

Human Development Indicators Analysis

Vaishaali Kondapalli - 23200337

HDIA

For the Human Development Indicators Analysis task the two countries whose data I have chosen to work with are **Malaysia** and **Netherlands**.

Data Preparation

Here we're just reading our datasets, merging them using `rbind` so that data analysis becomes easier further.

Data Exploration

Let's start off with doing some quick data exploration to understand our data.

Summary for Malaysia:

country_code	country_name	indicator_id	indicator_name
Length:862	Length:862	Length:862	Length:862
Class :character	Class :character	Class :character	Class :character
Mode :character	Mode :character	Mode :character	Mode :character

index_id	index_name	value	year
Length:862	Length:862	Min. : -11.00	Min. :1990
Class :character	Class :character	1st Qu.: 10.35	1st Qu.:1999
Mode :character	Mode :character	Median : 17.22	Median :2008
		Mean : 2132.75	Mean :2007
		3rd Qu.: 74.62	3rd Qu.:2015
		Max. :34983.47	Max. :2022

Structure for Malaysia:

Classes 'data.table' and 'data.frame': 862 obs. of 8 variables:

```
$ country_code : chr "MYS" "MYS" "MYS" "MYS" ...
$ country_name : chr "Malaysia" "Malaysia" "Malaysia" "Malaysia" ...
$ indicator_id : chr "abr" "abr" "abr" "abr" ...
$ indicator_name: chr "Adolescent Birth Rate (births per 1,000 women ages 15-19)" "Adolescent Birth Rate (births per 1,000 women ages 15-19)" "Adolescent Birth Rate (births per 1,000 women ages 15-19)" "Adolescent Birth Rate (births per 1,000 women ages 15-19)" ...
$ index_id      : chr "GII" "GII" "GII" "GII" ...
$ index_name    : chr "Gender Inequality Index" "Gender Inequality Index" "Gender Inequality Index" "Gender Inequality Index" ...
$ value        : num 19.4 19.3 19.1 18.4 18.1 ...
```

```
$ year      : int  1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 ...
- attr(*, ".internal.selfref")=<externalptr>
```

First few rows for Malaysia:

	country_code	country_name	indicator_id		indicator_name	index_id
1:	MYS	Malaysia	abr			
2:	MYS	Malaysia	abr			
3:	MYS	Malaysia	abr			
4:	MYS	Malaysia	abr			
5:	MYS	Malaysia	abr			
6:	MYS	Malaysia	abr			
1:	Adolescent Birth Rate (births per 1,000 women ages 15-19)				GII	
2:	Adolescent Birth Rate (births per 1,000 women ages 15-19)				GII	
3:	Adolescent Birth Rate (births per 1,000 women ages 15-19)				GII	
4:	Adolescent Birth Rate (births per 1,000 women ages 15-19)				GII	
5:	Adolescent Birth Rate (births per 1,000 women ages 15-19)				GII	
6:	Adolescent Birth Rate (births per 1,000 women ages 15-19)				GII	
	index_name	value	year			
1:	Gender Inequality Index	19.418	1990			
2:	Gender Inequality Index	19.329	1991			
3:	Gender Inequality Index	19.064	1992			
4:	Gender Inequality Index	18.422	1993			
5:	Gender Inequality Index	18.055	1994			
6:	Gender Inequality Index	16.597	1995			

Now lets do it for netherlands:

Summary for Netherlands:

country_code	country_name	indicator_id	indicator_name
Length:894	Length:894	Length:894	Length:894
Class :character	Class :character	Class :character	Class :character
Mode :character	Mode :character	Mode :character	Mode :character

index_id	index_name	value	year
Length:894	Length:894	Min. : -12.00	Min. :1990

Class :character	Class :character	1st Qu.: 11.34	1st Qu.:1998
Mode :character	Mode :character	Median : 18.34	Median :2007
		Mean : 5330.65	Mean :2007
		3rd Qu.: 79.37	3rd Qu.:2015
		Max. :71504.37	Max. :2022

Structure for Netherlands:

Classes 'data.table' and 'data.frame': 894 obs. of 8 variables:

```
$ country_code : chr "NLD" "NLD" "NLD" "NLD" ...
$ country_name : chr "Netherlands" "Netherlands" "Netherlands" "Netherlands" ...
$ indicator_id : chr "abr" "abr" "abr" "abr" ...
$ indicator_name: chr "Adolescent Birth Rate (births per 1,000 women ages 15-19)" "Adolescent Birth Rate
(births per 1,000 women ages 15-19)" "Adolescent Birth Rate (births per 1,000 women ages 15-19)" "Adolescent
Birth Rate (births per 1,000 women ages 15-19)" ...
$ index_id : chr "GII" "GII" "GII" "GII" ...
$ index_name : chr "Gender Inequality Index" "Gender Inequality Index" "Gender Inequality Index" "Gender
Inequality Index" ...
$ value : num 8.73 8.68 8.02 7.42 6.86 ...
$ year : int 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 ...
- attr(*, ".internal.selfref")=<externalptr>
```

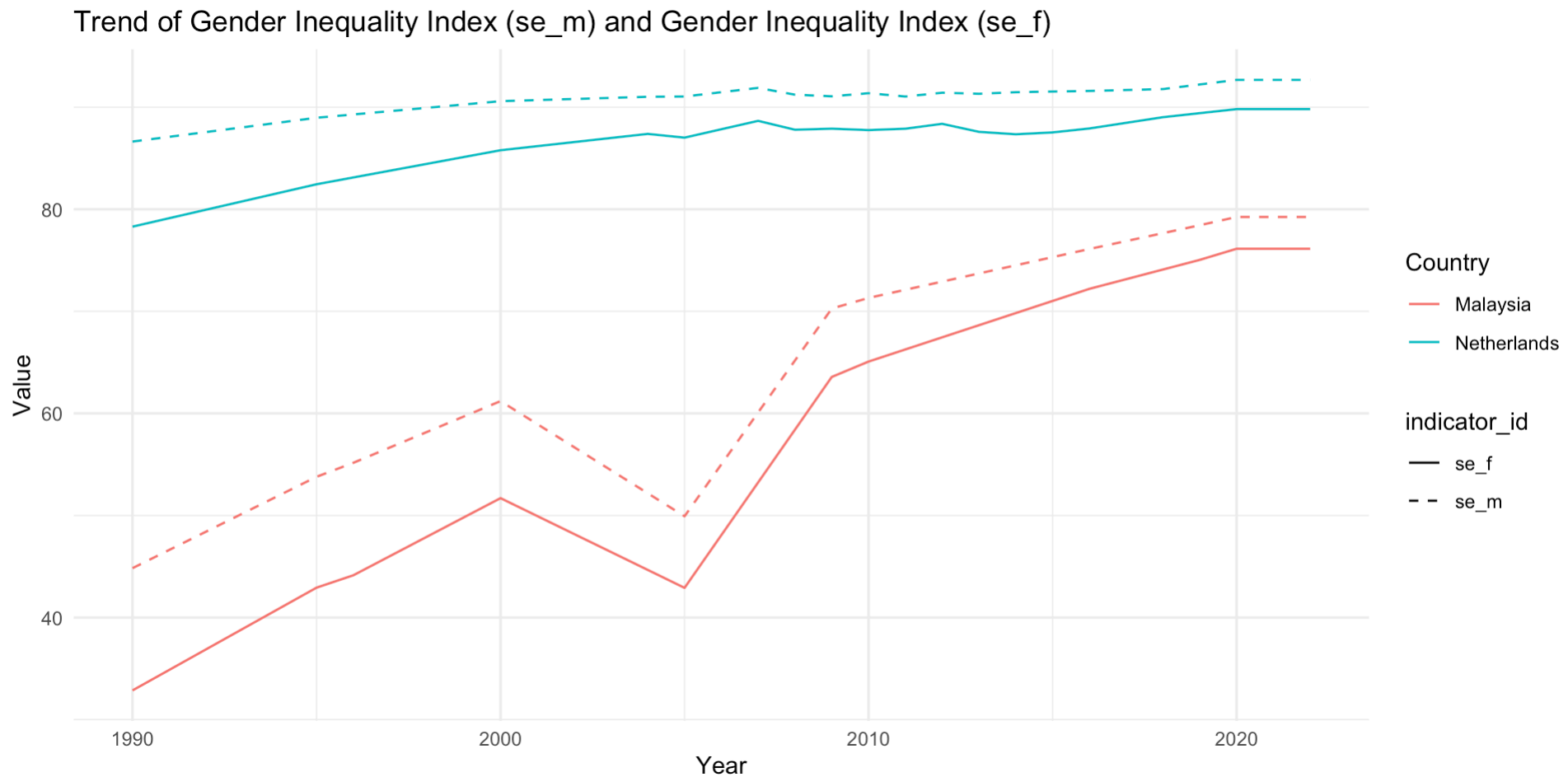
First few rows for Netherlands:

	country_code	country_name	indicator_id		indicator_name	index_id
1:	NLD	Netherlands	abr			
2:	NLD	Netherlands	abr			
3:	NLD	Netherlands	abr			
4:	NLD	Netherlands	abr			
5:	NLD	Netherlands	abr			
6:	NLD	Netherlands	abr			
1:				Adolescent Birth Rate (births per 1,000 women ages 15-19)	GII	
2:				Adolescent Birth Rate (births per 1,000 women ages 15-19)	GII	
3:				Adolescent Birth Rate (births per 1,000 women ages 15-19)	GII	
4:				Adolescent Birth Rate (births per 1,000 women ages 15-19)	GII	
5:				Adolescent Birth Rate (births per 1,000 women ages 15-19)	GII	
6:				Adolescent Birth Rate (births per 1,000 women ages 15-19)	GII	
				index_name	value	year
1:				Gender Inequality Index	8.729	1990
2:				Gender Inequality Index	8.679	1991

3:	Gender Inequality Index	8.023	1992
4:	Gender Inequality Index	7.420	1993
5:	Gender Inequality Index	6.861	1994
6:	Gender Inequality Index	5.884	1995

Data Analysis and Visualization

PLOT 1

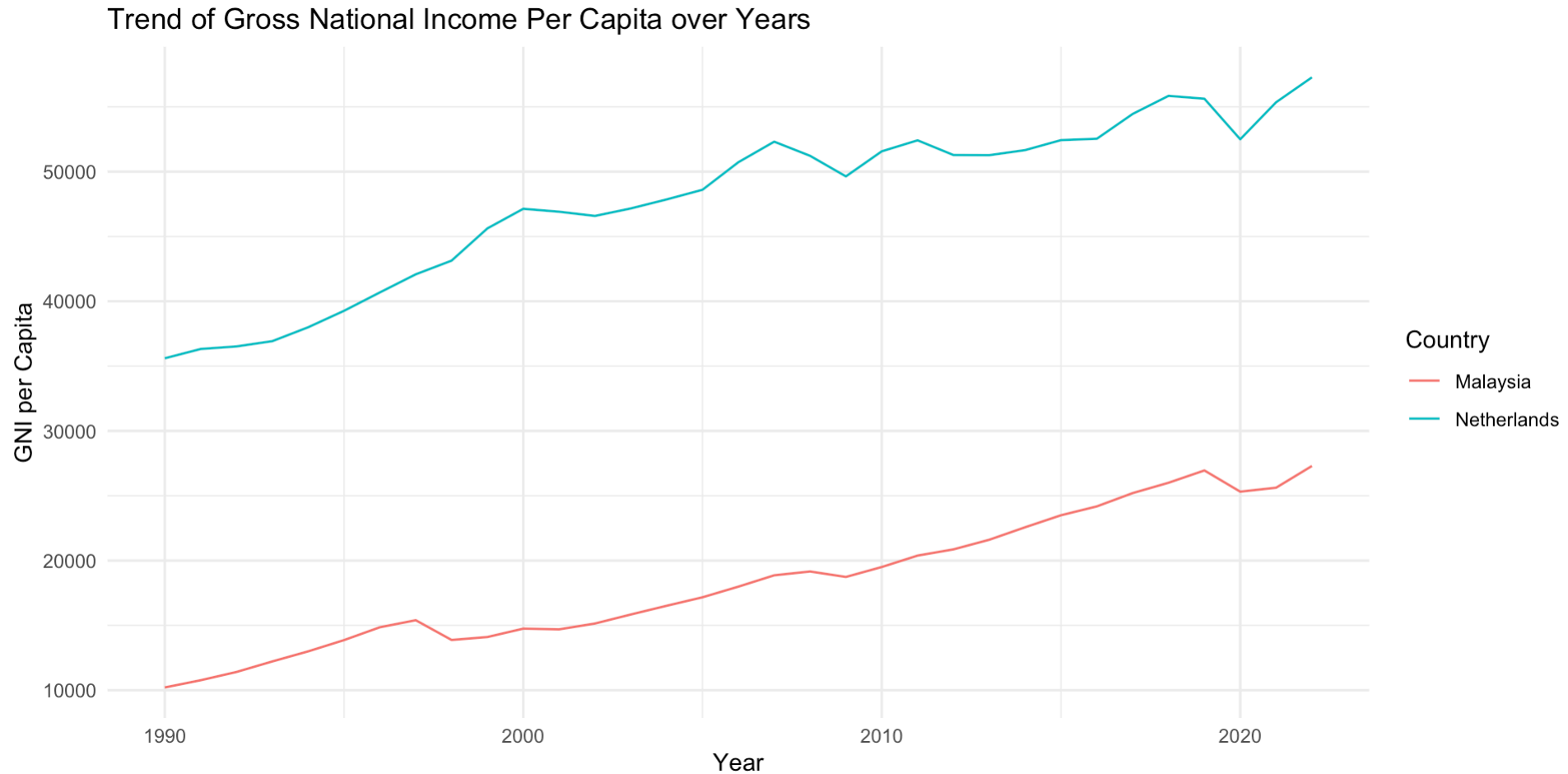


Here, the first observation we can make immediately is that the values present for Gender Inequality Index for se_m i.e, Population with at least some secondary education, male (% ages 25 and older) have always been higher than se_f i.e, Population with at least some secondary education, female (% ages 25 and older) at both countries Malaysia and Netherlands.

Also, the values for se_m and se_f for Netherlands were approximately (X2) compared to Malaysia {40 ish values at Malaysia but around 80 at Netherlands} in 1990 but by 2022, the values for Netherlands just went up 5 points, the values for Malaysia went up by soo much more. The trendlines portray how the values shot up from 40 to almost 80 for Malaysia by 2022.

So, even though Population with at least some secondary education(% ages 25 and older) did increase in both countries, its notable that a gap still exists amost the two genders in both countries.

PLOT 2



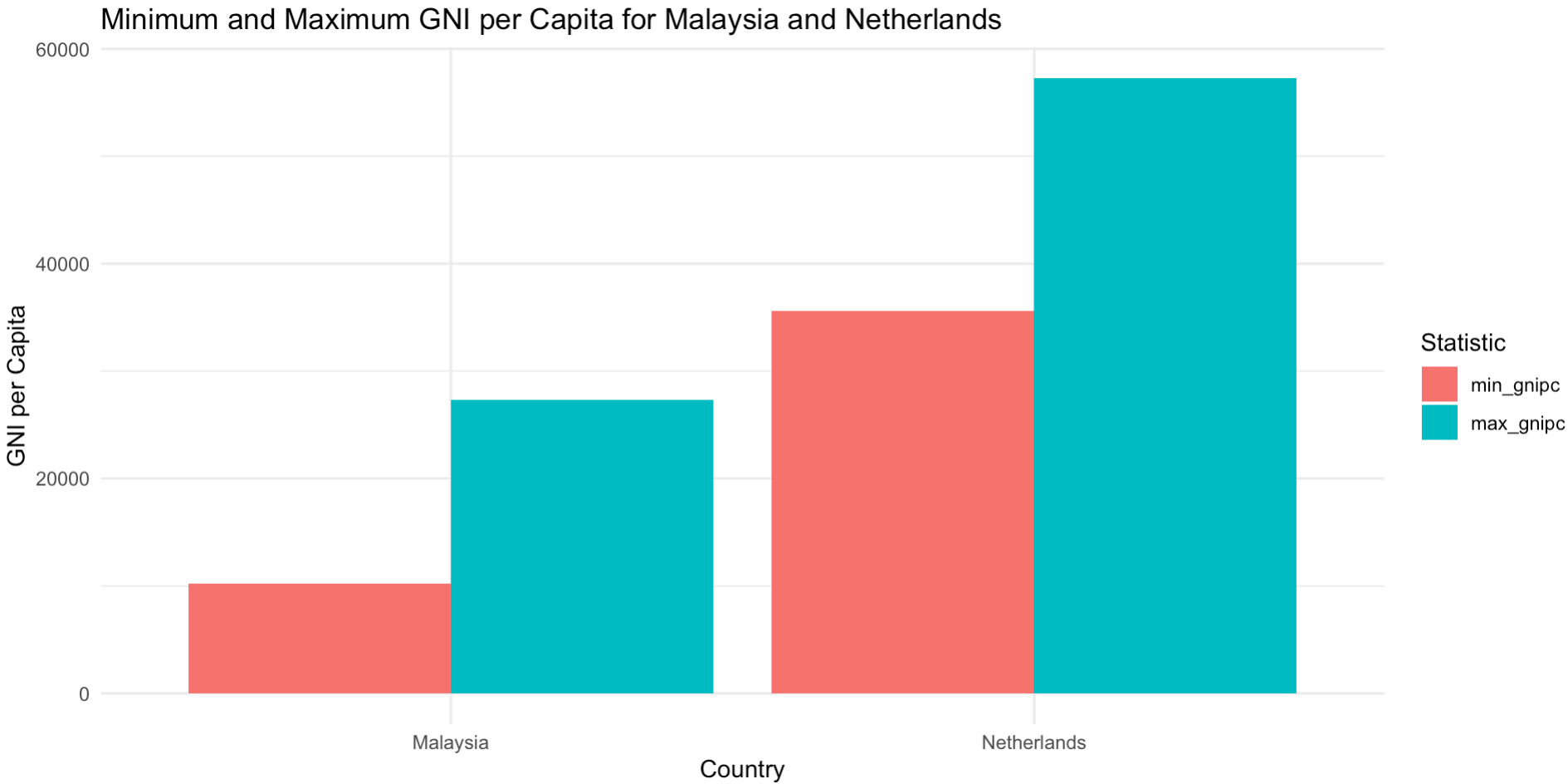
The observation we can make from this plot is that both the countries have a positive trendline, which is a good sign but Malaysia has always been behind Netherlands though.

Both of the countries do have a dip in their gnipc i.e, Gross National Income Per Capita (2017 PPP\$) in the year 2020 which could have taken place because of the pandemic but eventually both of them did recover from that dip by 2022.

PLOT 3

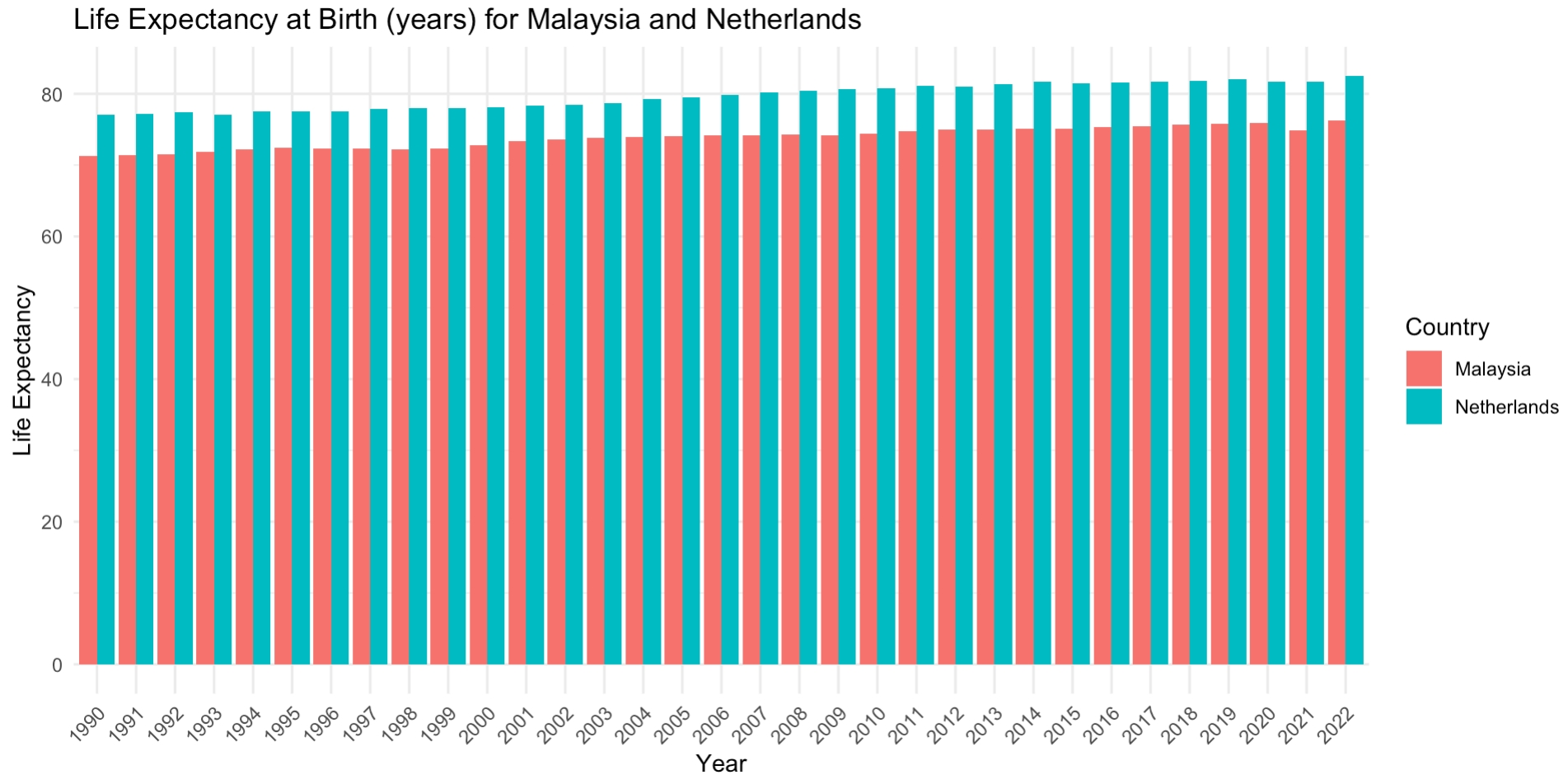
Minimum and Maximum GNI per Capita by Country:

	country_name	min_gnipc	max_gnipc
1:	Malaysia	10209.35	27295.41
2:	Netherlands	35604.22	57278.31



Even the minimum gnipc of Netherlands is more than the maximum gnipc of Malaysia :(

PLOT 4

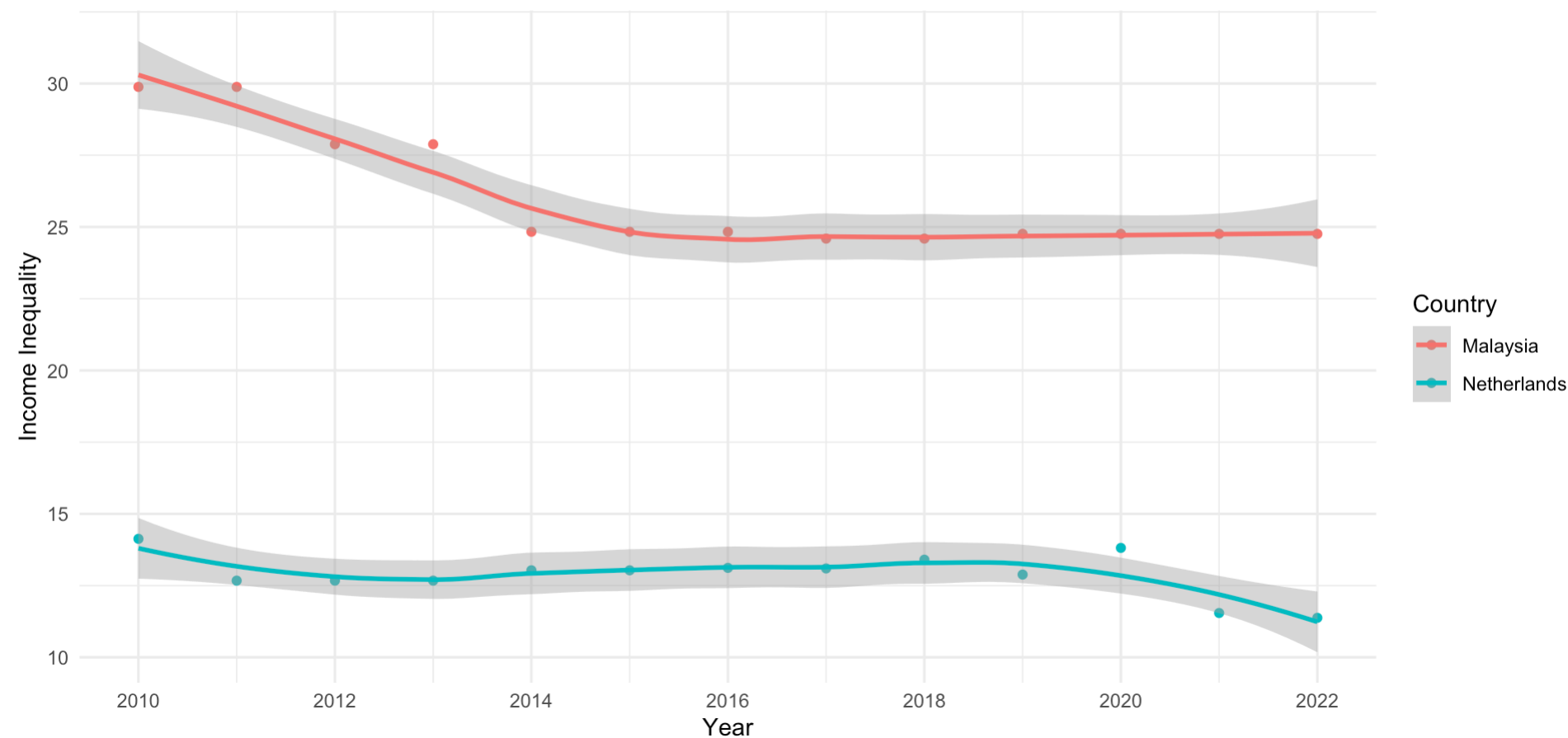


As we can see from above plot, 'le' i.e., the Life Expectancy at Birth (years) for both the countries are well above 70 and 'le' has always been higher at Netherlands as compared to Malaysia.

PLOT 5

Now for inequality in income, data has been provided only from year 2010, so we'll plot from 2010 to 2022.

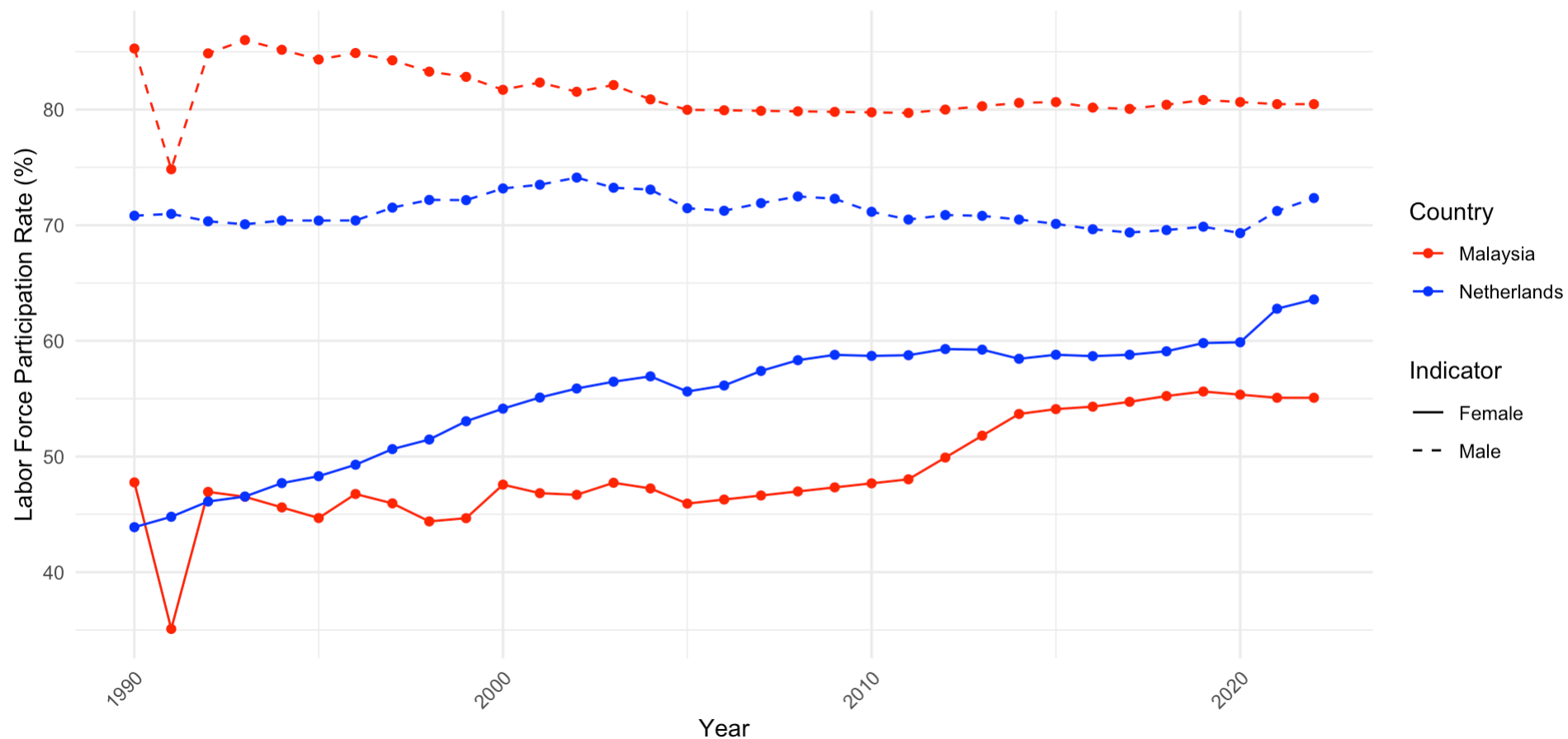
Trend of Income Inequality over Years



Now, in the plot shown above, it portrays how inequality in income has reduced in both countries and its clear than from 2010 to 2022, the income inequality value has always been been higher in Malaysia as compared to Netherlands.

PLOT 6

Year-wise Labor Force Participation Rate (Male and Female) in Malaysia and Netherlands



From the above plot we now know that value of lfpr_m i.e, Labour force participation rate, male (% ages 15 and older) has always been higher in Malaysia as compared to Netherlands.

As for lfpr_f i.e, Labour force participation rate, female (% ages 15 and older), we can see clearly that the value has increased significantly by 2022 as compared to 1990 for both countries.

Fin

