“ANALYSING IMPACT OF ECONOMIC GROWTH ON INCOME INEQUALITY”

A PROJECT REPORT

SUBMITTED FOR THE PARTIAL FULFILMENT OF

MASTER’S DEGREE IN STATISTICS

SESSION : 2019-2020

UNDER THE SUPERVISION OF:

DR . NIRPEKSH KUMAR

Department of Statistics

Institute of Science, B.H.U

SUBMITTED BY:

MAYANK SHANKAR DAYAL RAI

M.Sc. Statistics 4th Sem

Roll No- 18420STA035

DEPARTMENT OF STATISTICS

BANARAS HINDU UNIVERSITY

VARANASI 221005

CERTIFICATE

This is to certify that the secondary data given in this project has been tabulated, analyzed and presented by "MAYANK SHANKAR DAYAL RAI", student of M.Sc. STATISTICS under my supervision and guidance in the session 2019-2020 .

The title of project is :

"ANALYSING IMPACT OF ECONOMIC GROWTH ON INCOME INEQUALITY"

DR. NIRPEKSH KUMAR

Department of Statistics

Banaras Hindu University

Varanasi - 221005

DATE:………………….

Signature……………….

ACKNOWLEDGEMENT

It gives me immense pleasure to express my profound gratitude, gratefulness and indebtedness to reverend teacher and supervisor DR.NIRPEKSH KUMAR ,Department of Statistics, Institute of Science, Banaras Hindu University, for his untiring help, constant encouragement, worthy supervision without which it would not have been possible for me to complete this project work. The guidance and valuable criticism that I received from him, during the entire period of this work, has been a great help in the completion of this work.

I also thanks to the Friends who co-operated me in Project Work Analysis.

At last I would thank the Department of Statistics, B.H.U. which gave me the opportunity for this project work and for which I shall ever remain grateful.

Date:

MAYANK SHANKAR DAYAL RAI

M.Sc. Statistics 4th Semester

Department of Statistics

Institute of Science

Banaras Hindu University

CONTENTS

INTRODUCTION

Economic growth is the increase in the inflation-adjusted market value of the goods and services produced by an economy over time. It is conventionally measured as the percent rate of increase in real gross domestic product, or real GDP. Growth is usually calculated in real terms - i.e., inflation-adjusted terms – to eliminate the distorting effect of inflation on the price of goods produced. Measurement of economic growth uses national income accounting.[2] Since economic growth is measured as the annual percent change of gross domestic product (GDP), it has all the advantages and drawbacks of that measure. The economic growth rates of nations is commonly compared using the ratio of the GDP to population or per-capita income.

The "rate of economic growth" refers to the geometric annual rate of growth in GDP between the first and the last year over a period of time. Implicitly, this growth rate is the trend in the average level of GDP over the period, which implicitly ignores the fluctuations in the GDP around this trend.

An increase in economic growth caused by more efficient use of inputs (such as labor productivity, physical capital, energy or materials) is referred to as intensive growth. GDP growth caused only by increases in the amount of inputs available for use (increased population, new territory) is called extensive growth.

Income Inequality is the difference found in various measures of economic wellbeing among individuals in a group, among groups in a population, or among countries. Economic inequality sometimes refers to Economic Inequality, Wealth Inequality, or the Wealth Gap. Economists generally focus on economic disparity in three metrics: wealth, income, and consumption.[1] The issue of economic inequality is relevant to notions of equity, equality of outcome, and equality of opportunity.[2]

Economic inequality varies between societies, historical periods, economic structures and systems. The term can refer to cross-sectional distribution of income or wealth at any particular period, or to changes of income and wealth over longer periods of time.[3] There are various numerical indices for measuring economic inequality. A widely used index is the Gini coefficient, but there are also many other methods.

REGRESSION ANALYSIS

In statistical modeling, Regression Analysis is a set of statistical processes for estimating the relationships among variables. It includes many techniques for modeling and analyzing several variables, when the focus is on the relationship between a dependent variable and one or more independent variables (or 'predictors'). More specifically, regression analysis helps one understand how the typical value of the dependent variable (or 'criterion variable') changes when any one of the independent variables is varied, while the other independent variables are held fixed.Most commonly, regression analysis estimates the conditional expectation of the dependent variable given the independent variables – that is, the average value of the dependent variable when the independent variables are fixed. Less commonly, the focus is on a quantile, or other location parameter of the conditional distribution of the dependent variable given the independent variables. In all cases, a function of the independent variables called the Regression Function is to be estimated. In regression analysis, it is also of interest to characterize the variation of the dependent variable around the prediction of the regression function using a probability distribution. A related but distinct approach is Necessary Condition Analysis[1] (NCA), which estimates the maximum (rather than average) value of the dependent variable for a given value of the independent variable (ceiling line rather than central line) in order to identify what value of the independent variable is necessary but not sufficient for a given value of the dependent variable.Regression analysis is widely used for prediction and forecasting, where its use has substantial overlap with the field of machine learning. Regression analysis is also used to understand which among the independent variables are related to the dependent variable, and to explore the forms of these relationships. In restricted circumstances, regression analysis can be used to infer causal relationships between the independent and dependent variables. However this can lead to illusions or false relationships, so caution is advisable; for example, correlation does not prove causation.Many techniques for carrying out regression analysis have been developed. Familiar methods such as linear regression and ordinary least squares regression are parametric, in that the regression function is defined in terms of a finite number of unknown parameters that are estimated from the data. Nonparametric regression refers to techniques that allow the regression function to lie in a specified set of functions, which may be infinitedimensional.The performance of regression analysis methods in practice depends on the form of the data generating process, and how it relates to the regression approach being used. Since the true form of the data-generating process is generally not known, regression analysis often depends to some extent on making assumptions about this process. These assumptions are sometimes testable if a sufficient quantity of data is available. Regression models for prediction are often useful even when the assumptions are moderately violated, although they may not perform optimally. However, in many applications, especially with small effects or questions of causality based on observational data, regression methods can give misleading results. In a narrower sense, regression may refer specifically to the estimation of continuous response (dependent) variables, as opposed to the discrete response variables used in classification.[5] The case of a continuous dependent variable may be more specifically referred to as metric regression to distinguish it from related problems.

REGRESSIONS MODELS

Regression models involve the following parameters and variables:

- The Unknown Parameters denoted as β ,which may represent a Scalar or a Vector.
- The Independent Variables X.
- The Dependent Variable Y

SIMPLE LINEAR REGRESSION MODEL

We consider the modeling between the dependent and one independent variable. When there is only one independent variable in the linear regression model, the model is generally termed as simple linear regression model.

MULTIPLE LINEAR REGRESSION MODEL

We consider the problem of regression when study variable depends on more than one explanatory or independent variables, called as multiple linear regression model. This model generalizes the simple linear regression in two ways. It allows the mean function E(y) to depend on more than one explanatory variables and to have shapes other than straight lines, although it does not allow for arbitrary shapes.

STEPS IN REGRESSION ANALYSIS

Regression analysis includes the following steps:

- Statement of the problem under consideration
- Choice of relevant variables
- Collection of data on relevant variables
- Specification of model
- Choice of method for fitting the data
- Fitting of model
- Model validation and criticism
- Using the chosen model(s) for the solution of the posed problem.

ASSUMPTIONS

Classical assumptions for regression analysis include:

- The sample is representative of the population for the inference prediction.
- The error is a random variable with a mean of zero conditional on the explanatory variables.
- The independent variables are measured with no error.
- The independent variables (predictors) are linearly independent, i.e. it is not possible to express any predictor as a linear combination of the others.
- The errors are uncorrelated, that is, the variance–covariance matrix of the errors is diagonal and each non-zero element is the variance of the error.
- The variance of the error is constant across observations (homoscedasticity).

EXPLORATORY DATA ANALYSIS

In statistics, Exploratory Data Analysis (EDA) is an approach to analyzing data sets to summarize their main characteristics, often with visual methods. A statistical model can be used or not, but primarily EDA is for seeing what the data can tell us beyond the formal modeling or hypothesis testing task. Exploratory data analysis was promoted by John Tukey to encourage statisticians to explore the data, and possibly formulate hypotheses that could lead to new data collection and experiments. EDA is different from initial data analysis (IDA),[1] which focuses more narrowly on checking assumptions required for model fitting and hypothesis testing, and handling missing values and making transformations of variables as needed. EDA encompasses IDA.

The Purpose of Exploratory Data Analysis is to:

- Check for missing data and other mistakes.
- Gain maximum insight into the data set and its underlying structure.
- Uncover a parsimonious model, one which explains the data with a minimum number of predictor variables.
- Check assumptions associated with any model fitting or hypothesis test.
- Create a list of outliers or other anomalies.
- Find parameter estimates and their associated confidence intervals or margins of error.
- Identify the most influential variables.

TYPES OF EXPLORATORY DATA ANALYSIS

EDA falls into four main areas:

- Univariate non-graphical — looking at one variable of interest, like age, height, income level etc.
- Univariate graphical.
- Multivariate non-graphical — analysis of multiple variables at the same time.
- Multivariate graphical.

OBJECTIVE

To Apply Regression Analysis to our Data using R for Studying the following :

- To Study the Relationship Between Income Inequality and Economic Growth .
- To Study the Relationship Between Gross Savings and Growth of Gross Domestic Product.
- To Study the Relationship Between Fertility Rate and Growth of Gross Domestic Product.
- To Study the Relationship Between Unemployment Rate and Growth of Gross Domestic Product.
- To Study the Relationship Between Mean School Years and Growth of Gross Domestic Product.

ABOUT DATA

We chose the Gini coefficient (pre-tax) for the explanatory variable (x) in our simple regression line. The Gini coefficient was chosen for this model because it is a common measure of income inequality across many countries that represents the income distribution of a country's residents, where 0 represents perfect equality and 100 represents max inequality, and is recognized and used in much of the literature. Annual growth percentage of gross domestic product (GDP) was the dependent variable (y). The Gini coefficient and GDP growth datasets in this paper were obtained from the World Bank's Development Research Group (World Bank, 2011). We chose to regress GDP growth on the Gini coefficient because most of the literature we referenced found income inequality to have a more marked effect on GDP growth than GDP growth on income inequality. Our ultimate objective was to find the relationship between income inequality and economic growth. However, there are numerous variables that may affect economic growth, including urbanization ratio, population growth rate, financial development (M2/GDP), openness (export/GDP), etc (Li and Zou, 1998). In order to better understand and analyze the effect of income inequality on GDP growth, we controlled for other factors that had the most significant impacts on economic growth in an economy. These variables were gross savings, unemployment rate, education (mean school years), and fertility rate. Gross savings (World Bank, 2011) is one of the most common indicators of the growth of a country because it reflects the country's ability to consume and save. Fertility rate was included because research has shown that lower fertility rates lead to economic growth. Unemployment rate (World Bank, 2011) represents the long term unemployment rate, or natural rate of unemployment, in a country. Unemployment rate is an obvious indicator of a country's economic well-being. The mean school years are also expected to have an impact on economic growth. The more educated a country, the more growth is to be expected because of the capacity for highskilled laborers. Finally, a dummy variable was used to measure if the level of development of a country would affect their economic growth. These two categories (developed and developing) were classified according to the World Bank classification system.

The Description of Various Variables is given below-

VARIABLE	DESCRIPTION
Grgdp	Growth of Gross Domestic Product
Gini	Gini Coefficient (Measure of Inequality)

gsav	Gross Savings
fertil	Fertility Rate
unemp	Unemployment Rate
educ	Mean School Years

REGRESSION ANALYSIS
Data Cleaning

Load Data

First of all we copy our dataset in to our working directory and then use the following R code to read it.

CODE-

```
data<- read.csv('economic_data')
```

Removing Missing Values

Now we Remove the Missing Values from our Data. Firstly we extract each Variable from our Data and convert it in to Numeric as it is stored as Character Type.

CODE-
```
grgdp=datas$grgdp

grgdp=as.numeric(grgdp)

fertile=datas$fertile

fertile=as.numeric(fertile)

gini=datas$gini gini=as.numeric(gini)

educ=datas$educ

educ=as.numeric(educ)

unemp=datas$unemp

unemp=as.numeric(unemp)

gsav=datas$gsav gsav=as.numeric(gsav)
```

Now to Remove the Missing Values from Each Variables .

CODE-
```
grgdp1=na.omit(grgdp)

fertile1=na.omit(fertile)

gini1=na.omit(gini)

educ1=na.omit(educ)
```

```
unemp1=na.omit(unemp)

gsav1=na.omit(gsav)
```

Summary Statistics

Now we find the Summary of each variables (Mean , Minimum , Maximum , Standard Deviation) to know more about them. We use the following R code.

CODE-

```
summary(grgdp1)

dim(as.matrix(grgdp1)) sd(grgdp1)

summary(fertile1)

dim(as.matrix(fertile1)) sd(fertile1)

summary(gini1) dim(as.matrix(gini1))

sd(gini1)


summary(educ1)

dim(as.matrix(educ1)) sd(educ1)

summary(unemp1)

dim(as.matrix(unemp1)) sd(unemp1)

summary(gsav1)

dim(as.matrix(gsav1)) sd(gsav1)
```

Table below Shows the Summary Statistics for the data. This study was conducted using 225 countries. Because a country's economy can regress, the fact that the minimum of grgdp is a negative number is not a huge concern.

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
grgdp1	184	3.909663	3.158448	-7.304	12.615
gini1	105	38.0155	8.6441	24.70333	63.38
gsav1	160	21.36649	12.6191	-11.5887	60.00212
fertil1	184	2.876358	1.4462	1.205333	7.655
unemp1	170	8.658712	5.921921	0.3	31.46667
educ1	103	8.9109	2.980666	.055963	13.72269

Finding Correlation Among Variables

Now we find the Correlation between each pair of variables to study which Independent Variable is Highly Correlated with Dependent variable grgdp and to Check for presence of Multicollinearity by studying the Correlation between Independent Variables. CODE-

```
new=data.frame(grgdp,fertile,educ,unemp,gsav,gini) new=na.omit(new)

cor(new)
```

Correlation Matrix

	grgdp	gini	gsav	fertil	unemp	educ
grgdp	1.0000					
Gini	0.3943	1.0000				

gsav	0.3769	-0.0118	1.0000			
fertil	0.4549	0.3233	-0.0422	1.0000		
unemp	-0.3815	0.0274	-0.3224	-0.2558	1.0000	
educ	-0.4028	-0.4369	0.1541	-0.6355	0.1802	1.0000

INTERPRETATION

The above Correlation Matrix shows that there is No Perfect Collinearity Between any of the Independent Variables.

Exploratory Data Analysis

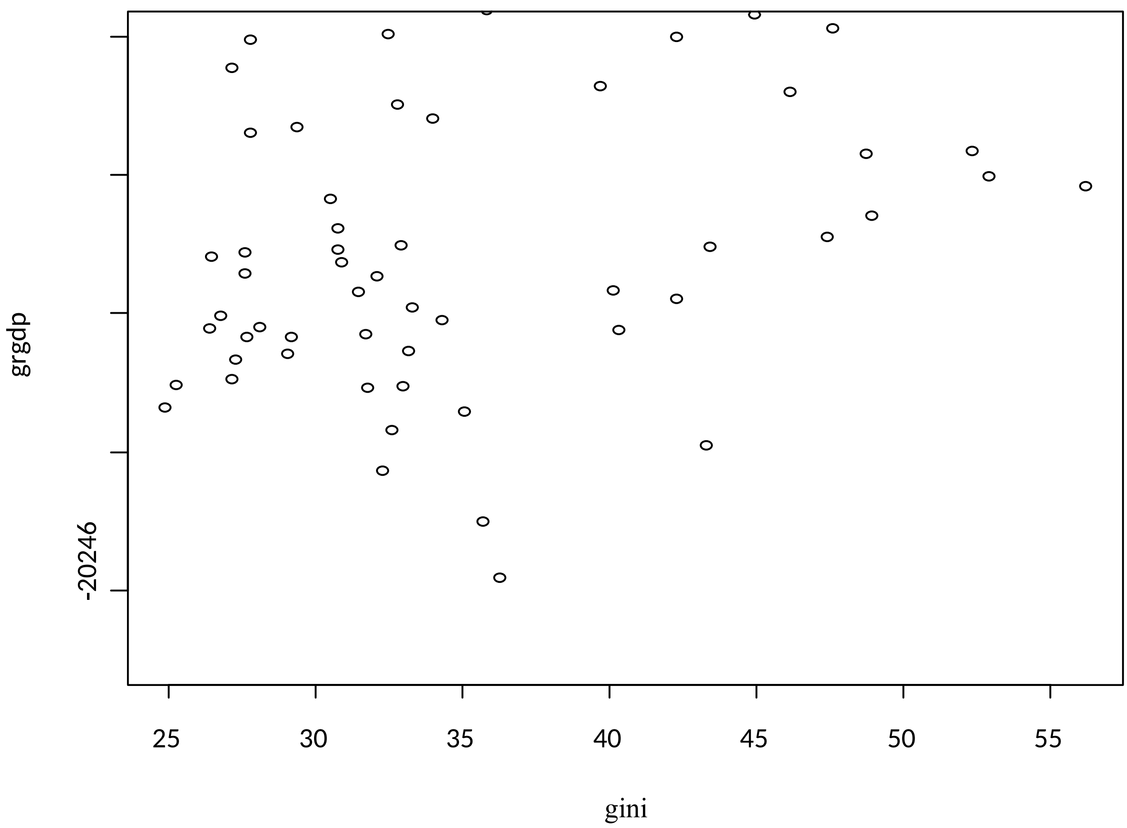
Now we draw some graphs between independent variables and dependent variable to study the relationship between them.

grgdp Vs gini

CODE-

```
plot(gini,grgdp,ylim = c(-3,6),main = 'Scatter Plot of
GDP Growth Vs Gini Coefficient')
```

Scatter Plot of GDP Growth Vs Gini Coefficient

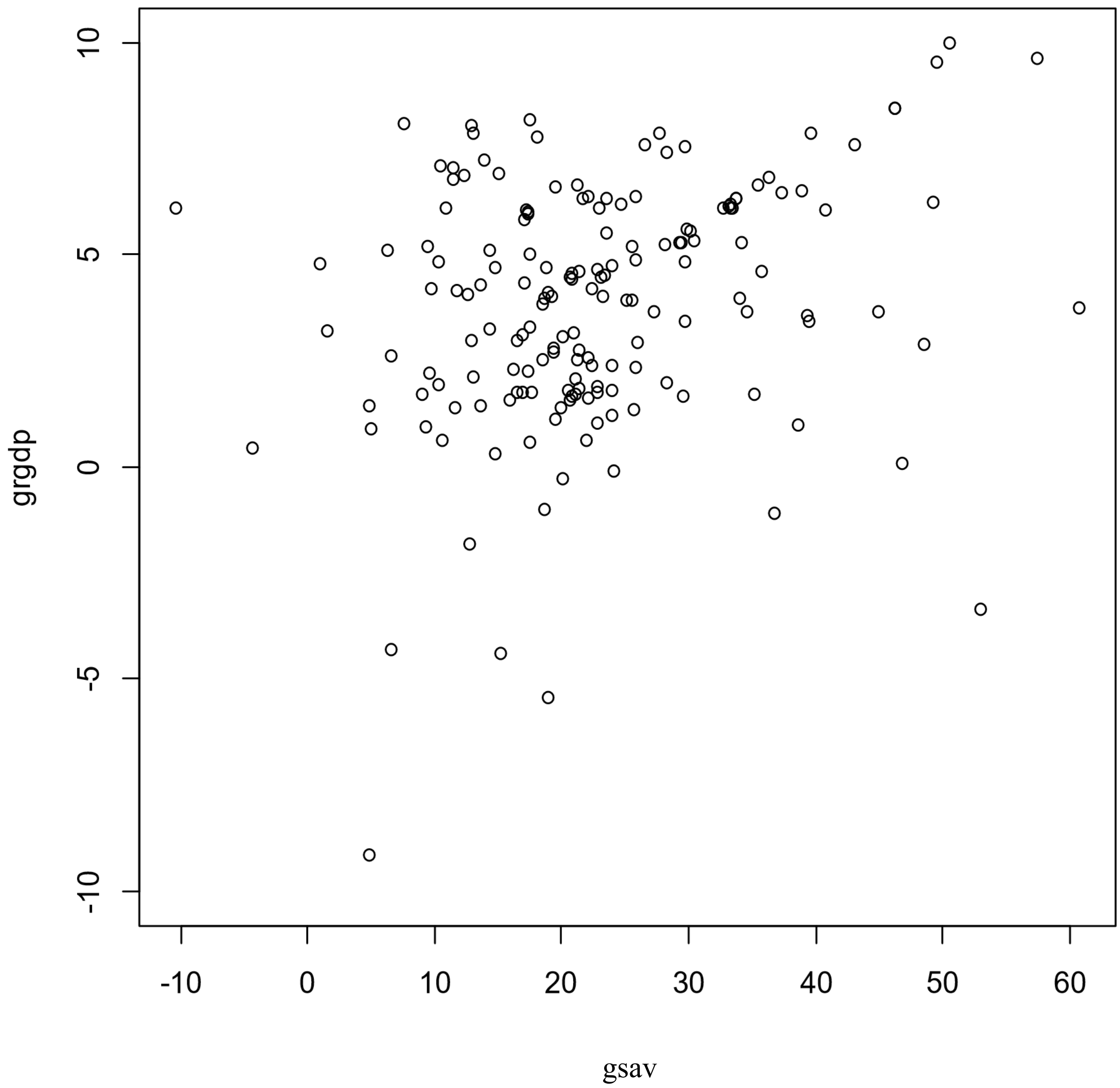


grgdp Vs gsav

CODE-

```
plot(gsav,grgdp,ylim= c(-10,10),main = 'Scatter Plot of
 GDP Growth Vs Gross Savings Rate')
```

Scatter Plot of GDP Growth Vs Gross Savings Rate

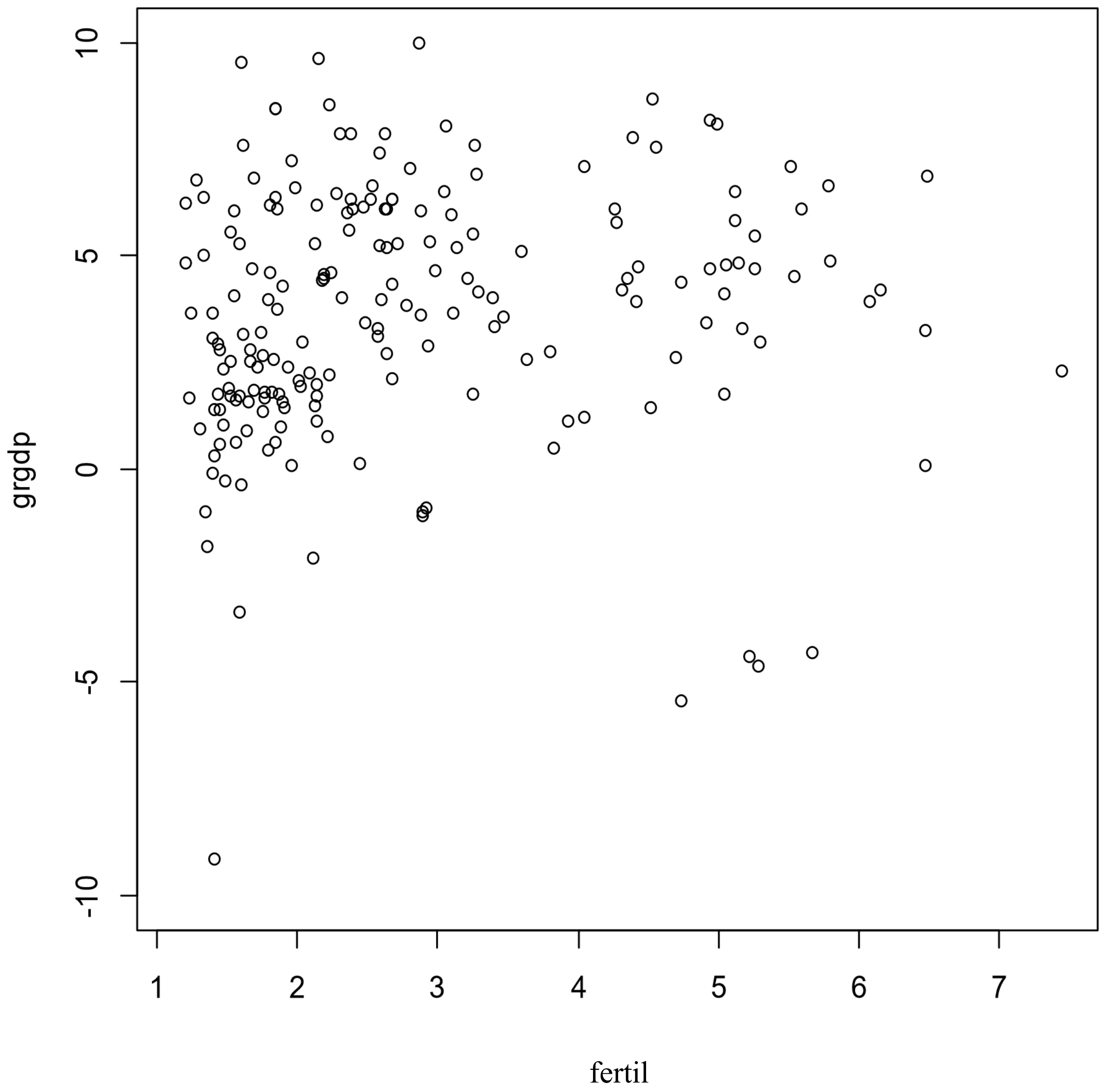


grgdp Vs fertil

CODE-

```
plot(fertil,grgdp,ylim= c(-10,10),main = 'Scatter Plot of GDP Growth Vs Fertility Rate')
```

Scatter Plot of GDP Growth Vs Fertility Rate

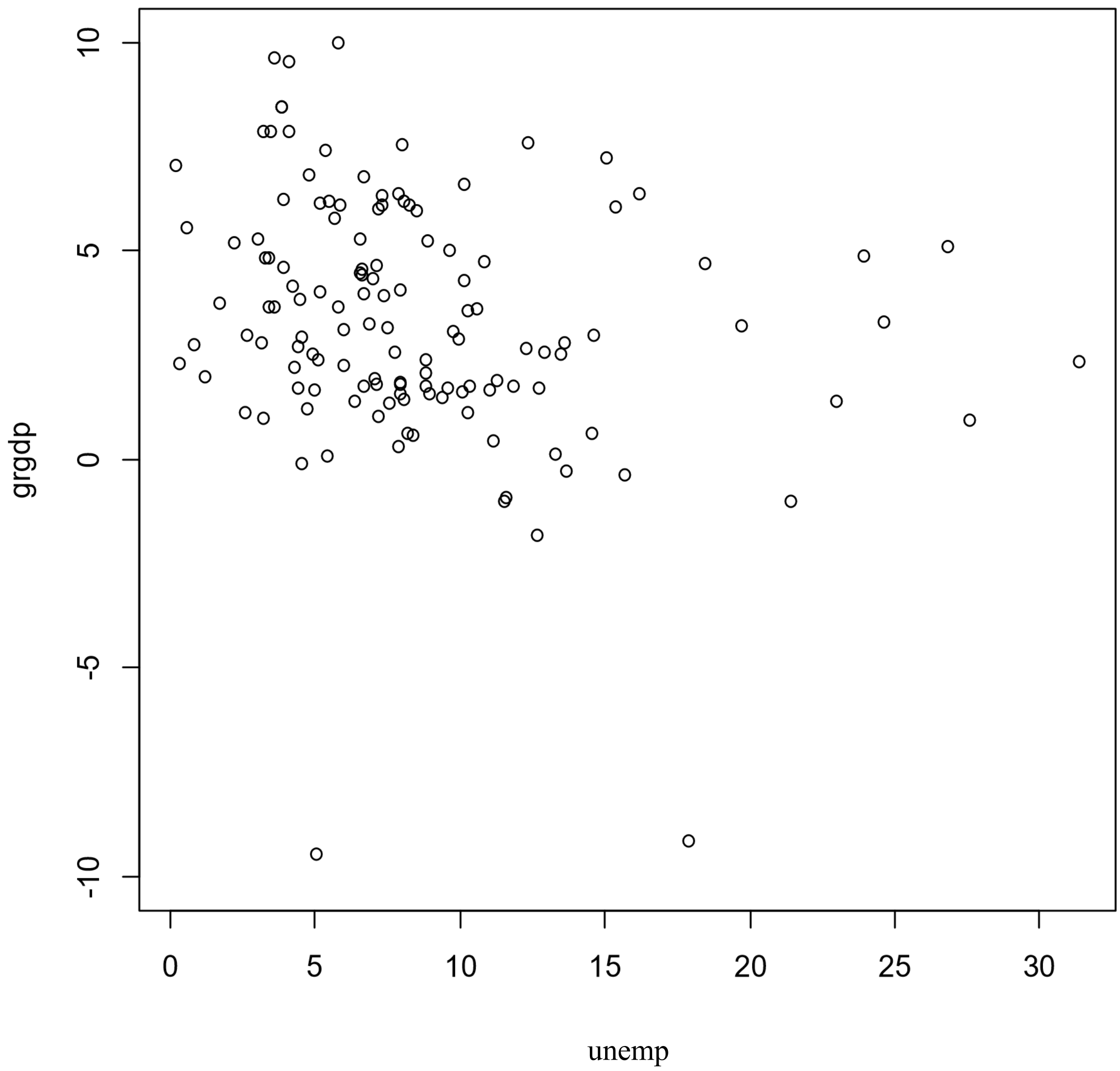


grgdp Vs unemp

CODE-

```
plot(unemp,grgdp,ylim=c(-10,10),main = 'Scatter Plot of
 GDP Growth Vs Unemployment Rate’)
```

Scatter Plot of GDP Growth Vs Unnempoyment Rate



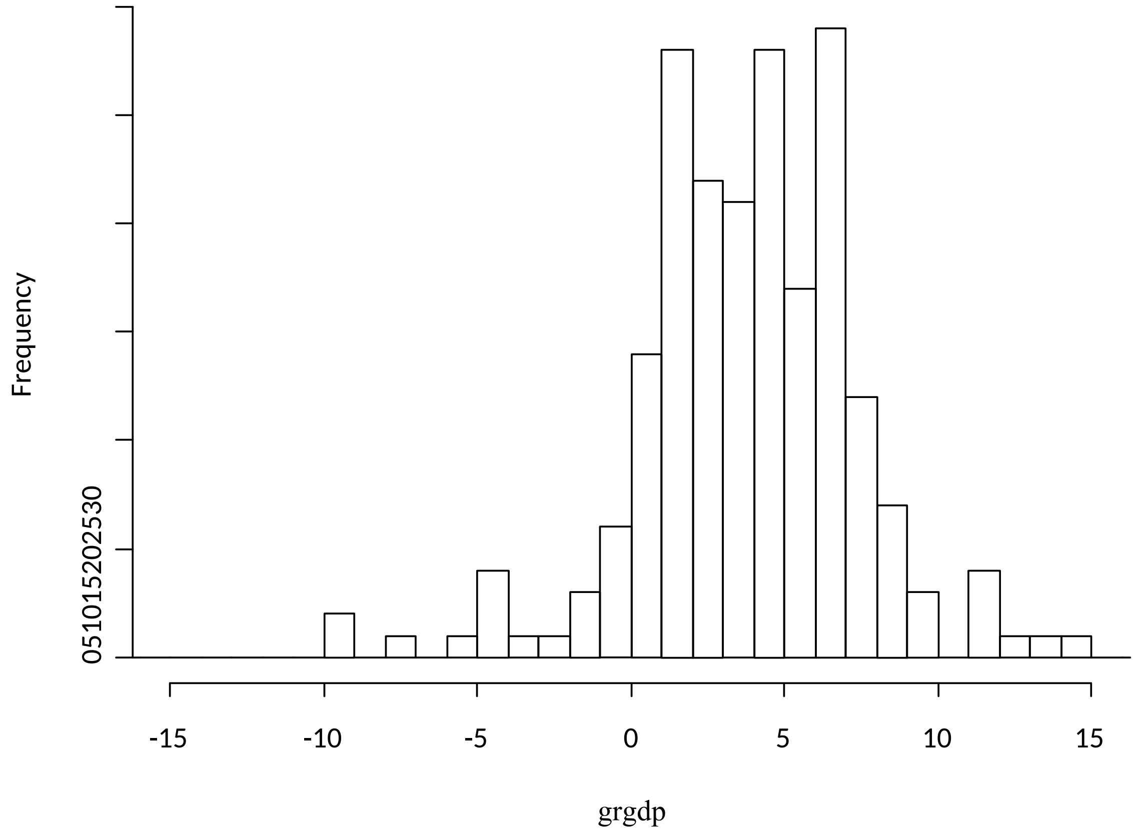
INTERPRETATION-From the above Scatterplots Between Independent and Dependent Variable we observe that there is approximately a Linear Relationship between them.

Histogram of GDP Growth

CODE-

```
hist(grgdp,xlim = c(-15,15),breaks=100,main="Histogram of GDP Growth")
```

Histogram of GDP Growth



INTERPRETATION- Our Histogram Looks Approximately Normal.

MODEL FITTING

Simple Linear Regression Model

The Purpose of the Simple Linear Regression Model given below is to test the Relationship between GDP growth and the Gini coefficient. To Test this relationship, GDP Growth was only Regressed on the Gini coefficient.

$$\text{Model 1: } grgdp = \beta_0 + \beta_1 gini + e$$

CODE- model1=lm(grgdp~gini)

summary(model1)

OUTPUT-

```
Call:
lm(formula = grgdp ~ gini)

Residuals:
Min        1Q     Median    3Q      Max
-12.6042  -1.6232  -0.6819   1.7059  13.9208

Coefficients:

          Estimate    Std. Error  t value  Pr(>|t|)
(Intercept) -0.3067456  1.287696    -0.24    0.001   gini        0.1105578  0.0330297
3.35    0.0343***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.9117 on 103 degrees of freedom
Multiple R-squared:  0.0981, Adjusted R-squared:  0.0893
F-statistic: 19.78 on 1 and 103 DF,  p-value: < 2.2e-16
```

INTERPRETATION

The Results showed a Positive Relationship between the Gini coefficient and GDP growth.. This Indicates that for One Unit Increase in Gini Coefficient, the GDP growth Rate Increases by 11.06 percent. Since the Intercept is Negative, this means that with zero Inequality (Gini equals zero), there would be Negative Growth. This is a reasonable inference because perfect inequality, which is what is assumed be no inequality, would allow the assumption of negative growth. The pvalue of Gini was 0.001, indicating a very high statistical significance. Also, the R^2 found is 0.0981, which means the Gini coefficient only explains 9.8 percent of the GDP growth in the model--a low value. We found this rather unsatisfactory. This could indicate a non-linear relationship. We add more variables to our Model.

Multiple Linear Regression Model

We Construt Several Multiple Regression Models to account for other variables with Economic Significance that may affect Economic Growth. The additional variables were gross savings, unemployment rate, years of education, and fertility rate.

Our First Multiple Regression Model includes the Gini Coefficient and the Gross Savings Rate.

Model 2:

$$grgdp = \beta_0 + \beta_1 gini + \beta_2 gsav + e$$

CODE- model2=lm(grgdp~gini+gsav)

summary(model2)

OUTPUT- Call:

lm(formula = grgdp ~ gini+gsav)

Residuals:
Min 1Q Median 3Q Max
-10.5789 -1.7414 -0.4986 1.8996 12.5815

Coefficients:
Estimate Std. Error t value Pr(>|t|) (Intercept) -2.644401 1.401742 -1.89 0.062* gini 0.114785 0.0316657 3.62 0.0005*** gsav 0.1027894 0.0290295 3.54 0.001****

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

Residual standard error: 2.73 on 96 degrees of freedom
Multiple R-squared: 0.1987, Adjusted R-squared: 0.1820
F-statistic: 11.90 on 2 and 96 DF, p-value: < 2.2e-16

INTERPRETATION

In Model 2, both independent variables were positive and significant at the 1% level. The R^2value was 0.199, which increased from the R^2value of 0.098 for the Simple Regression Model 1.

In the Following model , Model 3, the variable fertil, for fertility rate, was added to the pre existing variables Gini Coefficient and Gross Savings.

Model 3:

$grgdp = \beta_0 + \beta_1 gini + \beta_2 gsav + \beta_3 fertil + e$ CODE-

model3=lm(grgdp~gini+gsav+fertil) summary(model3)

lm(formula = grgdp ~ gini+gsav+fertil)

OUTPUT- Call:

Residuals:

Min	1Q	Median	3Q	Max
-10.3417	-1.6618	-0.6098	1.8748	12.4016

Coefficients:

Estimate Std. Error t value Pr(>|t|) (Intercept) -4.107141 1.26482 -3.52 0.002* gini 0.0670464 0.029288 2.29 0.0024** gsav 0.1337481 0.0262202 5.10 0.0005** fertil 1.032928 0.1928839 5.36 0.0005***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.4052 on 95 degrees of freedom
Multiple R-squared: 0.3845, Adjusted R-squared: 0.3650
F-statistic: 19.78 on 3 and 95 DF, p-value: < 2.2e-16

INTERPRETATION

The Results show that fertility rate was also a positive and significant relation to GDP growth. The Gini coefficient and the gross savings rate retained significance in Model 3.The Gini coefficient is now significant at the 5% level, while gross savings and fertility rate were significant at the 1% level. The R^2value increased to 0.385, which means that the variables explain 38.5% of the variation in grgdp.
This makes sense because as we control for more variables, the larger R^2will be.

Model 4:

$$grgdp = \beta_0 + \beta_1 Gini + \beta_2 gsav + \beta_3 fertil + \beta_4 unemp + e$$

CODE-

Model4=lm(grgdp~gini+gsav+fertil+unemp) summary(model4)

lm(formula = grgdp ~ gini+gsav+fertile+unemp)

Residuals:

Min 1Q Median 3Q Max

OUTPUT- Call:

-9.7163 -1.6784 -0.8092 1.3777 12.2039

Coefficients:

Estimate Std. Error t value Pr(>|t|) (Intercept) -2.750765 1.413114 -1.95 0.055* gini 0.07297654 0.0290482 2.51 0.014** gsav 0.1153335 0.0274353 4.20 0.0005*** fertil 0.8932998 0.2050265 4.36 0.0005*** unemp -0.0980788 0.0468862 -2.09 0.39**

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.362 on 92 degrees of freedom
Multiple R-squared: 0.4218, Adjusted R-squared: 0.3967 F-statistic: 16.78 on 4 and 92 DF, p-value: < 2.2e-16

INTERPRETATION

In Model 4 we added the Unemployment Rate, which although proved to be Significant alongside the other variables, had a negative relationship with GDP growth. The Gini coefficient maintained significance at the 5% level, like the unemployment rate, while gross savings and fertility rate remained significant at the 1% level. The R^2value for this model increased once more to 0.423.

Model 5:

grgdp = β_0+β_1Gini + $\beta_2$gsav + $\beta_3$fertil + $\beta_4$unemp + $\beta_5$educ+e CODE-

model5=lm(grgdp~gini+gsav+fertil+unemp+educ) summary(model5)

lm(formula = grgdp ~ gini+gsav+fertile+unemp+educ)

Residuals:
Min 1Q Median 3Q Max
-9.0960 -1.6686 -0.4202 1.7884 9.0293

OUTPUT- Call:

Coefficients:

Estimate Std. Error t value Pr(>|t|) (Intercept) -1.375652 2.421891 -0.57 0.572 gini 0.0885431 0.0353214 2.51 0.015** gsav 0.1182324 0.032796 3.61 0.001*** fertil 0.6557092 0.330266 1.99 0.051* unemp -0.0932868 0.0513511 -1.82 0.074* educ -0.1652503 0.1292013 -1.28 0.206

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

Residual standard error: 2.3612 on 64 degrees of freedom
Multiple R-squared: 0.4783, Adjusted R-squared: 0.4375
F-statistic: 11.73 on 5 and 65 DF, p-value: < 2.2e-16

INTERPRETATION

In Model 5 we incorporated the variable educ, which represents mean years of education. The mean years of education had a negative relationship with GDP growth. This new variable differed from all the other variables because it was not statistically significant at the 10%, 5%, or 1% levels. The Gini coefficient maintained significance at the 5% and 10% levels, while gsav, fertil, and unemp all retained their statistical significance at the 1% level. In addition, the intercept was not statistically significant at any level in this model, unlike the other models. Therefore, we can conclude from these results that the Gini coefficient, gross savings, fertility rate, and unemployment all have an impact on GDP growth, while no conclusions can be made about education. The R^2value rose once more to 0.478, which means that 47.8% of the variation can be explained by the model.

TABLE SHOWING SUMMARY OF ALL MODELS

Dependent Variable grgdp

Independent Variables	**SLR Model 1**	**MLR Model 2**	**MLR Model 3**	**MLR Model 4**	**MLR Model 5**
gini	0.111*** (0.033)	0.114*** (0.317)	0.067** (0.029)	0.073** (0.029)	0.0885** (0.035)
gsav		0.103*** (0.029)	0.134*** (0.026)	0.115*** (0.027)	0.118*** (0.033)
fertil			1.033*** (0.193)	0.893*** (0.205)	0.656* (0.330)
unemp				-0.098** (0.047)	-0.093* (0.051)
educ					-0.165 (0.129)
Intercept	-0.307 (1.287)	-2.644* (1.402)	-4.107*** (1.265)	-2.751* (1.413)	-1.376 (2.422)
No. of obs	105	99	99	97	70
R-square	0.098	0.199	0.385	0.423	0.478
Adj R-square	0.0893	0.1820	0.3650	0.3967	0.4375

The Quantities in Parentheses are Standard Errors. *, **, *** denotes Significance of Coefficients at 10%, 5%, and 1% respectively

CONCLUSIONS

- The Gini Coefficient is Positively Correlated to GDP Growth.
- The Gross Savings Rate is Positively Correlated to GDP Growth.
- The Fertility Rate is Positively Correlated to GDP Growth.
- Unemployment Rate is Negatively Correlated to GDP Growth.
- Mean School Years is Negatively Correlated to GDP Growth.
- The Gini Coefficient is Statistically Significant at 10 % , 5 % and 1 % Level of Significance.
- The Gini Coefficient is Statistically Significant at 5 % and 10 % Level of Significance when Independent Variables are added for the Construction of Models 3-5.
- The Gross Savings Rate is Statistically Significant at 10 % , 5 % and 1 % Level of Significance.
- The Fertility Rate is Statistically Significant at 10 % , 5 % and 1 % Level of Significance.
- The Unemployment Rate is Statistically Significant at 10 % and 5 % Level of Significance.
- The Significance of the Intercept varied widely throughout the models, so we cannot conclude much about its Statistical Significance.
- Overall, all the Independent Variables helps to explain the Variance in Dependent Variable GDP Growth.

REFERENCES

https://www.wikipedia.org/ http://statisticsbyjim.com/

https://stackoverflow.com/

https://www.tutorialspoint.com/r/index.htmL/

https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.Z

G http://www.r-tutor.com/