Analysis of value added services on GDP Growth Rate using Data Mining Techniques

Douglas KUNDA[1], Sipiwe CHIHANA[1]
[1]Department of Computer Science, School of Science Engineering and Technology,
Mulungushi University, PO box 80415 Kabwe, Zambia
dkunda@mu.edu.zm, sipiwechihana09@gmail.com

The growth of Information Technology has spawned large amount of databases and huge data in numerous areas. The research in databases and information technology has given rise to an approach to store and manipulate this data for further decision making. In this paper certain data mining techniques were adopted to analyze the data that shows relevance with desired attributes. Regression technique was adopted to help us find out the influence of Agriculture, Service and Manufacturing on the performance of gross domestic product (GDP). Trend and time series technique was applied to the data to help us find out what trend of GDP with respect to service, agriculture and manufacturing sector for the past decade has been. Finally Correlation was also used to help us analyze the relationship among the variables (service, agriculture and manufacturing sector). From the three techniques analyzed, service value added variable was the most prominent variable which showed the strong influence on GDP growth rate.

Keywords: *GDP, Regression, Time-series/trends analysis, Correlation, Data mining, Predictions*

1 Introduction

The development of Information Technology has produced large amounts of databases and huge data in numerous areas. The research in databases and information technology has given rise to an approach that stores and manipulate this data for further decision making. Data mining began its life in specialist applications such as geological research and meteorological research, more recently it has been applied in a number of areas of industry and commerce [1]. Data mining is a process of extraction of useful information and patterns from huge data, it is also called as knowledge discovery process, knowledge mining from data, knowledge extraction or data pattern analysis [2]. [3]Adds that Data mining generally refers to the process of extracting interesting hidden information from available chunks of data, which could otherwise be manually impossible. Despite data mining being relatively a new technology that has not completely matured, there are a number of areas that are already using it on a regular basis. Some of these organizations include retail stores, hospitals, banks, insurance companies. Many of these organizations are combining data mining with such things as statistics, pattern recognition, and other important tools [2]. Data mining in economics is an emerging field of high importance for providing prognosis and a deeper understanding of economic data. Researchers are using data mining techniques in the analysis and prediction of several economic indicators such as GDP. [3]) points out that GDP is one of the measures of national income and output for a given country's economy at a given period of time and adds that the definition of GDP is based on the total market value of all final goods and services produced within the country in a given period of time (normally one year). GDP is also an important statistic that indicates whether an economy is expanding or contracting. To know whether an economy is expanding or not and what sectors contribute the most to

annual GDP growth and to find out how GDP has been performing in a given period of time can be achieved by using data mining techniques. Additionally Successful data mining is based on various investigations of the data using different methods, parameters, and data to find most meaningful relations [4]. [2]Various algorithms and techniques like Classification, Clustering, Regression, Artificial Intelligence, Neural Networks, Association Rules, Decision Trees, Genetic Algorithm, Nearest Neighbor method etc., are used for knowledge discovery from databases. Data mining techniques such as Trend analysis, application of forecast in terms of time series model is found widely in economic development [5], Regression and Correlation helps in finding the patterns, relationships among different economic sectors on how they influence GDP growth to decide upon the future trends on how GDP is likely to perform.

Based on the above information this paper seeks to discuss the analysis that was conducted using data mining techniques on GDP with a view to answer some data mining questions that were proposed to gain a deeper understanding of the dataset, below are the questions that guided the authors on what techniques to use;

1. What is the influence of Agriculture, Service and Manufacturing on the performance of gross domestic product (GDP)?
2. What has been the trend of gross domestic product with respect to service, agriculture and manufacturing sector for the past decade?
3. What is the relationship among the variables service, agriculture and manufacturing sector with respect to GDP as a controlling variable?

1.1. The Significance of GDP

Economic growth of a country is measured by the Gross Domestic Product (GDP) [6]. Gross domestic product (GDP) is the backbone of Zambia's economy and plays a significant role in that it is our economic indicator that shows how the economy performs with respect to many sectors and it also shows whether as a country we are growing or not. According to ([7]) GDP is one of the measures of national income and output for a given country's economy at a given period of time. [8] Adds that the definition of GDP is based on the total market value of all final goods and services produced within the country in a given period of time (normally one year). In a related literature [9] adds that GDP is the total market value of all final goods and services produced in a country in a given year, equal to total consumption, investment and government spending, plus the value of exports, minus the value of imports.GDP is commonly used as an indicator of the economic health of a country, as well as a gauge of a country's standard of living. Since the mode of measuring GDP is uniform from country to country, GDP can be used to compare the productivity of various countries with a high degree of accuracy. Adjusting for inflation from year to year allows for the seamless comparison of current GDP measurements with measurements from previous years or quarters. In this way, a nation's GDP from any period can be measured as a percentage relative to previous periods. [10] GDP is total national income and outcome in relation to commodity and services in a period of time, GDP is one of main economic characteristic since it gives demonstration of the economic activity, higher value of GDP indicates better economic activity while lower value indicates the contrary, expressed in billions of national currency units An important statistic that indicates whether an economy is expanding or contracting, GDP can be tracked over long spans of time and used in measuring a nation's economic growth or decline, as well as in determining if an economy is in recession (generally defined as two

consecutive quarters of negative GDP growth). [11]Noted that GDP is still one of the most important indicators of macroeconomic statistics, it is an effective tool to make people understand and grasp a national (or regional) macroeconomic status, and it is a scientific and effective method to inspect economic policy and assess the important comprehensive index of economic situation.

1.1.1. Overview of Value added

The various sectors of any economy have a contribution to the development of any economy. This is to say that no matter how small the contribution of any sector to the national income of that economy, it adds up to the aggregate income of the economy and thus contributing directly or indirectly to the gross domestic earnings of such economy [12].The value-added measure of GDP adds together the value of output produced by each of the productive sectors in the economy using the concept of value added. Value added is simply the increase in the value of goods or services as a result of the production process. Our study chose the three among the many economic sectors that contribute towards the growth of GDP and these include Service sector, Agriculture and Manufacturing. This section briefly discusses the sectors.

1.1.2. Service Sector

The tertiary sector or service sector is the third of the three economic sectors of the three-sector theory. The others are the secondary sector (approximately the same as manufacturing), and the primary sector (raw materials). The service sector consists of the parts of the economy, i.e. activities where people offer their knowledge and time to improve productivity, performance, potential, and sustainability, which is termed as affective labor. The basic characteristic of this sector is the production of services instead of end products. Services (also known as "intangible goods") include attention, advice, access, experience, and discussion. The production of information has long been regarded as a service, but some economists now attribute it to a fourth sector, the quaternary sector.

The tertiary sector of industry involves the provision of services to other businesses as well as final consumers. Services may involve the transport, distribution and sale of goods from producer to a consumer, as May happen in wholesaling and retailing, or may involve the provision of a service, such as in pest control or entertainment. The goods may be transformed in the process of providing the service, as happens in the restaurant industry. However, the focus is on people interacting with people and serving the customer rather than transforming physical goods.

1.1.3. Manufacturing Sector

Manufacturing sector refers to those sectors which involve in the manufacturing and processing of items and indulge in either creation of new commodities or in value addition. The manufacturing sector is considered to be one of the prominent sectors for the revitalization of the economy in the strategy for Zambia's socio-economic development and poverty reduction. The country's manufacturing sector comprises of companies in food processing, beverages, textiles, leisure and sporting equipment. The activities majorly include the smelting and refining of copper and other metals and metal products, petroleum refining, the production of fertilizers, chemicals, explosives, cement, tobacco products and textiles. The manufacturing industry accounts for a significant share of the industrial sector in developed countries. The final products can either serve as a finished good for sale to customers or as intermediate goods used in the production process. According to [13] Zambia has a relatively diversified manufacturing sector, which is concentrated in the food, beverages and tobacco sub-sector, accounting for 63 per cent of manufacturing sector activities. Wood and wood products, which is the second dominant sub-sector, makes up 11 per cent of the manufacturing sector. It is

followed by the chemicals, rubber and plastics products, which constitutes 9 per cent of the sector.

1.1.4. Agriculture Sector

The Agriculture sector is a sector which comprises establishments primarily engaged in growing crops, raising animals, and harvesting fish and other animals from a farm, ranch, or their natural habitats. Additionally agricultural yield primarily depends on environmental factors such as rainfall, temperature and geographical topology of the particular region. These factors along with some other influence the crop cultivation [4].

2 Data Collection

The study made use of secondary data obtained from World Bank national accounts data, and OECD National Accounts data files which is available on their website.26 years interval (1990 – 2015) [14].Contribution of the identified sectors of the nation's GDP were also obtained from the same source of the same year interval. The three (3) identified sectors include: agriculture, manufacturing, and service value added. The dataset used was loaded in R for analysis, the dataset imported as CSV file, below is the dataset.

Year	GDP-growth	Services-value-added	Manufacturing-value-added	Agriculture-value-added
1990	-0.481072127	28.12360801	36.0610909	20.60479675
1991	-0.036133391	31.56585227	36.74848648	17.42768673
1992	-1.730922033	27.21993643	37.16310495	23.81226939
1993	6.797273924	24.00501499	27.94480912	34.10152067
1994	-8.625441835	43.63323144	10.42293014	13.56260604
1995	2.897668709	45.03132238	10.42252367	16.00057803
1996	6.218546514	47.18508198	12.39794018	15.10412944
1997	3.814007573	47.21561167	12.29002026	15.91910159
1998	-0.3857	50.96601704	12.24400329	18.03400006
	46225			
1999	4.650189789	53.66768149	11.53670967	20.55077743
2000	3.89732287	55.4134235	10.69943865	18.2786512
2001	5.316868319	55.80267884	10.3217762	17.64738381
2002	4.506014409	55.89433695	10.77533324	17.2675372
2003	6.944973873	55.57070429	11.342165	17.45460527
2004	7.032395131	54.18088113	11.17279892	17.29544049
2005	7.23559899	54.15917557	10.88597265	16.13852965
2006	7.903694512	52.14726684	10.32499946	14.4936135
2007	8.352436224	51.88007276	9.489240695	13.23450525
2008	7.773895812	53.59948565	9.241890106	12.52152896
2009	9.220348444	55.23794946	9.305427104	12.3781784
2010	10.29820581	55.89983928	8.022200348	9.972843012
2011	5.564624717	53.33181109	7.955301724	10.21165686
2012	7.597616969	56.27467003	7.48550623	9.861323998
2013	5.059376378	56.52608605	6.587368149	8.759118438
2014	4.695826373	57.39439898	7.314753999	7.271416345
2015	2.919881111	59.43534303	7.929997344	5.250774058

Fig.8.csv file

3 Data Mining Process

Data mining techniques and methods are used in the main related disciplines and technologies from the following areas: Statistical Methods, Decision Tree, Neural Network, Genetic Algorithm and Fuzzy Set, in [1]. The data mining techniques represent such a tool that solves different types of problems from banking and finance domains, by finding patterns, correlations, rules sets, causalities etc., and helps the human analyst in the process of analysis and prediction of some financial tasks evolution [15], furthermore, some data mining packages offer statistical methods, such as principal components, logistic regression, correspondence analysis etc., for financial predictions.

3.1 Statistical Methods

In data mining this technique often involves a certain degree of statistical process, as data sample and modeling to determine assumptions and error control. Including descriptive statistics, probability theory, regression analysis, time series, including many of the statistical methods, data mining plays an important role. For this study the data mining process was conducted with regression, time series/trend analysis and correlation techniques on the dataset.

3.2 Regression and Value Prediction Technique

Regression according to [16]is a data mining (machine learning) technique used to fit an equation to a dataset. [16] Adds that a regression algorithm estimates the value of the target (response) as a function of the predictors for each case in the build data. These relationships between predictors and target are summarized in a model, which can then be applied to a different data set in which the target values are unknown. In a related study [17] says that regression models make use of relationships between the variable of interest and one or more related predictor variables. Sometimes regression models are called causal forecasting models, because the predictor variables are assumed to describe the forces that cause or drive the observed values of the variable of interest. Regression analysis is a statistical technique for modeling and investigating the relationships between an outcome or response variable and one or more predictor or regressor variables. The end result of a regression analysis study is often to generate a model that can be used to forecast or predict future values of the response variable given specified values of the predictor variables .In this study however we used multiple regression which has been outlined below, multiple regression attempts to model the relationship between two or more explanatory variables and a response variable by fitting a linear equation to observed data [18].

3.2.1 Model Specification

In order to produce an observed study we established a functional relationship for Gross Domestic Product (GDP) as a proxy. As a result of this research, the functional model measurement was represented mathematically as follows:

GDP= f (SVA, MVA, AVA, μ)…….. (a)

Where;

GDP – Gross Domestic Product Growth Rate

SVA – Service Value Added

MVA – Manufacturing Value Added

AVA – Agriculture Value Added

μ- error term

In order to understand the relationship and significance of the variables above, the econometric model was expressed as follows from (a) above:

$GDP = \beta_0 + \beta_1 SVA + \beta_2 MVA + \beta_3 AVA + \mu$…… (b)

Where;

GDP – Gross Domestic Product Growth Rate

SVA – Service Value Added

MVA – Manufacturing Value Added

AVA – Agriculture Value Added

μ- error term β0= Constant term and

β1= the coefficient of the independent variable (SVA) which $\beta_1 > 0$

β2= the coefficient of the independent variable (MVA) which $\beta_2 > 0$

β3= the coefficient of the independent variable (AVA) which $\beta_3 > 0$

3.2.2 Identification of Variables

- The independent variables are Service Value Added, Manufacturing Value Added, and Agriculture Value Added.
- The dependent variable is Gross Domestic Product Growth Rate

Summary of the output is shown below

Table 1 Summary of Output

Regression					
Multiple R	0.53491				
R Square	0.286129				
Adjusted R Square	0.188782				
Standard Error	3.688029				
Observations	26				
	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	119.9368	39.97894	2.939291	0.055642
Residual	22	299.2343	13.60156		

Table 2 Regression Summary

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-14.4138	13.39847	-1.07578	0.293682
Services value added	0.303039	0.198561	1.526175	0.141216
Manufacturing value added	0.012526	0.199203	0.06288	0.95043
Agriculture value added	0.246642	0.182411	1.352122	0.190074

3.2.3 Interpretation of Results and Predictions

Constant Interpretation

The intercept value of -14.4138 was found, mechanically interpreted, means that if the values of SVA,MVA, and AVA rate were fixed at zero, the GDP growth rate would reduce by 14.4138 per year. This implies that, GDP growth rate will still decrease by 14.4138 when service value added, manufacturing value added and agriculture value added involved are held constant.

Coefficients Interpretation:

1. Service Value Added

0.303039 is the partial regression coefficient of SVA and tells us that with the influence of MVA and AVA held constant, as SVA increases by 1 percent, GDP growth rate in a year will increase by 0.303039. The coefficient was positive and statistically significantly with t-test value of 1.526175 at 0.05 level of significance. This means that, if service value added increases by 1 percent, GDP growth rate will increase by 30.3 percent. The service sector is a very promising sector to improve Zambian GDP growth rate, thus the government need to improve and invest more in the service industry. This implies that, if the service sector increases it value added by at least 1 percent, then the Zambian GDP growth rate is likely to increase by more than 30 percent

2. Manufacturing Value Added

In the model, the partial regression coefficient of MVA was found to be 0.012526 which implies that for every 1 percent increase in manufactured goods, GDP growth rate will increases by 0.012526. The coefficient was positive and statistically significantly with t-test value

of 0.199203 at 0.05 level of significance. This economically means that, if manufacturing industry continues to produce goods for domestic and foreign consumption then GDP growth rate will continue increasing by 1.2 percent. This means that, government has a lot to do in manufacturing industry because it is not contributing more to GDP according to its potential. If manufacturing industry is heavily funded, then it will bring more of foreign exchange to the country which will make the country to have surplus trade balance and in return boosting the national income (GDP). As a result, if manufacturing industry increases it value added by at least 1 percent, then the Zambian GDP growth rate is likely to increase by more than 1.2 percent

3. Agriculture Value Added

The partial coefficient of AVA was found to be 0.246642 and tells us that holding the influence of SVA, and MVA coefficient, constant, GDP growth rate in a year will increase by 24.6642 percent at 0.05 level of significance. Economically, it implies that GDP growth rate will still increase by 24.6642 percentage rate. The government of Zambia should invest more or improve budget allocation to agriculture sector as it is a viable sector which can improve the GDP growth rate. If agriculture industry is well financed through the processing of value added on raw materials, then more of foreign exchange will be brought to the country and increase the GDP growth rate. This shows that, if agriculture sector increases its value additional by more than 1 percent, the Zambian GDP growth rate will likely to increase by 24.6 percent or more than 24.6 percent.

4. Coefficient of Variation R-Squared (R^2)

The coefficient of variation for this econometric model was found to be 0.286129. This implies that 27 percent of the disparity on total trips traveled was attributed to the variation on service value added, manufacturing value added and agriculture value added and the remaining 73 percent of the variables in the model are not explained. The R^2 is greater than 0.2 as a result the fitted regression line is of good fit.

5. F-test and its significance:

The F-test shows that the relation of the disparity in the regression to the ratio of the variation of the residual or errors was found as 2.939291. This value found was greater than the tabulated F-value of 0.055642 at 0.05 level of significance. This implies that the econometric model was significantly correct.

4 Trend Analysis using time series

Time series data type, also called chronological series or simply time series represent results of measurements made on the characteristics of a unit of population studied, over time, at successive moments of its evolution in some time intervals [19]. Using this technique we wanted to find out the trend of gross domestic product with respect to service, agriculture and manufacturing sector for the past decade. Below is plotted graph showing the trend with respect to the aforementioned variables.

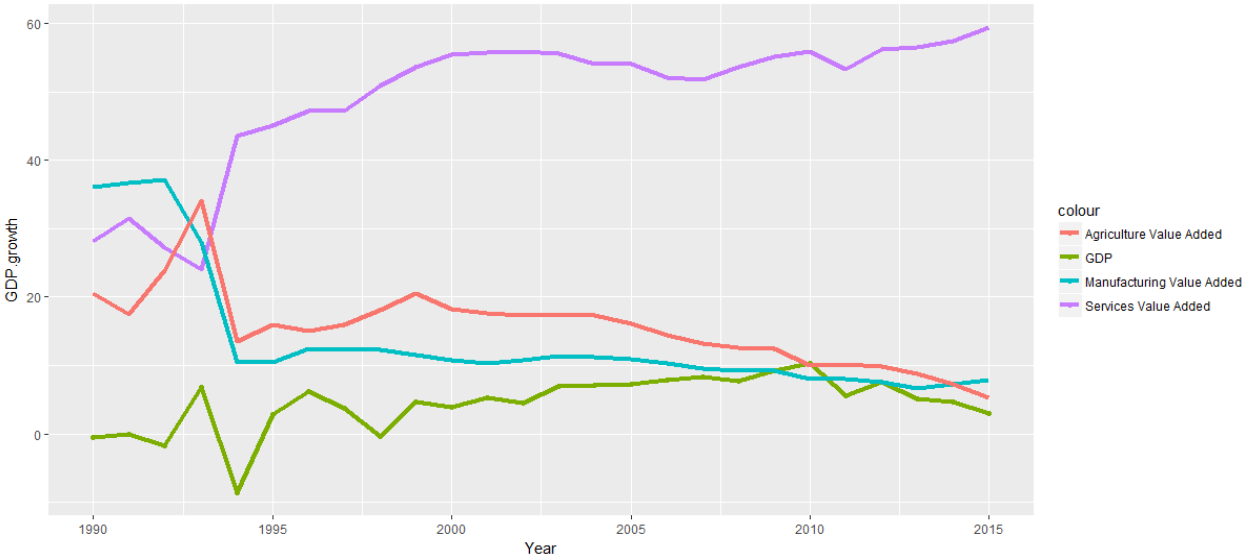

Fig. 9. GDP Growth, Service Value added, Manufacturing Value Added and Agriculture Value Added trend analysis

4.1. Interpretation and Results Prediction

4.1.1. Service Value Added

The service sector consists of the parts of the economy, i.e. activities where people offer their knowledge and time to improve productivity, performance, potential, and sustainability, which is termed as affective labor or a system supplying a public need such as transport, communications, or utilities such as electricity and water etc. The Zambian service sector has the potential to boost the Zambian GDP growth rate, because it has kept increasing for the past two decade. The Zambian government needs to start providing and training skilled human capital and improve on a number of services within the economy as it has shown potential to make GDP grow. The government can do more in the service sector to have a desirable GDP growth rate.

4.1.2. Manufacturing Value Added

The manufacturing sector in Zambia accounts for about 11 percent of the country's Gross Domestic Product (GDP) and has been growing at an average annual growth rate of three (3) percent in the last three years. Growth in the sector is largely driven by the agro processing (food and beverages), textiles and leather subsectors. Secondary processing of metals in another main activity in the sector, including the smelting and refining of copper, and this has led to the manufacturing of metal products. Fertilizers, chemicals, explosives and construction materials such as cement are also produced in the sector. Other activities include wood products and paper products. The manufacturing activities in the country are undertaken by the private sector with government playing a proactive role. The sector is of vital importance in relation to the country' macroeconomic strategy for encouraging broad based economic growth. In this regard, the Government has put in place measures to support manufacturing activities, such as the establishment of Multi-Facility Economic Zones (MFEZs) and Industrial Parks (these are industrial areas for both export orientated and domestic orientated industries, with the necessary support infrastructure installed), and provision of sector-specific investment incentives. Government also promotes small and medium enterprises in rural and urban areas so as to enhance labor intensive light manufacturing activities in these areas. In order for Zambia manufacturing industry to improve or growth, value additional should the key factor to implement as most of the goods are been sold in raw form or in unfinished

products. And if the Zambian government can have a deliberate policy to have the manufacturing industry add value to unfinished products before exporting them as to enjoy the full value price for the product which in turn will improve the GDP growth rate.

4.1.3. Agriculture Value Added

For the past two decades, the Zambian agriculture sector has not been doing well, as the result, the productivity has continued to decline because of a number of factors such as: over dependence of small scale farmers on farmers input support programme (FISP), government policies, climate and over dependence on copper. The Zambian government has continued to invest in the agriculture sector, according to Ministry of Agriculture and Livestock acting permanent secretary, who also praised the industry's role in the economy in a recent statement issued in Lusaka. The performance of the industry in the just ended year was significant with the growth of GDP about seven per cent. The Zambian government has announced it will implement plans to boost crop and livestock production to help further strengthen the agriculture sector. It said it would invest in extension services, irrigation, and research and development in order to improve crop yield. The government also claimed it was working to increase local production of key inputs such as fertilizers, seeds and pesticides, along with restocking the livestock sector with a view to improve its performance. If the government can continue boosting the agriculture sector, it can increase the gross domestic product by a bigger percentage, because this sector has a potential to boost the Zambian economy and a lot has to be done to improve the value creation or adding of value to the raw material as to boost the value creation in agriculture sector which in turn will boost the Zambian GDP.

4.1.4. GDP Growth

Gross domestic product (GDP) is the monetary value of all the finished goods and services produced within a country's borders in a specific period of time. Though GDP is usually calculated on an annual basis, it can be calculated on a quarterly basis as well. GDP includes all private and public consumption, government outlays, investments, private inventories, paid-in construction costs and the foreign balance of trade (exports are added, imports are subtracted). From the graph, we can deduce that, they have been some reduction in Zambian GDP growth rate for the past two decades. The Zambian GDP growth rate will continue to decrease, unless the Zambia government improves on the service, agriculture and manufacturing sector. In order to improve the Zambian GDP, the government should focus on service, agriculture and manufacturing sector.

5 Correlation Technique

Correlation technique was chosen to help us find out the relationship among the variables (service, agriculture and manufacturing sector) with respect to GDP as a controlling variable.

The table below shows correlation relationships amongst the variables

Table 3. Variable Correlation Relationship

	Year	**GDP growth**	**Services Value added**	**Manufacturing value added**	**Agriculture value added**
Year	1	0.4849692	0.7734849	-0.5000892	-0.2820188

GDP growth	0.484969	1	0.4491033	-0.4298382	-0.1623941
Services Value added	0.773485	0.4491033	1	-0.7928582	-0.5073032
Manufacturing value added	-0.50009	-0.4298382	-0.7928582	1	0.5588934
Agriculture value added	-0.28202	-0.1623941	-0.5073032	0.5588934	1

Table 4. Correlation Relationship among service, manufacturing and agriculture

Correlations Summary showing the relation among service, manufacturing and Agriculture variables					
Control Variables			Service	Manufacturing	Agriculture
GDP	Service	Correlation	1.000000	-0.792**	-0.507**
	Manufacturing	Correlation	-0.792**	1.000000	0.558**
	Agriculture	Correlation	-0.507**	0.558**	1.000000
**. Correlation is significant at 0.01 level					

In statistics, the correlation coefficient *r* measures the strength and direction of a linear relationship between two variables . The value of *r* is always between +1 and –1. To interpret its value, we follow values the correlation *r* is closest to:

- Exactly –1. A perfect downhill (negative) linear relationship
- –0.70. A strong downhill (negative) linear relationship
- –0.50. A moderate downhill (negative) relationship
- –0.30. A weak downhill (negative) linear relationship
- 0. No linear relationship
- +0.30. A weak uphill (positive) linear relationship
- +0.50. A moderate uphill (positive) relationship
- +0.70. A strong uphill (positive) linear relationship
- Exactly +1. A perfect uphill (positive) linear relationship

5.1. Interpretation of Results and Prediction

5.1.1 Service Value Added

There is a strong negative correlation (-0.792) between service value added and manufacturing, and at the same time, there is a weak negative correlation (-0.507**)

between service value added and agriculture value added.

5.1.2 Manufacturing Value Added

There is a strong negative correlation (-0.792) between service value added and manufacturing, and at the same time, there is a moderate correlation (0.558**) between manufacturing value added and agriculture value added.

5.1.3 Agriculture Value Added

There is a negative correlation (-0.507**) between agriculture value added and service value added, and at the same time, there is a moderate correlation (0.558) between manufacturing value added and agriculture value added.

From the above correlation analysis, manufacturing and agriculture value added seems to correlate and it's very possible that if these two sectors are boosted through value addition, the Zambian GDP will grow at a steady increasing rate. These two sectors are viable and government has to pump in a lot of funds (investment) as they will positively increase Zambian GDP growth rate.

6. Limitation of the techniques

6.1. Multiple Regression Model

Multiple Regression Model has a problem of problem of multicollinearity (One or more variables explaining the same factor i.e. Mother and a lady) as it may exist between or among variables. In such an event, one or more variables should be eliminated to reduce or eliminate multicollinearity. As a result, it helps to check or eliminate the highly correlated independent variables from the analysis, recognizing that the two variables essentially are measuring the same factors and there is no need of having both variables. Shanta Khumari in [12] also noted that multicollinearity causes major interpretative problems in regression analysis, such as wrong sign problem, produces unstable and inconsistent estimates of parameter, insignificant regression coefficients where in fact it is significant and it is thus very essential to investigate and detect the presence to reduce the destructive effects of multicollinearity. Furthermore, it is very difficult to explain the coefficient of variables in a multiple regression model as it may not be simplified in layman's language.

6.2. Time series analysis

Time series analysis helps to show the trends about the variables over a period of time (It shows the up and downs about a variable). The major limitation about time series analysis is that, it does not show the statistical figures about the variables and which makes it difficult for the researchers to forecast the future behaviour for the variables.

6.3. Correlation Analysis

Correlation analysis is an essential analysis when one wants to see the relationship between or among variables that are being analysed (It ranges between -1 to +1) and it helps to see how variables correlate among each other. The challenge which this techniques brings during correlation analysis is that, the analysis only shows the correlation among variables and makes it difficult for the researchers to predict the future behavior of variables based on correlation of variables. The analysis does not show the coefficients of variables or the statistical figure which may add more value or significance on the analysis.

Conclusions and Recommendations

7 Conclusion

In this paper certain data mining techniques were adopted to analyze the data that shows relevance with desired attributes. Regression technique was adopted to help us find out the influence of Agriculture, Service and Manufacturing value added on the performance of gross domestic product (GDP). Trend and time series technique was applied to the data to help us find out what trend of GDP with respect to service, agriculture and manufacturing sector for the past decade has been. Finally Correlation was also used to help us analyze the relationship among the variables (service, agriculture

and manufacturing sector) with respect to GDP as a controlling variable. Many data mining techniques were however not employed on the GDP dataset because the study had set out objectives that were only going to be solved with the aforementioned techniques. From the three techniques analyzed, service value added variable was the most prominent variable which showed the strong influence on GDP growth rate (Under multiple regression service value added had highest positive coefficient, under time series – it had a strong positive correlation and under correlation analysis – GDP growth rate and service value added were more correlated than any other variable). From the three techniques used, regression analysis was easy to use for predicament influence of independent variables on dependent variable as compared to the other two techniques used. There are however other variables which affect the GDP growth rate apart from the three independent variables which were analyzed. The researchers had to pick three independent variables as to reduce the level of complication in analysis and making it simpler for the reader to understand. Future work however will consider utilizing other data mining techniques that will help us gain a deeper understanding and mining of GDP data on a different perspective. Overall as scholars we have come to appreciate data-mining because it helps in making informed decisions, derive new knowledge, analyze the data and make predictions that in turn if well utilized as far as the growth of the economy is concerned can greatly contribute positively and certain measures can be put in place to help the economy grow for the betterment of our nation Zambia.

8Recommendations

Agriculture Sector

The limited financial and non-financial support available has curtailed the overall growth of some nonfarming sub-sectors of agriculture, which has reduced their overall contribution to wealth (income) and employment creation. Within these sectors as well, it is important to invest in forward and backward integration to increase the value of products exported and the number of jobs created. Furthermore, for certain sub-sectors, such as livestock and wood processing, there is a need to develop the skills endowment levels of the labor force [13]. Inorder to boost value addition in the agriculture sector, following should be considered by government:

i. Review the national agriculture policy to prioritize livestock development and fish farming by supporting both production and market infrastructure for value added production within these sub-sectors.
ii. Reform rural agricultural co-operatives into business-oriented entities focused on adding value to agricultural commodities. In particular, this should be done by pooling resources to upgrade or purchase new machinery and technology.
iii. Increase investments to expand rural irrigation infrastructure to lessen smallholder farmer dependency on rain-fed agriculture.
iv. Strengthen agricultural extension and veterinary services. Also, link livestock development research to the smallholder sector to help control animal disease outbreaks and enhance the productivity of this sector.
v. Resolve institutional issues, such as improving the collection levels and supply of quality raw hides and skins, and ensuring that producers earn competitive prices to stimulate growth and counter the rampant smuggling of these vital agricultural inputs.
vi. Improve inter-ministerial coordination across agriculture,

industry, trade, and employment, to ensure the development of policies and a regulatory framework that is conducive for sectoral growth and development.

vii. Accelerate the implementation of land reforms, especially under the Customary Tenure. These reforms are critical to guaranteeing land rights which can, in turn, allow for capital investment in land and provision of credit.

viii. Set aside funds or provide an enabling regulatory environment to attract investments in suitable R&D that would enable Zambian manufacturers to gain increasing access to competitive global markets.

ix. Prioritize entrepreneurship development, technical skills, and management training across this sector to promote productivity gains and accelerate the commercialization of smallholder agriculture.

x. Realign the agricultural sector budget to give equal emphasis to crop agriculture, horticulture, and livestock development.

xi. Promote sustainable exploitation of fisheries resources and increased fish production [20].

Manufacturing Sector

Under the manufacturing sector for value addition, there is need for strengthening the sector and putting in place certain measures and viable strategies such as the ones outlined below:

i. Accelerating the policy and enactment of the commerce, trade and industrial related policies to prioritize wealth and employment creation through value added manufacturing.

ii. Parallel diversification of the mining activities towards industrial minerals such as iron ore and steel milling, to introduce a local capital goods sector. The policy goal should be to hasten the growth of a vibrant and competitive iron and steel industry to become the anchor of industrial policy. This should support the growth of automotive components, and medium and heavy industry commercial vehicles to support the growth and expansion of mining and mineral beneficiation.

iii. Embark on strategic partnerships and joint ventures with public and private, as well as foreign and domestic investors to provide skills, technical expertise, and technology to target strategic sectors and integrate them into the global value chains.

iv. Prioritize research and development and provide smarter subsidies to accelerate the development of the leather and leather products subsector.

v. Establish specialized research and training Institutes and technical colleges to provide the necessary technology and local capability to support the development of a competitive agro-industry, textile and clothing, iron and steel, and leather and leather products industry.

vi. Prioritize the development of the forest and forestry sector through consistent funding of tree planting to provide the much-needed raw material to the wood and furniture, and paper and paper products, sub-sectors.

vii. Invest in the expansion of green and energy-saving industries, such as hydroelectricity, solar water heating, concentrated solar power, and improvements in energy efficiency [13].

viii. Promote entrepreneurship development and training at all levels of the education system.

Service sector

The Zambian government needs to start providing and training skilled human capital and improve on a number of services within the economy as it has shown potential to make GDP grow. The government can do more in the service sector to have a desirable GDP growth rate. Additionally Government needs to develop and adopt a comprehensive strategy that will see the creation of an environment for new industries to spring up, add-value and diversify the products for export

References

[1] D. Bhardwaj, "ANALYSIS OF DATA MINING TRENDS, APPLICATIONS, BENEFITS AND ISSUES," International Journal of Computer Science and Communication Engineering, vol. 5, no. 1, pp. 1-5, 2016.

[2] M. R. Bharati, "DATA MINING TECHNIQUES AND APPLICATIONS," Indian Journal of Computer Science and Engineering, vol. 1, no. 4, pp. 301-305, 2011.

[3] Z. M. Fathimath and A. Geetha, "Analysis of Data Mining Techniques and its Applications," International Journal of Computer Applications, vol. 140, no. 3, pp. 6-14, 2016.

[4] A. B. Vinayak, P. A. Prachi, S. P. Ravina, S. P. Sonal, U. N. Tejaswini and M. J. Anaya, "Analysis and Prediction in the Agriculture data using data mining techniques," International Journal of Research In Science & Engineering, pp. 386-393, 2017.

[5] L. Yang and H. Bin, "The Establishment and Analysis of Time Series Model of Per Capita GDP in Yunnan, China," pp. 1-3, 2010.

[6] A. Naveed, A. Kartheek, P. Tapan and V. Meghana, "Predicting Economic Recession using Data Mining Techniques".

[7] R. K. Alex, "The Factors Affecting Gross Domestic Product (GDP) in Developing Countries: The Case of Tanzania," European Journal of Business and Management, vol. 5, no. 4, pp. 148-158, 2013.

[8] M. Gregory, Principles of Economics, Cengage Learning, 2014, p. 880.

[9] R. Barbara and Z. Andrea, "EXCESS MONEY GROWTH AND INFLATION DYNAMICS," pp. 1-42, may 2007.

[10] S. Siti, H. ,. The and A. , "Prediction System of Economic Crisis in Indonesia Using Time Series Analysis and System Dynamic Optimized by Genetic Algorithm," International Conference on System Engineering and Technology, pp. 1-6, 2012.

[11] G. LEI and Z. HUI-BEN, "THE ANALYSIS OF AFFECTING GDP GROWTH FACTORS BASED ON EVIEWS ECONOMETRIC MODEL," 2013.

[12] V. E. Ikenna, Olugbenga and O. K, "Sectoral Contribution to Nigeria's Gross Domestic Product (GDP) Growth Rate: A Study of Multicollinearity in Aggregated Time Series Data," Journal of Scientific Research & Reports, vol. 11, no. 1, pp. 1-13, 2016.

[13] ZHDR, "Zambia Human Development Report 2016," UNDP, Lusaka, 2016.

[14] a. O. N. A. d. f. World Bank national accounts data, "GDP growth (annual %)," [Online]. Available: http://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=ZM. [Accessed May 2017].

[15] M.Cărbureanu, "The Analysis of Currency Exchange Rate Evolution using a Data Mining Technique," Economic Sciences, vol. LXIII, no. 3, pp. 105-112, 2011.

[16] G. Swati, "A Regression Modeling Technique on Data Mining," International Journal of Computer Applications, vol. 116, no. 9, pp. 27-29, 2015.

[17] S. SAIGAL and D. MEHROTRA, "PERFORMANCE COMPARISON OF TIME SERIES DATA USING PREDICTIVE DATA MINING TECHNIQUES," Advances in Information Mining, vol. 4, no. 1, pp. 57-66, 2012.

[18] C. Hao and Y. Lixin, "Using Multi-regression to Analyze and Predict Road Traffic Safety Level in China," in The 3rd International Conference on Transportation Information and Safety, Wuhan,, 2013.
[19] J. Popeanga, "Data Mining Smart Energy Time Series," Database Systems Journal, vol. 1, no. 1, pp. 14-22, 2015.
[20] M. o. F. a. N. Planning, "SIXTH NATIONAL DEVELOPMENT PLAN 2011-2015," Ministry of Finance and National Planning, Lusaka, 2011.
[21] G. Matjaž and J. Krivec, "Demographic Analysis of Fertility Using Data Mining Tools," Informatica , pp. 147-156, 2008.