midterm_project

YuzheHuang

2019/10/18

Research on Female Enrollment Rate in Eight Main Areas across the World between 1986 and 2018

1.Setup

```r
setwd("C:/Users/huang/Desktop/midterm_project_615")
data<-read.csv("education_gender.csv",header = TRUE)
library(tidyr)
library(stringr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(ggplot2)
```

2.Data wrangling

```r
n<-colnames(data)
nn<-n[5:37]
data1<-gather(data,year,value,nn)
data2<-unite(data1,"varname",Series,Series.Code, sep="-",remove = TRUE)
data3<-spread(data2,varname,value)

data3$year<-as.numeric(str_sub(data3$year,2,5))

a<-c("country_name","country_code","year","pre_primary_both","pre_primary_female","primary_female","sec
colnames(data3)<-a
```

3.Calculation

3.1 Calculate total female education ratio each year

```r
data4<-data3%>%group_by(year)%>%mutate(total_pre_edu_ratio=(sum(as.numeric(pre_primary_female)))/(sum(as
```

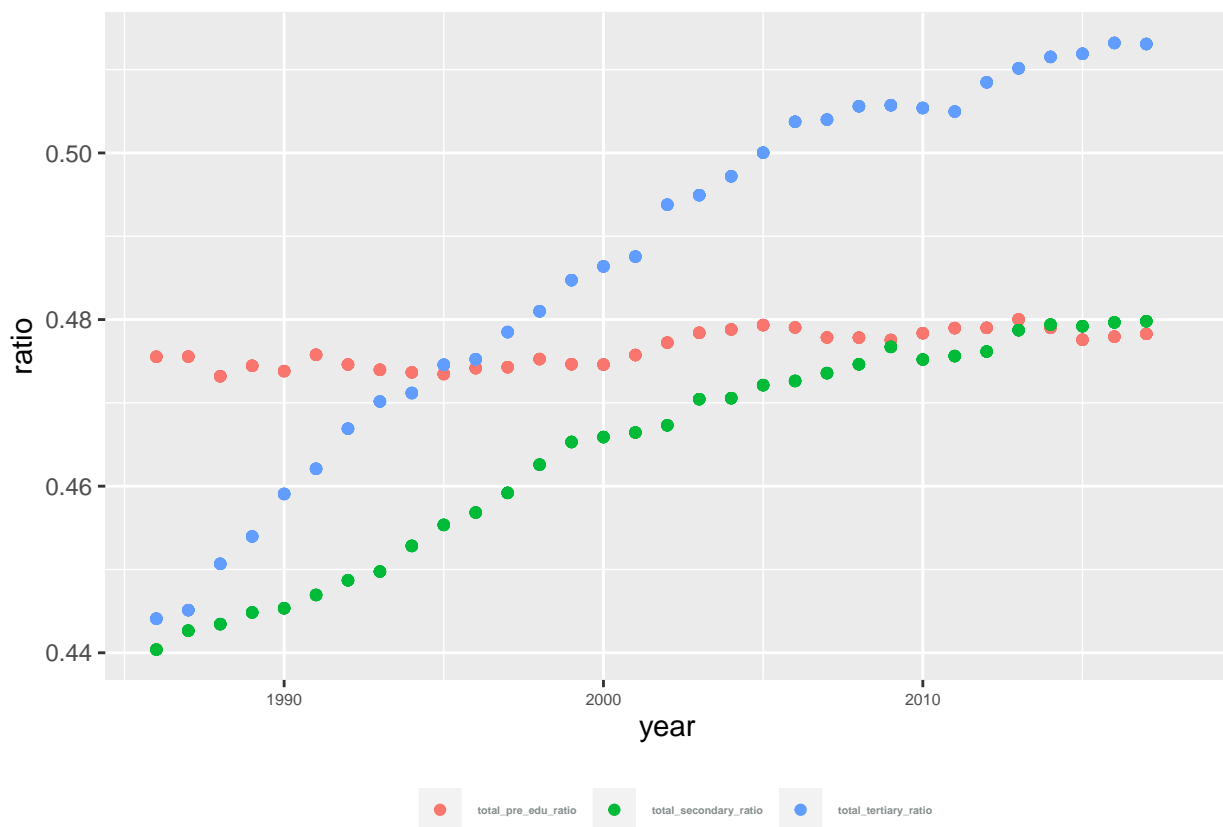### 3.2 Calculate female education ratio by country each year

```
data5<-data4%>%mutate(pre_edu_ratio=as.numeric(pre_primary_female)/as.numeric(pre_primary_both),seconda
```

4. Exploratory Data Analysis

4.1 Overview of female ratio in pre-primary, secondary and tertiary education between 1986 and 2018.

We can see that more females enrolled in tertiary schools compared with pre-primary and secondary ones after 1995.

```
total<-data5[,c(3,11:13)]
totalrev<-gather(total,totaledu,ratio,-year)
ggplot(data=totalrev)+geom_point(mapping=aes(x=year,y=ratio,color=totaledu))+guides(color=guide_legend(
```
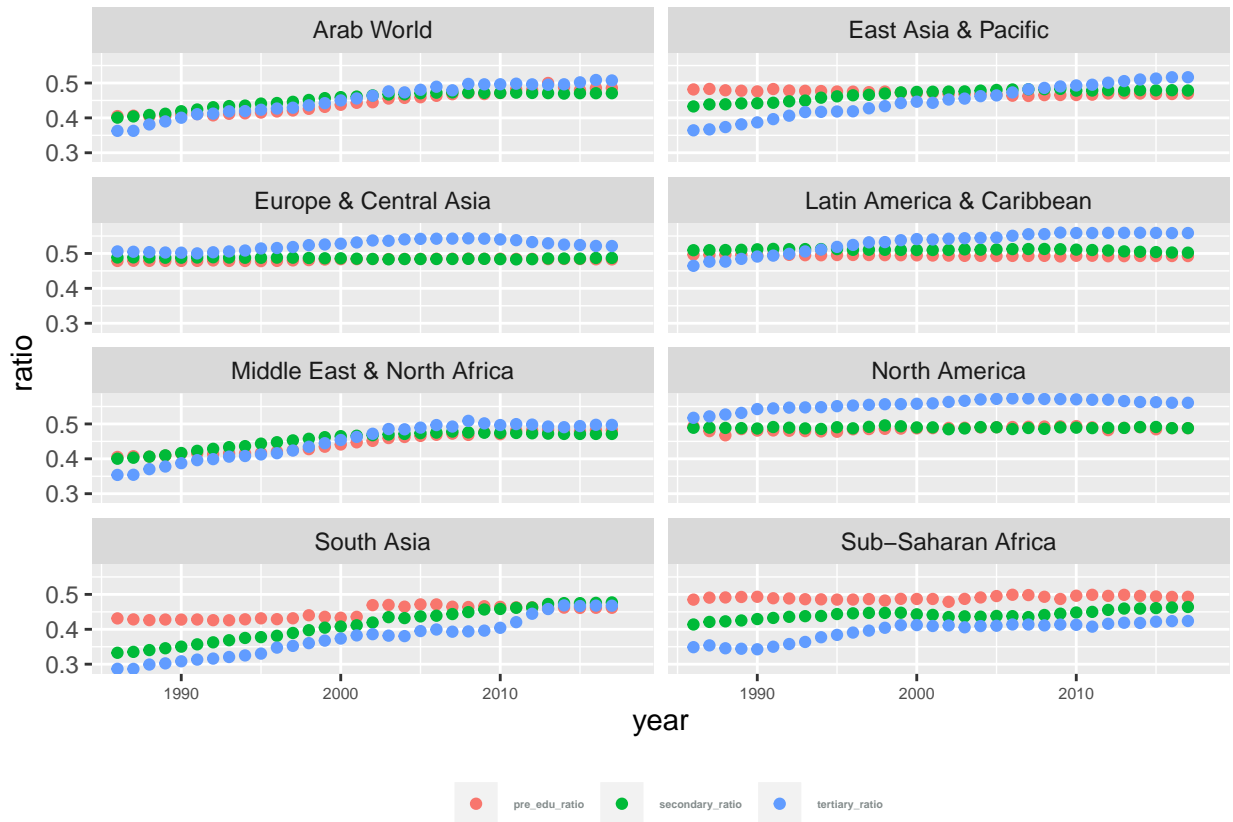

4.2 Overview of female ratio in pre-primary, secondary and tertiary education in each main area across the world between 1986 and 2018.

As is shown below, female enrollment rate in tertiary education is outstanding in most areas except Sub-Saharan Africa.

```
a<-colnames(data5)
aa<-a[14:16]
data5rev<-gather(data5,edulevel,ratio,aa)

ggplot(data=data5rev)+geom_point(mapping=aes(x=year,y=ratio,color=edulevel))+guides(color=guide_legend(
```
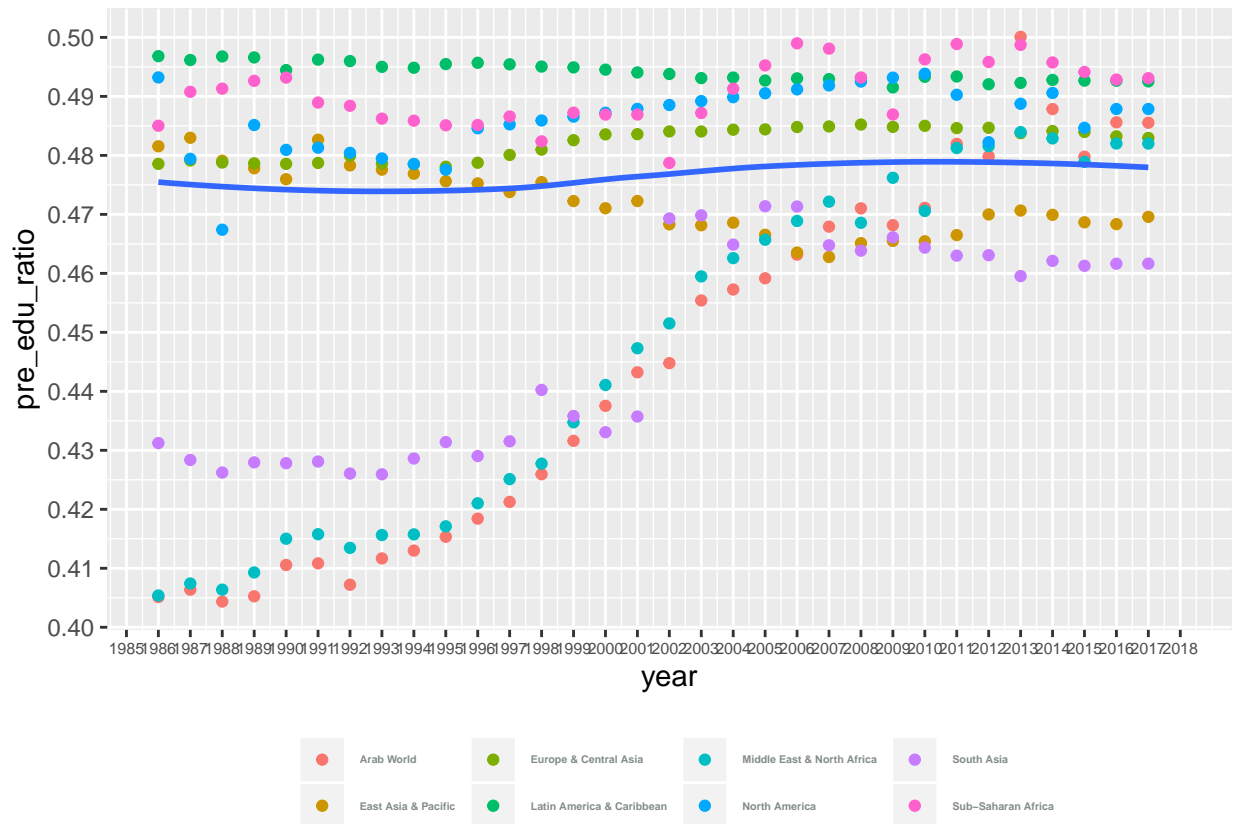
4.3 Female enrollment rates in different education level year by year.

4.31 Female enrollment rates in pre-primary education It shows that female enrollment rates of pre-primary education in four areas (Europe & central Asia, Latin America & Caribbean, Sub-Saharan Africa and North America) are above the world's average level.
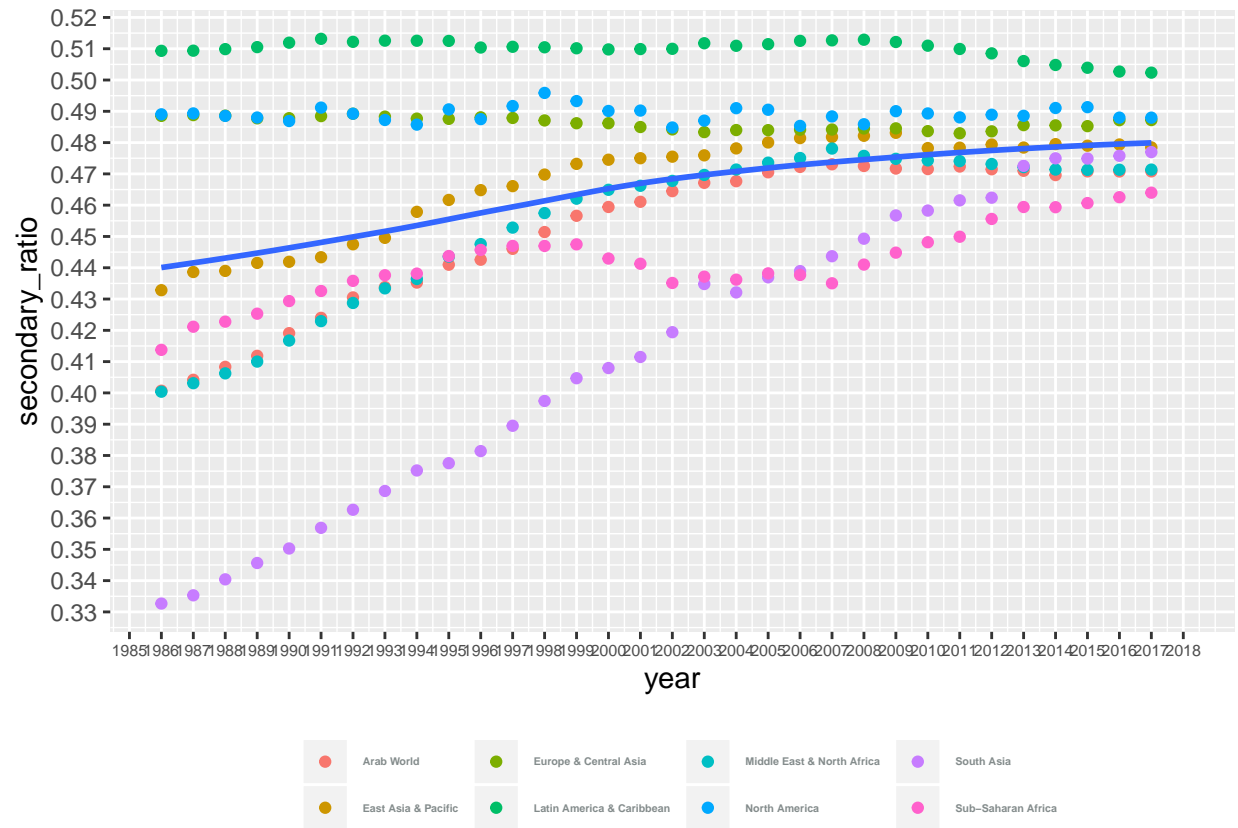
```
ggplot(data=data5)+geom_point(mapping=aes(x=year,y=pre_edu_ratio,color=data5$country_name))+scale_y_con
scale_x_continuous(breaks=seq(1985, 2018, 1))+guides(color=guide_legend(title=NULL))+theme(legend.text =
```

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

4.32 Female enrollment rates in secondary education It shows that female enrollment rates of secondary education in four areas (Europe & central Asia, Latin America & Caribbean, East Asia & Pacific and North America) are above the world's average level.
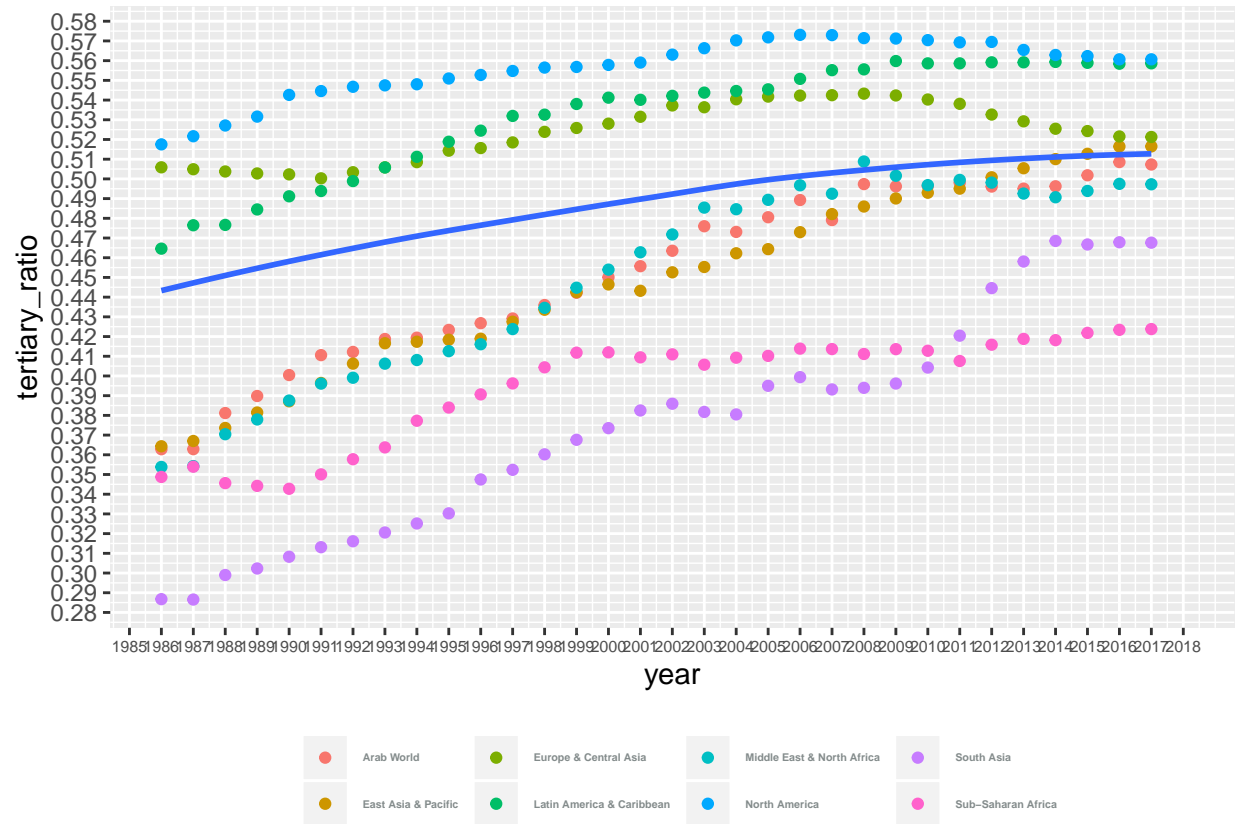
```
ggplot(data=data5)+geom_point(mapping=aes(x=year,y=secondary_ratio,color=data5$country_name))+scale_y_c
scale_x_continuous(breaks=seq(1985, 2018, 1))+guides(color=guide_legend(title=NULL))+theme(legend.text
```

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

4.33 Female enrollment rate in tertiary edutation It shows that female enrollment rates of tertiary education in three areas (Europe & central Asia, Latin America & Caribbean and North America) are above the world's average level.

```
ggplot(data=data5)+geom_point(mapping=aes(x=year,y=tertiary_ratio,color=data5$country_name))+scale_y_cor
scale_x_continuous(breaks=seq(1985, 2018, 1))+guides(color=guide_legend(title=NULL))+theme(legend.text =
```

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

5. Conclusion

1) The female enrollment rates of three education level in most areas are around 40-50% and have an upward trend as time goes by.

2) The female enrollment rates of three education level in Europe & Central Asia, Latin America & Caribbean and North America always outstrip the world's average from 1986 to 2018.