



# Analysis on Gender Statistics

Team 6- The Outsiders

Hsu, Quintero, Wadhwa

## Contents

0.1	Inflation . . . . .	3
0.2	Gender distribution for labour force . . . . .	9
0.3	Life expectancy for female and male at birth . . . . .	10
0.4	Employment Analysis by Country . . . . .	11
0.5	Gender Workforce Distribution by Country . . . . .	14
0.6	Gender Workforce Comparison by Industry. . . . .	14
0.7	Analysis on total unemployed males and females in developed and developing countries . . . .	20
0.8	Analysis on unemployed males and females on the bases of qualifications . . . . .	20

## 1

#1 Janice

```

# Libraries
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.0 --

## v ggplot2 3.3.3      v purrr  0.3.4
## v tibble  3.1.0      v dplyr  1.0.5
## v tidyr   1.1.3      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

library(ggplot2)
library(readr)
library(broom)
library(stringr)
library(patchwork)
library(kableExtra)

##
## Attaching package: 'kableExtra'

## The following object is masked from 'package:dplyr':
##
##   group_rows

library(knitr)
library(bookdown)
library(naniar)
library(GGally)

## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg      ggplot2

Genderstatistics <- read_csv("Data/Genderstatistics.csv")%>%
  rename('2011' = '2011 [YR2011]',
        '2012' = '2012 [YR2012]',
        '2013' = '2013 [YR2013]',
        '2014' = '2014 [YR2014]',
        '2015' = '2015 [YR2015]',
        '2016' = '2016 [YR2016]',
        '2017' = '2017 [YR2017]',
        '2018' = '2018 [YR2018]',
        '2019' = '2019 [YR2019]') %>%
  mutate(`2019` = as.numeric(`2019`))

```

```
##
## -- Column specification -----
## cols(
##   'Series Name' = col_character(),
##   'Series Code' = col_character(),
##   'Country Name' = col_character(),
##   'Country Code' = col_character(),
##   '2011 [YR2011]' = col_double(),
##   '2012 [YR2012]' = col_double(),
##   '2013 [YR2013]' = col_double(),
##   '2014 [YR2014]' = col_double(),
##   '2015 [YR2015]' = col_double(),
##   '2016 [YR2016]' = col_double(),
##   '2017 [YR2017]' = col_double(),
##   '2018 [YR2018]' = col_double(),
##   '2019 [YR2019]' = col_character()
## )

## Warning in mask$eval_all_mutate(quo): NAs introduced by coercion

analysis<- Genderstatistics %>%
  filter(`Series Name` %in% c("Inflation, consumer prices (annual %)", "Population ages 15-64, female"),
  select(-c(`Series Code`,`Country Code`)) %>%
  pivot_longer(cols = -c(`Country Name`,`Series Name`),
    names_to = "Year",
    values_to = "count") %>%
  pivot_wider(names_from = "Series Name",
    values_from = "count")

analysis <- analysis %>%
mutate(Inflation = as.numeric(`Inflation, consumer prices (annual %)`)) %>%
  mutate(Population_ages_15_64_female = as.numeric(`Population ages 15-64, female`)) %>%
  mutate(Population_ages_15_64_male = as.numeric(`Population ages 15-64, male`)) %>%
  mutate(labour_force = `Population_ages_15_64_female`+`Population_ages_15_64_male`) %>%
  mutate(Year = as.numeric(Year)) %>%
  mutate(Life_expectancy_at_birth_female = as.numeric(`Life expectancy at birth, female (years)`)) %>%
  mutate(Life_expectancy_at_birth_male = as.numeric(`Life expectancy at birth, male (years)`)) %>%
  select(-c(`Inflation, consumer prices (annual %)` , `Population ages 15-64, female`, `Population ages 15-
```

## 0.1 Inflation

```
library(purrr)
library(ggplot2)
library(patchwork)
countries <- c("Colombia", "United States", "France", "Egypt, Arab Rep.")

infla_labour <- function(countries){
  p1 <- analysis %>%
  filter(`Country Name`== countries)%>%
  na.omit()%>%
  ggplot(aes(x=Year, y=Inflation)) +
```

```

geom_line(color="#69b3a2", size=2) +
scale_x_continuous(breaks = c(2011:2019))+
  scale_y_continuous(labels = scales::comma)+
ggtitle("Inflation rate") +
  labs(title = countries)

p2 <- analysis %>%
  filter(`Country Name`== countries)%>%
  na.omit()%>%
ggplot(aes(x=Year, y=labour_force)) +
  geom_line(color="grey",size=2) +
  scale_x_continuous(breaks = c(2011:2019))+
  scale_y_continuous(labels = scales::comma)+
  ggtitle("number of labour force") +
  labs(title = countries)

p1 + p2
}

Q1 <- map(countries, infla_labour)
print(Q1)

```

```
## [[1]]
```

```
##
## [[2]]
```

```
##
## [[3]]
```

```
##
## [[4]]
```

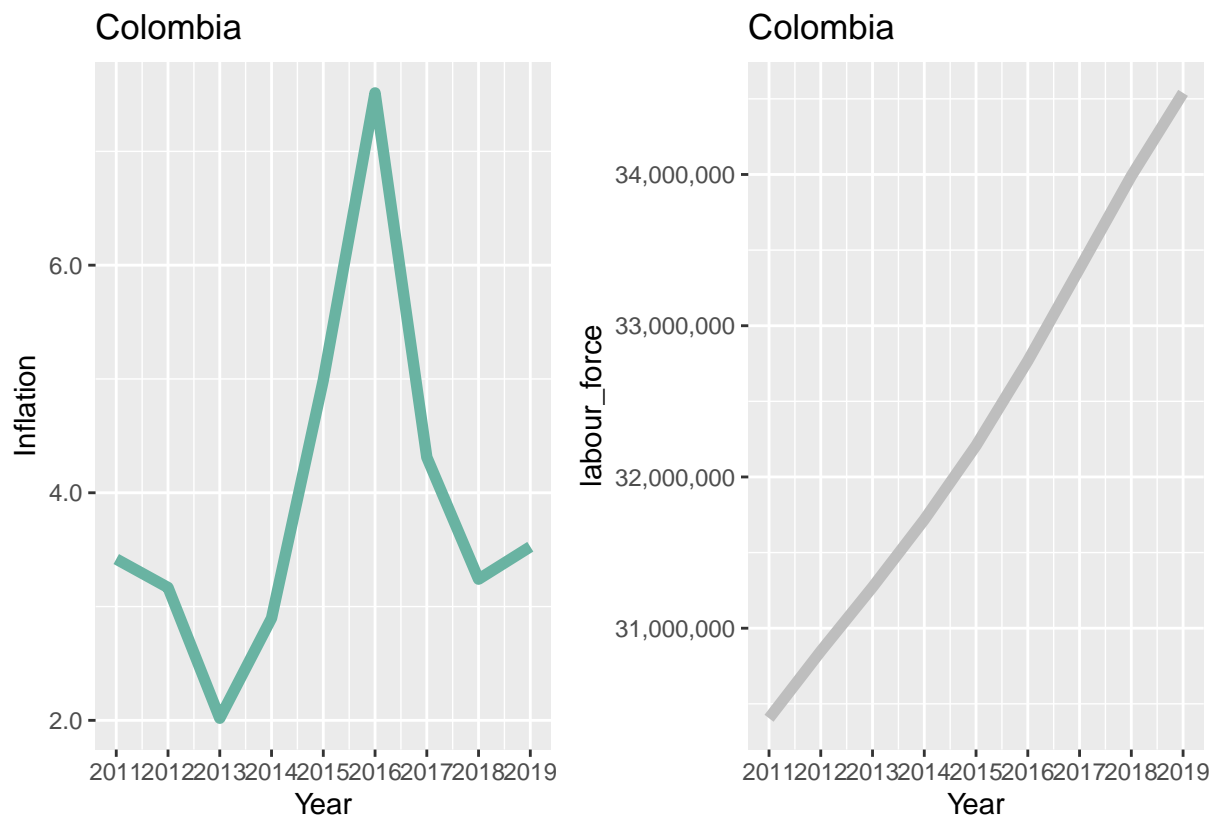


Figure 1: Inflation vs Labour force

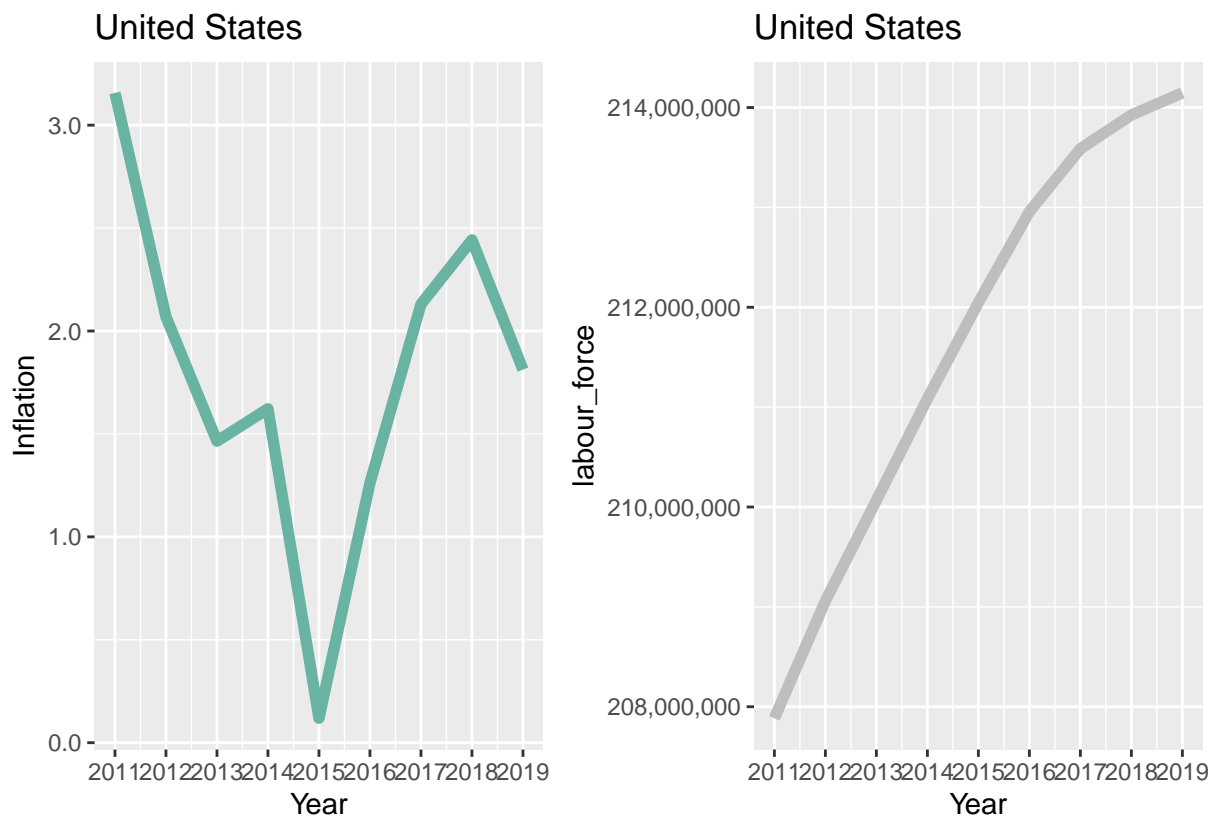


Figure 2: Inflation vs Labour force

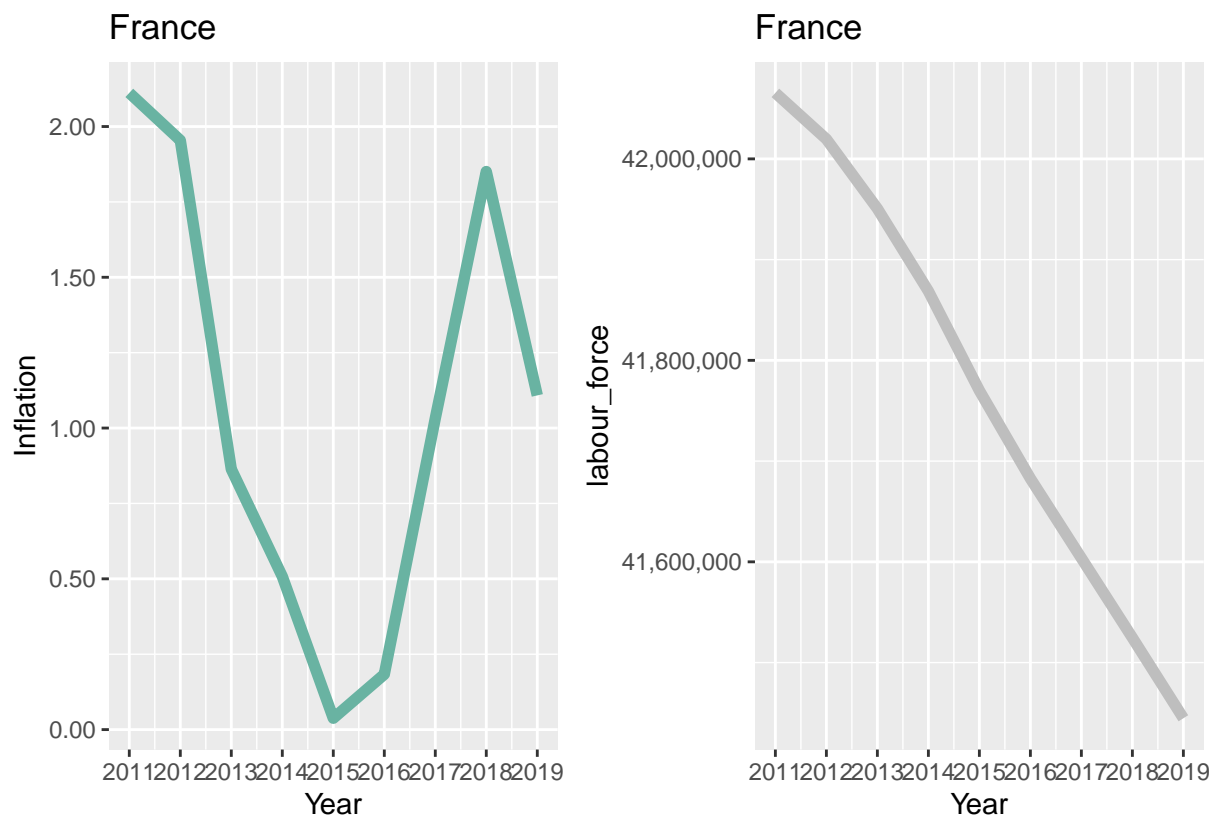
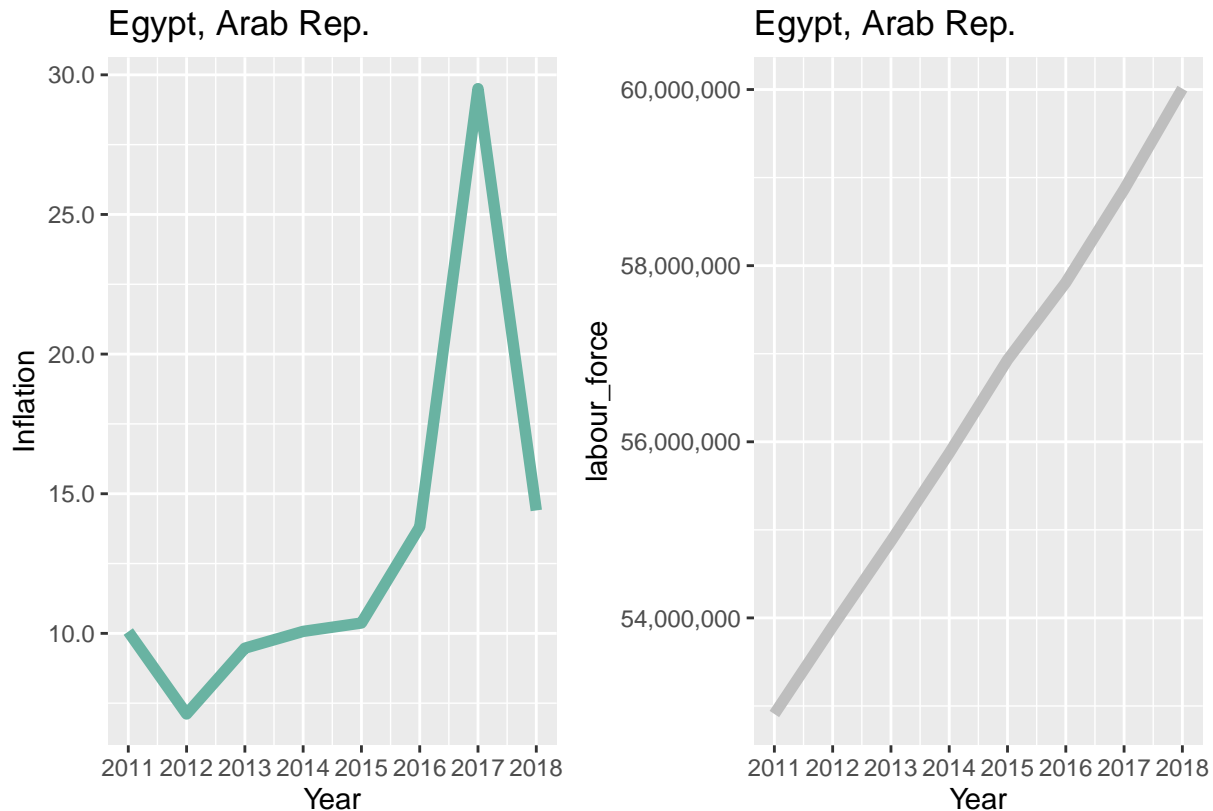


Figure 3: Inflation vs Labour force



According to Figure@ref(fig:A1), developed countries such as United States and France both have the lowest inflation rate in 2015. This is because in 2015, the crude oil price collapsed. (<https://blogs.worldbank.org/developmenttalk/what-triggered-oil-price-plunge-2014-2016-and-why-it-failed-deliver-economic-impetus-eight-charts>), and the global economy has not recovered from the GFC yet. (<https://www.reuters.com/article/us-usa-economy-inflation-idUSKCN0UY1LH>)).

Moreover, low inflation rate does not mean the currency is more valuable. On the contrary, it signals demand for goods and services is lower than expected and will then result in recession and the an increase in unemployment. (<https://www.weforum.org/agenda/2019/06/inflation-is-healthy-for-the-economy-but-too-much-can-trigger-a-recession-7d37501704>)

Additionally, developed countries usually have more stable inflation rate than developing countries. This is to keep the economy and the currency stable.

```
library(kableExtra)
t1 <- analysis %>%
  select(`Country Name`,Year,Inflation,labour_force) %>%
  filter(Year == "2015") %>%
  knitr::kable(
    caption = "Inflation and the labour force in 2015"
  ) %>%
  kable_styling(c("hover", "striped"))
t1
```

Table @ref(tab:A2) shows the 2015 inflation rate.



Table 1: Inflation and the labour force in 2015

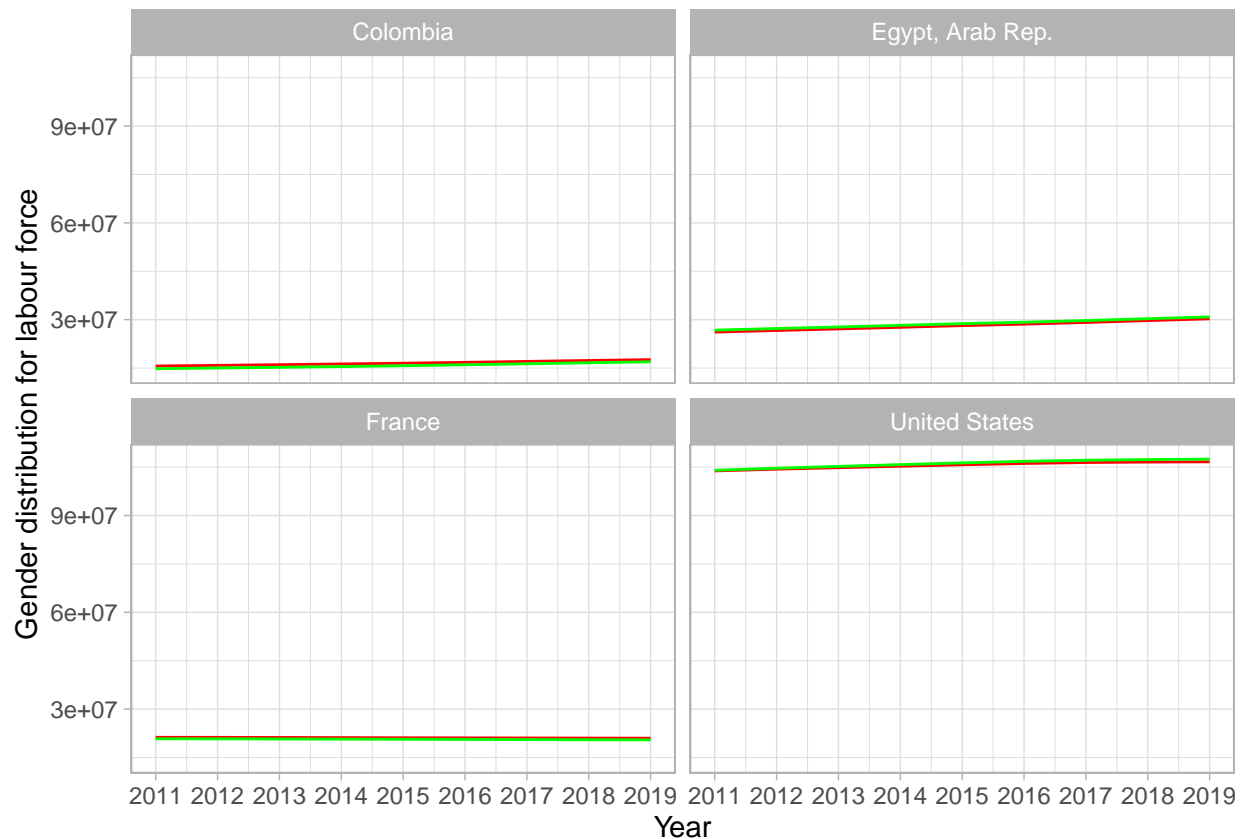
Country Name	Year	Inflation	labour_force
Colombia	2015	4.9902343	32207438
United States	2015	0.1186271	212046898
France	2015	0.0375144	41770007
Egypt, Arab Rep.	2015	10.3704903	56930104

## 0.2 Gender distribution for labour force

Moreover, the labour force in each country is increasing but decreasing in France. We will take a deeper look in the employment and unemployment in the following sections and try to conclude why the labour force for France is decreasing.

Let's also look at the gender distribution in the labour force.

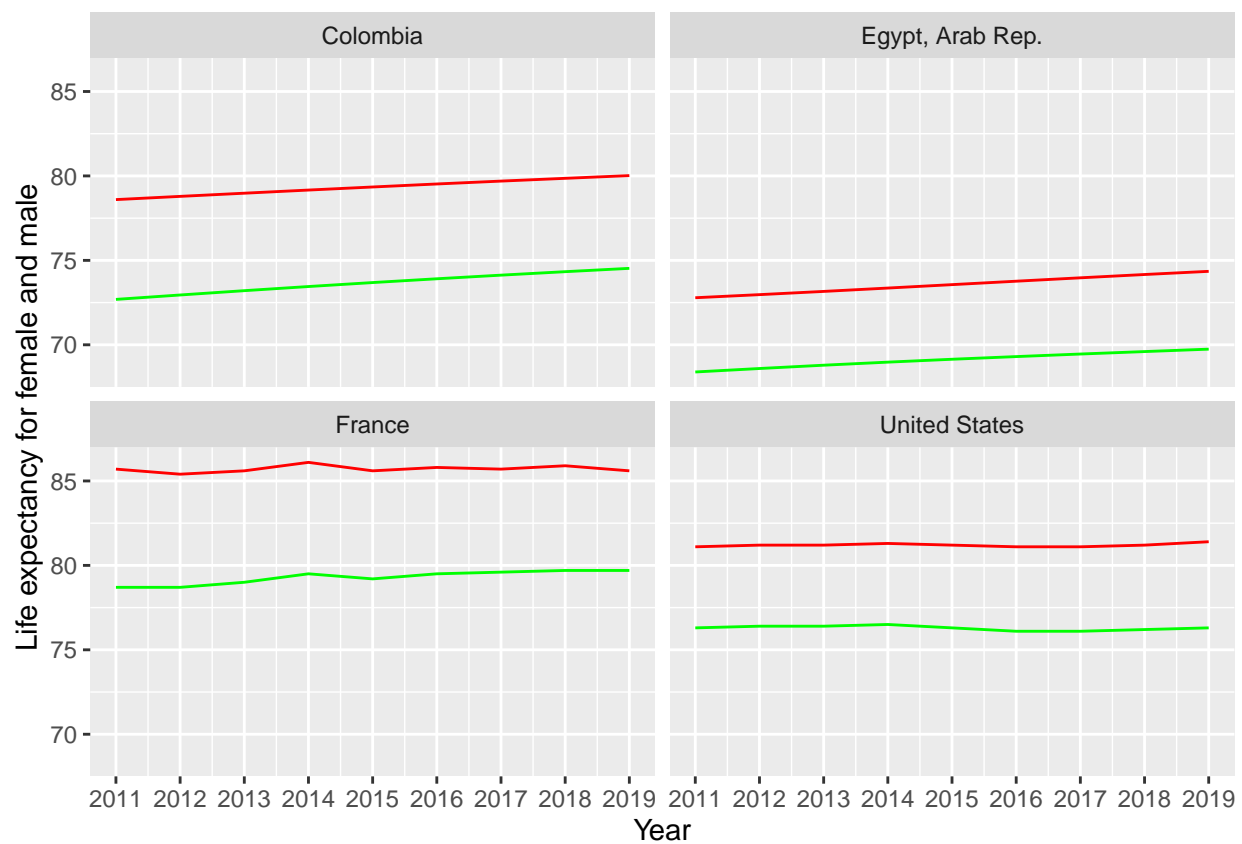
```
analysis %>%
  ggplot(aes(x = Year))+
  geom_line(aes(y = Population_ages_15_64_female), color = "red")+
  geom_line(aes(y = Population_ages_15_64_male), color = "green")+
  ylab("Gender distribution for labour force")+
  scale_x_continuous(breaks = c(2011:2019))+
  theme_light()+
  facet_wrap(~`Country Name`)
```



Figure@ref(fig:A3) shows the lines for female labour force and the male labour force are almost overlapped with each other, meaning the gender distribution for labour force is fairly equal in these countries.

### 0.3 Life expectancy for female and male at birth

```
analysis %>%
  ggplot(aes(x = Year))+
  geom_line(aes(y = Life_expectancy_at_birth_female), color = "red")+
  geom_line(aes(y = Life_expectancy_at_birth_male), color = "green")+
  ylab("Life expectancy for female and male")+
  scale_x_continuous(breaks = c(2011:2019))+
  facet_wrap(~`Country Name`)
```



Figure@ref(fig:A4) shows the life expectancy for female is obviously higher than male. More interestingly, US and France are having a stable life expectancy from 2011 til 2019, while in Egypt and Colombia, the life expectancy for both genders is increasing.

## 2

```

readingdata <- read_csv("data/088121e0-ea83-4b15-be2a-1bcc36ea893f_Data.csv")

byyear <- readingdata %>% select(c(`Series Name`, `Country Name`, `2011 [YR2011]`:`2019 [YR2019]`)) %>%
  filter(`Series Name` %in% c("Employment in agriculture, female (% of female employment) (modeled ILO est.",
    "Employment in agriculture, male (% of male employment) (modeled ILO est.",
    "Employment in industry, female (% of female employment) (modeled ILO est.",
    "Employment in industry, male (% of male employment) (modeled ILO est.",
    "Employment in services, female (% of female employment) (modeled ILO est.",
    "Employment in services, male (% of male employment) (modeled ILO est."))
  mutate(`Series Name` = case_when(`Series Name` == "Employment in agriculture, female (% of female employment) (modeled ILO est." ~ "Employment in agriculture, female (% of female employment)",
    `Series Name` == "Employment in agriculture, male (% of male employment) (modeled ILO est." ~ "Employment in agriculture, male (% of male employment)",
    `Series Name` == "Employment in industry, female (% of female employment) (modeled ILO est." ~ "Employment in industry, female (% of female employment)",
    `Series Name` == "Employment in industry, male (% of male employment) (modeled ILO est." ~ "Employment in industry, male (% of male employment)",
    `Series Name` == "Employment in services, female (% of female employment) (modeled ILO est." ~ "Employment in services, female (% of female employment)",
    `Series Name` == "Employment in services, male (% of male employment) (modeled ILO est." ~ "Employment in services, male (% of male employment)",
    TRUE ~ `Country Name`),
    `Country Name` = case_when(`Country Name` == "Egypt, Arab Rep." ~ "Egypt",
    TRUE ~ `Country Name`),
    "2011" = as.double(`2011 [YR2011]`),
    "2012" = as.double(`2012 [YR2012]`),
    "2013" = as.double(`2013 [YR2013]`),
    "2014" = as.double(`2014 [YR2014]`),
    "2015" = as.double(`2015 [YR2015]`),
    "2016" = as.double(`2016 [YR2016]`),
    "2017" = as.double(`2017 [YR2017]`),
    "2018" = as.double(`2018 [YR2018]`),
    "2019" = as.double(`2019 [YR2019]`)) %>%
  filter(`Country Name` %in% c("United States", "Egypt", "Colombia", "France")) %>%
  select(`Series Name`,
    `Country Name`,
    `2011`:`2019`) %>%
  pivot_longer(names_to = "Year",
    values_to = "Percentage",
    cols = c(-`Series Name`,
      -`Country Name`))

```

## 0.4 Employment Analysis by Country

The core analysis of this report is to analyze the different workforce distribution among high and low income countries and the gender distribution inside them. High income countries such as United States and France and low income such as Colombia and Egypt were taken into account to evaluate the labor force condition and the general trends of the citizens performing jobs in agriculture, industry and services jobs.

Having a closer look to the data, the distribution in the job market according to the gender and country it is taking part in, tends to variate according to the economy of each country. High income countries such as United States or France manages a similar trend in every industry according to the gender. But also, it can be seen that the rates are different compared to the low income countries.

### 0.4.1 High Income Countries Workforce

The higher income countries have, the higher participation in services by males. On the other hand, there is lower concentration in the participation of female in industries that are considered “Masculine” for the time being, agriculture and industry.

For high income countries the similarity in the allocation of workforce among the studies industries is surprisingly similar. in the figure @ref(fig:highinc) USA and France have an average of 67% of male workforce, also, have a similar percentage by 2019 in industry of 30% and in agriculture of 4%.

Generally, females are on top of the chart with about 90% working in services, and with similar rates for agriculture and industry of 1% and 9% respectively.

```
usafraemp <- byyear %>%
  filter(`Country Name` %in% c("United States",
                                "France")) %>%

  ggplot(aes(x = `Year`,
             y = `Percentage`,
             group = `Series Name`,
             color = `Series Name`))+
  facet_wrap(~`Country Name`)+
  geom_point(size = 2.5)+
  geom_line(aes(linetype=`Series Name` %in% c("Females in Agriculture (%)",
                                              "Females in Industry (%)",
                                              "Females in Services (%)" ),
               size = 1.5,
               show.legend = F))+
  theme(legend.position = "bottom")+
  labs(title = "Gender Workforce in High Income Countries")+
  scale_y_continuous(labels = function(x) paste0(x*1, "%"),
                    breaks = seq(0,100,10),
                    limits = c(0,100))

usafraemp
```



### 0.4.2 Low Income Countries Workforce

Males in the job market for low income countries keep similar trends for the jobs in agriculture. In 2019, in Colombia and Egypt got in average 25% of male participation. In industry, Colombia and Egypt have a notorious difference of 25% and 32% respectively.

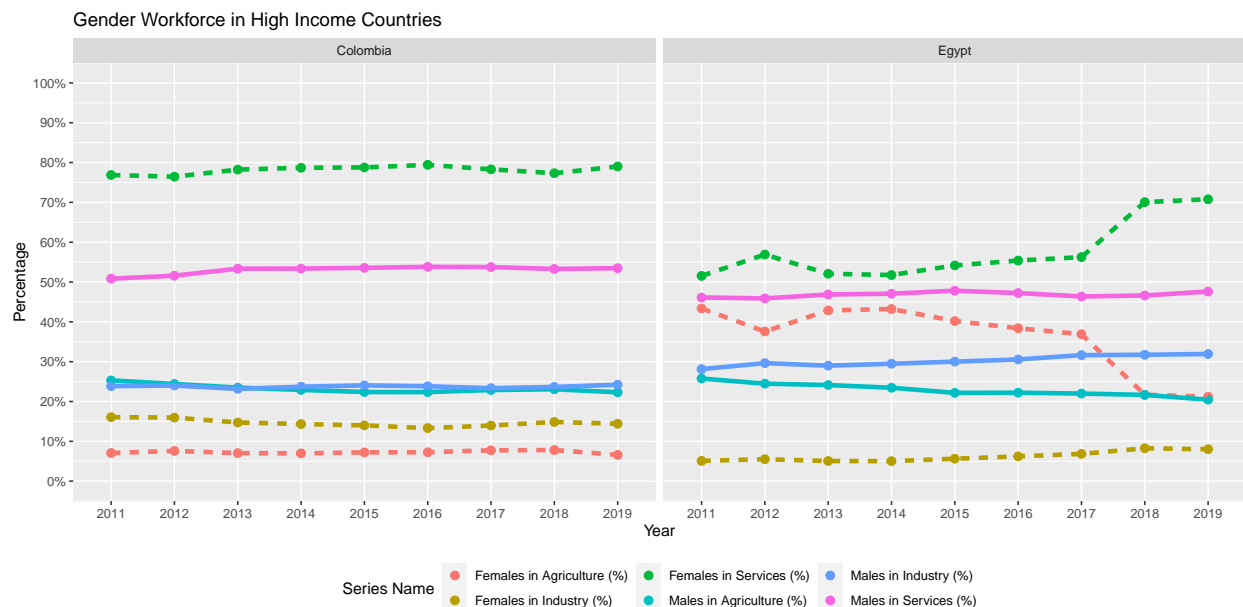
In the case of females, they keep an average from 10% to 20% in “Masculine” jobs, as per in services, they keep from 50% to 80%. Egypt had a huge decline in the agricultural jobs for females from a 43% to 21%. It is remarkable this variation along the previous 9 years as well as that the woman workforce seemed to move in the same rate to the services industry as seen in the figure @ref(fig:lowinc) below.

```
colegyemp <- byyear %>%
  filter(`Country Name` %in% c("Colombia",
                                "Egypt")) %>%

  ggplot(aes(x = `Year`,
             y = `Percentage`,
             group = `Series Name`,
             color = `Series Name`))+
  facet_wrap(~`Country Name`)+
  geom_point(size = 2.5)+
  geom_line(aes(linetype=`Series Name` %in% c("Females in Agriculture (%)",
                                              "Females in Industry (%)",
                                              "Females in Services (%)"),

              size= 1.5,
              show.legend = F))+
  theme(legend.position = "bottom")+
  labs(title = "Gender Workforce in High Income Countries")+
  scale_y_continuous(labels = function(x) paste0(x*1, "%"),
                    breaks = seq(0,100,10),
                    limits = c(0,100))

colegyemp
```

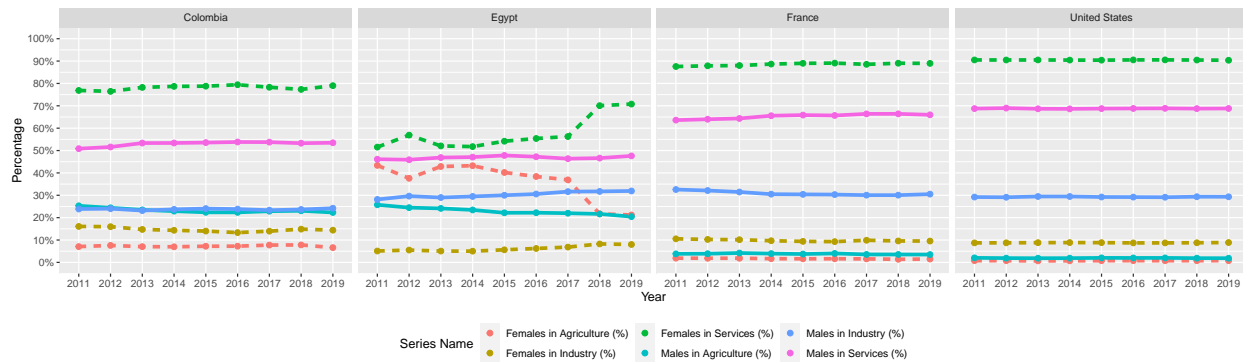


## 0.5 Gender Workforce Distribution by Country

As seen in the current analysis the economical capacity of the countries can infer in the workforce distribution. In a general view of the selected samples, females have in all of them the highest rate of employment in services as well as the lowest in agriculture, this does not apply in Egypt but the current trend is showing that there is a moving out of that industry.

Similarly, males have the highest level of employment in services, but they keep leading the industry and agricultural workforce. It is noticeable that if a country is wealthy, there is a higher level of participation of females in service jobs compared to low income countries. On the other hand, despite the income level of the country males keep the same percentage among the countries. See figure @ref(fig:allcountries).

```
byyear %>% ggplot(aes(x = `Year`,
                      y = `Percentage`,
                      group = `Series Name`,
                      color = `Series Name`))+
  facet_wrap(~`Country Name`, ncol = 4)+
  theme(legend.position = "bottom")+
  scale_y_continuous(labels = function(x) paste0(x*1, "%"),
                    breaks = seq(0,100,10),
                    limits = c(0,100))+
  geom_line(aes(linetype=`Series Name` %in% c("Females in Agriculture (%)",
                                              "Females in Industry (%)",
                                              "Females in Services (%)"),
              size = 1.5,
              show.legend = F))+
  geom_point(size=2)
```



## 0.6 Gender Workforce Comparison by Industry.

From other point of view and analyzing the variables across all the countries it can be seen that the genders maintain a similar level of employment according to the selected industries.

### 0.6.1 Female and Male Employment in Agriculture.

in the figure @ref(fig:agriculture) women keep the lowest participation in agriculture and industry jobs in the selected countries except Egypt which have had a decrease around of 50% during the past 9 years, keeping lowest rates compared to male results. The level of jobs have been steady for males and females during the analyzed years.

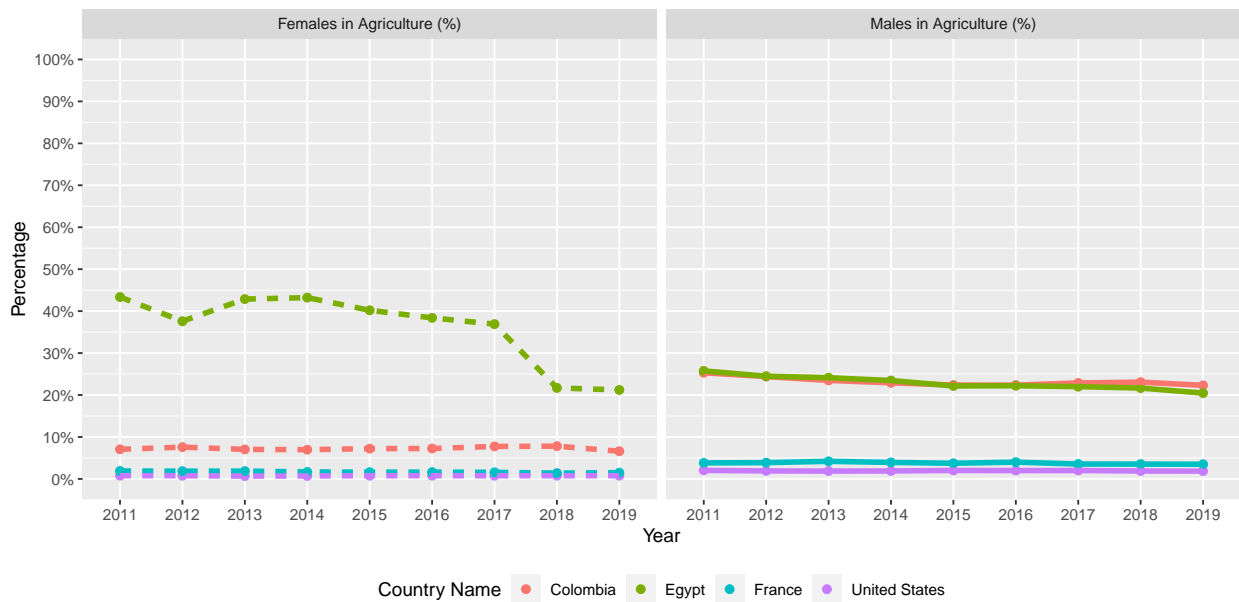
```

byyear %>%
  filter(`Series Name` %in% c("Males in Agriculture (%)",
                              "Females in Agriculture (%)"),
         `Country Name` %in% c("Egypt",
                              "Colombia",
                              "United States",
                              "France")) %>%

  ggplot(aes(x = `Year`,
             y = `Percentage`,
             group = `Country Name`,
             color = `Country Name`))+
  geom_point(size = 2)+
  facet_wrap(~`Series Name`,
            ncol = 2)+
  theme(legend.position = "bottom")+
  scale_y_continuous(labels = function(x) paste0(x*1, "%"),
                    breaks = seq(0,100,10),
                    limits = c(0,100))+
  geom_line(aes(linetype=`Series Name` %in% c("Females in Agriculture (%)",
                                              "Females in Industry (%)",
                                              "Females in Services (%)")),

            size= 1.5,
            show.legend = F)

```



### 0.6.2 Female and Male Employment in Industry.

In the case of industry jobs, male keep a highest rate compared to female across the analyzed countries. In 2019 the number of female rose by 5% in Egypt, while in Colombia decreased by 2% and in USA and France maintained the same levels.

In the case of males, all the countries maintained about the same levels since 2011 as seen in the figure @ref(fig:industry).

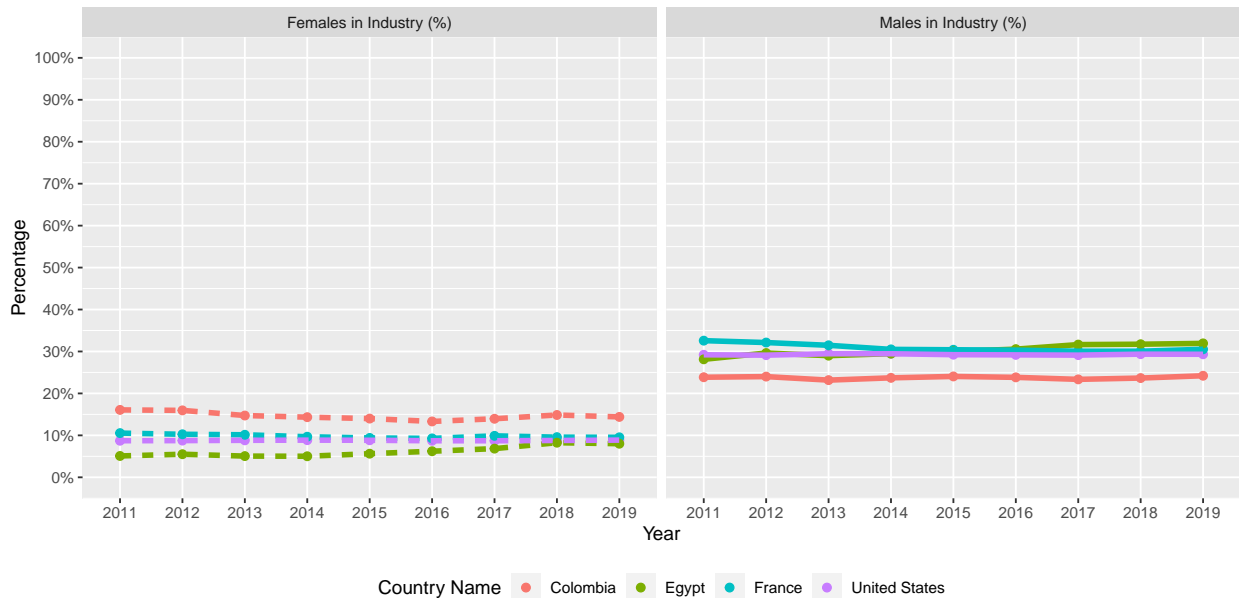
```

byyear %>%
  filter(`Series Name` %in% c("Males in Industry (%)",
                              "Females in Industry (%)"),
         `Country Name` %in% c("Egypt",
                              "Colombia",
                              "United States",
                              "France")) %>%

  ggplot(aes(x = `Year`,
             y = `Percentage`,
             group = `Country Name`,
             color = `Country Name`))+
  geom_point(size = 2)+
  facet_wrap(~`Series Name`,
            ncol = 2)+
  theme(legend.position = "bottom")+
  scale_y_continuous(labels = function(x) paste0(x*1, "%"),
                    breaks = seq(0,100,10),
                    limits = c(0,100))+
  geom_line(aes(linetype=`Series Name` %in% c("Females in Agriculture (%)",
                                              "Females in Industry (%)",
                                              "Females in Services (%)")),

            size= 1.5,
            show.legend = F)

```



### 0.6.3 Female and Male Employment in Services.

Female have a highest participation in the services sector compared to males and across all the industries.

In general, all the countries kept the same average levels since 2011 and they are in a similar range despite the income level of the country. But, Egypt has an interesting variation of the the jobs allocation. In this case, females in services have rose over 30% in the last years, maintaining the leading over their male peers. In this Industry Egypt has the lowest of people, but the trend keeps a future positive path as well as France. See figure @ref(fig:services).



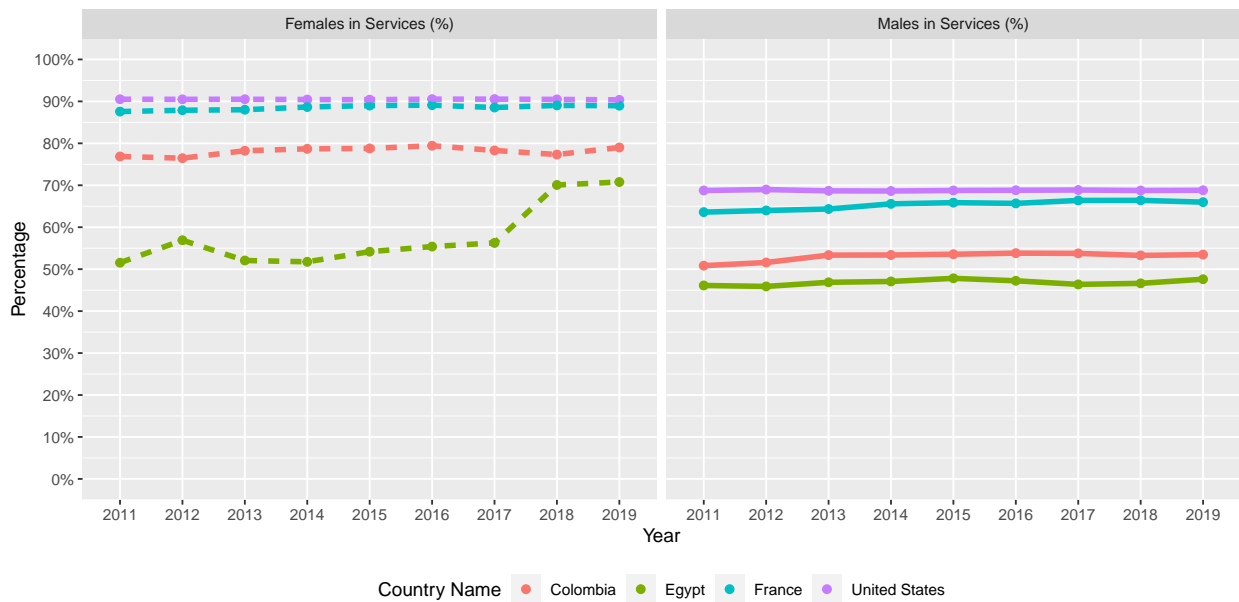
```

byyear %>%
  filter(`Series Name` %in% c("Males in Services (%)",
                              "Females in Services (%)"),
         `Country Name` %in% c("Egypt",
                              "Colombia",
                              "United States",
                              "France")) %>%

  ggplot(aes(x = `Year`,
             y = `Percentage`,
             group = `Country Name`,
             color = `Country Name`))+
  geom_point(size = 2)+
  facet_wrap(~`Series Name`,
            ncol = 2)+
  theme(legend.position = "bottom")+
  scale_y_continuous(labels = function(x) paste0(x*1, "%"),
                    breaks = seq(0,100,10),
                    limits = c(0,100))+
  geom_line(aes(linetype=`Series Name` %in% c("Females in Agriculture (%)",
                                              "Females in Industry (%)",
                                              "Females in Services (%)")),

            size= 1.5,
            show.legend = F)

```



#### 0.6.4 Gender Workforce by Industry

In summary, females across all countries have lower levels than male occupations in agriculture and industry. On the other hand females have the lead in the services job market, the most notorious case of growth in this industry was Egypt that in 2019 the percentage of females in this industry was 71%.

```

byyear %>% ggplot(aes(x = `Year`,
                     y = `Percentage`,
                     group = `Country Name`,
                     color = `Country Name`))+
  facet_wrap(~fct_relevel(`Series Name`, c("Females in Agriculture (%)",
                                           "Males in Agriculture (%)",
                                           "Females in Industry (%)",
                                           "Males in Industry (%)",
                                           "Females in Services (%)"),
            ncol = 2))+
  theme(legend.position = "bottom")+
  scale_y_continuous(labels = function(x) paste0(x*1, "%"),
                    breaks = seq(0,100,10),
                    limits = c(0,100))+
  geom_line(aes(linetype=`Series Name` %in% c("Females in Agriculture (%)",
                                              "Females in Industry (%)",
                                              "Females in Services (%)"),
              size= 1.5,
              show.legend = F,
              alpha = 0.8))+
  geom_point(size=2)

```

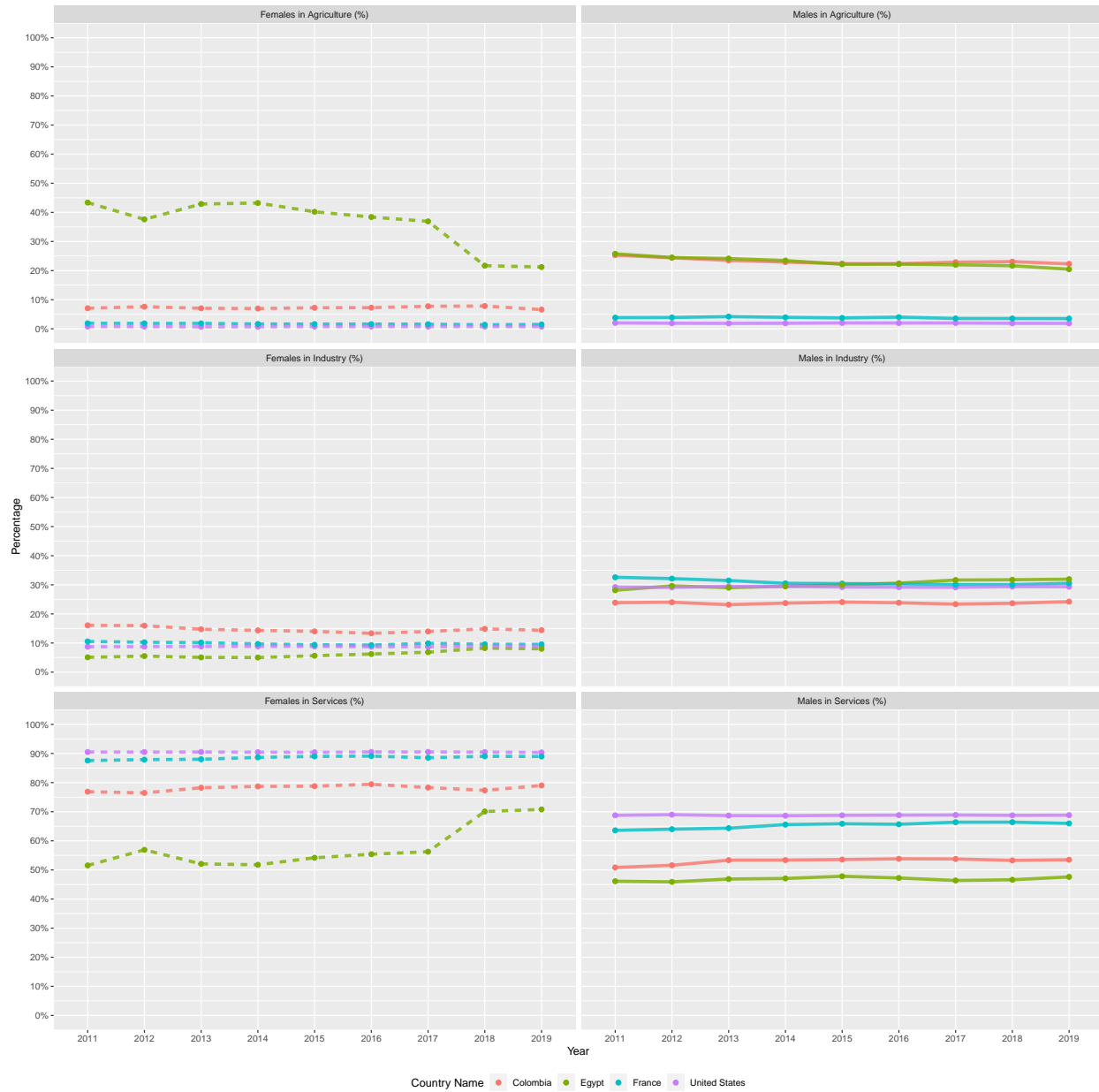


Table 2: Unemployed percentage of males and females in Developed countries

Country Name	Year	Unemployed_F	Unemployed_M
France	2011	9.12	8.530000
France	2012	9.36	9.440000
France	2013	9.79	10.040000
France	2014	10.03	10.540000
France	2015	9.91	10.770001
France	2016	9.84	10.220000
France	2017	9.37	9.440000
France	2018	9.05	8.990000
France	2019	8.38	8.500000
United States	2011	8.46	9.370000
United States	2012	7.89	8.229999
United States	2013	7.08	7.640000
United States	2014	6.06	6.260000
<b>United States</b>	<b>2015</b>	<b>5.18</b>	<b>5.370000</b>
United States	2016	4.79	4.940000
United States	2017	4.31	4.400000
United States	2018	3.84	3.950000
United States	2019	3.61	3.720000

### 3

## 0.7 Analysis on total unemployed males and females in developed and developing countries

### United States & France

#### Colombia & Egypt, Arab Rep.

From table @ref(tab:tabref) and table @ref(tab:tabref1) we can summarize that clearly the percentage of female unemployment is way *more in developing countries* than that of the *developed countries* like US and France. For instance, in **2015** the reported percentage of females unemployed in United States was just 5% where as it was 24 percentage in developing countries like Egypt.

The gap in participation rates between men and women is narrowing in developed countries but continues to widen in developing countries, as we can observe that the percentage is almost equal for both males and females in US and France where as in Columbia and Egypt, the employment rate is more for men than women.(?)

Another interesting observation from table @ref(tab:tabref) and table @ref(tab:tabref1) was seen that the overall unemployment trend in the developing countries is more than that of developed countries. The basic cause of this can be the deficiency of the availability of essential consumer goods, often called wage goods (?).

## 0.8 Analysis on unemployed males and females on the bases of qualifications

### Advance Education

Figure @ref(fig:Advancedeved) shows the unemployment percentage of males and females with advanced education in developed countries like United states and France.

Figure @ref(fig:Advancedeving) shows the unemployment percentage of males and females with advanced education in developing countries like Colombia and Egypt.

Table 3: Unemployed percentage of males and females in Developing countries

Country Name	Year	Unemployed_F	Unemployed_M
Colombia	2011	13.10	7.910000
Colombia	2012	12.66	7.550000
Colombia	2013	11.67	7.070000
Colombia	2014	11.03	6.720000
Colombia	2015	10.84	6.360000
Colombia	2016	11.21	6.780000
Colombia	2017	11.51	6.870000
Colombia	2018	11.79	7.090000
Colombia	2019	12.75	7.880000
Egypt, Arab Rep.	2011	22.44	8.770001
Egypt, Arab Rep.	2012	24.01	9.229999
Egypt, Arab Rep.	2013	24.17	9.800000
Egypt, Arab Rep.	2014	24.00	9.729999
<b>Egypt, Arab Rep.</b>	<b>2015</b>	<b>24.81</b>	<b>9.390000</b>
Egypt, Arab Rep.	2016	23.58	8.840000
Egypt, Arab Rep.	2017	23.01	8.220000
Egypt, Arab Rep.	2018	21.34	6.770000
Egypt, Arab Rep.	2019	NA	NA

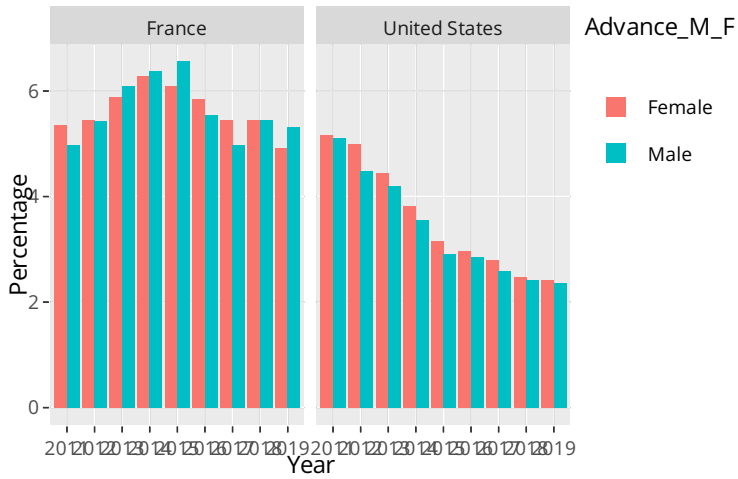


Figure 4: Unemployment with advanced education in developed countries

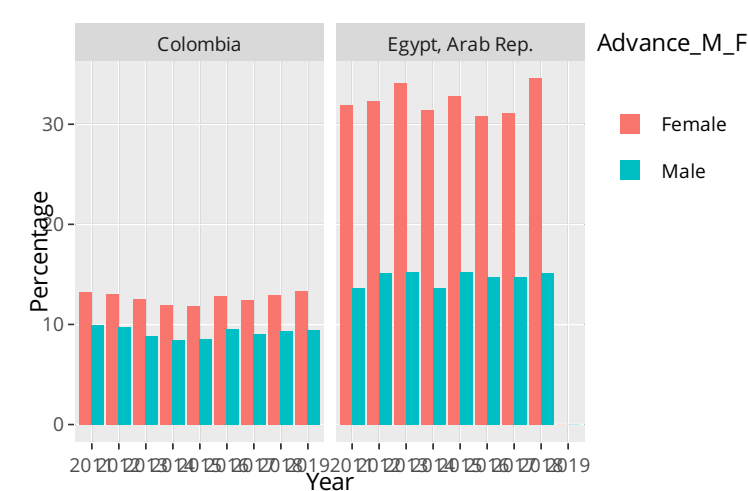


Figure 5: Unemployment with advanced education in developing countries

Basic Education

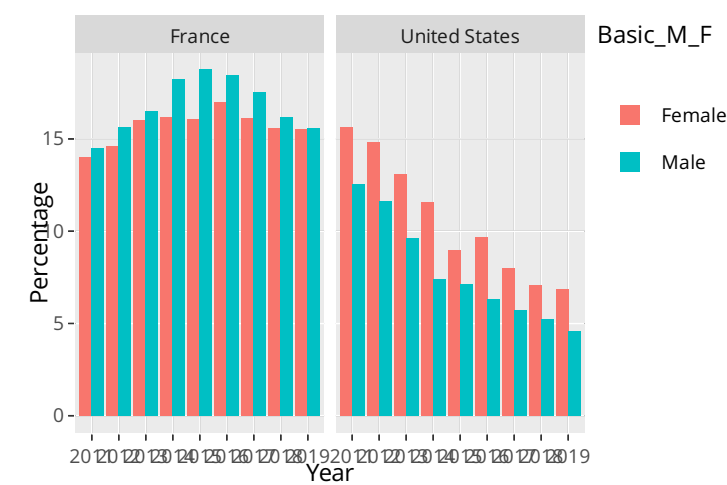


Figure 6: Unemployment with basic education in developed countries

Figure @ref(fig:Basicdeved) shows the unemployment percentage of males and females with basic education in developed countries like United states and France.

Figure @ref(fig:Basicdeving) shows the unemployment percentage of males and females with basic education in developing countries like Colombia and Egypt.

Intermediate Education

Figure @ref(fig:Intermediatedeved) shows the unemployment percentage of males and females with intermediate education in developed countries like United states and France.

Figure @ref(fig:Intermediatedeving) shows the unemployment percentage of males and females with intermediate education in developing countries like Colombia and Egypt.

Conclusion

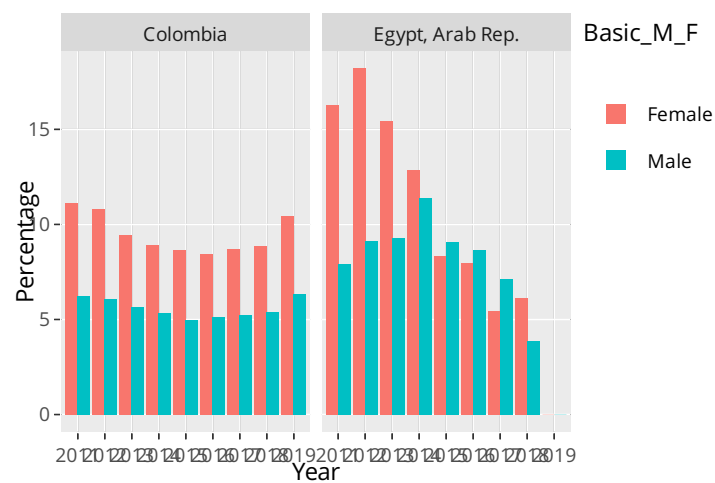


Figure 7: Unemployment with basic education in developing countries



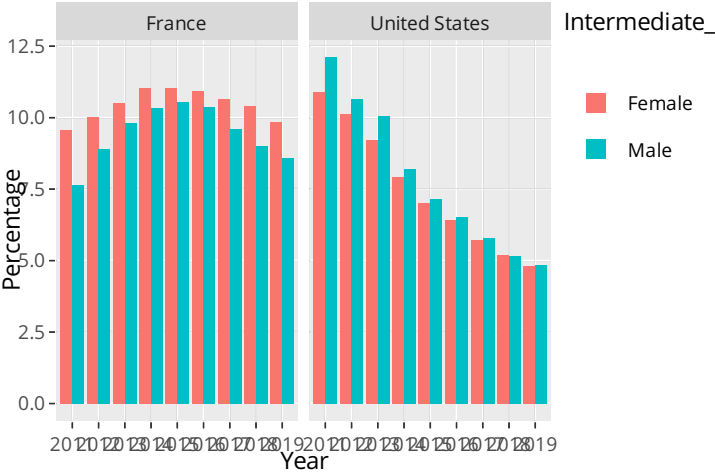


Figure 8: Unemployment with intermediate education in developed countries

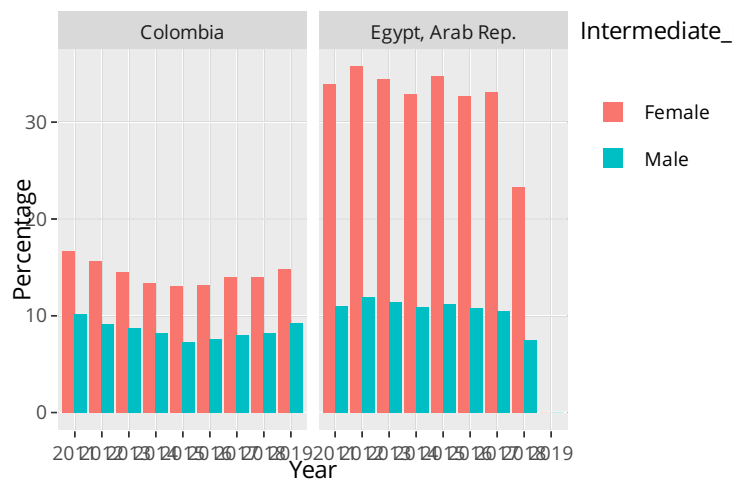


Figure 9: Unemployment with intermediate education in developing countries