
Istanbul Technical University

Faculty of Science and Letters

Relationship between Crime Rates and Income
Levels

Final Report

Melisa Kütük
Çağrı Gökpınar
Team Name: ananas

Mathematical Engineering
MAT381E
Asst. Prof. Gül İnan
January 2022

Project Description

- To define whether crimes rates are related to the income levels or not, generating ideas for avert these crimes if it is.

Project goal & social problem addressed

- Show that there is a relationship between the income levels and crime rates. The social problem in this project is the low income distribution and the high crime rate. As a result of this project, it is important to prevent crimes and increase income level.

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --
## v ggplot2 3.3.5      v purrr  0.3.4
## v tibble  3.1.6      v dplyr  1.0.7
## v tidyr   1.1.4      v stringr 1.4.0
## v readr   2.0.2      v forcats 0.5.1

## Warning: package 'tibble' was built under R version 4.1.2

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

options(warn=-1)
```

Project data

*Crime:

1. Homicide rates
2. Sexual abuse and rape rates

*Economy:

1. GDP rates
2. Employment
3. Gini index

Variables in crime: Year, gender, cause, countries, location etc.

Variables in economy: Work hours, countries, GDP, Inequality Index etc.

Dataset Analyzing

First of all, data on the subject were collected. Then, unnecessary data from these data were eliminated. Detailed information about dataset analyzing was given in the interim report.

- The datasets about crime rates are:
- Murders2005
- Murders2006
- Murders2007
- Murders2008
- Murders2009
- Murders2010
- Rapes2005
- Rapes2006
- Rapes2007

- Rapes2008
- Rapes2009
- Rapes2010

Importing datasets about income levels and renaming them

```
employment <- read.csv("datasets/ANHRS_31102021130035527.csv")
gpd <- read.csv("datasets/API_NY.GDP.MKTP.CD_DS2_en_csv_v2_3158872.csv")
gini <- read.csv("datasets/API_SI.POV.GINI_DS2_en_csv_v2_3158890.csv")
homicide <- read.csv("datasets/API_VC.IHR.PSRC.P5_DS2_en_csv_v2_3161002.csv")
sexualviolence <- read.csv("datasets/cts_sexual_violence_2012.csv")
```

Dataframes

```
head(employment,5)
```

```
##      i..COUNTRY   Country EMPSTAT Employment.status FREQUENCY Frequency TIME Time
## 1      AUS Australia      TE Total employment      A      Annual 2010 2010
## 2      AUS Australia      TE Total employment      A      Annual 2011 2011
## 3      AUS Australia      TE Total employment      A      Annual 2012 2012
## 4      AUS Australia      TE Total employment      A      Annual 2013 2013
## 5      AUS Australia      TE Total employment      A      Annual 2014 2014
##      Unit.Code   Unit PowerCode.Code PowerCode Reference.Period.Code
## 1      HOUR Hours           0      Units      NA
## 2      HOUR Hours           0      Units      NA
## 3      HOUR Hours           0      Units      NA
## 4      HOUR Hours           0      Units      NA
## 5      HOUR Hours           0      Units      NA
##      Reference.Period Value Flag.Codes Flags
## 1      NA      1778      NA      NA
## 2      NA      1774      NA      NA
## 3      NA      1771      NA      NA
## 4      NA      1766      NA      NA
## 5      NA      1755      NA      NA
```

```
head(gpd,5)
```

```
##      i..Data.Source World.Development.Indicators      Column1
## 1
## 2 Last Updated Date      2021-10-28
## 3
## 4      Country Name      Country Code      Indicator Name
## 5      Aruba      ABW GDP (current US$)
##      X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8 X_9 X_10 X_11 X_12 X_13
## 1      NA NA NA NA NA NA NA NA NA NA NA NA NA
## 2      NA NA NA NA NA NA NA NA NA NA NA NA NA
## 3      NA NA NA NA NA NA NA NA NA NA NA NA NA
## 4 Indicator Code 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971
## 5 NY.GDP.MKTP.CD NA NA NA NA NA NA NA NA NA NA NA NA NA
##      X_14 X_15 X_16 X_17 X_18 X_19 X_20 X_21 X_22 X_23 X_24 X_25 X_26 X_27
## 1      NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 2      NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 3      NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 4 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985
```

```
## 5 NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## X_28 X_29 X_30 X_31 X_32 X_33 X_34
## 1 NA NA NA NA NA NA NA
## 2 NA NA NA NA NA NA NA
## 3 NA NA NA NA NA NA NA
## 4 1986 1987 1988 1989 1990 1991 1992
## 5 405463417 487602458 596423607 695304363 764887117 872138715 958463184
## X_35 X_36 X_37 X_38 X_39 X_40 X_41
## 1 NA NA NA NA NA NA NA
## 2 NA NA NA NA NA NA NA
## 3 NA NA NA NA NA NA NA
## 4 1993 1994 1995 1996 1997 1998 1999
## 5 1082979721 1245688268 1320474860 1379960894 1531944134 1665100559 1722798883
## X_42 X_43 X_44 X_45 X_46 X_47 X_48
## 1 NA NA NA NA NA NA NA
## 2 NA NA NA NA NA NA NA
## 3 NA NA NA NA NA NA NA
## 4 2000 2001 2002 2003 2004 2005 2006
## 5 1873452514 1920111732 1941340782 2021229050 2228491620 2330726257 2424581006
## X_49 X_50 X_51 X_52 X_53 X_54 X_55
## 1 NA NA NA NA NA NA NA
## 2 NA NA NA NA NA NA NA
## 3 NA NA NA NA NA NA NA
## 4 2007 2008 2009 2010 2011 2012 2013
## 5 2615083799 2745251397 2498882682 2390502793 2549720670 2534636872 2727849721
## X_56 X_57 X_58 X_59 X_60 X_61 X_62 X_63
## 1 NA NA NA NA NA NA NA
## 2 NA NA NA NA NA NA NA
## 3 NA NA NA NA NA NA NA
## 4 2014 2015 2016 2017 2018 2019 2020 NA
## 5 2790849162 2962905028 2983636872 3092430168 3202188607 NA NA NA
```

```
head(gini,5)
```

```
## i..Data.Source World.Development.Indicators
## 1
## 2 Last Updated Date 2021-10-28
## 3
## 4 Country Name Country Code
## 5 Aruba ABW
## Column1 X_1 X_2 X_3 X_4 X_5 X_6 X_7
## 1 NA NA NA NA NA NA NA
## 2 NA NA NA NA NA NA NA
## 3 NA NA NA NA NA NA NA
## 4 Indicator Name Indicator Code 1960 1961 1962 1963 1964 1965
## 5 Gini index (World Bank estimate) SI.POV.GINI NA NA NA NA NA NA
## X_8 X_9 X_10 X_11 X_12 X_13 X_14 X_15 X_16 X_17 X_18 X_19 X_20 X_21 X_22
## 1 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 2 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 3 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 4 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980
## 5 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## X_23 X_24 X_25 X_26 X_27 X_28 X_29 X_30 X_31 X_32 X_33 X_34 X_35 X_36 X_37
## 1 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 2 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
```

```
## 3 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 4 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995
## 5 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## X_38 X_39 X_40 X_41 X_42 X_43 X_44 X_45 X_46 X_47 X_48 X_49 X_50 X_51 X_52
## 1 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 2 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 3 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 4 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## X_53 X_54 X_55 X_56 X_57 X_58 X_59 X_60 X_61 X_62 X_63
## 1 NA NA NA NA NA NA NA NA NA NA NA NA NA
## 2 NA NA NA NA NA NA NA NA NA NA NA NA NA
## 3 NA NA NA NA NA NA NA NA NA NA NA NA NA
## 4 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 NA
## 5 NA NA NA NA NA NA NA NA NA NA NA NA NA
```

```
head(homicide,5)
```

```
## i..Data.Source World.Development.Indicators
## 1
## 2 Last Updated Date 2021-10-28
## 3
## 4 Country Name Country Code
## 5 Aruba ABW
## Column1 X_1 X_2 X_3 X_4 X_5
## 1 NA NA NA NA NA
## 2 NA NA NA NA NA
## 3 NA NA NA NA NA
## 4 Indicator Name Indicator Code 1960 1961 1962 1963
## 5 Intentional homicides (per 100,000 people) VC.IHR.PSRC.P5 NA NA NA NA
## X_6 X_7 X_8 X_9 X_10 X_11 X_12 X_13 X_14 X_15 X_16 X_17 X_18 X_19 X_20
## 1 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 2 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 3 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 4 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978
## 5 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## X_21 X_22 X_23 X_24 X_25 X_26 X_27 X_28 X_29 X_30 X_31 X_32 X_33 X_34 X_35
## 1 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 2 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 3 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## 4 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993
## 5 NA NA NA NA NA NA NA NA NA NA NA 0 NA NA NA
## X_36 X_37 X_38 X_39 X_40 X_41 X_42 X_43 X_44
## 1 NA NA NA NA NA NA NA NA NA
## 2 NA NA NA NA NA NA NA NA NA
## 3 NA NA NA NA NA NA NA NA NA
## 4 1994 1995.000000 1996 1997 1998 1999.000000 2000 2001.000000 2002.000000
## 5 NA 7.469748 NA NA NA 2.246964 NA 4.306076 5.263601
## X_45 X_46 X_47 X_48 X_49 X_50
## 1 NA NA NA NA NA NA
## 2 NA NA NA NA NA NA
## 3 NA NA NA NA NA NA
## 4 2003.000000 2004.000000 2005.000000 2006.000000 2007.000000 2008.000000
## 5 4.123031 2.02544 5.99832 4.958842 2.963665 4.932815
## X_51 X_52 X_53 X_54 X_55 X_56 X_57
```

```
## 1      NA      NA      NA      NA      NA      NA      NA
## 2      NA      NA      NA      NA      NA      NA      NA
## 3      NA      NA      NA      NA      NA      NA      NA
## 4 2009.000000 2010.000000 2011.000000 2012.000000 2013.000000 2014.000000 2015
## 5    3.942751    3.934491    1.959824    3.899966    5.815926    1.927228    NA
##   X_58 X_59 X_60 X_61 X_62 X_63
## 1   NA   NA   NA   NA   NA   NA
## 2   NA   NA   NA   NA   NA   NA
## 3   NA   NA   NA   NA   NA   NA
## 4 2016 2017 2018 2019 2020   NA
## 5   NA   NA   NA   NA   NA   NA
```

```
head(sexualviolence,5)
```

```
##   country.territory  date sexual.violence      rate
## 1   #country+name #date      #affected  #meta+rate
## 2           Albania 2003             <NA>      <NA>
## 3           Albania 2004             <NA>      <NA>
## 4           Albania 2005              89 2.832771023
## 5           Albania 2006              87 2.756123901
```

- GDP, Gini and Homicide's headers needs to be replaced with 4th row. Sexual violence's first row needs to be deleted.

```
names(gpd) <- as.matrix(gpd[4, ])
gpd <- gpd[-4, ]
gpd[] <- lapply(gpd, function(x) type.convert(as.character(x)))
gpd <- gpd[-1, ]
gpd <- gpd[-1, ]
gpd <- gpd[-1, ]
row.names(gpd) <- NULL
head(gpd)
```

```
##           Country Name Country Code   Indicator Name Indicator Code
## 1                Aruba          ABW GDP (current US$) NY.GDP.MKTP.CD
## 2 Africa Eastern and Southern          AFE GDP (current US$) NY.GDP.MKTP.CD
## 3                Afghanistan          AFG GDP (current US$) NY.GDP.MKTP.CD
## 4 Africa Western and Central          AFW GDP (current US$) NY.GDP.MKTP.CD
## 5                 Angola          AGO GDP (current US$) NY.GDP.MKTP.CD
## 6                Albania          ALB GDP (current US$) NY.GDP.MKTP.CD
##           1960      1961      1962      1963      1964      1965
## 1           NA      NA      NA      NA      NA      NA
## 2 19342484576 19753490586 21526615650 25772356399 23563232195 26851350246
## 3   537777811   548888896   546666678   751111191   800000044   1006666638
## 4 10407321640 11131302981 11946843969 12680220415 13842621612 14866816737
## 5           NA      NA      NA      NA      NA      NA
## 6           NA      NA      NA      NA      NA      NA
##           1966      1967      1968      1969      1970      1971
## 1           NA      NA      NA      NA      NA      NA
## 2 29196502382 30219070807 32927067005 37801761961 40377109505 44544318707
## 3 1399999967 1673333418 1373333367 1408888922 1748886596 1831108971
## 4 15837474343 14430648807 14884699923 16887028428 23511477700 20838908163
## 5           NA      NA      NA      NA      NA      NA
## 6           NA      NA      NA      NA      NA      NA
##           1972      1973      1974      1975      1976      1977
## 1           NA      NA      NA      NA      NA      NA
```

## 2	48374959174	63079306619	78369918525	83562484550	83337002757	95133441245
## 3	1595555476	1733333264	2155555498	2366666616	2555555567	2953333418
## 4	25272340678	31282962686	44227412162	51459772973	62147555474	65334104528
## 5	NA	NA	NA	NA	NA	NA
## 6	NA	NA	NA	NA	NA	NA
##	1978	1979	1980	1981	1982	1983
## 1	NA	NA	NA	NA	NA	NA
## 2	106508000000	124688000000	156751000000	160622000000	154905000000	160001000000
## 3	3300000109	3697940410	3641723322	3478787909	NA	NA
## 4	71220525033	88654314398	112064000000	211065000000	187218000000	138156000000
## 5	NA	NA	5934073604	5553824464	5553824464	5787823809
## 6	NA	NA	NA	NA	NA	NA
##	1984	1985	1986	1987	1988	1989
## 1	NA	NA	405463417	487602458	596423607	695304363
## 2	146244000000	130638000000	147249000000	180013000000	189291000000	194839000000
## 3	NA	NA	NA	NA	NA	NA
## 4	114296000000	116541000000	107529000000	110354000000	108975000000	101799000000
## 5	6135166254	7558613008	7076793823	8089279285	8775116269	10207922517
## 6	1857338012	1897050133	2097326250	2080796250	2051236250	2253090000
##	1990	1991	1992	1993	1994	1995
## 1	764887117	872138715	958463184	1082979721	1245688268	1320474860
## 2	212659000000	221100000000	220554000000	220950000000	225100000000	253136000000
## 3	NA	NA	NA	NA	NA	NA
## 4	121838000000	117491000000	118317000000	97186773684	85693055814	107403000000
## 5	11236275843	NA	NA	NA	3390500000	5561222222
## 6	2028553750	1099559028	652174991	1185315468	1880951520	2392764853
##	1996	1997	1998	1999	2000	2001
## 1	1379960894	1531944134	1665100559	1722798883	1873452514	1920111732
## 2	252550000000	265549000000	250378000000	247067000000	268315000000	242105000000
## 3	NA	NA	NA	NA	NA	NA
## 4	119044000000	119983000000	122621000000	130199000000	134150000000	141863000000
## 5	7526963964	7649716157	6506619145	6152936539	9129634978	8936063723
## 6	3199642580	2258515610	2545967253	3212119044	3480355189	3922099471
##	2002	2003	2004	2005	2006	2007
## 1	1941340782	2021229050	2228491620	2330726257	2424581006	2615083799
## 2	247657000000	326744000000	405860000000	471743000000	533533000000	613164000000
## 3	4055179566	4515558808	5226778809	6209137625	6971285595	9747879532
## 4	170532000000	197384000000	245856000000	302111000000	384336000000	451866000000
## 5	15285594828	17812705294	23552052408	36970918699	52381006892	65266452081
## 6	4348070165	5611492283	7184681399	8052075642	8896073938	10677321490
##	2008	2009	2010	2011	2012	2013
## 1	2745251397	2498882682	2390502793	2549720670	2534636872	2727849721
## 2	668037000000	670986000000	805795000000	898605000000	915590000000	930086000000
## 3	10109225814	12439087077	15856574731	17804292964	20001598506	20561069558
## 4	553031000000	492546000000	580217000000	658428000000	716935000000	807819000000
## 5	88538611205	70307163678	83799496611	111790000000	128053000000	136710000000
## 6	12881354104	12044223353	11926928506	12890765324	12319830252	12776217195
##	2014	2015	2016	2017	2018	2019
## 1	2790849162	2962905028	2983636872	3092430168	3202188607	NA
## 2	958825000000	895440000000	856992000000	964791000000	986611000000	980372000000
## 3	20484885120	19907111419	18017749074	18869945678	18353881130	19291104008
## 4	846943000000	757492000000	687485000000	680989000000	738131000000	792079000000
## 5	145712000000	116194000000	101124000000	122124000000	101353000000	89417190341
## 6	13228144008	11386846319	11861200797	13019693451	15147020535	15286612573

```
##          2020 NA
## 1          NA NA
## 2 900829000000 NA
## 3 19807067268 NA
## 4 786585000000 NA
## 5 62306913444 NA
## 6 14799615097 NA

names(gini) <- as.matrix(gini[4, ])
gini <- gini[-4, ]
gini[] <- lapply(gini, function(x) type.convert(as.character(x)))
gini <- gini[-1, ]
gini <- gini[-1, ]
gini <- gini[-1, ]
row.names(gini) <- NULL
head(gini)
```

##		Country Name	Country Code	Indicator Name											
## 1		Aruba	ABW	Gini index (World Bank estimate)											
## 2		Africa Eastern and Southern	AFE	Gini index (World Bank estimate)											
## 3		Afghanistan	AFG	Gini index (World Bank estimate)											
## 4		Africa Western and Central	AFW	Gini index (World Bank estimate)											
## 5		Angola	AGO	Gini index (World Bank estimate)											
## 6		Albania	ALB	Gini index (World Bank estimate)											
##	Indicator Code	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971		
## 1	SI.POV.GINI	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA		
## 2	SI.POV.GINI	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA		
## 3	SI.POV.GINI	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA		
## 4	SI.POV.GINI	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA		
## 5	SI.POV.GINI	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA		
## 6	SI.POV.GINI	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA		
##	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986
## 1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
## 2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
## 3	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
## 4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
## 5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
## 6	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
##	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
## 1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
## 2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
## 3	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
## 4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
## 5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	52	NA
## 6	NA	NA	NA	NA	NA	NA	NA	NA	NA	27	NA	NA	NA	NA	NA
##	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
## 1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
## 2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
## 3	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
## 4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
## 5	NA	NA	NA	NA	NA	NA	42.7	NA	NA	NA	NA	NA	NA	NA	NA
## 6	31.7	NA	NA	30.6	NA	NA	30.0	NA	NA	NA	29	NA	34.6	32.9	33.7
##	2017	2018	2019	2020	NA										
## 1	NA	NA	NA	NA	NA										
## 2	NA	NA	NA	NA	NA										


```
## 3    NA    NA    NA    NA NA
## 4    NA    NA    NA    NA NA
## 5    NA 51.3    NA    NA NA
## 6 33.2    NA    NA    NA NA
```

```
names(homicide) <- as.matrix(homicide[4, ])
homicide <- homicide[-4, ]
homicide[] <- lapply(homicide, function(x) type.convert(as.character(x)))
homicide <- homicide[-1, ]
homicide <- homicide[-1, ]
homicide <- homicide[-1, ]
row.names(homicide) <- NULL
head(homicide)
```

```
##          Country Name Country Code
## 1                Aruba          ABW
## 2 Africa Eastern and Southern      AFE
## 3            Afghanistan          AFG
## 4 Africa Western and Central      AFW
## 5                Angola          AGO
## 6            Albania          ALB
##
##          Indicator Name Indicator Code 1960 1961 1962 1963
## 1 Intentional homicides (per 100,000 people) VC.IHR.PSRC.P5    NA    NA    NA    NA
## 2 Intentional homicides (per 100,000 people) VC.IHR.PSRC.P5    NA    NA    NA    NA
## 3 Intentional homicides (per 100,000 people) VC.IHR.PSRC.P5    NA    NA    NA    NA
## 4 Intentional homicides (per 100,000 people) VC.IHR.PSRC.P5    NA    NA    NA    NA
## 5 Intentional homicides (per 100,000 people) VC.IHR.PSRC.P5    NA    NA    NA    NA
## 6 Intentional homicides (per 100,000 people) VC.IHR.PSRC.P5    NA    NA    NA    NA
## 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978
## 1    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 2    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 3    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 4    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 5    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 6    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991    1992
## 1    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 2    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 3    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 4    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 5    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 6    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA 4.005069
## 1993 1994 1995 1996 1997 1998 1999 2000
## 1    NA    NA 7.469748    NA    NA    NA 2.246964    NA
## 2    NA    NA    NA    NA    NA    NA    NA    NA
## 3    NA    NA    NA    NA    NA    NA    NA    NA
## 4    NA    NA    NA    NA    NA    NA    NA    NA
## 5    NA    NA    NA    NA    NA    NA    NA    NA
## 6 5.977724 3.55949 8.320154 8.777877 43.03577 21.2171 17.036539 4.186312
## 2001 2002 2003 2004 2005 2006 2007 2008
## 1 4.306076 5.263601 4.123031 2.025440 5.998320 4.958842 2.963665 4.932815
## 2    NA    NA    NA    NA    NA    NA    NA    NA
## 3    NA    NA    NA    NA    NA    NA    NA    NA
## 4    NA    NA    NA    NA    NA    NA    NA    NA
## 5    NA    NA    NA    NA    NA    NA    NA    NA
```

```
## 6 7.029425 6.909384 5.323897 4.219147 4.988969 3.101513 3.460786 3.097230
##      2009      2010      2011      2012      2013      2014      2015      2016
## 1 3.942751 3.934491 1.959824 3.899966 5.815926 1.927228      NA      NA
## 2      NA      NA      NA 15.100000      NA 10.500000 10.300000      NA
## 3 3.926774 3.368110 4.087337 6.251328      NA      NA 9.783922 6.551163
## 4      NA      NA      NA 13.300000      NA 8.100000 8.100000      NA
## 5      NA      NA 4.355786 4.847075      NA      NA      NA      NA
## 6 2.859023 4.307963 4.848732 5.387615 4.270284 4.039627 2.214131 2.736948
##      2017      2018 2019 2020 NA
## 1      NA      NA  NA  NA NA
## 2      NA      NA  NA  NA NA
## 3 6.678401 6.655561  NA  NA NA
## 4      NA      NA  NA  NA NA
## 5      NA      NA  NA  NA NA
## 6 2.010978 2.289492  NA  NA NA
```

```
sexualviolence <- sexualviolence[-1, ]
head(sexualviolence)
```

```
## country.territory date sexual.violence      rate
## 2      Albania 2003      <NA>      <NA>
## 3      Albania 2004      <NA>      <NA>
## 4      Albania 2005      89 2.832771023
## 5      Albania 2006      87 2.756123901
## 6      Albania 2007     101 3.186456613
## 7      Albania 2008      76 2.388887649
```

Importing datasets about crime rates and renaming them

```
Murders2005 <- read.csv("data/Murders2005.csv")
Murders2006 <- read.csv("data/Murders2006.csv")
Murders2007 <- read.csv("data/Murders2007.csv")
Murders2008 <- read.csv("data/Murders2008.csv")
Murders2009 <- read.csv("data/Murders2009.csv")
Murders2010 <- read.csv("data/Murders2010.csv")

Rapes2005 <- read.csv("data/Rapes2005.csv")
Rapes2006 <- read.csv("data/Rapes2006.csv")
Rapes2007 <- read.csv("data/Rapes2007.csv")
Rapes2008 <- read.csv("data/Rapes2008.csv")
Rapes2009 <- read.csv("data\\Rapes2009.csv")
Rapes2010 <- read.csv("data/Rapes2010.csv")
```

Joining datasets by using merge function

```
df05= merge(x = Murders2005, y = Rapes2005, by = "Country")
head(df05)
```

```
## Country Amount.x Date.x Amount.y Date.y
## 1 Albania      131  2005      49  2005
## 2 Argentina  2,115  2005   3,154  2005
## 3 Armenia     58  2005     15  2005
## 4 Austria     54  2005    678  2005
## 5 Bahrain      4  2005     25  2005
```

```
## 6 Bangladesh      3,592    2005    11,291    2005
df06= merge(x = Murders2006, y = Rapes2006, by = "Country")
head(df06)
```

```
##      Country Amount.x Date.x Amount.y Date.y
## 1   Albania        87   2006         40   2006
## 2  Argentina    2,052   2006     3,264   2006
## 3   Armenia        79   2006          7   2006
## 4   Austria        61   2006         700   2006
## 5   Bahrain         7   2006          17   2006
## 6 Bangladesh    4,166   2006    11,682   2006
```

```
df07= merge(x = Murders2007, y = Rapes2007, by = "Country")
head(df07)
```

```
##      Country Amount.x Date.x Amount.y Date.y
## 1   Albania       105   2007         41   2007
## 2   Algeria       438   2007        836   2007
## 3  Argentina    2,071   2007     3,276   2007
## 4   Armenia        77   2007          5   2007
## 5   Austria        45   2007         710   2007
## 6 Azerbaijan     189   2007         34   2007
```

```
df08= merge(x = Murders2008, y = Rapes2008, by = "Country")
head(df08)
```

```
##      Country Amount.x Date.x Amount.y Date.y
## 1   Albania        93   2008         39   2008
## 2   Algeria       516   2008        812   2008
## 3  Argentina    2,305   2008     3,367   2008
## 4   Armenia        83   2008         15   2008
## 5  Australia     261   2008     6,382   2008
## 6   Austria        45   2008        693   2008
```

```
df09= merge(x = Murders2009, y = Rapes2009, by = "Country")
head(df09)
```

```
##      Country Amount.x Date.x Amount.y Date.y
## 1   Albania        85   2009         34   2009
## 2   Armenia        83   2009         15   2009
## 3  Australia     263   2009     6,343   2009
## 4   Austria        43   2009        779   2009
## 5  Barbados        19   2009         68   2009
## 6   Belarus     473   2009        218   2009
```

```
df10= merge(x = Murders2010, y = Rapes2010, by = "Country")
head(df10)
```

```
##      Country Amount.x Date.x Amount.y Date.y
## 1   Albania       127   2010         24   2010
## 2   Armenia        44   2010          11   2010
## 3  Australia     229   2010     6,378   2010
## 4   Austria        53   2010        875   2010
## 5  Azerbaijan     206   2010         16   2010
## 6   Belgium     180   2010     2,991   2010
```

Changing the column names and removing the year

```
MR2005<-subset(df05, select=-c(Date.x,Date.y))
MR2006<-subset(df06, select=-c(Date.x,Date.y))
MR2007<-subset(df07, select=-c(Date.x,Date.y))
MR2008<-subset(df08, select=-c(Date.x,Date.y))
MR2009<-subset(df09, select=-c(Date.x,Date.y))
MR2010<-subset(df10, select=-c(Date.x,Date.y))
MR0506<-merge(MR2005,MR2006,by='Country')
MR0708<-merge(MR2007,MR2008,by='Country')
MR0910<-merge(MR2009,MR2010,by='Country')
MR1<-merge(MR0506,MR0708,by='Country')
MR<-merge(MR1,MR0910,by='Country')
MR<-rename(MR, Murders2005=Amount.x.x.x, Rape2005=Amount.y.x.x,
Murders2006=Amount.x.y.x, Rape2006=Amount.y.y.x,
Murders2007=Amount.x.x.y, Rape2007=Amount.y.x.y,
Murders2008=Amount.x.y.y, Rape2008=Amount.y.y.y,
Murders2009=Amount.x.x, Rape2009=Amount.y.x,
Murders2010=Amount.x.y, Rape2010=Amount.y.y)
```

The final results

```
head(MR)
```

##		Country	Murders2005	Rape2005	Murders2006	Rape2006	Murders2007	
## 1		Albania	131	49	87	40	105	
## 2		Armenia	58	15	79	7	77	
## 3		Austria	54	678	61	700	45	
## 4		Belgium	221	3,029	221	3,194	211	
## 5		Bolivia	596	1,137	530	1,137	723	
## 6		Bosnia and Herzegovina	70	39	73	40	66	
##		Rape2007	Murders2008	Rape2008	Murders2009	Rape2009	Murders2010	Rape2010
## 1		41	93	39	85	34	127	24
## 2		5	83	15	83	15	44	11
## 3		710	45	693	43	779	53	875
## 4		3,232	200	3,126	186	2,953	180	2,991
## 5		1,437	724	1,596	673	1,989	884	2,587
## 6		44	66	35	67	29	56	46

- The economical data

```
employmentcopy<-employment
employmentcopy<-filter(employment, Employment.status == 'Total employment')
EMP<-subset(employmentcopy, select=c(Country,Time,Value))
EMP2010 <- filter(EMP, Time == 2010)
EMP2010<-subset(EMP2010, select=-c(Time))
EMP2010<-rename(EMP2010, Unemployment2010=Value)

GPD2005<-data.frame(gpd[, c(1,50)])
GPD2005<-rename(GPD2005, Country=Country.Name, GPD2005=X2005)
```

One can see that only years 2005 and 2010 are common in merged data frames. Only the data of 2005 and 2010 were kept and the data of the other years were deleted

```
MR0510<-subset(MR, select=c(Country,Murders2005,Rape2005,Murders2010,Rape2010))
head(MR0510)
```

```
##           Country Murders2005 Rape2005 Murders2010 Rape2010
## 1      Albania         131         49         127         24
## 2      Armenia          58          15          44          11
## 3      Austria          54         678          53         875
## 4      Belgium        221       3,029         180       2,991
## 5      Bolivia        596       1,137         884       2,587
## 6 Bosnia and Herzegovina    70          39          56          46
```

Merging with the economical values

- This is the final data frame to make analyzes from.

```
prefinal<-merge(EMP2010,MR0510, by='Country')
FINAL<-merge(prefinal,GPD2005, by='Country')
head(FINAL)
```

```
## Country Unemployment2010 Murders2005 Rape2005 Murders2010 Rape2010
## 1 Austria          1552.0          54         678          53         875
## 2 Belgium          1574.0         221       3,029         180       2,991
## 3 Canada           1715.0         663         565         554         576
## 4 Estonia           1785.0         113         179          70          81
## 5 Finland           1585.0         119         593         118         818
## 6 Germany           1425.7         869       8,133         690       7,724
##           GPD2005
## 1 3.159740e+11
## 2 3.855710e+11
## 3 1.173110e+12
## 4 1.410679e+10
## 5 2.048090e+11
## 6 2.845800e+12
```

Before continuing, checking the types of data stored in FINAL

```
sapply(FINAL, mode)
```

```
##           Country Unemployment2010 Murders2005 Rape2005
## "character"      "numeric"         "character" "character"
## Murders2010      Rape2010          GPD2005
## "character"      "character"         "numeric"
```

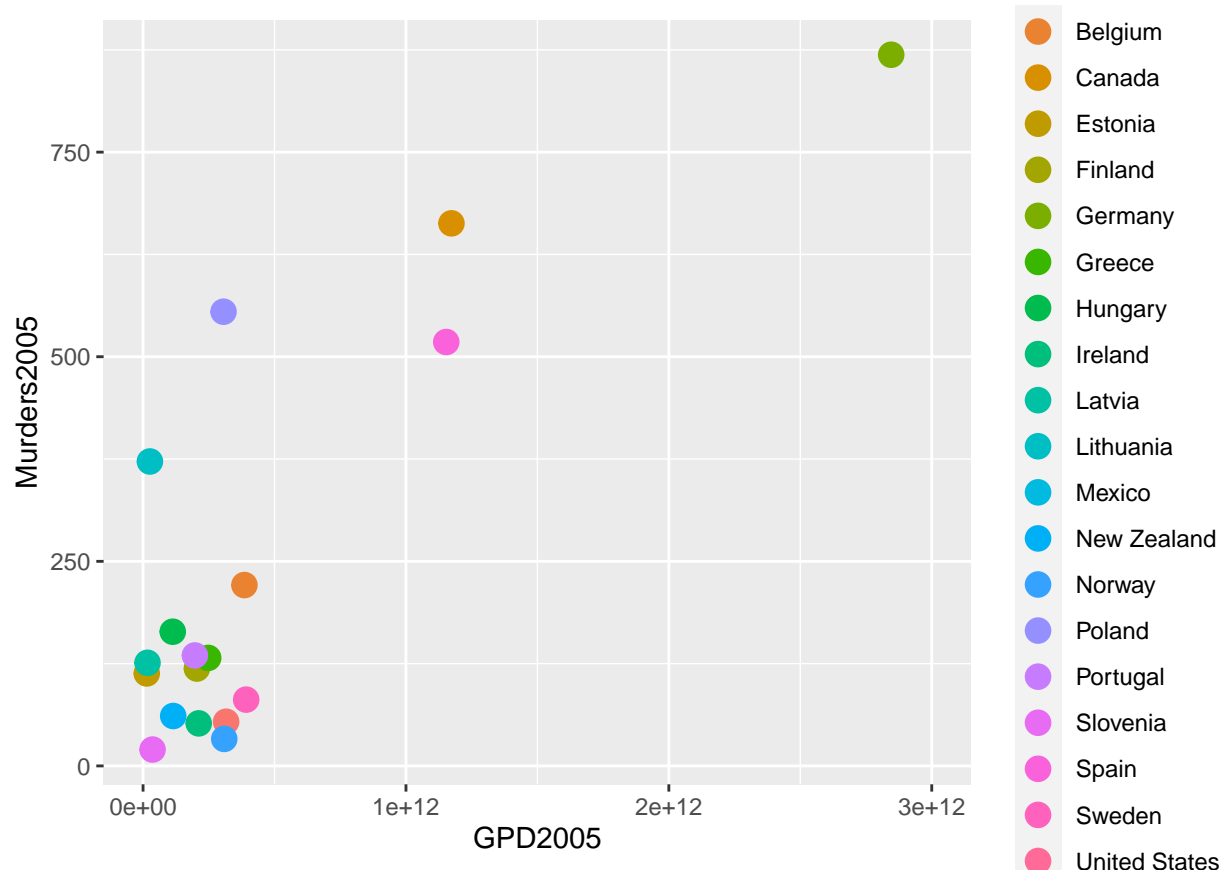
Converting characters to numerical values

```
cols.num <- c("Murders2005","Rape2005","Murders2010","Rape2010")
FINAL[cols.num] <- sapply(FINAL[cols.num],as.numeric)
sapply(FINAL, mode)
```

```
##           Country Unemployment2010 Murders2005 Rape2005
## "character"      "numeric"         "numeric"    "numeric"
## Murders2010      Rape2010          GPD2005
## "numeric"        "numeric"         "numeric"
```

- Result Although the data is concentrated in a certain region, a positive correlation is observed between murder counts and GPD. That is, as the number of murders increases, so does the GPD.

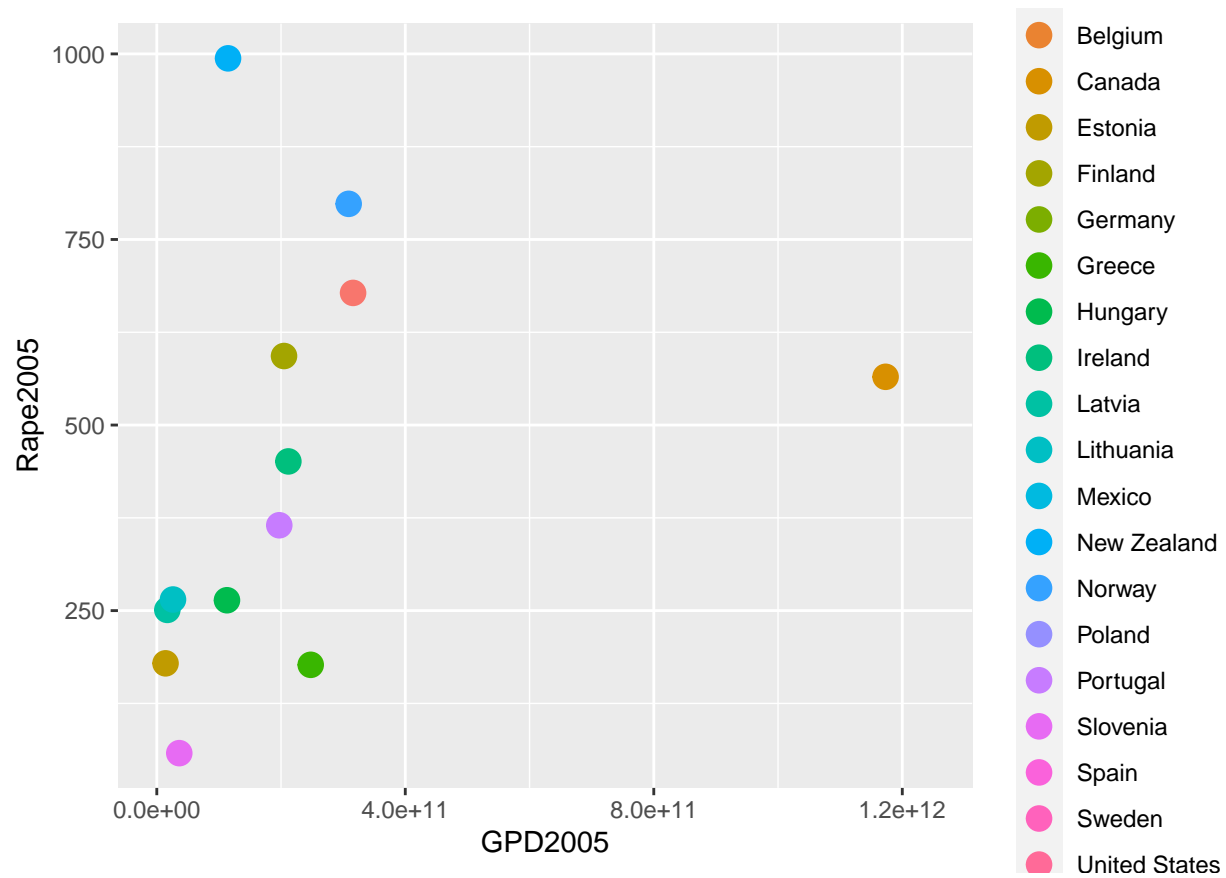
```
ggplot(FINAL, aes(x=GPD2005, y=Murders2005, color=Country)) +
  geom_point(size=4) + xlim(0, 3*1012)
```



Graphing 2

- Result Although there is not much difference between the GDPs, positive correlation is also seen in this graph. Canada is located here in a separate position from other countries in terms of GDP.

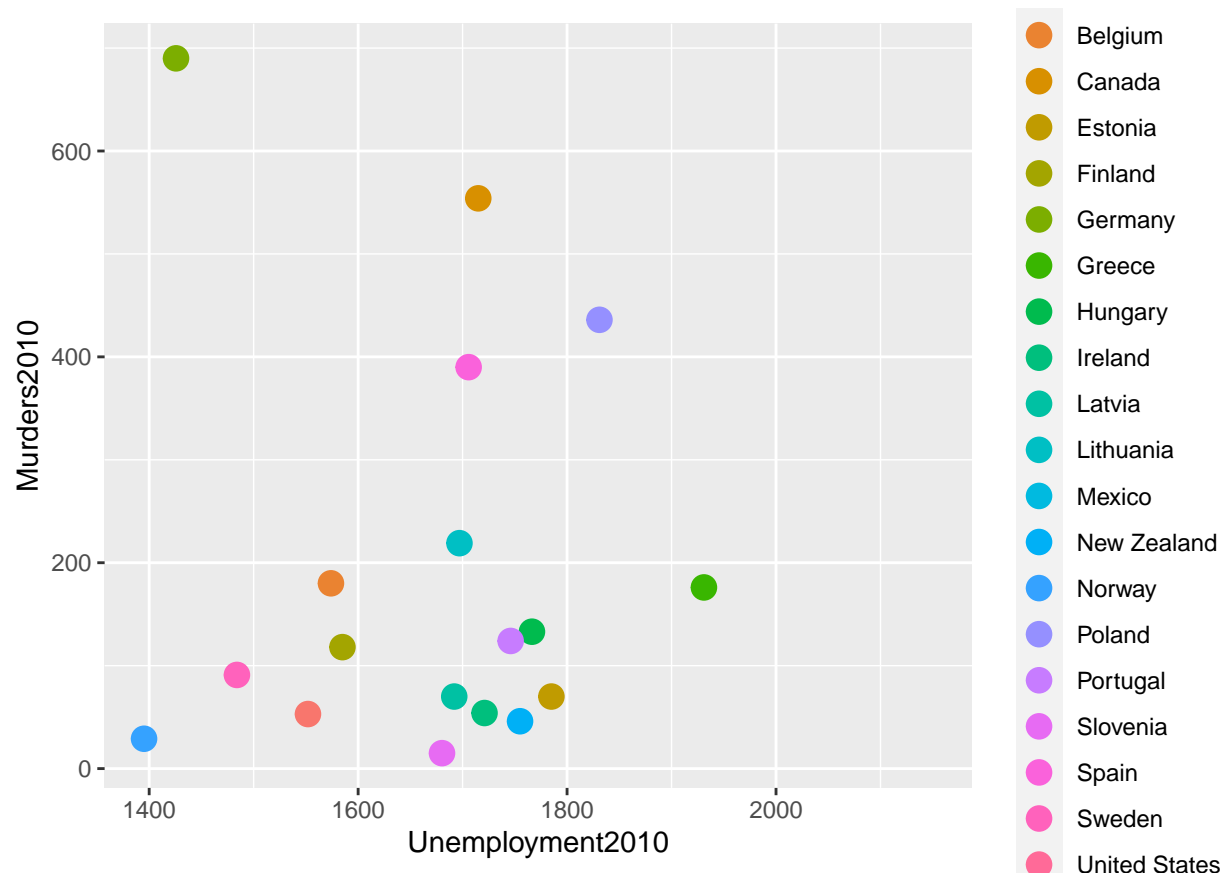
```
ggplot(FINAL, aes(x=GPD2005, y=Rape2005, color=Country)) +
  geom_point(size=4) + xlim(0, 125*1010)
```



Graphing 3

- Result Positive correlation is seen in this graph. Germany has a very high value in homicides and low in unemployment. For this reason, it is in a different position from other countries.

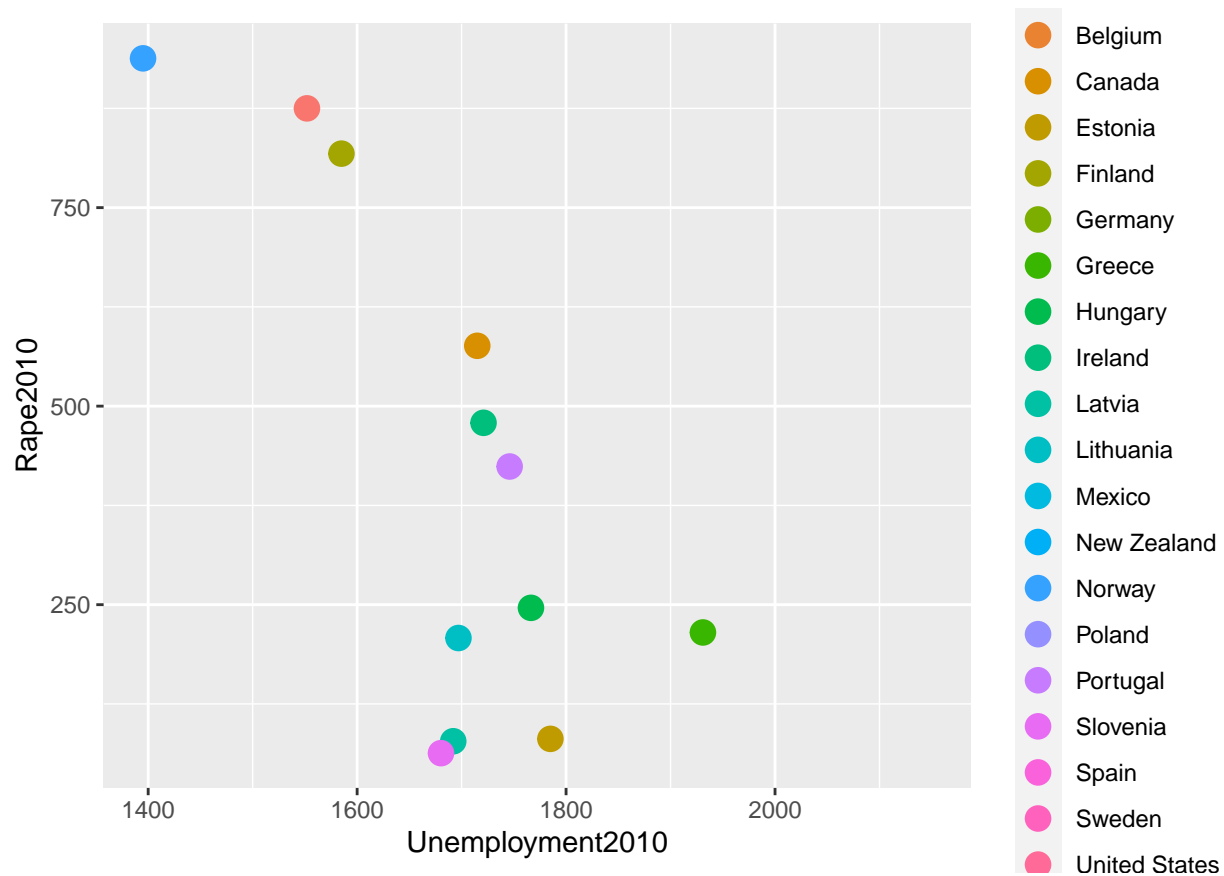
```
ggplot(FINAL, aes(x=Unemployment2010, y=Murders2010, color=Country)) +
  geom_point(size=4)
```



Graphing 4

- Result Unlike other graphs, there is no positive correlation in this graph. The employment rate increases as the rape rate decreases.

```
ggplot(FINAL, aes(x=Unemployment2010, y=Rape2010, color=Country)) +  
  geom_point(size=4)
```

Conclusion.

With the exception of the third chart, similar results are obtained with the other charts. The relationship between GPD and crime rates is surprising. In 2010, it was observed that as unemployment rates increased, murder increased, but rape rates decreased. In the past 5 years, it has been observed that there is an awareness of rape in high-income countries.

References.

- https://www.numbeo.com/crime/rankings_by_country.jsp?region=150&title=2019
- <https://data.worldbank.org/indicator/SI.POV.GINI>
- <https://microdata.worldbank.org/index.php/catalog/1790/get-microdata>