A Bivariate Analysis Project by students of 4th semester, Statistics Department .

India's Annual Gross Domestic Product and Population Size

A Correlation and Regression Analysis

Department of Statistics, Surendranath College, Kolkata
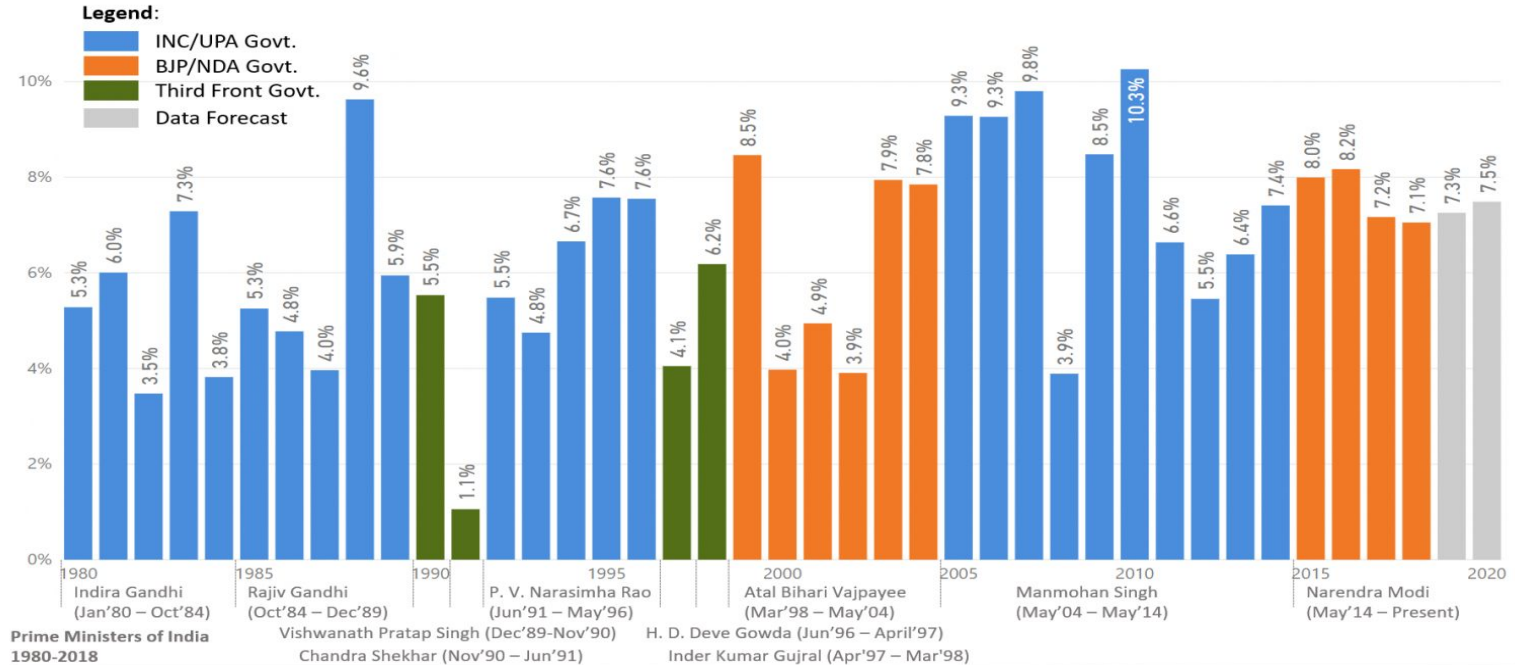
What is Gross Domestic Product?

Gross domestic product (GDP) is the total monetary or market value of all the finished goods and services produced within a country's borders in a specific time period.
As a broad measure of overall domestic production, it functions as a comprehensive scorecard of a given country's economic health.

This graph shows India's real GDP from 1980-2020 (Source: mgmresearch.com)

Population of India from 1800 to 2020 (in millions)
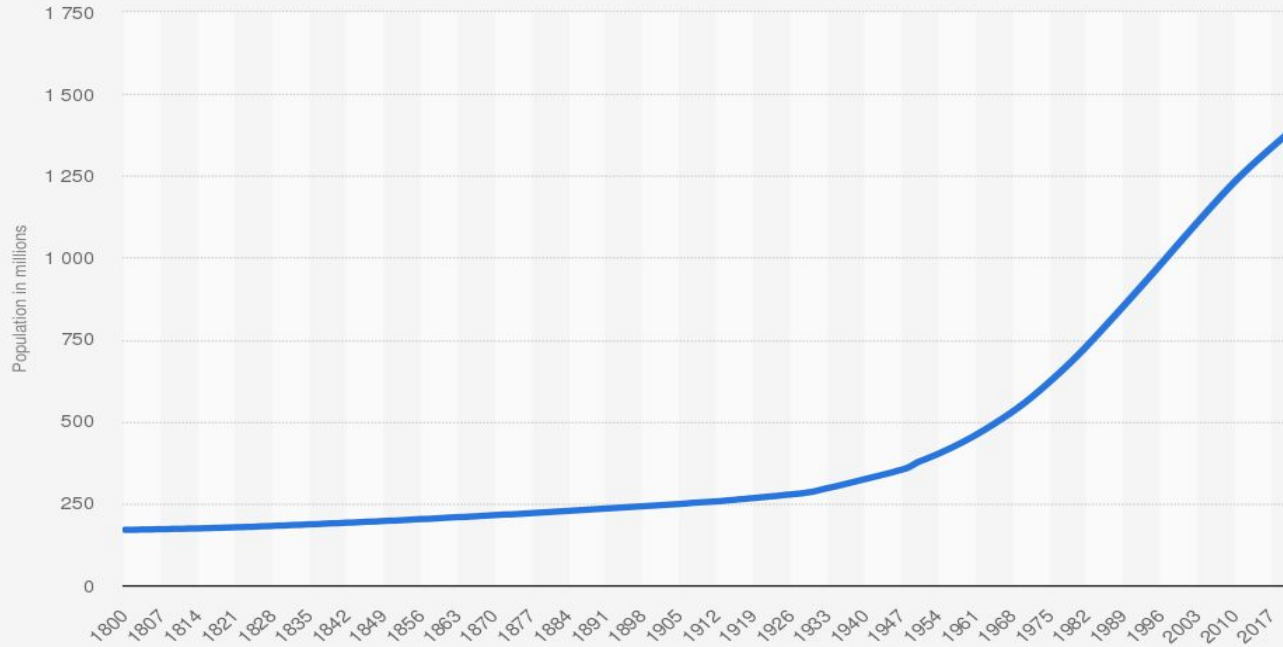
This graph shows India's population growth from 1800-2020 (Source: statista.com)

Effect of Population Growth on Economy

The increase in population leads to the increase in demand for goods. Thus growing population means the growing market for goods is enlarged, they can be produced on a large scale and thus economies of large-scale production can be reaped.

The economic history of the USA and European countries shows that growth of population and labour force contributed a good deal to the increase in their national output.

In India where supplies of other economic resources, especially capital equipment, are relatively scarce, increase in population or labour force does not lead to the employment of all due to scarcity of capital resources.

As for the argument that population growth leads to increase in demand or market for goods, it may be noted that the mere growth of unemployed or paupers cannot lead to greater demand for goods or expansion in their markets.

Objective of the Project

- ❏ The Objective of the Project is to analyze the relationship between the rate of population growth of India to India's GDP through statistical measures like correlation and linear regression.
- ❏ The Goal of the Project is to analyze whether the increasing population growth has been a cause of India's recent GDP.
- ❏ We have taken data of 20 years (2001-2020) to study the relation between Population and Economic growth of India.
 *(Source: macrotrends.net)*

Data for Analysis

Year	GDP (%)	Population	Year	GDP (%)	Population
2001	4.82	1,075,000,085	2011	5.24	1,250,287,943
2002	3.80	1,093,317,189	2012	5.46	1,265,780,247
2003	7.86	1,111,523,144	2013	6.39	1,280,842,125
2004	7.92	1,129,623,456	2014	7.41	1,295,600,772
2005	7.92	1,147,609,927	2015	8.00	1,310,152,403
2006	8.06	1,165,486,291	2016	8.26	1,324,517,249
2007	7.66	1,183,209,472	2017	7.04	1,338,676,785
2008	3.09	1,200,669,765	2018	6.12	1,352,642,280
2009	7.86	1,217,726,215	2019	4.18	1,366,417,754
2010	8.50	1,234,281,170	2020	-7.3	1,380,004,385
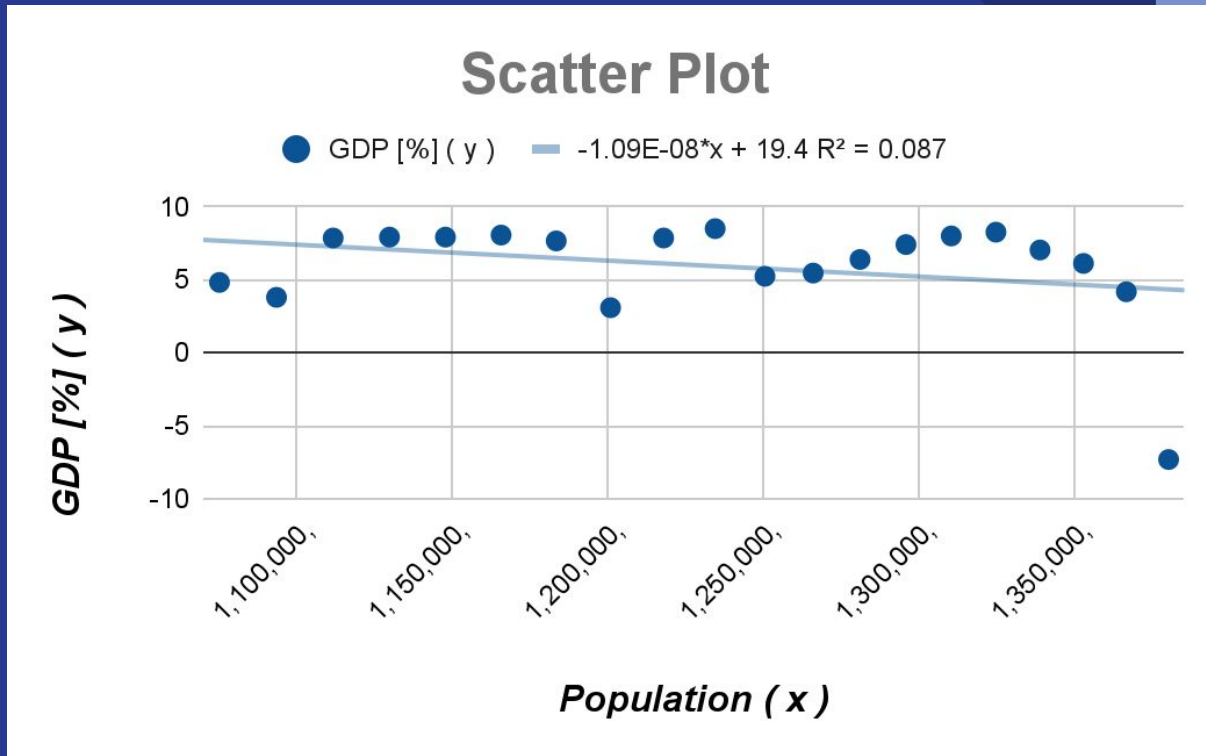
Here we have done a scatter plot of our data keeping GDP in the y axis and Population growth in the x axis. The best fitted regression line and correlation coefficient has also been plotted using the following calculations.

Calculations

❏ Correlation Coefficient determines the linear relationship between two variables. And it is given by :

$$r = \frac{\sum \left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum \left(x_i - \bar{x}\right)^2 \sum \left(y_i - \bar{y}\right)^2}} \implies r = \frac{Covariance(x,y)}{SD(x)*SD(y)}$$

❏ Regression Equation of *y* on *x* determines the predicted value of y (GDP) with a specified given value of x (Population Growth)

Here, our regression model is : $y = a + bx$

where $a = \bar{y} - b\bar{x}$; $b = r\frac{Sy}{Sx}$

Calculation Table

Population (x)	GDP [%] (y)	(x-mean(x))(y-mean(y))	(y - mean(y))^2	(x - mean(x))^2
1,075,000,085	4.82	176398756.9	1.19793025	2.59752E+16
1,093,317,189	3.8	302058955.4	4.47111025	2.04065E+16
1,111,523,144	7.86	-242497409.7	3.78497025	1.55364E+16
1,129,623,456	7.92	-213675951.4	4.02203025	1.13518E+16
1,147,609,927	7.92	-177604083.8	4.02203025	7.84261E+15
1,165,486,291	8.06	-151648535.7	4.60317025	4.99597E+15
1,183,209,472	7.66	-92439866.43	3.04677025	2.80465E+15
1,200,669,765	3.09	100265987.8	7.97780025	1.26016E+15
1,217,726,215	7.86	-35879335.12	3.78497025	340115404759524
1,234,281,170	8.5	-4879518.487	6.68481025	3561761631169
1,250,287,943	5.24	-9523609.495	0.45495025	199360562640100
1,265,780,247	5.46	-13458569.46	0.20657025	876859528370596
1,280,842,125	6.39	21242340.55	0.22610025	1.99574E+15
1,295,600,772	7.41	88881062.97	2.23652025	3.5322E+15
1,310,152,403	8	154293569.4	4.34931025	5.47363E+15

Calculation Table

	Population (x)	GDP [%] (y)	(x-mean(x))(y-mean(y))	(y - mean(y))^2	(x - mean(x))^2
	1,324,517,249	8.26	207222147.9	5.50137025	7.80551E+15
	1,338,676,785	7.04	115373150.2	1.26675025	1.0508E+16
	1,352,642,280	6.12	23935375.56	0.04223025	1.35662E+16
	1,366,417,754	4.18	-225917447.3	3.00849025	1.69649E+16
	1,380,004,385	-7.3	-1900720188	174.6230103	2.06888E+16
Total Sum	24,723,368,657	118.29	-1878573168	235.510895	1.72128E+17
Mean	1236168433	5.9145	-93928658.39	11.77554475	8.60641E+15
Standard Deviation	92770723.2	3.431551362			
Correlation Coefficient (rxy)	-0.2950507164		Value of 'b'	-0.0000001091380613	
			Value of 'a'	19.40580262	

Results

❏ The **correlation coefficient** between **India's GDP and Population Growth** for the last 20 years is **-0.295.**

❏ The **linear regression model** of **GDP on Population Growth** is :

$$y = -1.0913 \times 10^{-8} x + 19.405$$

Conclusions

- ❏ -0.295 is the value of the correlation coefficient which indicates that there is a very low negative correlation between India's GDP and Population Growth for the last 20 years (2001-2020).

- ❏ This vaguely shows that the Country's growing population hasn't been much of a factor for the fluctuating GDP.

- ❏ The calculated Linear Regression Model can approximately calculate the Country's GDP for a given population of a certain year.

A PROJECT BY: Agnibh Pathak, Pabitra Patra, Shreya Kundu, Shalini Paul, Partha Sarathi Ghosh, Swagata Ghosh, Swapnomoy Pal