippines	Male	81347	107524	161904	133023	107933	109783	113239
Tunisia	Female	19084	18822	18040	17040	20936	27279	27711
Burundi	Female	169843	130688	64534	88069	119436	147020	162855
Albania	Female	35434	38070	40705	31746	25883	25514	24106
Montenegro	Female	NA	NA	NA	NA	46873	43541	42985
Burkina Faso	Female	181818	240948	269808	311787	353499	369348	376508
Czechia	Male	57405	88316	119227	181719	236160	239228	293245
Croatia	Female	252907	358352	310475	307628	304782	300620	278402
Fiji	Female	6436	6275	6118	5871	6214	6330	6462

Table 3:

Year	Sex	Proportion
1990	Female	0.4924
1995	Female	0.4936
2000	Female	0.4929
2005	Female	0.4893
2010	Female	0.4834
2015	Female	0.4822
2019	Female	0.4791
1990	Male	0.5076
1995	Male	0.5064
2000	Male	0.5071
2005	Male	0.5107
2010	Male	0.5166
2015	Male	0.5178
2019	Male	0.5209

```
proportions_by_year <- grouped_total %>%
  select(-Total_m) %>%
  spread(Sex, Total) %>%
  mutate(Total = Female + Male, Female = round(Female/Total,4), Male = round(Male/Total,4))  %>%
  select(Year, Female, Male) %>%
  gather("Female":"Male", key = Sex, value = Proportion)
```

```
showtable(proportions_by_year, "")
```

```
diff <- migrants %>%
  spread(Sex, MigrantCount) %>%
  select(1:2, Male, Female) %>%
  mutate(Delta = Male - Female, Delta_m = Delta/1000000, Total = round(Male + Female,4), Prop = round(Ma
```

```
showtable(sample_n(diff, 10),"")
```

Here we randomly sample around 9% of the observations that we will use in our statistical analysis. We

Table 4:

Country	Year	Male	Female	Delta	Delta_m	Total	Prop
Jamaica	2015	11766	11401	365	0.000365	23167	0.5079
Ecuador	2010	167773	157593	10180	0.010180	325366	0.5156
Nepal	2010	189518	389139	-199621	-0.199621	578657	0.3275
Sudan	2000	404618	400468	4150	0.004150	805086	0.5026
Denmark	1990	114411	120778	-6367	-0.006367	235189	0.4865
Poland	1990	482176	645595	-163419	-0.163419	1127771	0.4275
Bosnia and Herzegovina	2000	39691	43261	-3570	-0.003570	82952	0.4785
Nepal	2000	241959	475941	-233982	-0.233982	717900	0.3370
Romania	2010	87667	89544	-1877	-0.001877	177211	0.4947
Vanuatu	2005	1398	1402	-4	-0.000004	2800	0.4993

choose 9% to respect the condition for inference stated below. The two following tables show an example of the randomly selected countries for 1990 and 2019.

```
set.seed(606)
# determine number of observations
n <- length(clean$Country)
# sample 9%
y1990 <- diff %>% filter(Year == "1990") %>% sample_n(round(0.09*n))
y2019 <- diff %>% filter(Year == "2019") %>% sample_n(round(0.09*n))
```

```
showtable(y1990, "Sampled Data for 1990")
```

```
showtable(y2019, "Sampled Data for 2019")
```

Exploratory Data Analysis

Population

The following analysis is based on the true population data.

Below is a summary of the data. We see in particular that the greatest net positive of female migrants was ~1.7m while it is nearly 4.9m for men. In terms of proportion, the parameter of interest (proportion of males) was as high as 87% and as low as 29%. The summary statistics just stated are from all observations, meaning here that these numbers are individual observations so they could be from any country or any year.

```
summary(diff)
```

```
##    Country              Year                Male
##  Length:1608        Length:1608        Min.   :      61
##  Class :character   Class :character   1st Qu.:   13160
##  Mode  :character   Mode  :character   Median :   63032
##                                        Mean   :  453763
##                                        3rd Qu.:  286557
##                                        Max.   :24488382
```

Table 5: Sampled Data for 1990

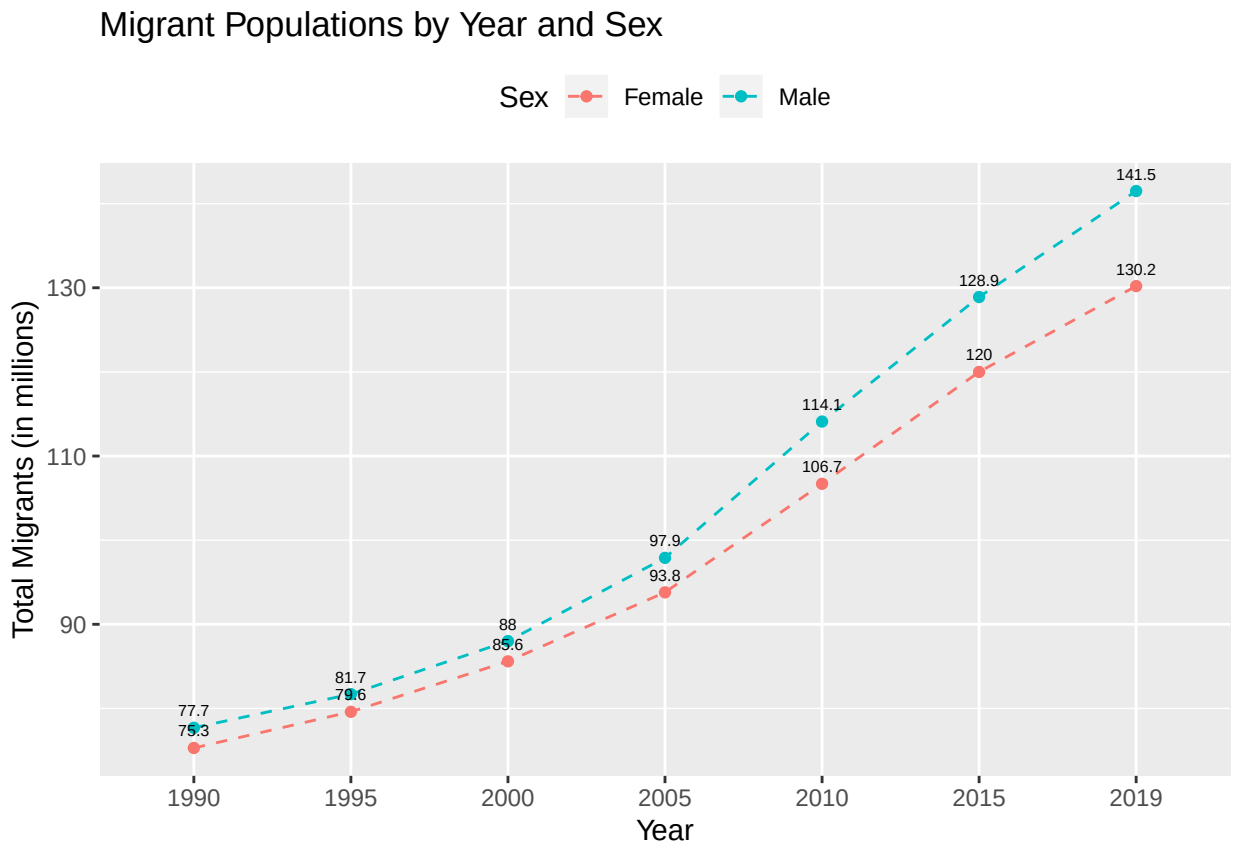
Country	Year	Male	Female	Delta	Delta_m	Total	Prop
Grenada	1990	2106	2157	-51	-0.000051	4263	0.4940
Armenia	1990	270535	388254	-117719	-0.117719	658789	0.4107
Algeria	1990	150234	123720	26514	0.026514	273954	0.5484
Afghanistan	1990	32558	25128	7430	0.007430	57686	0.5644
Anguilla	1990	1258	1312	-54	-0.000054	2570	0.4895
Samoa	1990	1771	1586	185	0.000185	3357	0.5276
El Salvador	1990	22218	25142	-2924	-0.002924	47360	0.4691
Finland	1990	31682	31573	109	0.000109	63255	0.5009
Saint Helena	1990	110	68	42	0.000042	178	0.6180
Turkmenistan	1990	134174	172326	-38152	-0.038152	306500	0.4378
Gibraltar	1990	4528	4181	347	0.000347	8709	0.5199
Haiti	1990	10606	8478	2128	0.002128	19084	0.5558
Portugal	1990	209922	225860	-15938	-0.015938	435782	0.4817
Syrian Arab Republic	1990	364077	350063	14014	0.014014	714140	0.5098
Lebanon	1990	267922	255771	12151	0.012151	523693	0.5116
Cyprus	1990	20463	23342	-2879	-0.002879	43805	0.4671
Kuwait	1990	654878	419513	235365	0.235365	1074391	0.6095
Georgia	1990	133285	171185	-37900	-0.037900	304470	0.4378
Sweden	1990	383085	405682	-22597	-0.022597	788767	0.4857
Serbia	1990	46712	52557	-5845	-0.005845	99269	0.4706
Zimbabwe	1990	356240	278381	77859	0.077859	634621	0.5613

Table 6: Sampled Data for 2019

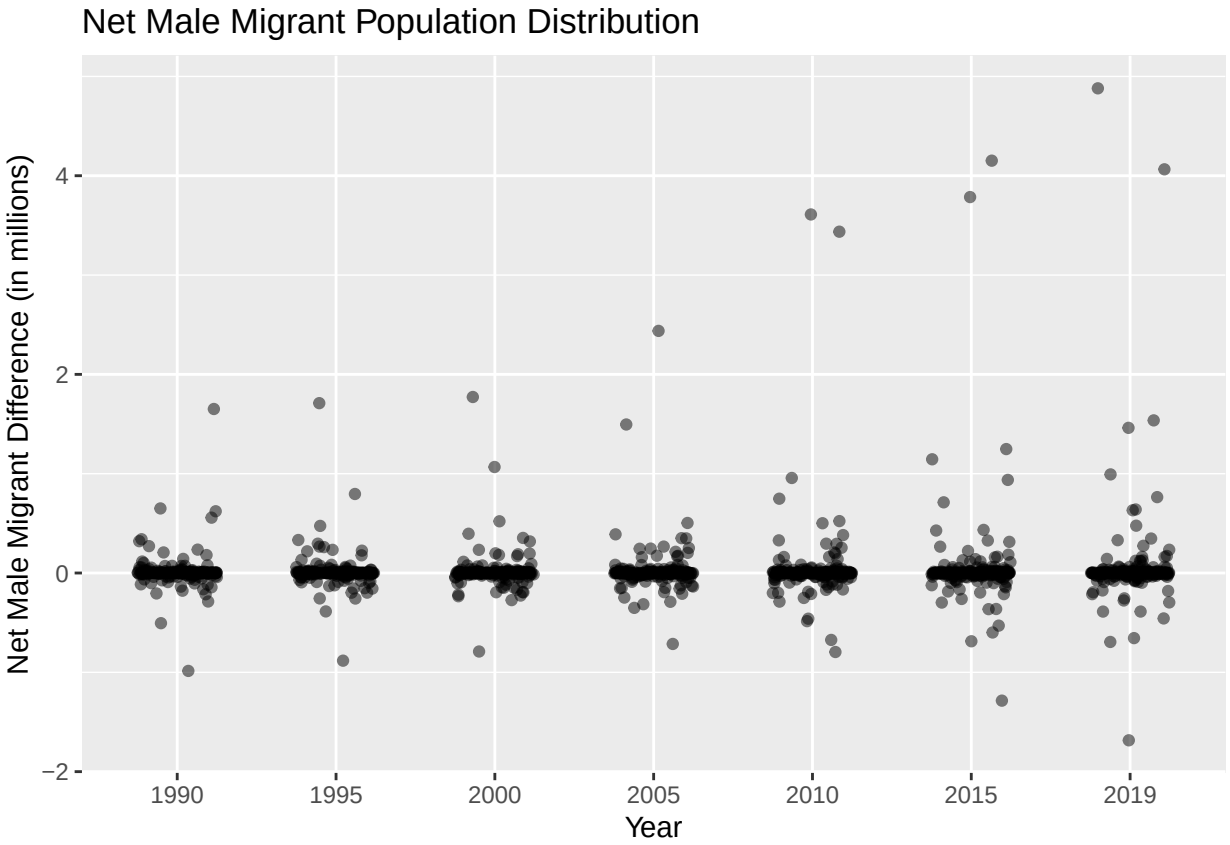
Country	Year	Male	Female	Delta	Delta_m	Total	Prop
Serbia	2019	361216	459096	-97880	-0.097880	820312	0.4403
Guadeloupe	2019	44537	55493	-10956	-0.010956	100030	0.4452
Mali	2019	237428	230802	6626	0.006626	468230	0.5071
Slovenia	2019	142626	110496	32130	0.032130	253122	0.5635
Belize	2019	30180	29818	362	0.000362	59998	0.5030
Seychelles	2019	9049	3877	5172	0.005172	12926	0.7001
Timor-Leste	2019	5086	3331	1755	0.001755	8417	0.6043
Niue	2019	319	269	50	0.000050	588	0.5425
Haiti	2019	10426	8330	2096	0.002096	18756	0.5559
Colombia	2019	575805	566514	9291	0.009291	1142319	0.5041
Cook Islands	2019	1748	1743	5	0.000005	3491	0.5007
Malawi	2019	117932	129720	-11788	-0.011788	247652	0.4762
Venezuela (Bolivarian Republic of)	2019	685975	689715	-3740	-0.003740	1375690	0.4986
Tokelau	2019	242	262	-20	-0.000020	504	0.4802
Algeria	2019	131596	117479	14117	0.014117	249075	0.5283
Anguilla	2019	2694	2985	-291	-0.000291	5679	0.4744
Benin	2019	183593	206519	-22926	-0.022926	390112	0.4706
Burkina Faso	2019	341830	376508	-34678	-0.034678	718338	0.4759
Iraq	2019	214288	153774	60514	0.060514	368062	0.5822
India	2019	2640513	2514224	126289	0.126289	5154737	0.5122
Bulgaria	2019	82718	85798	-3080	-0.003080	168516	0.4909

```
##      Female              Delta             Delta_m
##  Min.   :      47   Min.   :-1684385   Min.   :-1.684385
##  1st Qu.:   12203   1st Qu.:   -5937   1st Qu.:-0.005937
##  Median :   61813   Median :     224   Median : 0.000224
##  Mean   :  429830   Mean   :   23933   Mean   : 0.023933
##  3rd Qu.:  279594   3rd Qu.:    6722   3rd Qu.: 0.006722
##  Max.   :26172767   Max.   : 4880026   Max.   : 4.880026
##      Total               Prop
##  Min.   :     108   Min.   :0.2930
##  1st Qu.:   26197   1st Qu.:0.4800
##  Median :  125646   Median :0.5079
##  Mean   :  883593   Mean   :0.5163
##  3rd Qu.:  588974   3rd Qu.:0.5392
##  Max.   :50661149   Max.   :0.8767
```

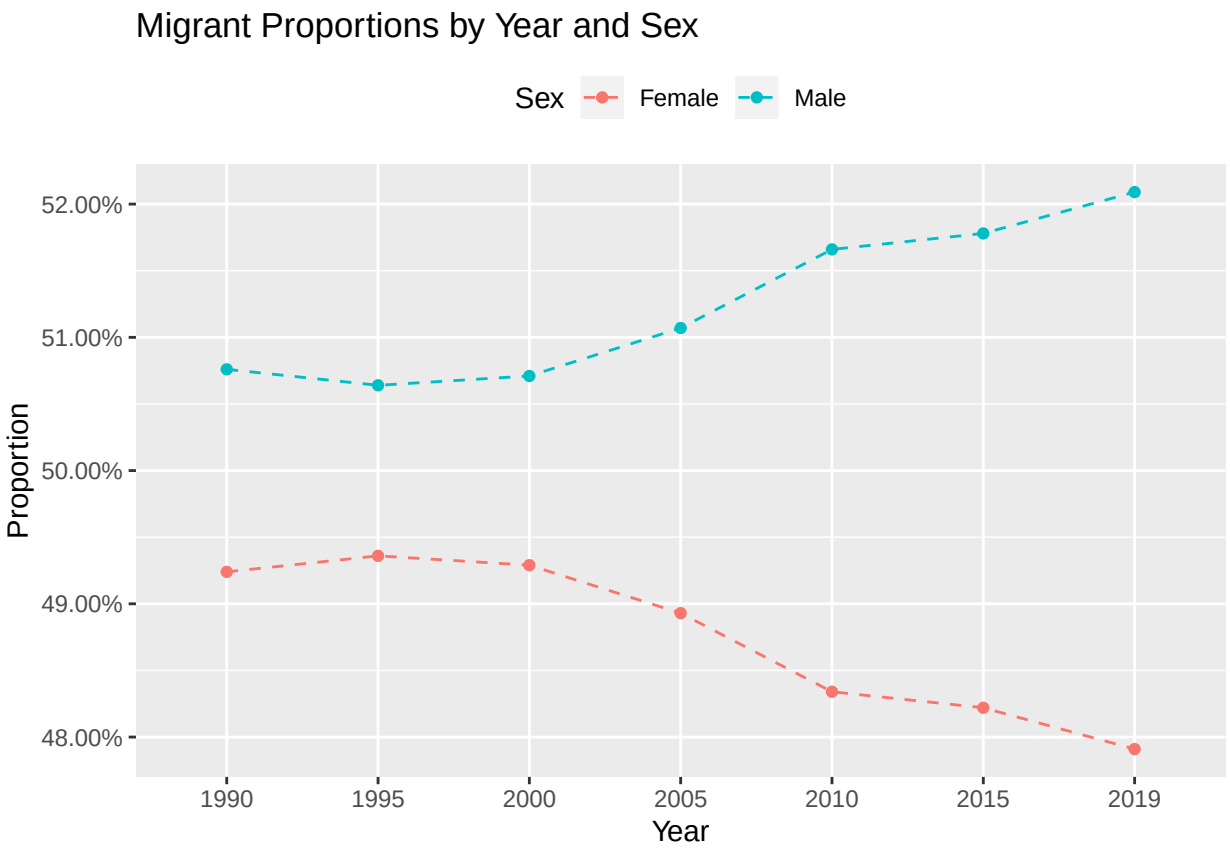
The plots below reveals that the total population of migrants is increasing over time. We notice that the difference between the number of male and female migrants is increasing. Also shown is the change in the gender proportion over time.



Here we see how the difference between male and female migrant population is dsitributed over time. We can identify the max and min observations from the summary above in 2019.

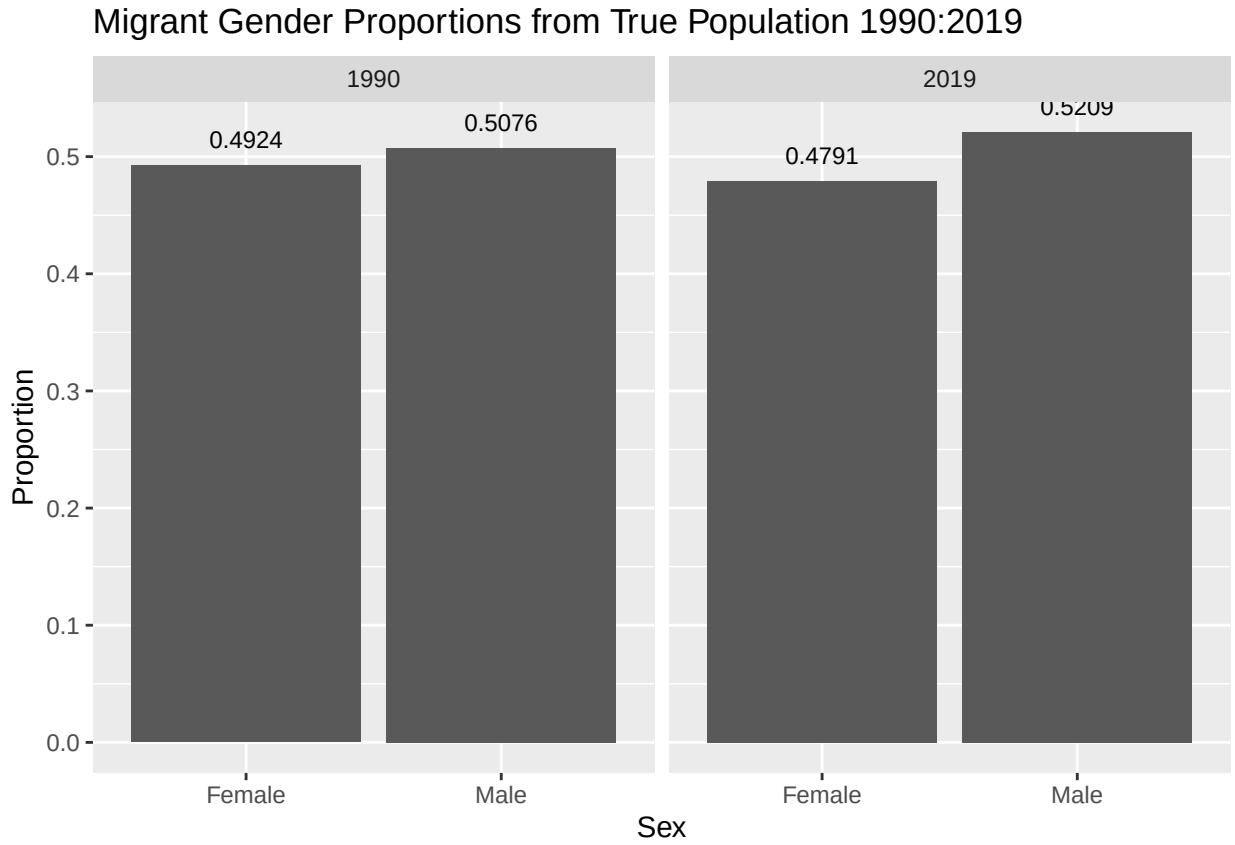


As seen below, there is averall increase in the proportion of male migrants.



The side-by-side plot shown below looks at the cases of interest for the statistical study.

```
ggplot(pop, aes(Sex, Proportion)) + geom_bar(stat="identity") + facet_wrap(~Year) +
  geom_text(aes(label=Proportion), vjust=-1, size=3) +
  ggtitle("Migrant Gender Proportions from True Population 1990:2019")
```



By looking at the population data, we can say that there has been an increase in the number of male migrants relative to females.

Sampled Data

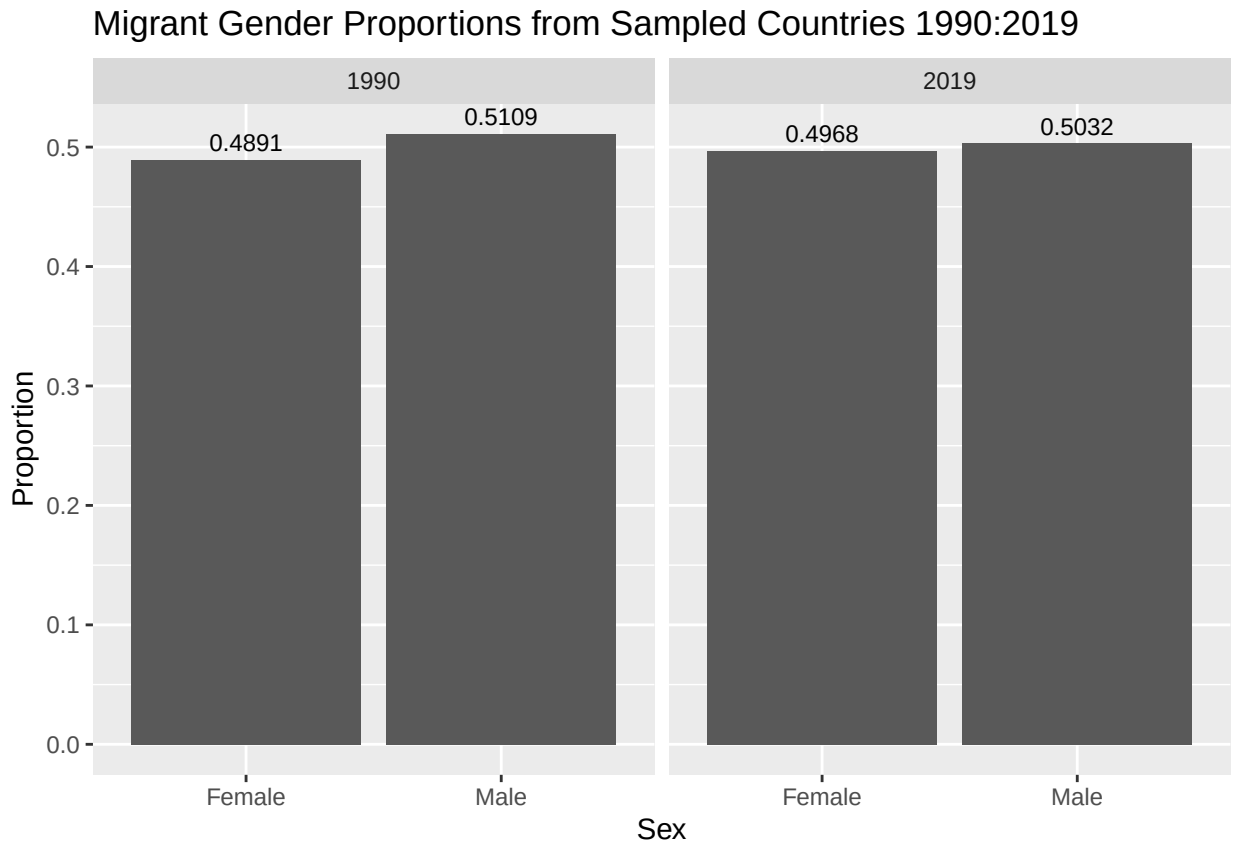
This side-by-side plot below shows the proportions from the sampled data. We observe very minor change from 1990 to 2019.

```
summary(y1990)
```

```
##    Country              Year               Male              Female       
##  Length:21          Length:21          Min.   :   110   Min.   :    68  
##  Class :character   Class :character   1st Qu.: 10606   1st Qu.:  8478  
##  Mode  :character   Mode  :character   Median : 46712   Median : 52557  
##                                        Mean   :147541   Mean   :141251  
##                                        3rd Qu.:267922   3rd Qu.:255771  
##                                        Max.   :654878   Max.   :419513  
##      Delta             Delta_m             Total              Prop       
##  Min.   :-117719   Min.   :-0.117719   Min.   :    178   Min.   :0.4107  
##  1st Qu.:  -5845   1st Qu.:-0.005845   1st Qu.:  19084   1st Qu.:0.4706  
##  Median :     42   Median : 0.000042   Median :  99269   Median :0.5009  
##  Mean   :   6290   Mean   : 0.006290   Mean   : 288792   Mean   :0.5082  
##  3rd Qu.:   7430   3rd Qu.: 0.007430   3rd Qu.: 523693   3rd Qu.:0.5484  
##  Max.   : 235365   Max.   : 0.235365   Max.   :1074391   Max.   :0.6180
```

```
summary(y2019)
```

```
##    Country               Year                Male              Female
##  Length:21          Length:21          Min.   :    242   Min.   :    262
##  Class :character   Class :character   1st Qu.:   9049   1st Qu.:   3877
##  Mode  :character   Mode  :character   Median : 117932   Median : 110496
##                                        Mean   : 277133   Mean   : 273655
##                                        3rd Qu.: 237428   3rd Qu.: 230802
##                                        Max.   :2640513   Max.   :2514224
##      Delta            Delta_m              Total              Prop
##  Min.   :-97880   Min.   :-0.097880   Min.   :    504   Min.   :0.4403
##  1st Qu.: -3740   1st Qu.:-0.003740   1st Qu.:  12926   1st Qu.:0.4762
##  Median :    50   Median : 0.000050   Median : 247652   Median :0.5030
##  Mean   :  3478   Mean   : 0.003478   Mean   : 550788   Mean   :0.5170
##  3rd Qu.:  6626   3rd Qu.: 0.006626   3rd Qu.: 468230   3rd Qu.:0.5425
##  Max.   :126289   Max.   : 0.126289   Max.   :5154737   Max.   :0.7001
```



Inference

Methodology

The distribution of the sample proportion is described by the following:

$$\hat{p} \sim N\left(mean = p, SE = \sqrt{\frac{p(1-p)}{n}}\right)$$

The confidence interval is described by the following:

$confidence\ interval = \hat{p} \pm z * SE$

In this study, we will be using a 95% significance level.

Since we are looking at the distribution of the difference of two independent sample proportions the standard error takes the following form:

$$SE_{(\hat{p}_1 - \hat{p}_2)} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Conditions

We then verify the conditions for inference on proportions.

1. Individual observations should be independent (the sample size should be $< 10\%$ of the population if sampling without replacement)

This condition is respected as we sampled 9% of the population.

2. The sample distribution should be approximately normal.

We are on the border of acceptable with the check below:

$np = 21 \times 0.5 = 10.5 > 10$
$n(1-p) = 21 \times 0.5 = 10.5 > 10$

3. Randomness: the data must come from a random sample or a randomized experiment

The data was randomly sampled.

Confidence Intervals

To determine is there is a statistical significance in the different of proportions, we study the overlap of the confidence intervals for the years of interest.

```
# year 1990
total1990 <- sum(y1990$Male) + sum(y1990$Female)
p1990 <- sum(y1990$Male) / total1990
se1990 <- sqrt((p1990*(1-p1990))/total1990)
me1990 <- qnorm(0.975) * se1990
ci1990 <- c(p1990-me1990,p1990+me1990)
```

```
# year 2019
total2019 <- sum(y2019$Male) + sum(y2019$Female)
p2019 <- sum(y2019$Male) / total2019
se2019 <- sqrt((p2019*(1-p2019))/total2019)
me2019 <- qnorm(0.975) * se2019
ci2019 <- c(p2019-me2019,p2019+me2019)
```

Table 7: Confidence Intervals

year	proportion	se	ci_lower	ci_upper
1990	0.5108898	0.000203	0.5104919	0.5112876
2019	0.5031577	0.000147	0.5028696	0.5034459

Here is a summary of the analysis. What it reveals is that there is no overlap in the confidence intervals of the male migrant proportion between 1990 and 2019. At a significance level of 95%, we reject the null hypothesis and draw the conclusion that the gender proportion has changed.

```
summary_df <- data.frame("year" = c(1990,2019),
                         "proportion" = c(p1990,p2019),
                         "se" = c(se1990,se2019),
                         "ci_lower" = c(ci1990[1],ci2019[1]),
                         "ci_upper" = c(ci1990[2],ci2019[2]))
showtable(summary_df, "Confidence Intervals")
```

Here we reach the same conclusion differently by looking at the differences in proportion and recognizing that the confidence interval does not contains the null hypothesis value of 0.

```
p_diff <- p2019-p1990
se <- sqrt((p1990*(1-p1990))/total1990+(p2019*(1-p2019))/total2019)
me <- qnorm(0.975) * se
ci_diff<- c(p_diff-me,p_diff+me)
ci_diff
```

```
## [1] -0.008223262 -0.007240805
```

Conclusion

We return to the research question: **Is there convincing evidence that the world has seen a change in its gender proportion of migrants between 1990 and 2019?**

From the exploration and analysis above, we conclude that based on our sample, there is enough evidence to reject the null hypothesis and conclude that the gender proportions have changed from 1990 to 2019.

We must note that with our sample size of 21 countries we are on the lower bound of the normality assumption.

A 95% confidence interval tells us that in 95% of the cases, we expect to capture the true population mean. This can be confirmed by simulation.

To improve upon this study we could pursue the following research:

- Run a simulation to collect many samples
- Instead of overall proportion, we could investigate the number of countries in a sample that have a male proportion greater than 50%
- Investigate the average male proportion of the sampled countries instead of the male proportion of the total migrant population of the sample countries.
- Investigate the year to year differential instead of the cummulative differences.