Volume 3, Issue 8, August 2013 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Data Mining and Analysis of Economic Data

Rinkesh Jain
M. Tech. Scholar
BUIT, Bhopal, India

Divakar Singh
Associate Professor,
HOD, CSE, BUIT, Bhopal, India

Abstract:-Gross State Domestic Product (GSDP) is a benchmark for economic production conditions of an Indian state. It is also a measure of standard of living of people in a particular region. Estimates of economic growth in an Indian state has important roles, among others as a touchstone in determining business plans for business entities, and also the basis for devising government policies. This study has been attempted to shed light on the issues such as forecasting growth rates of GSDP of Madhya Pradesh. In this study data has been collected about various sectors like Primary Sector, Secondary Sector, Tertiary Sector (Services) at current and constant prices and showing the current inclination in these sectors. GSDP is the summation of these sectors, thus also showing inclination of GSDP and its growth over previous year. After observing inclination in these sectors, some tools and techniques for forecasting growth of GSDP of Madhya Pradesh has been discussed.

Keywords: GSDP, Growth Rate Primary Sector, Secondary Sector, Tertiary Sector

I. INTRODUCTION

State Domestic Product is a measure in monetary terms of the volume of all goods and services produced by an economy during a given period of time accounted without duplication [1, 2, 3, 6]. SDP is a reflection of economic development of the State and its determinate Per Capita Income is a suitable measure of the well-being of its people [1]. The measure obviously has to be in value terms as the different units of production and different measures of services are not directly additive [1, 2].

This is the most important single economic indicator used to measure the growth and to study the structural changes taking place in the economy. SDP estimates over a period of time reveal the extent and direction of the changes in the level of economic development. Sectoral composition of SDP tells about the relative position of different sectors in the economy over a period of time. This indicates the real structural changes taking place in the economy, and also helps in the formulation of the plans for overall economic development [3]. These estimates depicts the level of development and are useful for proper planning, Evaluating progress and inter-state comparisons and as such are a valuable tool in the hands of planners and policy makers[6]. The estimates of SDP at current prices are obtained by evaluating the product at current prices existing during a particular year. The valuation of SDP at current prices, do not reveal the actual economic growth over a period of time, because it contains a combined effect of (i) the changes in volume of goods and services and (ii) the changes in the prices of goods and services. In order to minimize the effect of inflation, the estimates of SDP can also be prepared by evaluating the goods and services at the prices prevailing in the fixed year known as base year and are known as the estimates of State Domestic Product at constant prices [1, 7].

The State Domestic Product (SDP) commonly known as State Income is one of the important indicators to measure the economic development of the State/UT. In the context of planned economic development of the State/UT, State Income and Per Capita Income (PCI) plays a vital role in formulation of policies by policy makers, governing bodies, administrators and planners. Such estimates serve as an indicator to assess and analyze the status of the economy among the States/UTs in the country as well as overall impact of various developmental programmes implemented by the Government. It gives an overall picture of the economy over a period of time [7].

The estimates of State Domestic Product are prepared both in terms of Gross and Net basis, for all sectors of economy. The difference between gross and net basis of estimating SDP is that, in the gross estimates, no deduction is made for the Consumption of Fixed Capital (CFC) which takes place in the process of production, whereas in the net estimates, CFC is subtracted from the gross value figures. Net State Domestic Product is also called State Income. Again, capital is one of the primary factors used in production and this results in the consumption of the fixed capital and hence, a reduction in the economic life of the capital. In nutshell, it can be said that, the capital undergoes depreciation as a result of its use in the process of production. The Consumption of Fixed Capital (CFC) measures the replacement value of that part of the capital stock, which has been utilized in the production process during the year[1,2]. The SDP estimates are prepared with respect to a base year and this base year gets revised from time to time to take into account the structural changes which have been taking place in the economy and to depict a true picture of the economy through macro aggregates like GSDP, NSDP and Per Capita Income etc. For examining the performance of the economy in real terms through the macro economic aggregates like

Gross State Domestic Product (GSDP), State Income, consumption, expenditure, capital formation etc., estimates of these aggregates are prepared at the prices of selected year known as base year[1].

First series of estimates was compiled for the period 1960-61 to 1977-78 with base year 1960-61. While, 1970-71 to 1986-87 covers the second series, with base year 1970-71. The third series was 1980-81 to 1996-97 with base year 1980-81. Fourth revision took place in nineties with base year 1993-94 and continued up to 2004-05. Fifth revision took place in 2006 with base year 1999-00 and continued up to 2008-09. For the present series the base year has been revised to 2004-05. The estimates at the existing prices of the current year are termed "at current prices", while "at constant prices" is termed for those prepared at base year prices. The comparison of the estimates at constant prices, which means "in real terms", over the year gives the measure of real growth. Therefore, the base year of the present SDP series has been shifted to 2004-05[1].

Mentioned below are the three major components which influences the present revision exercise :

(i) revision to the base year to a more recent year (for meaningful analysis of the structural changes in the economy), (ii) complete review of the existing data base and methodology employed in the estimation of various macro-economic aggregates including choice of the alternative databases on individual subjects and (iii) to the extent feasible, implementing the recommendations of the System of National Accounts (1993 and 2008) prepared under the auspices of the Inter Secretariat Working Group on National Accounts comprising of the European Communities (EUROSTAT), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations and World Bank[1,2].

II. LITERATURE SURVEY

21. Estimation of SDP

For the purpose of estimation of SDP, the whole economy of the state has been divided into seventeen (17) sub-sectors in conformity with the procedure followed at the national level to estimate the National Income estimates. The procedure is followed by each state and union territory of the country to make the estimates comparable. The sectors have been further categorized into three broad sectors as follows[1.6,7,8].

(i) Primary Sector

- Agriculture and Livestock
- Forestry and logging
- Fishing
- Mining and Quarrying

(ii) Secondary Sector

- Manufacturing (Registered)
- Manufacturing (Un-Registered)
- Electricity, Gas and Water Supply
- Construction

(iii) Tertiary Sector

- Trade, Hotels and Restaurants
- Railways
- Transport by Other Means
- Storage
- Communication
- Banking and Insurance
- Real Estate, ownership of dwellings, business and legal services
- Public Administration
- Other services

Estimates of these sectors are prepared individually by adopting one or more of the following approaches[1,6,7,8].

2.1.1 Production Approach

In this method, the sum of economic value of all goods and services produced within the State during the year is considered after deducting the inputs consumed in the process of production. This approach is followed in Agriculture, Livestock, Forestry, Fishing, Mining and Quarrying and Manufacturing (Registered) Sectors [1, 6, 7, and 8].

2.1.2 Income Approach

The income accrued to the factors of production namely land, labour, capital and entrepreneurship in form of rent, salaries and wages, interest and profit is taken into consideration in estimation of value added. This approach is being followed in Manufacturing (un-registered), Electricity, Gas and water supply, Trade, Hotels and restaurants, Transport, Storage and Communication, Financing, Insurance, Real Estate, Business services, Public Administration and other services[1,6,7,8].

2.1.3 Expenditure Approach

This method is based on the measurement of income at the stage of disposal. The produce is either ultimately consumed or part of it is saved for further consumption or future production of goods and services. Thus, the money value of consumption expenditure plus the savings gives the income. The mentioned approach is used in estimating income from construction sector [1, 6, and 8].

2.2. Current Inclination in different Sectors

Table 1. GSDP from Primary Sector (Rs. In Lakh.) [17,18]

Prices	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13
Current	3668764	4032155	4453525	4833096	5882362	6978528	7382005	9039933	10656405
Constant	3668764	3893567	4021305	4037264	4377124	4762149	4737459	5445334	6071018

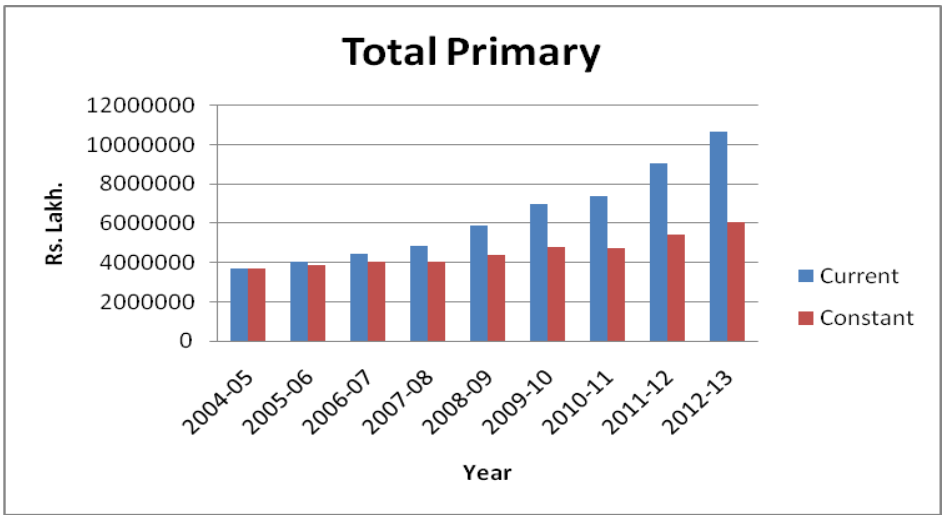

Fig1: GSDP from Primary Sector (Rs. In Lakh.)

Table 2. GSDP from Secondary Sector (Rs. In Lakh) [17, 18]

Prices	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13
Current	2520874	2794829	3505561	3922659	5113222	5689237	6705587	7793485	8824077
Constant	2520874	2660196	3147314	3302769	4014461	4325285	4796747	5187559	5460855

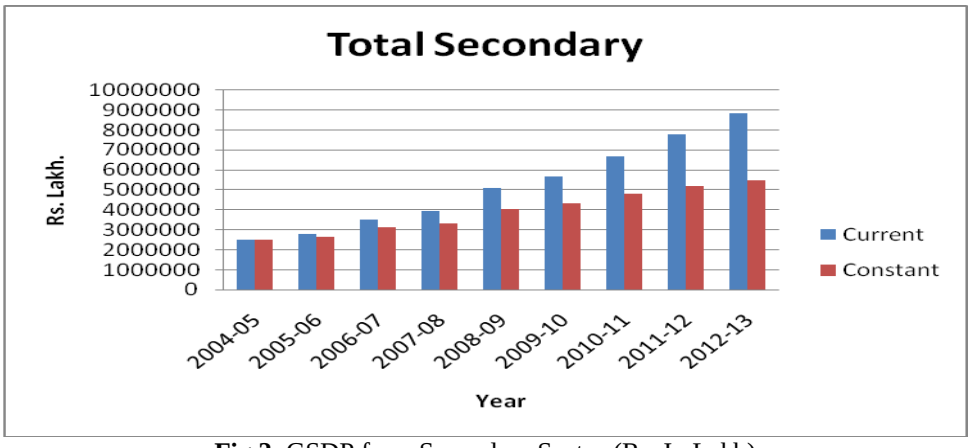

Fig 2: GSDP from Secondary Sector (Rs. In Lakh)

Table 3. GSDP from Services (Rs. In Lakh.) [17,18]

Prices	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13
Current	5103051	5600615	6498595	7392184	8732036	10130610	11932255	14135273	16706909
Constant	5103051	5338183	5820993	6258538	6903031	7718328	8469149	9496094	10614413

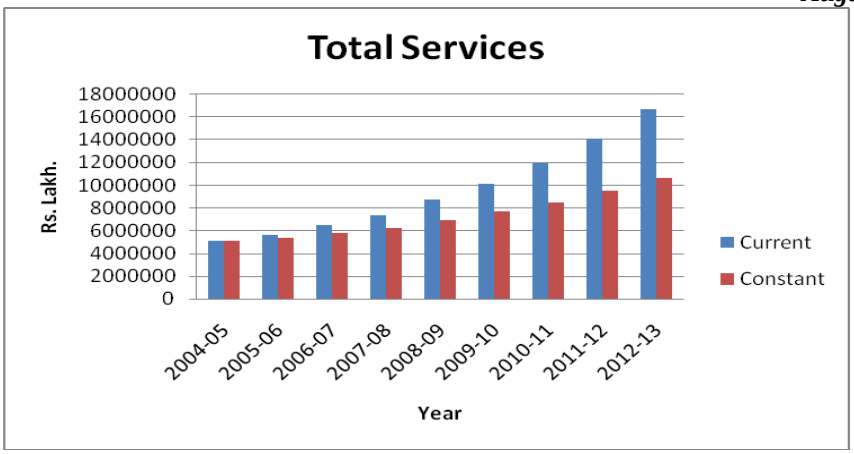

Fig 3: GSDP from Services (Rs. In Lakh.)

Table 4. GSDP at Current and Constant Prices (Rs. In Lakh.) [17,18]

Prices	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13
Current	11292689	12427599	14457681	16147939	19727620	22798375	26019847	30968691	36187391
Constant	11292689	11891946	12989612	13598571	15294616	16805762	18003355	20128987	22146286

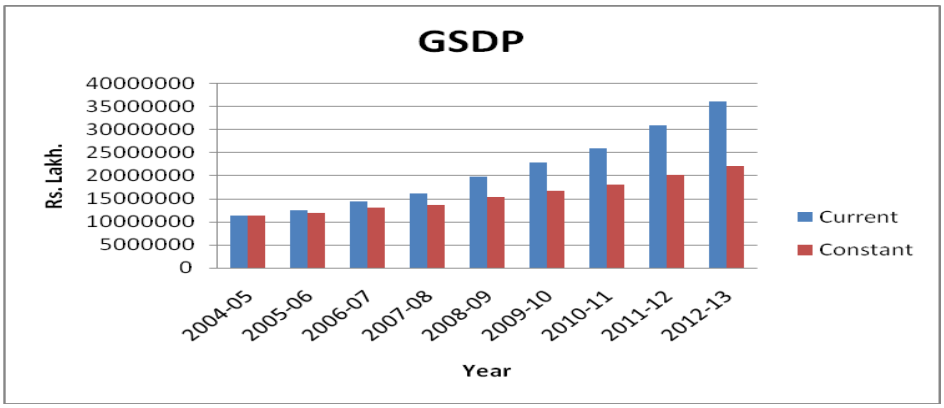

Fig 4: GSDP at Current and Constant Prices (Rs. In Lakh.)

Table 5. GSDP growth rate over previous year.[17,18]

GSDP growth rate	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13
Current	9.81	10.05	16.34	11.69	22.17	15.57	14.13	19.02	16.85
Constant	3.08	5.31	9.23	4.69	12.47	9.88	7.13	11.81	10.02

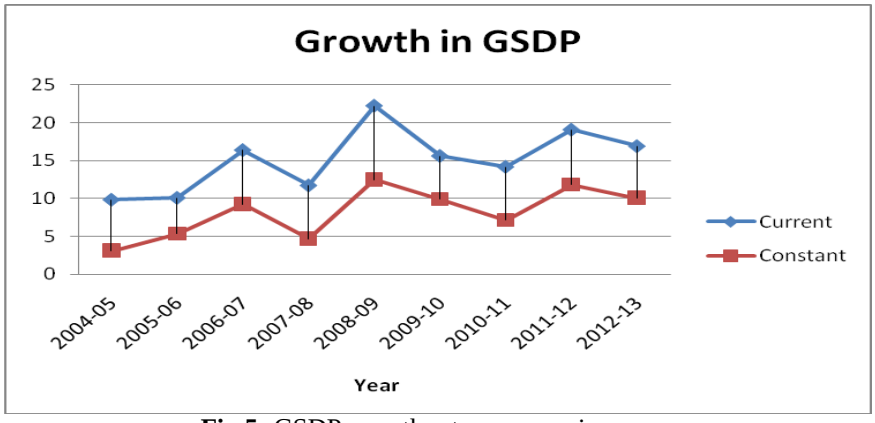

Fig 5: GSDP growth rate over previous year

This document aims to analyze the relationship between real estate investment and economic growth by using the principle of liner regression, the collected data about the investment of the real estate and GDP of the city of Nanjing from 1998 to 2007. The result shows that real estate investment has a significant impact on economic growth. Furthermore, the development of real estate plays a promotional role on the growth of GDP [12]. Large amounts of data are readily available and collected daily by global networks worldwide. This research presents a comparative study of some data mining tools for large scale time series dataset for analysis and mining. In this project Weka tool [14] and Rapid Miner tool [15] are used to predict future weather events. In an effort to find patterns association rules algorithms are used. The purpose of an association rules algorithm is to find relationships between different columns of data (data types) [13].

This study has been attempted to shed light on the issues such as forecasting growth rates of GDP of India. Data on GDP have been collected over a period of 60 years from various publications of Reserve Bank of India. A very simple tentative ARIMA (1, 2, 2) model has been fitted on data to estimate the parameters of autoregressive and moving average components of this model. The obtained data and its results suggest that only one period of autoregressive and moving average terms are statistically significant. Further, absolute values of forecasted GDP indicate an increasing trend and its respective growth rates reveal an opposite trend in future. Findings will assist the policy makers and managers to formulate economic and business strategies in turn more precisely [4].

Weka (>= 3.7.3) now has a dedicated time series analysis environment. This allows the evaluation and visualization of forecasting models. This environment takes the form of a plug-in tab in Weka's graphical "Explorer" user interface and can be installed via the package manager. Weka's time series framework takes a machine learning/data mining approach to modeling time series by transforming the data into a form that standard propositional learning algorithms can process. It does this by removing the temporal ordering of individual input examples by encoding the time dependency via additional input fields. These fields are sometimes referred to as "lagged" variables. After the data has been transformed, any of Weka's regression algorithms can be applied to learn a model. An obvious choice is to apply multiple linear regressions, but any method capable of predicting a continuous target can be applied. This approach to time series analysis and forecasting is often more powerful and more flexible than classical statistical techniques such as ARMA and ARIMA [16].

III. Proposed Work

It is one of the novel work, which has been carried out in this research project, which establishes a consequent relationship between an actual growth and GSDP of an indian state. To the obtained dataset, association rule algorithm can be applied in order to generate some rules as a basis for further research. The purpose of an association rules algorithm is to find relationships between different attributes of data. As it is known that forecasting is a statistic based approach thus results obtained may or may not be consistent and accurate always, thus, other tools and techniques can be used including Weka tool, Rapid Miner tool and Microsoft Excel for time series forecasting using Regression analysis, Exponential smoothing etc. These analytical tools might help in predicting future values of Primary sector, Secondary sector, Tertiary sector, summation of these sectors results in GSDP. Concurrently, forecasted values of GSDP can be used to calculate per capita GSDP which is a suitable measure of well being of its people.

IV. Conclusion

Several graphs shows that GSDP of Madhya Pradesh is increasing during the period from 2004-05 to 2012-13. Several sectors like Primary sectors, Secondary sector, Tertiary sector also showing same increasing inclination at current as well as constant prices. So this study depicts that for the next five or six years GSDP of Madhya Pradesh continue to increase. The findings of this study have some important implications for policy makers and managers. Policy makers-who deal with macro variables would find the results of this work are very helpful to formulate better policy. On the other hand, the manager who is planning to invest in the expansion of existing business or in new project will be benefitted tremendously since the findings will help them to portrait the picture of economic condition of Madhya Pradesh more precisely in advance.

References

[1] R. N. Sharma, Sh. Shan-E-Alam, “Estimates of State Domestic Product of Delhi 2012-13”, Directorate of Economics & Statistics, Government of N. C. T. of Delhi, April 2013 Delhi.

[2] B. K. Sharma, “Estimates of State Domestic Product 2006-07”, Directorate of Economics & Statistics, Government of N. C. T. of Delhi, April 2008 Delhi.

[3] D. C. Patel, “State Domestic Product Gujarat State 2009-10”, Directorate of Economics & Statistics, Government of Gujarat, April 2011 Gandhinagar.

[4] Bipasha Maity, Bani Chatterjee, “FORECASTING GDP GROWTH RATES OF INDIA: An Empirical Study”, *International Journal of Economics and* Management Sciences, Vol. 1, No. 9, 2012, pp. 52-58.

[5] Web reference: http://mospi.nic.in/mospi_new/upload/national_accounta.pdf “National Accounts Statistics: Manual On Estimation Of State And District Income, 2008” as accessed on [05/08/2013].

[6] M. L. Sharma, “State Domestic Product of Punjab (2004-05 to 2009-10)”, Economic Adviser Government of Punjab, Publication No. 931, Chandigarh.

7] Web reference: http://www.and.nic.in/stats/Economic%20survey%20ANI/CHAPTER%20-%203.pdf “Chapter 3 State Domestic Product“ as accessed on [05/08/2013].

[8] Web reference: http://himachal.nic.in/economics/REPORTS/GrossDomesticProduct2013.pdf "Chapter-I Gross State Domestic Product Of Himachal Pradesh (Gsdp) Base=2004-05" as accessed on [05/08/2013].

[9] Web reference: http://delhiplanning.nic.in/Economic%20Survey/ES2007-08/C2.PDF "Chapter 2 State Domestic Product "as accessed on [05/08/2013].

[10] Anand Sherkhane, "State Domestic Product in Goa 2004-05 to 2011-12(Q)", Directorate of Planning, Statistics and Evaluation, Government of Goa, October 2012 Panaji-Goa.

[11] G.A. Qureshi, "State Domestic Product of Jammu & Kashmir (1999-00 to 2007-08)", Directorate of Economics & Statistics Planning & Development Department, Government of Jammu & Kashmir, No: DES/SI/ 2009.

[12] Zhang Yun-ning1, Yu Qian2, "Regression Analysis of the Real Estate Investment and Economy Increased", 2013 Third International Conference on Intelligent System Design and Engineering Applications, 2012., pp. 1099-1101.

[13] Patricia Morreale, Steve Holtz, Allan Goncalves, "Data Mining and Analysis of Large Scale Time Series Network Data", 2013 27th International Conference on Advanced Information Networking and Applications Workshops, 2013, pp. 39-43.

14] Web reference: http://www.cs.waikato.ac.nz/ml/weka/ "Weka's website" as accessed on [05/08/2013].

[15] Web reference: http://rapid-i.com/content/view/ 181/190/ "RapidMiner's website" as accessed on [05/08/2013].

[16] Web reference: http://wiki.pentaho.com/display/DATAMINING/Time+Series+Analysis+and+Forecasting+with+Weka "Time Series Analysis and Forecasting with Weka" as accessed on [05/08/2013].

[17] Web reference: http://www.indiastat.com/table/economy/8/grossstatedomesticproductgsdp/12081/537548/data.aspx "Gross State Domestic Product (GSDP) at Factor Cost by Industry of Origin in Madhya Pradesh {(At Constant 2004-05 Prices) (2004-2005 to 2012-2013)" as accessed on [05/08/2013].

[18] Web reference: http://www.indiastat.com/table/economy/8/grossstatedomesticproductgsdp/12081/537547/data.aspx "Gross State Domestic Product (GSDP) at Factor Cost by Industry of Origin in Madhya Pradesh {(At Current Prices) (2004-2005 to 2012-2013)} as accessed on [05/08/2013].

