

About this site

```
library(magrittr)
library(ggthemes)
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

```
library(tidyr)
```

Attaching package: 'tidyr'

The following object is masked from 'package:magrittr':

extract

```
library(ggplot2)
```

## Data Visualization of Poverty Rate Data

```
library(ggplot2)
library(tidyr)
library(dplyr)
library(readxl)
# Poverty rate data set
poverty_rate_19_20 <- read_excel("C:/Users/Necati/Desktop/Poverty Rate 2020-2019.xlsx")

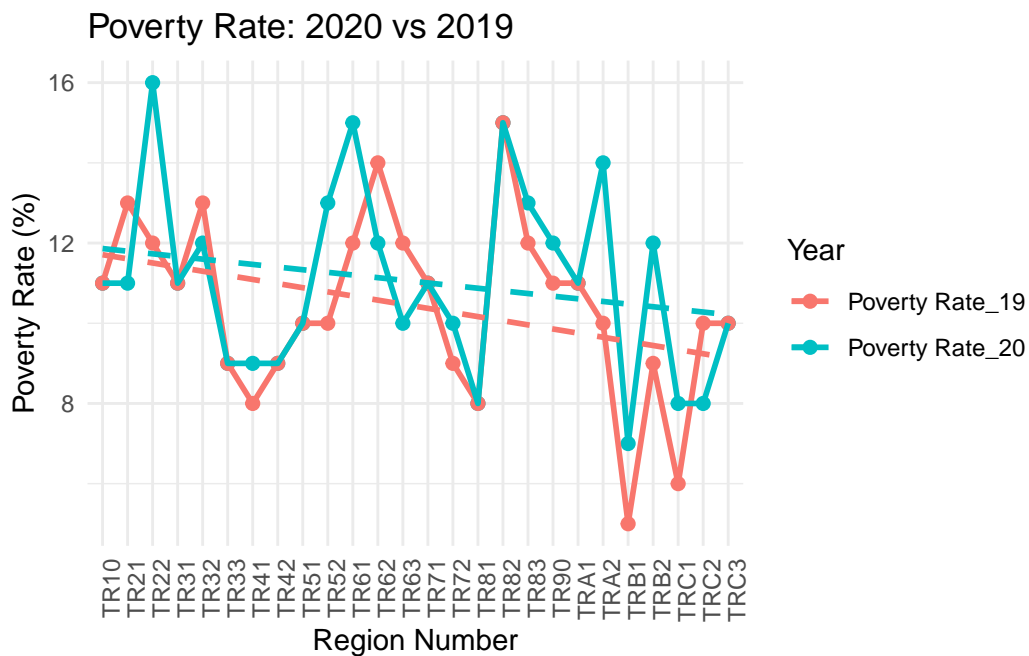
# We convert the Poverty Rate columns to long format.
poverty_rate_long <- poverty_rate_19_20 %>%
  pivot_longer(
    cols = c('Poverty Rate_20', 'Poverty Rate_19'), # Artık alt çizgili, boşluksuz sütun adları
    names_to = "Year",
    values_to = "Poverty_Rate"
  )
```

```
#Data visualization of Poverty rate data
b <- ggplot(poverty_rate_long,
            aes(x = Region_Number,
                y = Poverty_Rate,
                colour = Year,
                group = Year)) +
  geom_point(size = 2) +
  geom_line(size = 1) +
  geom_smooth(method = "lm", se = FALSE, linetype = "dashed", size = 1) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
  labs(
    title = "Poverty Rate: 2020 vs 2019",
    x = "Region Number",
    y = "Poverty Rate (%)",
    colour = "Year"
  )
)
```

Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.  
i Please use `linewidth` instead.

b

`geom\_smooth()` using formula = 'y ~ x'



## Poverty Rate vs Homicide for 2020 Male Data

```
library(ggplot2)
library(readxl)
library(ggthemes)
data_collect_erkek <- read_excel("C:/Users/Necati/Desktop/Data_collection_erkek.xlsx")
temiz_veri_erkek <- na.omit(data_collect_erkek)

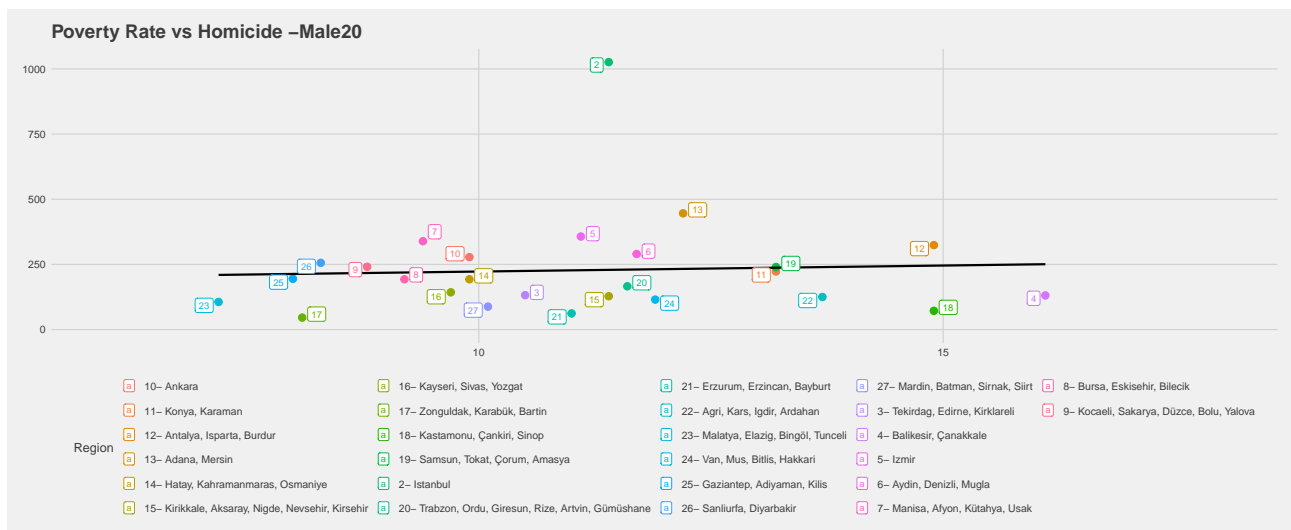
veri_temiz_erkek <- data_collect_erkek[rowSums(is.na(data_collect_erkek)) != ncol(data_collect_erkek)]

data_son_temiz_erkek <- veri_temiz_erkek[-c(1,2,5,6,7,9,10,13,14,15,17,18,19,20,22,23,24,26,27)]

b <- ggplot(data_son_temiz_erkek, aes(x = `Poverty Rate (%)`, y = Homicide, color = Region)) +
  geom_point(size = 3) +
  geom_smooth(method = "lm", se = FALSE, color = "black") +
  theme_minimal() +
  labs(title = "Poverty Rate vs Homicide -Male20",
       x = "Poverty Rate (%)",
       y = "Homicide") +
  scale_x_continuous(limits = c(6, 18)) +
  scale_y_continuous(limits = c(0, 1026)) +
  geom_label_repel(aes(label = Region_Number), size = 3, max.overlaps = 50) +
  theme_fivethirtyeight()
```

b

`geom\_smooth()` using formula = 'y ~ x'



## Correlation Analysis for Male 2020

```
#(Poverty Rate vs Homicide -Male 20)
correlation_result <- cor.test(data_son_temiz_erkek$`Poverty Rate (%)`, data_son_temiz_erkek$H
print(correlation_result)
```

Pearson's product-moment correlation

```
data: data_son_temiz_erkek$`Poverty Rate (%)` and data_son_temiz_erkek$Homicide
t = 0.27118, df = 24, p-value = 0.7886
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.3393465  0.4333449
sample estimates:
      cor
0.05527056
```

## Poverty Rate vs Homicide for 2020 Female Data

```
data_collect_kadin <- read_excel("C:/Users/Necati/Desktop/Data_collection_kadin.xlsx")
temiz_veri_kadin <- na.omit(data_collect_kadin)

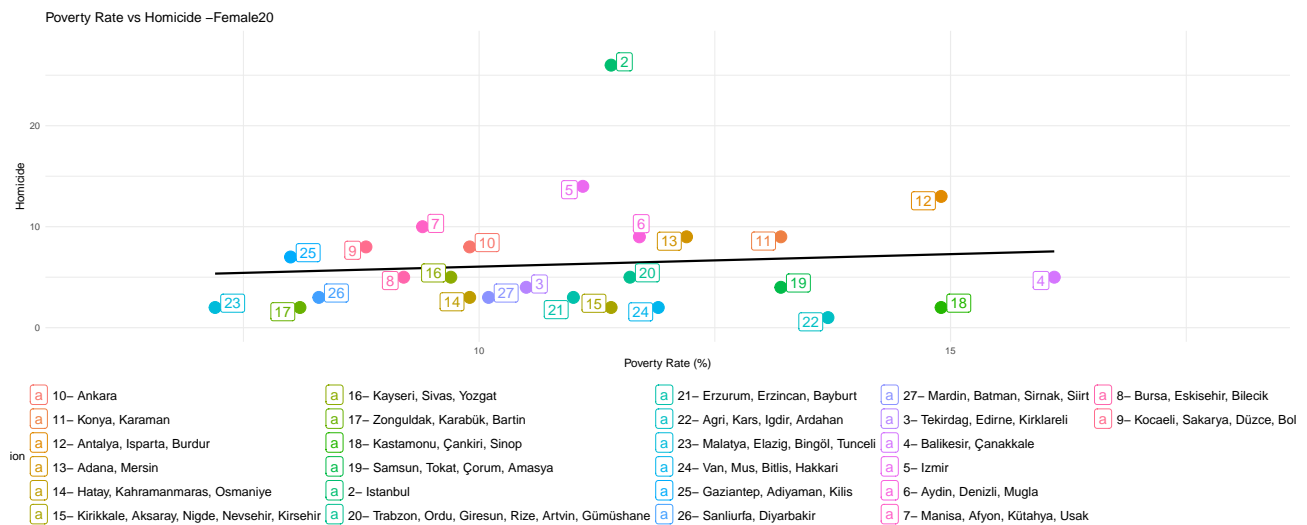
veri_temiz_kadin <- data_collect_kadin[rowSums(is.na(data_collect_kadin)) != ncol(data_collect_kadin)]

data_son_temiz_kadin <- veri_temiz_kadin[-c(1,2,5,6,7,9,10,13,14,15,17,18,19,20,22,23,24,26,27)]
b <- ggplot(data_son_temiz_kadin, aes(x = `Poverty Rate (%)`, y = Homicide, color = Region)) +
  geom_point(size = 5) +
  geom_smooth(method = "lm", se = FALSE, color = "black") +
  theme_minimal() +
  labs(title = "Poverty Rate vs Homicide -Female20",
       x = "Poverty Rate (%)",
       y = "Homicide") +
  scale_x_continuous(limits = c(6,18)) +
  scale_y_continuous(limits = c(0,28)) +
  geom_label_repel(aes(label = Region_Number,size = 5), size = 5, max.overlaps = 50) +

  theme(
    legend.position = "bottom",
    legend.box = "horizontal",
    legend.text = element_text(size = 12),
    legend.key.size = unit(0.5,"lines")
  )
```

```
print(b)
```

```
`geom_smooth()` using formula = 'y ~ x'
```



## Correlation Analysis for Female 2020

```
#Poverty Rate vs Homicide -Female 20
correlation_result <- cor.test(data_son_temiz_kadin$`Poverty Rate (%)`, data_son_temiz_kadin$Homicide)
print(correlation_result)
```

Pearson's product-moment correlation

```
data: data_son_temiz_kadin$`Poverty Rate (%)` and data_son_temiz_kadin$Homicide
t = 0.5227, df = 24, p-value = 0.606
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.2933119 0.4739678
sample estimates:
cor
0.1060939
```

## Poverty Rate vs Homicide for 2019 Male Data

```
data_collect_erkek_2019 <- read_excel("C:/Users/Necati/Desktop/Data_collection_2019_erkek.xlsx")
temiz_veri_erkek_2019 <- na.omit(data_collect_erkek_2019)

veri_temiz_erkek_2019 <- data_collect_erkek_2019[rowSums(is.na(data_collect_erkek_2019)) != nc

data_son_temiz_erkek_2019 <- veri_temiz_erkek_2019[-c(1,2,5,6,7,9,10,13,14,15,17,18,19,20,22,2

b <- ggplot(data_son_temiz_erkek_2019, aes(x = `Poverty Rate (%)`, y = Homicide, color = Region
  geom_point(size = 3) +
  geom_smooth(method = "lm", se = FALSE, color = "black") +
  theme_minimal() +
  labs(title = "Poverty Rate vs Homicide -Male19",
        x = "Poverty Rate (%)",
        y = "Homicide") +
  scale_x_continuous(limits = c(4, 16)) +
  scale_y_continuous(limits = c(0, 1605)) +
  geom_label_repel(aes(label = Region_Number), size = 3, max.overlaps = 50) +
  theme_fivethirtyeight()

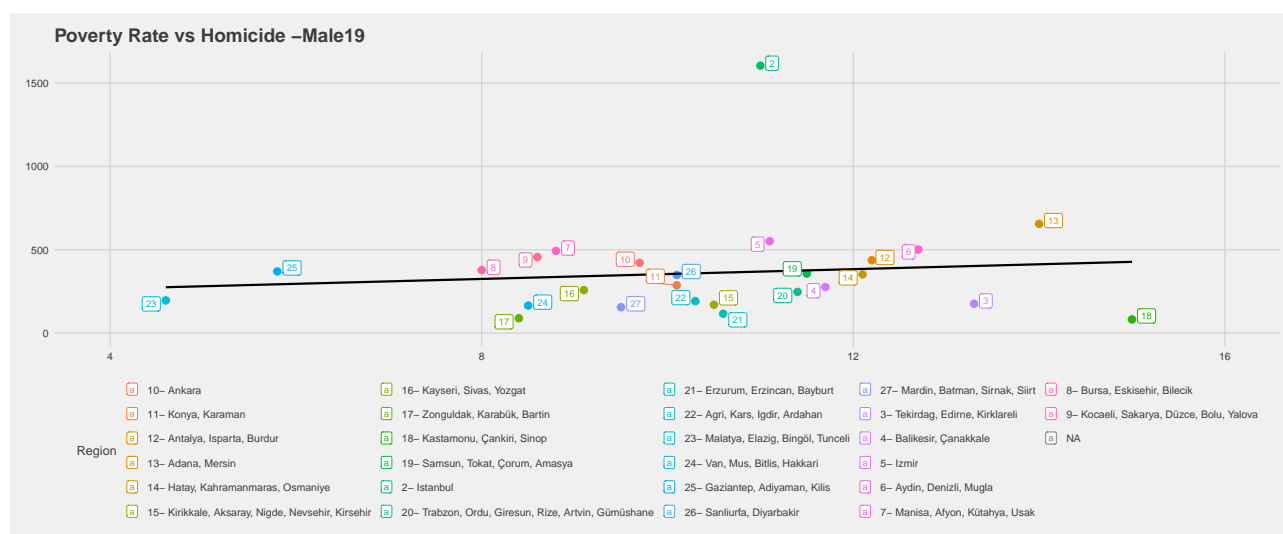
print(b)
```

`geom\_smooth()` using formula = 'y ~ x'

Warning: Removed 4 rows containing non-finite outside the scale range  
(`stat\_smooth()`).

Warning: Removed 4 rows containing missing values or values outside the scale range  
(`geom\_point()`).

Warning: Removed 4 rows containing missing values or values outside the scale range  
(`geom\_label\_repel()`).



## Correlation Analysis for Male 2019

```
# Poverty Rate vs Homicide -Male 19
correlation_result <- cor.test(data_son_temiz_erkek_2019$`Poverty Rate (%)`, data_son_temiz_erkek_2019$Homicide, method = "spearmanr")
print(correlation_result)
```

Pearson's product-moment correlation

```
data: data_son_temiz_erkek_2019$`Poverty Rate (%)` and data_son_temiz_erkek_2019$Homicide
t = 0.57003, df = 24, p-value = 0.574
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.2845124  0.4813780
sample estimates:
      cor
0.1155768
```

## Poverty Rate vs Homicide for 2019 Female Data

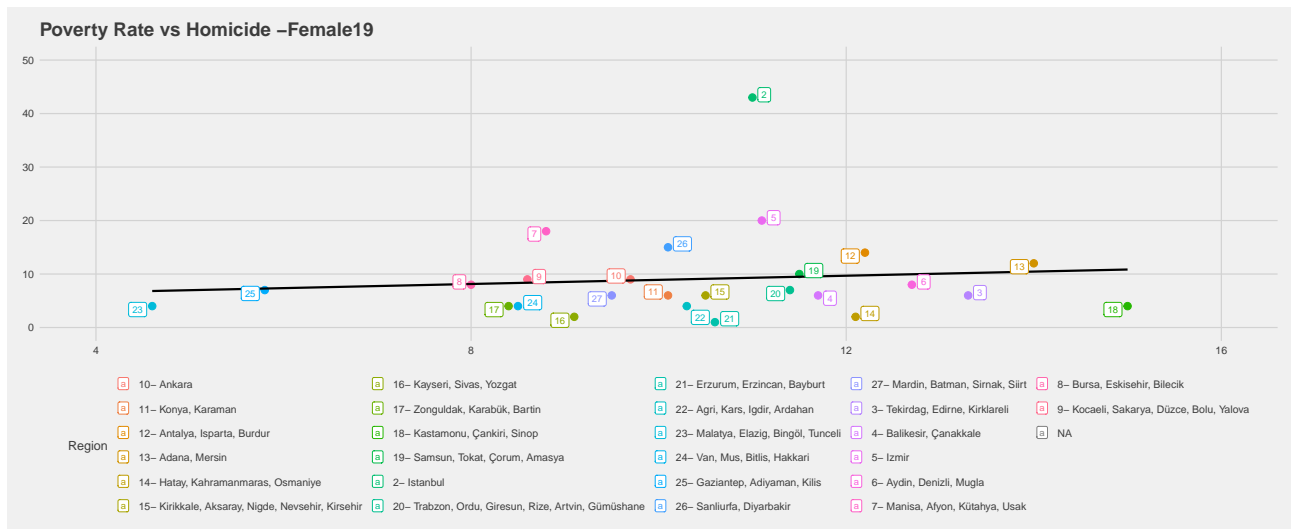
```
data_collect_kadin_2019 <- read_excel("C:/Users/Necati/Desktop/Data_collection_2019_kadin.xlsx")
temiz_veri_kadin_2019 <- na.omit(data_collect_kadin_2019)
veri_temiz_kadin_2019 <- data_collect_kadin_2019[rowSums(is.na(data_collect_kadin_2019)) != 0, ]
data_son_temiz_kadin_2019 <- veri_temiz_kadin_2019[-c(
  1, 2, 5, 6, 7, 9, 10, 13, 14, 15,
  17, 18, 19, 20, 22, 23, 24, 26, 27, 28,
  29, 30, 33, 34, 36, 37, 38, 40, 41, 43,
  44, 45, 47, 48, 49, 50, 51, 53, 54, 55,
  57, 58, 59, 61, 62, 63, 65, 66, 67, 68,
  70, 71, 72, 73, 74, 75, 77, 78, 79, 81,
  82, 83, 84, 86, 87, 88, 89, 91, 92, 93,
  94, 96, 97, 98, 100, 101, 103, 104, 105, 106
), ]
b<- ggplot(data_son_temiz_kadin_2019, aes(x = `Poverty Rate (%)`, y = Homicide, color = Region)) +
  geom_point(size = 3) +
  geom_smooth(method = "lm", se = FALSE, color = "black") +
  theme_minimal() +
  labs(title = "Poverty Rate vs Homicide -Female19",
       x = "Poverty Rate (%)",
       y = "Homicide") +
  scale_x_continuous(limits = c(4,16))+
  scale_y_continuous(limits = c(0,50))+ geom_label_repel(aes(label = Region_Number), size = 3)
b
```

```
`geom_smooth()` using formula = 'y ~ x'
```

```
Warning: Removed 4 rows containing non-finite outside the scale range
(`stat_smooth()`).
```

```
Warning: Removed 4 rows containing missing values or values outside the scale range
(`geom_point()`).
```

```
Warning: Removed 4 rows containing missing values or values outside the scale range
(`geom_label_repel()`).
```



## Correlation Analysis for Female 2019

```
# Poverty Rate vs Homicide -Female 19
correlation_result <- cor.test(data_son_temiz_kadin_2019$`Poverty Rate (%)`, data_son_temiz_kadin_2019$Homicide Rate (%))
print(correlation_result)
```

Pearson's product-moment correlation

```
data: data_son_temiz_kadin_2019$`Poverty Rate (%)` and data_son_temiz_kadin_2019$Homicide Rate (%)
t = 0.52722, df = 24, p-value = 0.6029
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.2924727 0.4746792
sample estimates:
cor
0.1070014
```



## Poverty Rate vs Robbery for 2020 Male Data

```
library(ggrepel)
library(readxl)
library(ggthemes)
data_collect_erkek <- read_excel("C:/Users/Necati/Desktop/Data_collection_erkek.xlsx")
temiz_veri_erkek <- na.omit(data_collect_erkek)

veri_temiz_erkek <- data_collect_erkek[rowSums(is.na(data_collect_erkek)) != ncol(data_collect_erkek)]

data_son_temiz_erkek <- veri_temiz_erkek[-c(1,2,5,6,7,9,10,13,14,15,17,18,19,20,22,23,24,26,27)]
data_son_temiz_erkek$`Poverty Rate (%)` <- as.numeric(data_son_temiz_erkek$`Poverty Rate (%)`)

b <- ggplot(data_son_temiz_erkek, aes(x = `Poverty Rate (%)`, y = Robbery, color = Region)) +
  geom_point(size = 3) +
  geom_smooth(method = "lm", se = FALSE, color = "black") +
  theme_minimal() +
  labs(title = "Poverty Rate vs Robbery -Male20",
       x = "Poverty Rate (%)",
       y = "Robbery") +
  scale_x_continuous(limits = c(6, 18)) +
  scale_y_continuous(limits = c(0, 600)) +
  geom_label_repel(aes(label = Region_Number), size = 3, max.overlaps = 50) +
  theme_fivethirtyeight()

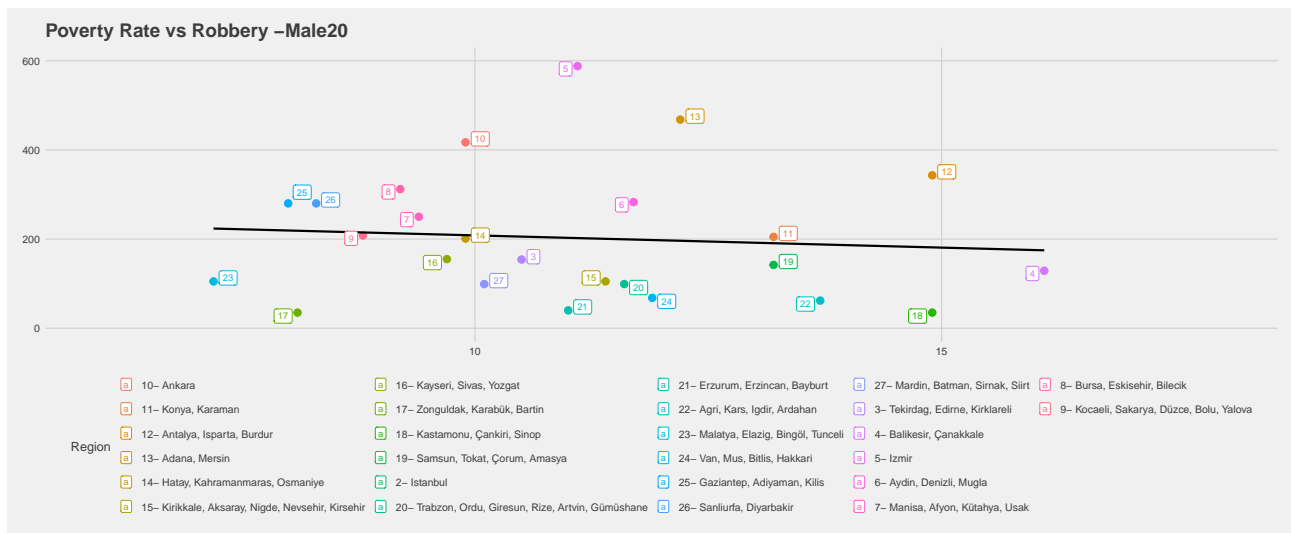
b
```

```
`geom_smooth()` using formula = 'y ~ x'
```

```
Warning: Removed 1 row containing non-finite outside the scale range
(`stat_smooth()`).
```

```
Warning: Removed 1 row containing missing values or values outside the scale range
(`geom_point()`).
```

```
Warning: Removed 1 row containing missing values or values outside the scale range
(`geom_label_repel()`).
```



## Correlation Analysis for Male 2020

```
# Poverty Rate vs Robbery -Male 20
correlation_result <- cor.test(data_son_temiz_erkek$`Poverty Rate (%)`, data_son_temiz_erkek$Robbery)
print(correlation_result)
```

Pearson's product-moment correlation

```
data: data_son_temiz_erkek$`Poverty Rate (%)` and data_son_temiz_erkek$Robbery
t = -0.020354, df = 24, p-value = 0.9839
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.3908777 0.3838148
sample estimates:
 cor
-0.004154802
```

## Poverty Rate vs Robbery for 2020 Female Data

```
data_collect_kadin <- read_excel("C:/Users/Necati/Desktop/Data_collection_kadin.xlsx")
temiz_veri_kadin <- na.omit(data_collect_kadin)

veri_temiz_kadin <- data_collect_kadin[rowSums(is.na(data_collect_kadin)) != ncol(data_collect_kadin)]
```

```
data_son_temiz_kadin <- veri_temiz_kadin[-c(1,2,5,6,7,9,10,13,14,15,17,18,19,20,22,23,24,26,27)]
b <- ggplot(data_son_temiz_kadin, aes(x = `Poverty Rate (%)`, y = Robbery, color = Region)) +
  geom_point(size = 5) +
  geom_smooth(method = "lm", se = FALSE, color = "black") +
  theme_minimal() +
  labs(title = "Poverty Rate vs Robbery -Female20",
       x = "Poverty Rate (%)",
       y = "Robbert") +
  scale_x_continuous(limits = c(6,18)) +
  scale_y_continuous(limits = c(0,28)) +
  geom_label_repel(aes(label = Region_Number,size = 5), size = 5, max.overlaps = 50) +

  theme(
    legend.position = "bottom",
    legend.box = "horizontal",
    legend.text = element_text(size = 12),
    legend.key.size = unit(0.5,"lines")
  )

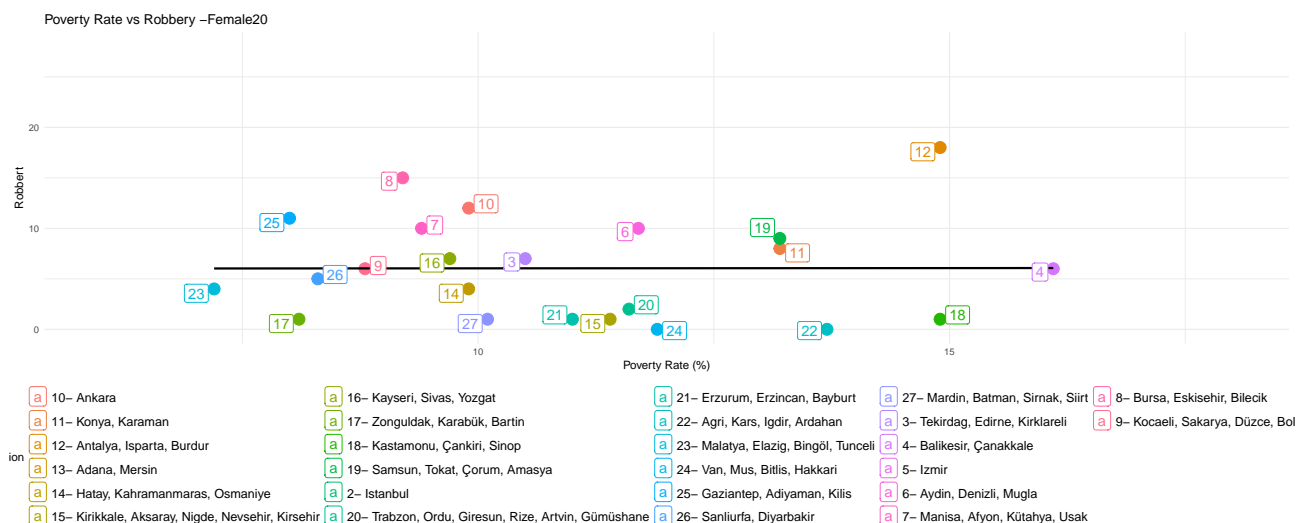
print(b)
```

`geom\_smooth()` using formula = 'y ~ x'

Warning: Removed 3 rows containing non-finite outside the scale range  
(`stat\_smooth()`).

Warning: Removed 3 rows containing missing values or values outside the scale range  
(`geom\_point()`).

Warning: Removed 3 rows containing missing values or values outside the scale range  
(`geom\_label\_repel()`).



## Correlation Analysis for Female 2020

```
#Poverty Rate vs Robbery -Female 20
correlation_result <- cor.test(data_son_temiz_kadin$`Poverty Rate (%)`, data_son_temiz_kadin$R
print(correlation_result)
```

Pearson's product-moment correlation

```
data: data_son_temiz_kadin$`Poverty Rate (%)` and data_son_temiz_kadin$Robbery
t = 0.31123, df = 24, p-value = 0.7583
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.3321056  0.4399498
sample estimates:
      cor
0.0634026
```

## Poverty Rate vs Robbery for 2019 Male Data

```
data_collect_erkek_2019 <- read_excel("C:/Users/Necati/Desktop/Data_collection_2019_erkek.xlsx")
temiz_veri_erkek_2019 <- na.omit(data_collect_erkek_2019)

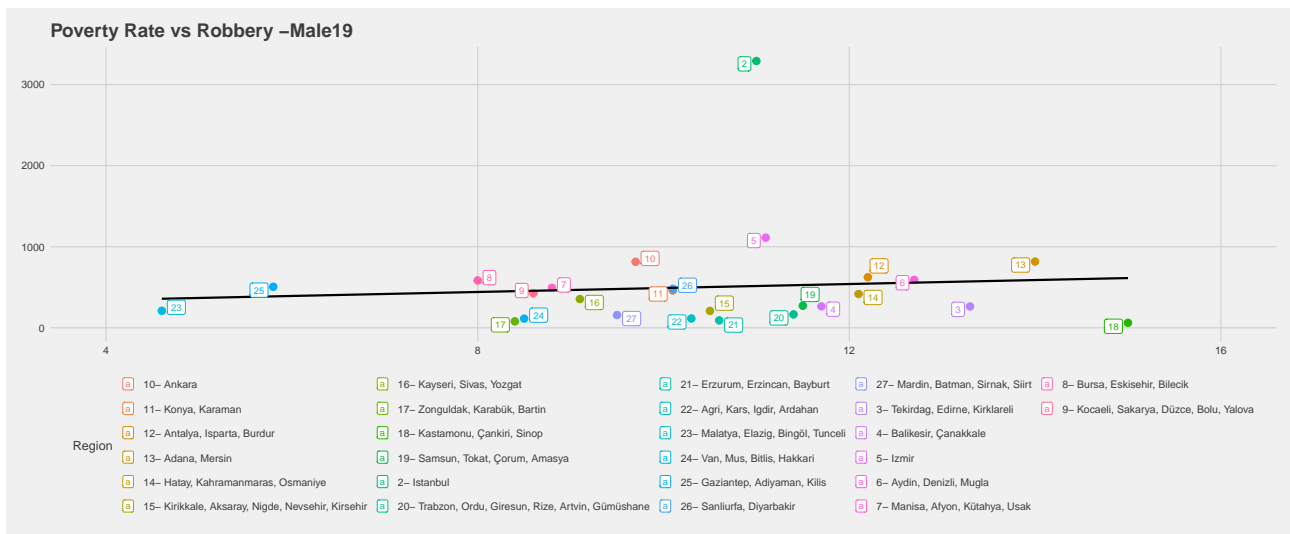
veri_temiz_erkek_2019 <- data_collect_erkek_2019[rowSums(is.na(data_collect_erkek_2019)) != nc

data_son_temiz_erkek_2019 <- veri_temiz_erkek_2019[-c(1,2,5,6,7,9,10,13,14,15,17,18,19,20,22,2
data_son_temiz_erkek_19<-data_son_temiz_erkek_2019[-c(27,28,29,30),]

b <- ggplot(data_son_temiz_erkek_19, aes(x = `Poverty Rate (%)`, y = Robbery, color = Region))
  geom_point(size = 3) +
  geom_smooth(method = "lm", se = FALSE, color = "black") +
  theme_minimal() +
  labs(title = "Poverty Rate vs Robbery -Male19",
       x = "Poverty Rate (%)",
       y = "Robbery") +
  scale_x_continuous(limits = c(4, 16)) +
  scale_y_continuous(limits = c(0, 3300)) +
  geom_label_repel(aes(label = Region_Number), size = 3, max.overlaps = 50) +
  theme_fivethirtyeight()

print(b)
```

```
`geom_smooth()` using formula = 'y ~ x'
```



## Correlation Analysis for Male 2019

```
# Poverty Rate vs Robbery -Male 19
correlation_result <- cor.test(data_son_temiz_erkek_19$`Poverty Rate (%)`, data_son_temiz_erkek_19$Robbery)
print(correlation_result)
```

Pearson's product-moment correlation

```
data: data_son_temiz_erkek_19$`Poverty Rate (%)` and data_son_temiz_erkek_19$Robbery
t = 0.44878, df = 24, p-value = 0.6576
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.3069748 0.4622426
sample estimates:
cor
0.09122443
```

## Poverty Rate vs Robbery for 2019 Female Data

```
data_collect_kadin_2019 <- read_excel("c:/Users/Necati/Desktop/Data_collection_2019_kadin.xlsx")
temiz_veri_kadin_2019 <- na.omit(data_collect_kadin_2019)
veri_temiz_kadin_2019 <- data_collect_kadin_2019[rowSums(is.na(data_collect_kadin_2019)) != nc
```

```
data_son_temiz_kadin_2019 <- veri_temiz_kadin_2019[-c(
  1, 2, 5, 6, 7, 9, 10, 13, 14, 15,
  17, 18, 19, 20, 22, 23, 24, 26, 27, 28,
  29, 30, 33, 34, 36, 37, 38, 40, 41, 43,
  44, 45, 47, 48, 49, 50, 51, 53, 54, 55,
  57, 58, 59, 61, 62, 63, 65, 66, 67, 68,
  70, 71, 72, 73, 74, 75, 77, 78, 79, 81,
  82, 83, 84, 86, 87, 88, 89, 91, 92, 93,
  94, 96, 97, 98, 100, 101, 103, 104, 105, 106
), ]
data_son_temiz_kadin_19<-data_son_temiz_kadin_2019[-c(27,28,29,30),]

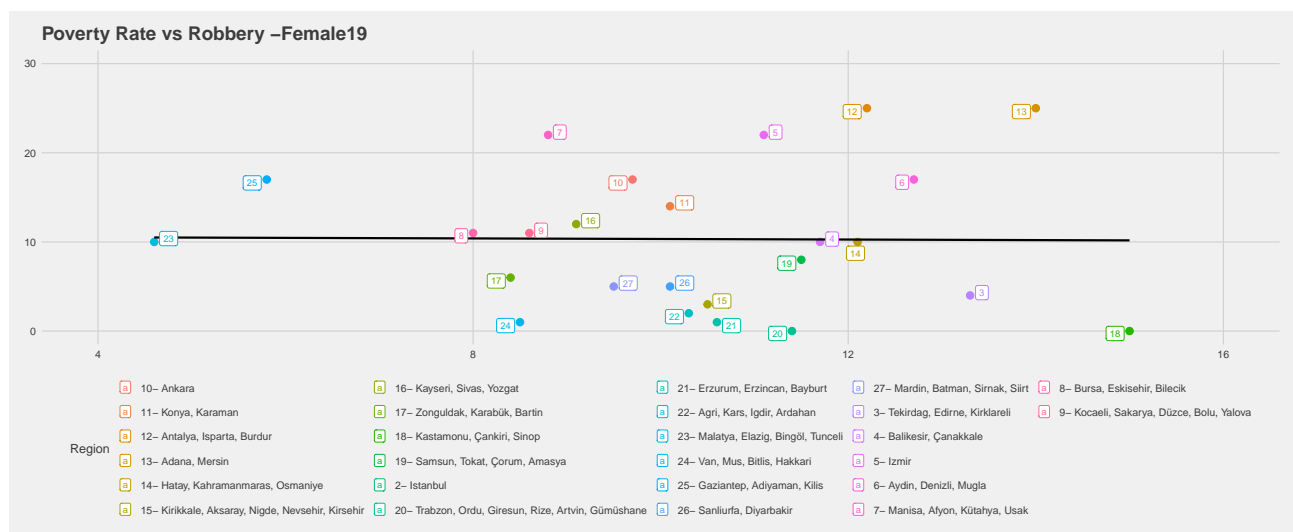
b<- ggplot(data_son_temiz_kadin_19, aes(x = `Poverty Rate (%)`, y = Robbery, color = Region))
  geom_point(size = 3) +
  geom_smooth(method = "lm", se = FALSE, color = "black") +
  theme_minimal() +
  labs(title = "Poverty Rate vs Robbery -Female19",
       x = "Poverty Rate (%)",
       y = "Robbery") +
  scale_x_continuous(limits = c(4,16))+
  scale_y_continuous(limits = c(0,30))+ geom_label_repel(aes(label = Region_Number), size = 3)
b
```

`geom\_smooth()` using formula = 'y ~ x'

Warning: Removed 1 row containing non-finite outside the scale range  
(`stat\_smooth()`).

Warning: Removed 1 row containing missing values or values outside the scale range  
(`geom\_point()`).

Warning: Removed 1 row containing missing values or values outside the scale range  
(`geom\_label\_repel()`).



## Correlation Analysis for Female 2019

```
#Poverty Rate vs Robbery -Female 19
correlation_result <- cor.test(data_son_temiz_kadin_19$`Poverty Rate (%)`, data_son_temiz_kadi
print(correlation_result)
```

Pearson's product-moment correlation

```
data: data_son_temiz_kadin_19$`Poverty Rate (%)` and data_son_temiz_kadin_19$Robbery
t = 0.20968, df = 24, p-value = 0.8357
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.3503947  0.4231048
sample estimates:
      cor
0.04276097
```

## Key Takeaways

### Poverty Rate 2019 vs 2020

In 2020, poverty rates **increased** in regions such as **TR21** (Tekirdağ, Edirne, Kırklareli), **TR62** (Adana, Mersin) and **TRA** (Ağrı, Kars, Iğdır, Ardahan), while **decreasing** in **TRC2** (Şanlıurfa, Diyarbakır) and **TRC3** (Mardin, Batman, Şırnak, Siirt). The **highest** poverty rates were observed in **TR63** (Hatay, Kahramanmaraş, Osmaniye) and **TRA1** (Erzurum, Erzinçan, Bayburt) regions, while rates remained **lower** in metropolitan areas such as **TR10** (Istanbul) and **TR51** (Ankara), but showed an upward trend in 2020.

## Detailed Evaluation of Correlation Analysis Results

### General View

The results of the presented correlation analysis show that there is not as strong a relationship between poverty rate and homicide and robbery rates as expected. Analyses for both men and women did not reveal any significant correlation between these two variables, even when looking at data from different years.

### Detailed Analysis and Comments

- **Weak Correlation Coefficients:** In most analyses, correlation coefficients were below 0.15. This indicates a very weak relationship between the two variables. In other words, it can be said that an increase in the poverty rate does not cause a significant increase in homicide or robbery rates.

- **Statistical Insignificance:** P-values are mostly above 0.05, indicating that the observed correlations are likely to have arisen by chance. This means that we cannot conclude that there is a causal relationship between poverty and crime based on these results.
- **Gender Differences:** Although there are some gender differences in the analyses, these differences are not statistically significant. That is, there is no significant difference between the results obtained for men and women.
- **Year-to-Year Comparison:** Analyses conducted for data from different years yielded similar results. This indicates that this relationship has not changed significantly over time.

### **Why might the results not meet our expectations?**

- **Other Factors:** There are many other factors that may affect the relationship between poverty and crime (e.g. education level, unemployment rate, social inequality, family structure). Since this analysis only looks at the relationship between poverty and crime, the impact of other factors is ignored.
- **Data Limitations:** The scope and quality of the data set used can affect the results. For example, the data set may not be large or representative enough.
- **Complex Relationship:** The relationship between poverty and crime may be non-linear or have a more complex structure. In this case, Pearson correlation analysis may not fully capture this relationship.
- **Insufficient Economic Indicators:** Only the poverty rate may have been taken into account. The impact of other economic indicators such as income inequality and unemployment rate on crime can also be analyzed.