

COVID-19 Impact To U.S. Economy



- 01** 2602190816 – Kimberly Kayla Dewi
- 02** 2602197305 – Callista Serena Ekaputri
- 03** 2602199582 – Fiona Maharani Nugraha
- 04** 2602203794 – Janice Ashley Kurniawan
- 05** 2602227775 – Keitaro Alexander Herman

Group 5 – LB09 – Data Mining and Visualization

Outline

01 Introduction

02 General Characteristics

03 Descriptive Statistics

04 Exploratory Visualization

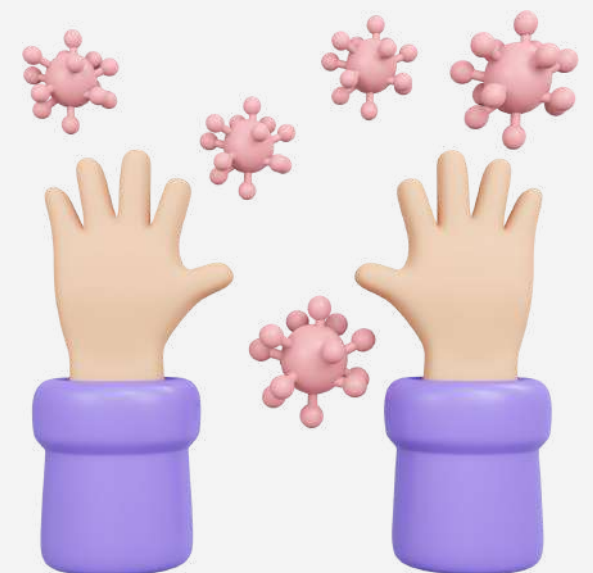
05 Explanatory Visualization

06 Predictive Model

07 Linear Regression

08 Discussion

09 Conclusion



01 Introduction

1.1 Prediction

--- How will the U.S. economy develop after COVID-19?

1.2 Problem Statement

--- How will U.S. economic developments occur in the future?



01 Introduction

1.2.1 Problem Statement

The spread of COVID-19 has had severe economic consequences. Therefore, it is necessary to carry out an **analysis regarding the impact of COVID-19 on the U.S. economy** based on Year, Gross Domestic Products, GDP per Capita, and Growth.

Knowing the effect of these variables on the economy can assist in **planning economic policies and improving people's welfare** during the COVID-19 pandemic.

That concludes our main problem, **how will U.S. economic developments occur in the future?**



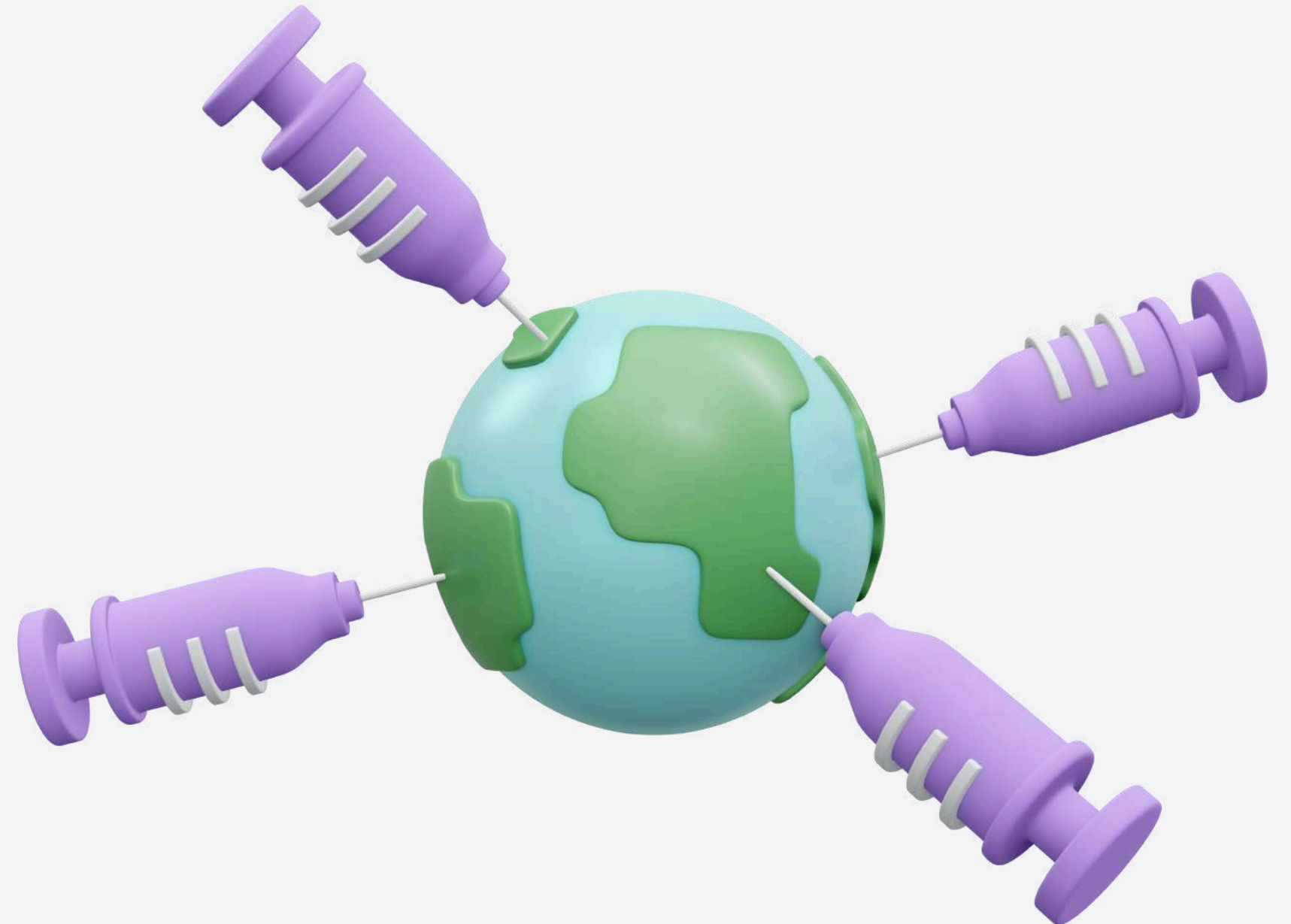
01 Introduction

1.3 Solution

--- We have to make predictions on how this economy will develop after COVID-19.

1.4 Objective

To understand the impact of COVID-19 on the U.S. economy. It is hoped that this analysis can provide helpful information for planning economic policies and improving people's welfare during the COVID-19 pandemic.





01 Introduction

1.5 Data Description

The data used is U.S. economic data taken from a trusted and valid source, **data.world**.

This data includes the value of the variable **year**, **Gross Domestic Products (gdp)**, **GDP per capita (gdp_per_capita)**, **growth**.

Data were obtained in the form of **tables and CSV files** which were then processed and analyzed with **RStudio**. This analysis will involve **Exploratory and Explanatory Data Analysis, Regression (Predictive Model), Correlation, and Data Visualization**.

02 General Characteristics

Head Data

year <dbl>	gdp <dbl>	gdp_per_capita <dbl>	growth <dbl>
2020	20893	63028	-0.0340
2019	21372	65095	0.0229
2018	20527	62805	0.0292
2017	19479	59915	0.0226
2016	18695	57867	0.0167
2015	18206	56763	0.0271



02 General Characteristics

Tail Data

year <dbl>	gdp <dbl>	gdp_per_capita <dbl>	growth <dbl>
1998	9062	32854	0.0448
1997	8577	31459	0.0445
1996	8073	29968	0.0377
1995	7639	28691	0.0268
1994	7287	27695	0.0403
1993	6858	26387	0.0275



02 General Characteristics

The data type for each column

```
      year      gdp  gdp_per_capita  
"numeric" "numeric" "numeric"  
      growth  
"numeric"
```

Data Characteristics

```
tibble [17 × 4] (S3: tbl_df/tbl/data.frame)  
$ year      : num [1:17] 2020 2019 2018 2017 2016 ...  
$ gdp       : num [1:17] 20893 21372 20527 19479 18695 ...  
$ gdp_per_capita: num [1:17] 63028 65095 62805 59915 57867 ...  
$ growth    : num [1:17] -0.034 0.0229 0.0292 0.0226 0.0167 0.0271 0.028 0.017 0.0095 0.0408 ...
```



02 General Characteristics

Data Attribute

```
$class
[1] "tbl_df"      "tbl"        "data.frame"

$row.names
[1]  1  2  3  4  5  6  7  8  9 10 11 12 13 15 16 17

$names
[1] "year"      "gdp"      "gdp_per_capita"
[4] "growth"
```



02 General Characteristics

Unique data in each column

year <dbl>	gdp <dbl>	gdp_per_capita <dbl>	growth <dbl>
2020	20893	63028	-0.0340
2019	21372	65095	0.0229
2018	20527	62805	0.0292
2017	19479	59915	0.0226
2016	18695	57867	0.0167
2015	18206	56763	0.0271
2003	11456	39490	0.0280
2002	10929	37998	0.0170
2001	10581	37134	0.0095
2000	10250	36330	0.0408



02 General Characteristics

Basic Summary

variable <chr>	type <chr>	levels <dbl>	topLevel <chr>
year	tbl_df	17	1993
gdp	tbl_df	17	6858
gdp_per_capita	tbl_df	17	26387
growth	tbl_df	17	-0.034

topCount <dbl>	topFrac <dbl>	missFreq <dbl>	missFrac <dbl>
1	0.059	0	0
1	0.059	0	0
1	0.059	0	0
1	0.059	0	0



02 General Characteristics

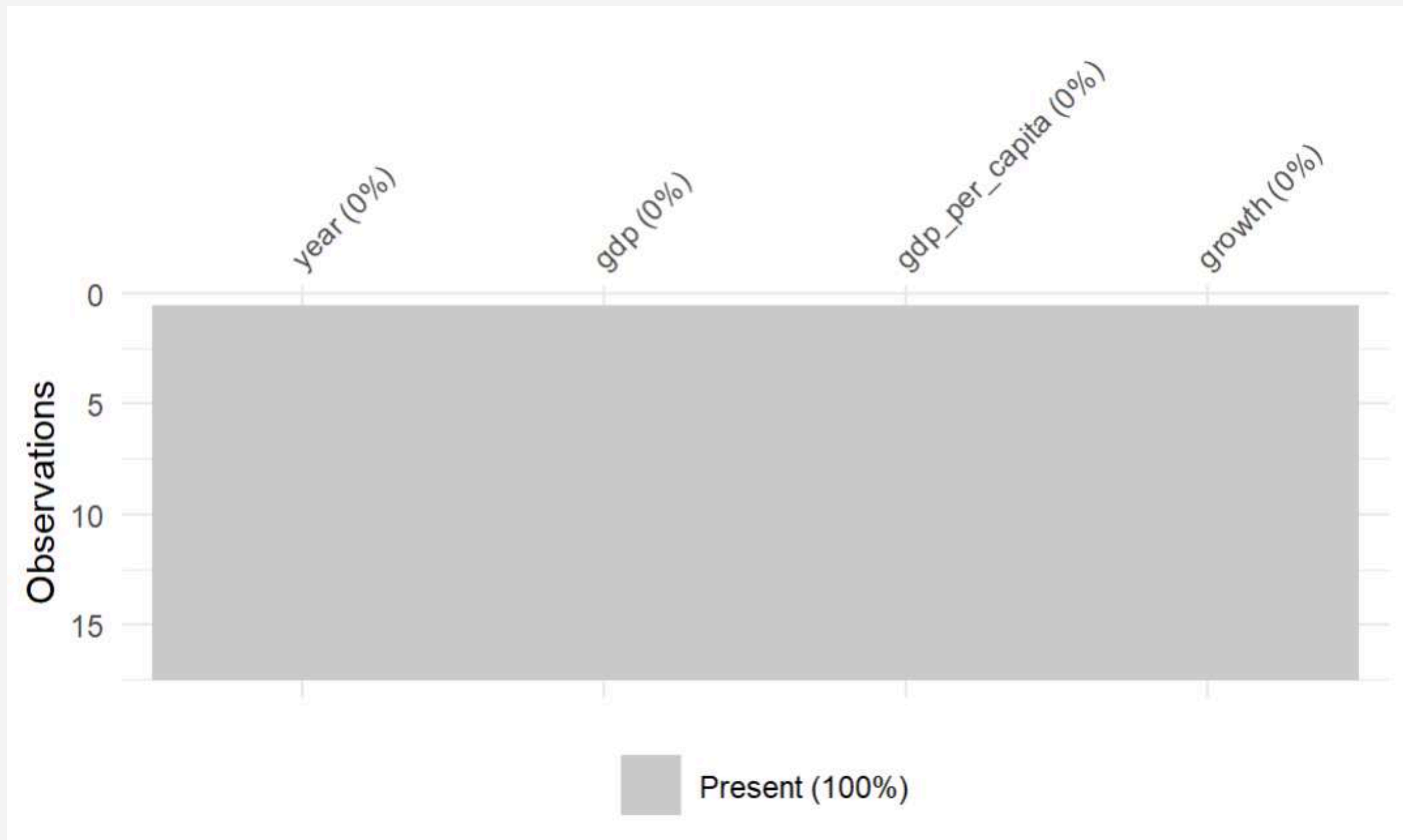
The amount of empty data per column

year	gdp	gdp_per_capita	growth
0	0	0	0



02 General Characteristics

Percentage of data completeness for each column



02 General Characteristics

Percentage of data completeness for each column

variable <chr>	type <chr>	levels <dbl>	topLevel <chr>	topCount <dbl>	
year	tbl_df	17	1993	1	
gdp	tbl_df	17	\$10,250.95B	1	
gdp_per_capita	tbl_df	17	26387	1	
growth	tbl_df	17	-0.034	1	

4 rows | 1-5 of 8 columns



03 Descriptive Statistics

Summary

year	gdp	gdp_per_capita
Min. :1993	Length:17	Min. :26387
1st Qu.:1997	Class :character	1st Qu.:31459
Median :2001	Mode :character	Median :37134
Mean :2005		Mean :42823
3rd Qu.:2016		3rd Qu.:57867
Max. :2020		Max. :65095

growth
Min. : -0.03400
1st Qu.: 0.02260
Median : 0.02750
Mean : 0.02643
3rd Qu.: 0.04030
Max. : 0.04790



03 Descriptive Statistics

Variance

	year	gdp
year	1.002353e+02	5.475771e+04
gdp	5.475771e+04	3.001968e+07
gdp_per_capita	1.432623e+05	7.850447e+07
growth	-1.073526e-01	-5.698717e+01

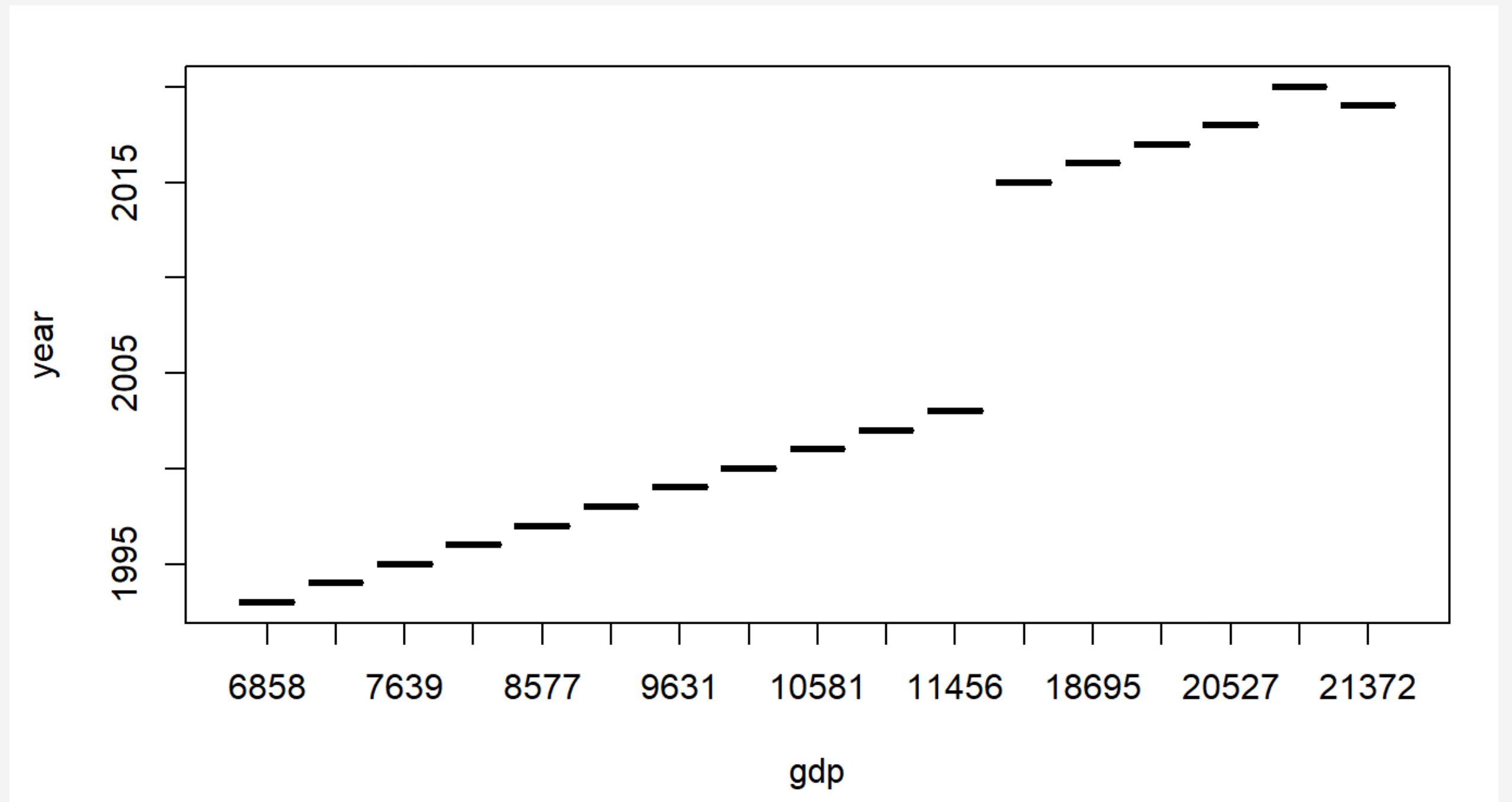
	gdp_per_capita	growth
year	1.432623e+05	-1.073526e-01
gdp	7.850447e+07	-5.698717e+01
gdp_per_capita	2.054229e+08	-1.467406e+02
growth	-1.467406e+02	3.632397e-04

Standard Deviation

year	gdp	gdp_per_capita	growth
1.001176e+01	5.479022e+03	1.433258e+04	1.905885e-02

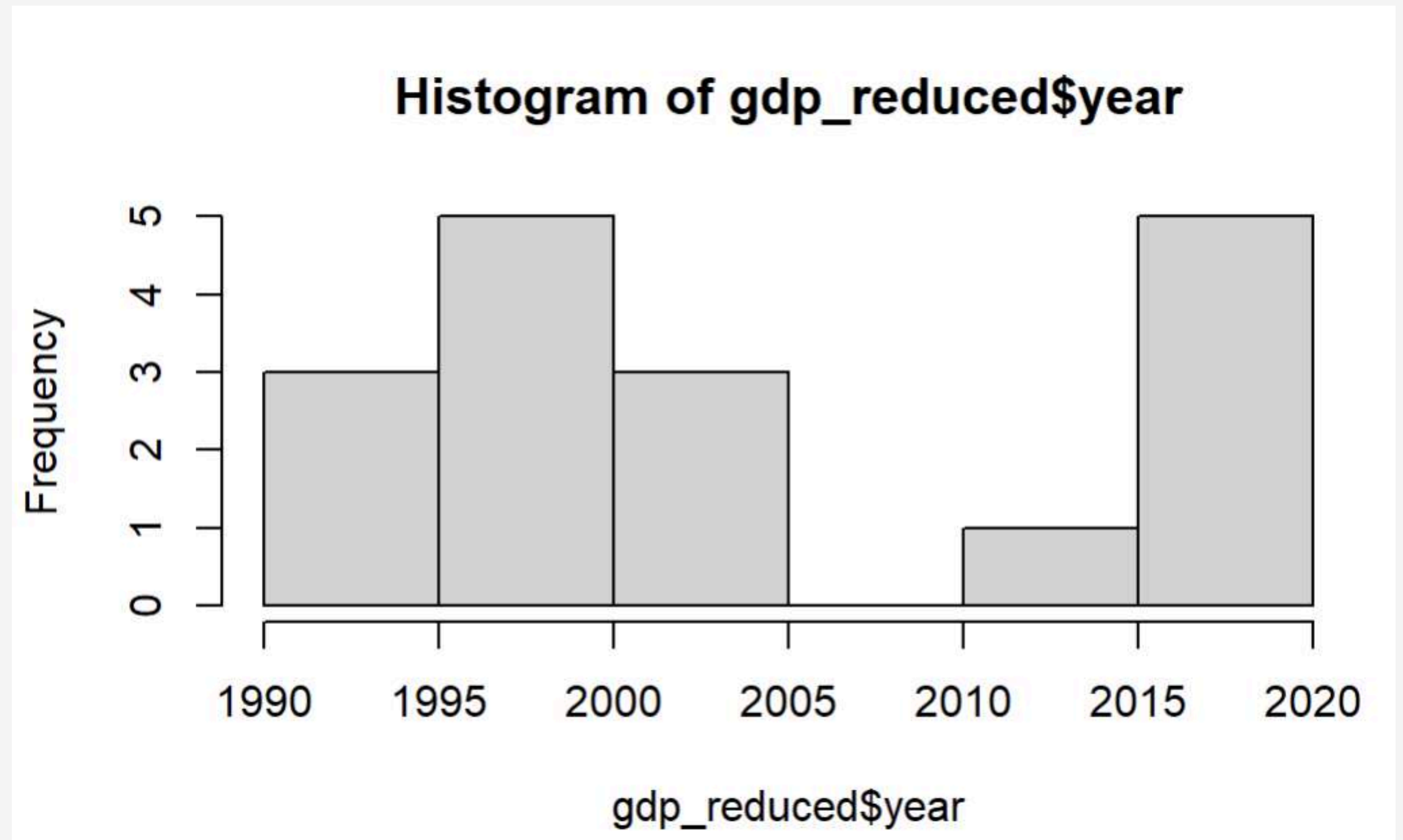
03 Descriptive Statistics

Data Anomalies/Outlier



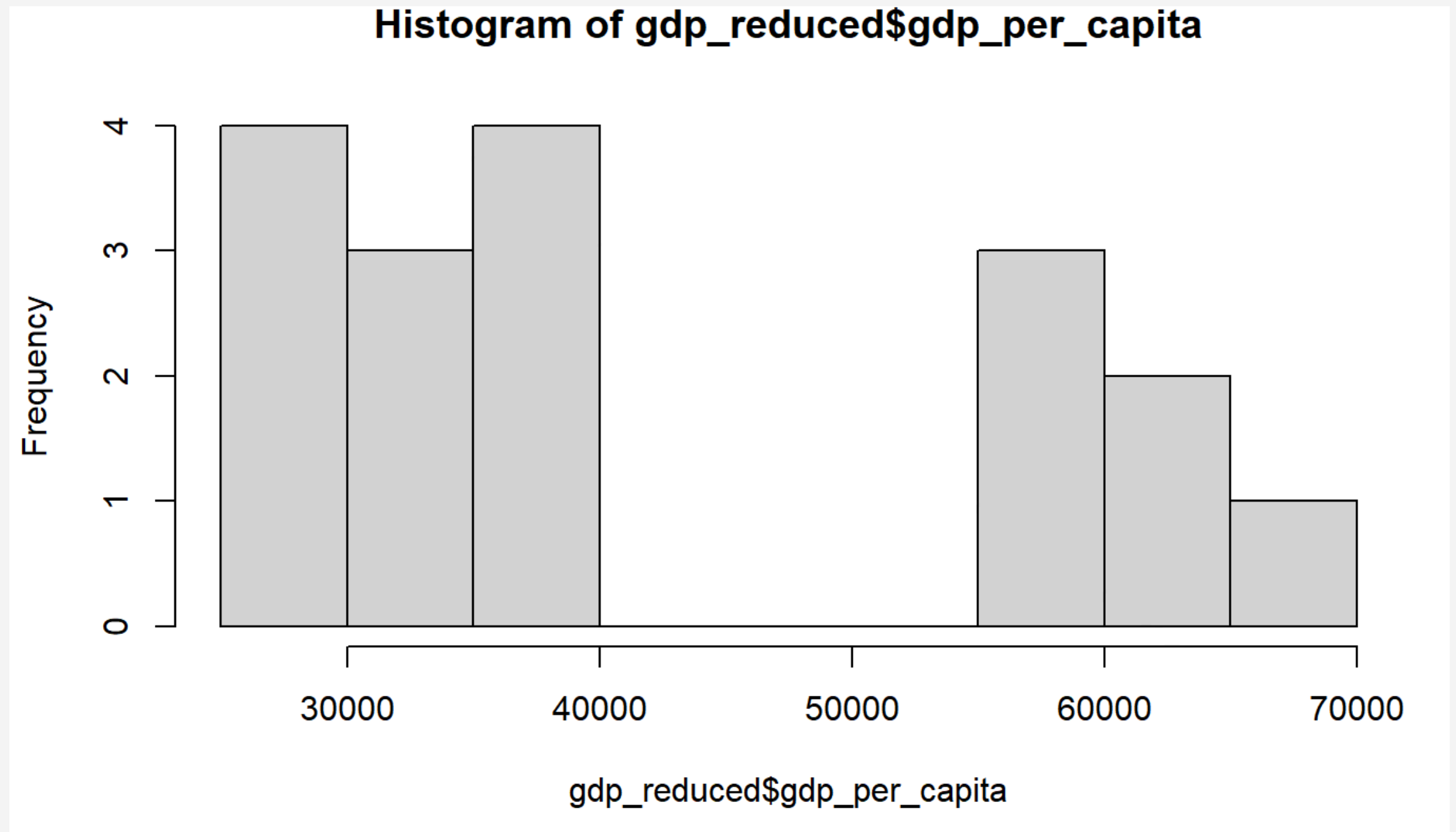
04 Exploratory Visualization

Histogram 1



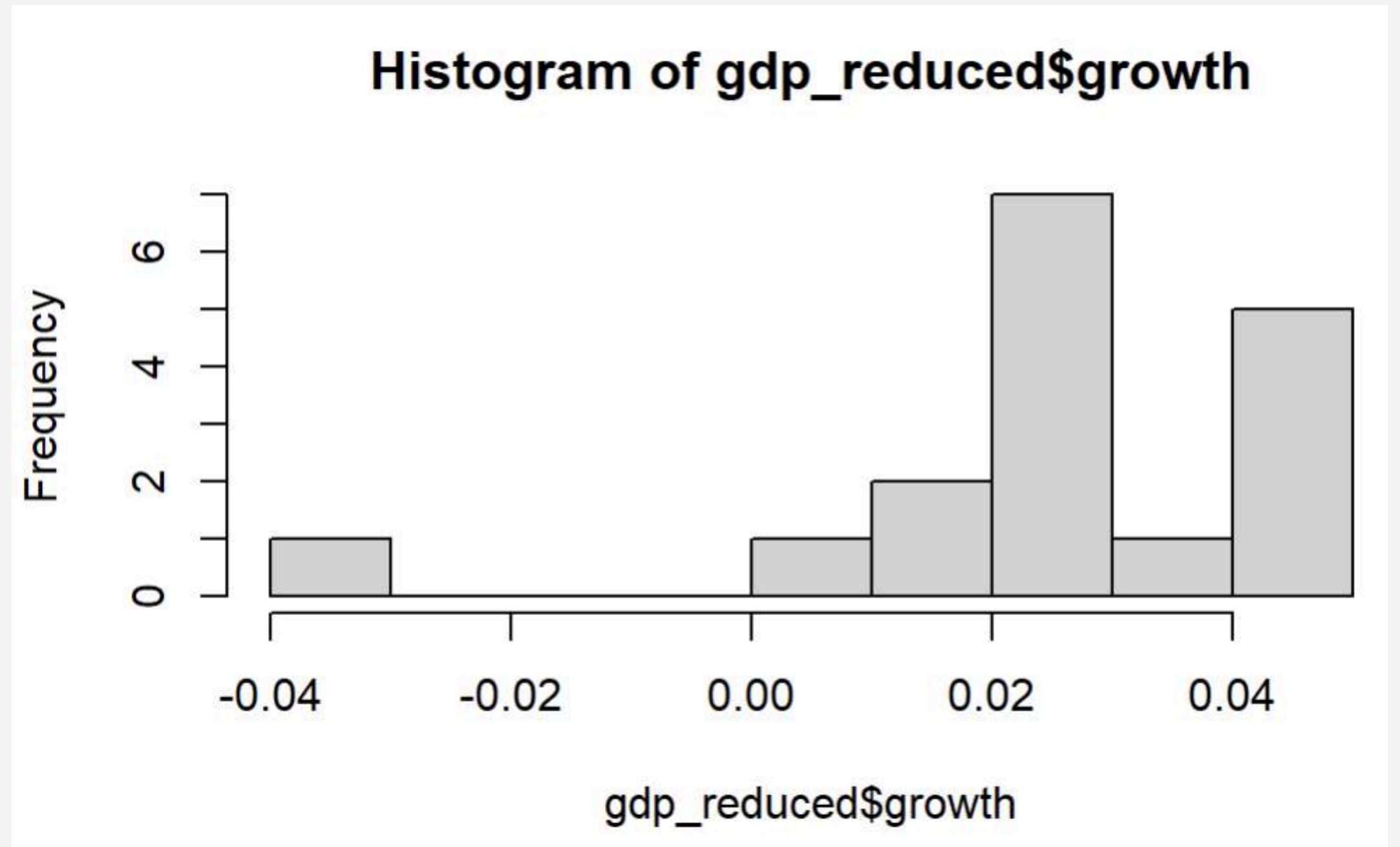
04 Exploratory Visualization

Histogram 2



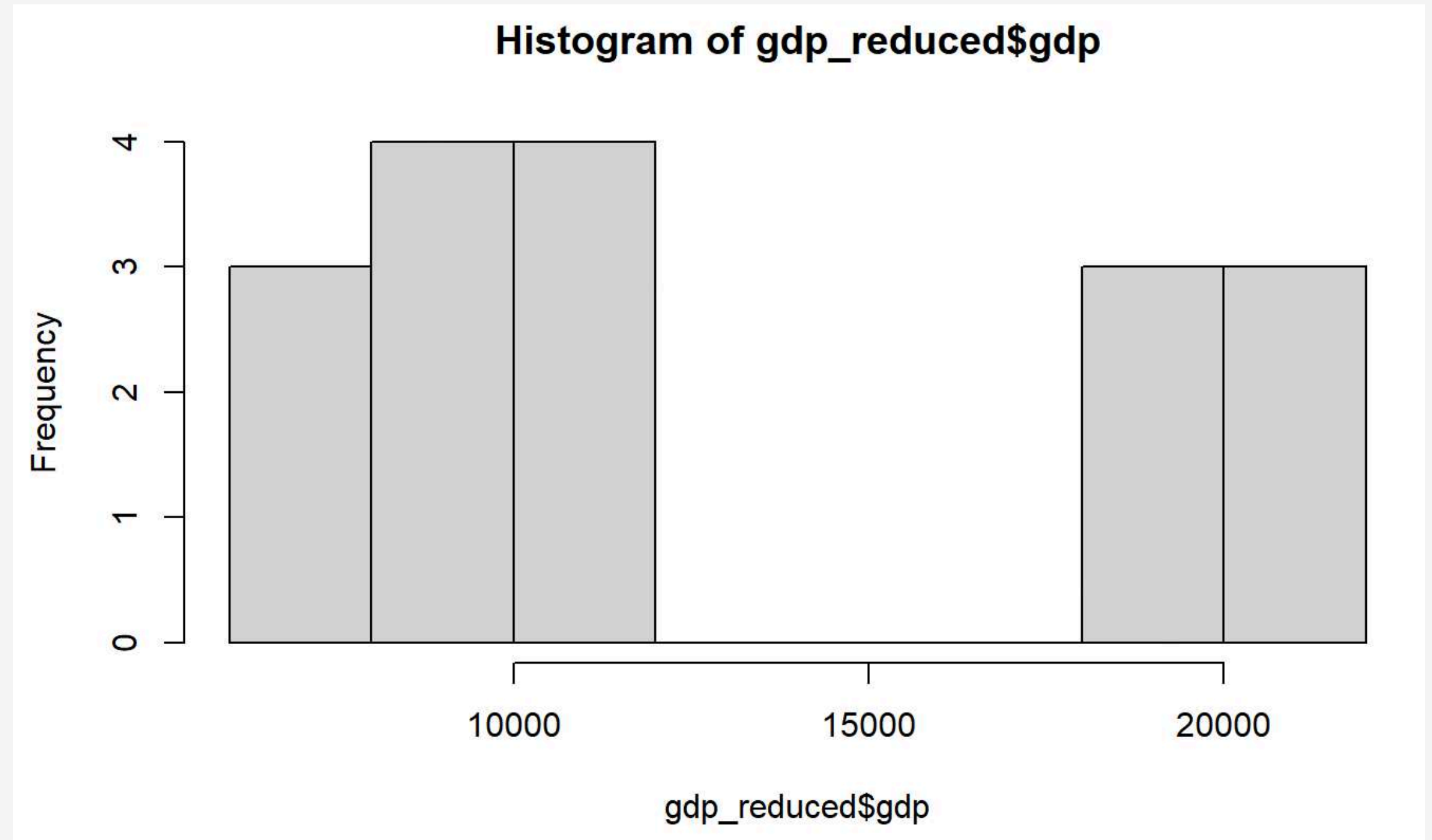
04 Exploratory Visualization

Histogram 3

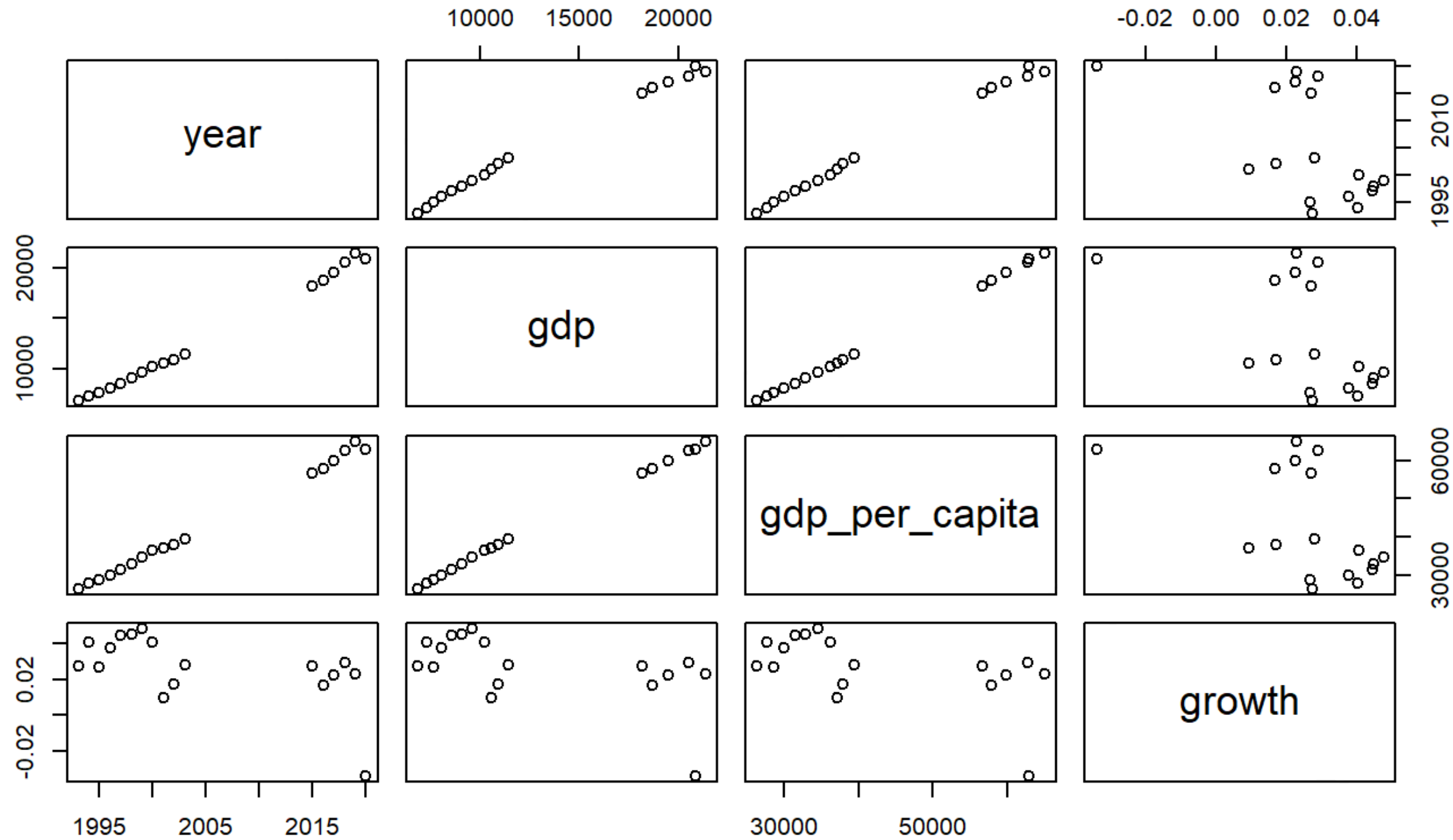


04 Exploratory Visualization

Histogram 4



05 Explanatory Visualization



06 Predictive Model

Full Summary Logistic for gdp_per_capita

```
call:
glm(formula = gdp ~ ., family = "poisson", data = gdpTrain)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.9912  -1.6298  -0.3952   0.7722   4.9350

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -8.201e+01  8.885e+00  -9.231  < 2e-16 ***
year         4.562e-02  4.494e-03  10.152  < 2e-16 ***
gdp_per_capita -2.531e-06  3.025e-06  -0.837    0.403
growth       1.183e+00  1.735e-01   6.820  9.12e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 26988.7  on 11  degrees of freedom
Residual deviance:   73.9  on  8  degrees of freedom
AIC: 217.86

Number of Fisher Scoring iterations: 3
```



06 Predictive Model

Exponentiating the Logistic Full Outcome

(Intercept)	year	gdp_per_capita	growth
2.410675e-36	1.046682e+00	9.999975e-01	3.264810e+00



06 Predictive Model

Exponentiating the Logistic Ref Outcome

```
call:
glm(formula = gdp ~ growth, family = "poisson", data = gdpTrain)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-57.77	-37.83	-17.55	36.71	55.62

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	9.783270	0.003161	3095.32	<2e-16 ***
growth	-9.207385	0.101617	-90.61	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 26989 on 11 degrees of freedom
Residual deviance: 19602 on 10 degrees of freedom
AIC: 19742

Number of Fisher Scoring iterations: 4



06 Predictive Model

Exponentiating the Logistic Ref Outcome

(Intercept)	growth
$1.773455e+04$	$1.002960e-04$



06 Predictive Model

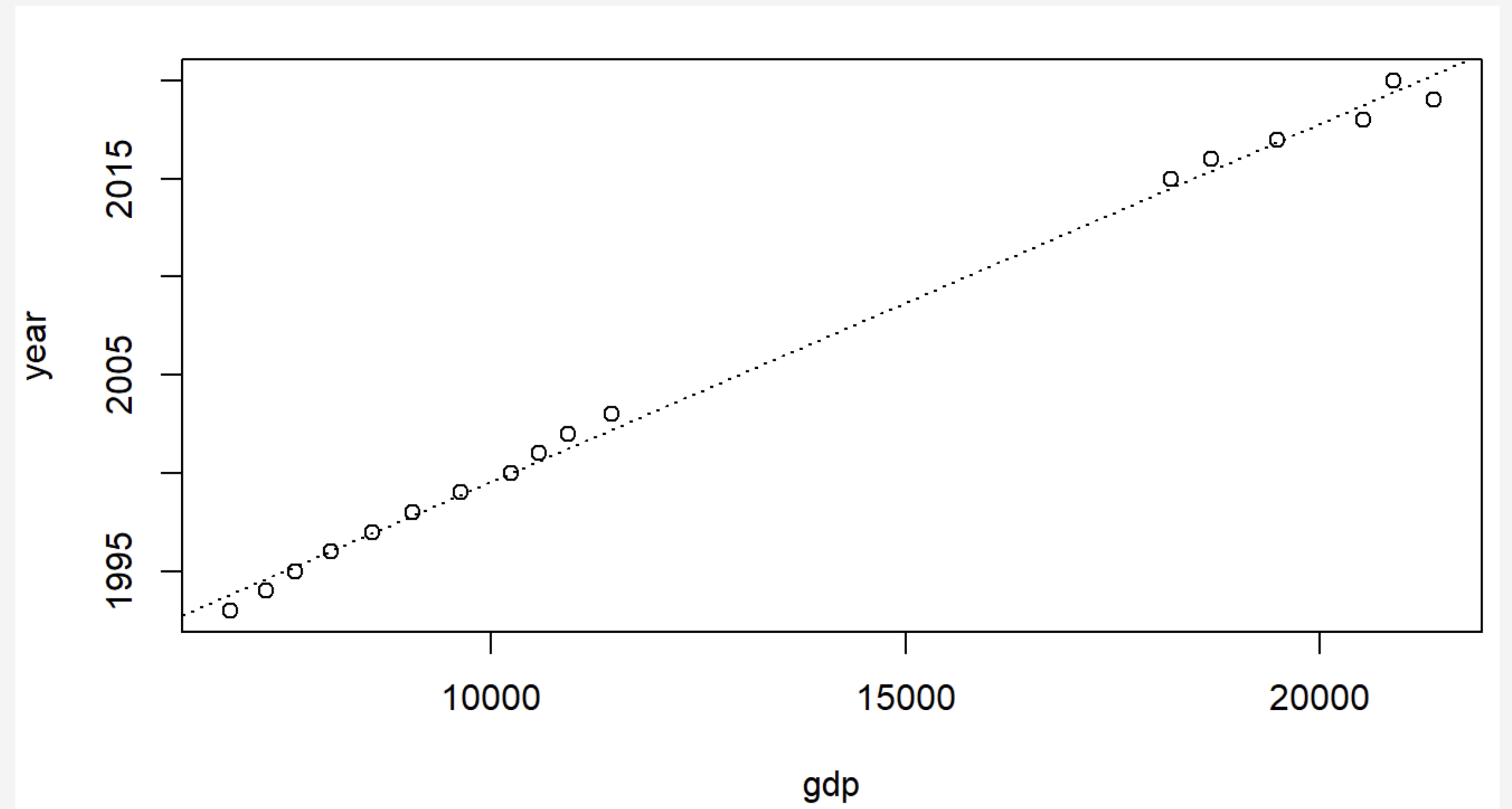
Accuracy Model

```
"Accuracy model 1 is 0.2"  
[1] "Accuracy model 2 is 0.2"
```



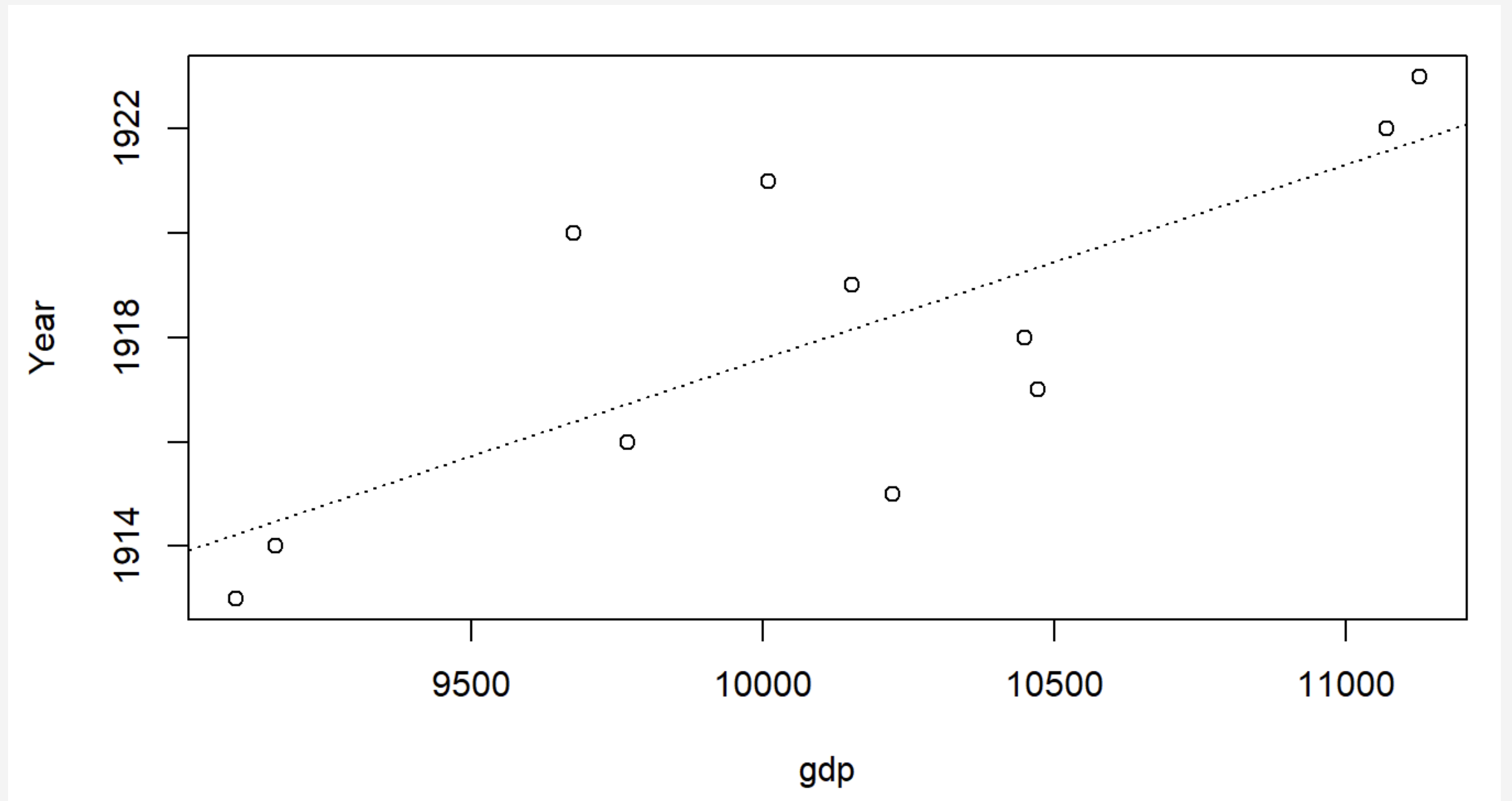
07 Linear Regression

GDP in America in 1993-2020



07 Linear Regression

GDP in America in 1913-1923



08 Discussion

In the predictive model, in 1998–2000 there was a decline in the economy as happened with the dataset above, which will have the same behavior, namely **returning to recover in 2 years.**

At the very end wherein 2020, where the initial outbreak of COVID-19 also experienced **the same decline as the previous year.**

This forms a **habit in economic data** where if there is a crisis related to the economy, it will recover within 2 years.



09 Conclusion

Based on data obtained from 100 years ago, if there was a crisis that shook a country enough, recovery would occur in approximately 2 years.

This occurs from the evidence of the Spanish flu 100 years ago and the economic crisis in 1998.

According to predictions, the U.S. economy will recover in mid-2023.



THANK

YOU

