RUNNING HEAD: REGRESSION ANALYSIS

Regression Analysis

 Author

[Name of the Institution]

Regression Analysis

***1. Purpose***

 This report is aimed at finding whether there is a significant relationship between life satisfaction and GDP. The findings of this report will be used to develop some policies to increase the living standard of people if there is a significant relationship between the two variables. If there is no significant relationship between these variables, the policy makers will have to work further to find the factors which affect the life satisfaction. Some researchers have studied the relationship between life satisfaction and some of the economic variables such as GDP. Well being of people is the ultimate aim of any economic policy which helps to run the market economy and democracy at the same time. It is important to study if any policy is successful in attaining a higher life satisfaction for the people. A government will have to continuously monitor its policies to find the best option available to it.

***2. Background***

 A study analysed the connection between life satisfaction and GDP keeping the individual countries as subjects. The researchers found that there is a positive relationship between life satisfaction and the GDP in almost all the countries studied. There is a difference in nature of this relationship in the poor and rich countries. In the poorer countries, the relationship between life satisfaction and GDP is found to be positive whereas in the richer countries, this relationship is found to be negative. The study used a regression model to analyse the relationship (Proto, 2013).

 Another study analysed the connection between personal life satisfaction and GDP. The applicability of the Easterlin’s paradox was checked. This paradox states that there is no effect of increase in the country’s income on people’s life satisfaction. The study rejected the concept put forward by Easterlin’s paradox and finds a positive relationship between living standard and total product of a country. There is a cross country analysis to see if there is a difference between countries with respect to this relationship (Gaižutis, 2010).

***3. Method***

 Two variables are included in this study namely the average life satisfaction and actual GDP per unit of population. The dependent variable is the average life satisfaction which means that we will check the impact of change in GDP on life satisfaction of people. The economic variable is considered to be independent which means that there will be many values of this variable (Groebner, P.W.Shannon, & Fry). The impact of change in the values of this variable will tell us about the relationship between the two variables. The sources of data are the semi-annual Euro barometer surveys which were undertaken in all the EU countries. There are many techniques used to select the people for this study and their interviews have been undertaken. GDP data has been taken from Eurostat. Purchasing power terms are used to gather the GDP data so that the difference between their respective price levels is catered. Data analysis will be undertaken with the help of regression because this will help us to identify the direction and quantity of the relationship between two variables.

***4. Results***

|  |  |  |
| --- | --- | --- |
|  | **LS** | **GDP** |
| **Means**  | 6.591 | $39,011.51 |
| **SD** | 0.745327 | $14,006.21 |
| **Min** | 5.2 | $17,122.53 |
| **Max** | 7.5 | $86,788.14 |

 Table 1

There are countries which have very high or very low values for the life satisfaction. The highest value for the variable is observed in more than one country. Norway, Switzerland and Iceland are the countries which show highest values for life satisfaction whereas Greece and Portugal have shown the lowest values are for Greece and Portugal. When we see the GDP data, the highest values for the variable correspond to Luxembourg and the lowest to Mexico. Arithmetic mean is used to represent the average for both variables and is calculated by dividing the sum of values by the number of values (Healey). The major drawback of this calculation is that it is affected by any extreme values present in the data. Standard Deviation shows the level of dispersion in the data which measures the distance of each value from the mean (Janes, 1999). A higher standard deviation can show the presence of outliers in the data. The standard deviation for the life satisfaction variable shows that the dispersion is low. The difference between highest value and mean for the GDP variable is very high. This shows a possibility of outliers in the data.

 Fig 1.

The above diagram depicts the scatter plot between the two variables. The GDP per capita is shown on the x axis and life satisfaction index is shown on the Y axis. It can be seen that there is a positive relationship between the two variables. The trend line in this diagram passes through some of the points and there are some points which are away from the trend line. These points can be considered as outliers. A scatter plot which shows widely scattered points shows that there is no relationship between the two variables (Friendly & Denis, 2005). The following equation shows regression relationship between the two variables.

$$Satisfaction= ∝ + β GDP$$

The above regression equation has two important elements other than the variables. The alpha is known as the intercept. This is the value of dependent variable when the value of independent variable is zero. The other value of regression equation is beta which is also called slope of the equation (A.Lind, G.Marchall, & A.Wathen, 2012). A unit change in the independent variable will bring a change in the dependent variable which is equal to the value of beta. The analysis can be done to get a value of level of satisfaction for any given value of GDP per capita. The major regression statistics are shown in the table below,

|  |
| --- |
| ***Regression Statistics*** |
| Multiple R | 0.590689172 |
| R Square | 0.348913697 |
| Adjusted R Square | 0.329183809 |
| Standard Error | 0.610448073 |
| Observations | 35 |

 Table 2

The table above shows various statistics related to the regression model. The power of a model explains the variation in the dependent variable caused by a change in the independent variable. These statistics suggest that almost 35% variation in the life satisfaction is explained by the GDP. The following table shows the magnitude of the relationship between the level of life satisfaction and GDP.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* |
| Intercept | 5.365181482 | 0.309314 | 17.34542 | 3.78E-18 |
| X Variable 1 | 0.314329537 | 7.47E-06 | 4.205297 | 0.000187 |

 Table 3.

 The above table shows that there is a positive relationship between life satisfaction and GDP per capita. The value of intercept shows that life satisfaction scale of 5.36 corresponds to a value of 0 in GDP. The value of coefficient shows that there will be a 0.314 unit of change in the life satisfaction as a result of a unit change in the GDP.

 Satisfaction = 5.365+0.314 GDP

 The relationship between the two variables is statistically significant because the p-value is less than the level of significance, i.e. 0.05.

 A good fit is a situation when a regression line passes through a majority of data points in a scatter diagram (Miles & Shevlin, 2001). We can see that the trend line passes through most of the data points which makes the model a good fit. There are equal number of points above and below the trend line which shows an equal spread.

 When the outliers are removed for GDP, the results are shown with the help of following table and diagram.

Fig 2.

 The goodness of fit is improved when outliers from GDP are removed because there are lesser values which lie away from the trend line (Forero, 2010). There is also an increase in the power of model to 0.49 which means that half of the variation in life satisfaction is explained by GDP. The following equation shows the relationship between satisfaction and GDP.

 Satisfaction = 4.55 + 0.551 GDP

***5. Discussion***

 The relationship between life satisfaction and GDP is significant in terms of statistics. This method allows us to know the magnitude and direction of any given relationship. This method has been used by the past researchers to study the same relationship. The downside of this method is that it is affected by the presence of outliers. Intercept, coefficient and goodness of fit are all affected by the presence of outliers in the data. These findings are consistent with the literature presented in the literature review section. Efforts to increase GDP will result in an increase of life satisfaction of the people.

***6. Recommendations***

 The first recommendation is to add some more variables to the regression equation so that it can be more meaningful. The assumptions of regression analysis should also be checked (Osborne & Waters, 2002).

 There can be an analysis in which time is involved. This will allow the researcher to see if there is any set pattern regarding the relationship.

 The analysis should also consider deletion of outliers which lie at the lower end of the data to get an even better regression model.

 A major issue in the regression analysis is the presence of any significant relationship between the independent variables. This will affect the relationship between dependent and independent variable. If there is a relationship between the independent variables, a researcher will have to delete any one of them to get better results. One variable can have significant relationship with more than one variable at the same time.

# ***Bibliography***

A.Lind, D., G.Marchall, W., & A.Wathen, S. (2012). *Basic Statistics for business and economics.* McGrawhill educaton.

Forero, A. M.-O. (2010). *https://www.researchgate.net/publication/283839108\_Goodness-of-Fit\_Testing/link/59dfed0045851537160159af/download.* Retrieved from https://www.researchgate.net: https://www.researchgate.net/publication/283839108\_Goodness-of-Fit\_Testing/link/59dfed0045851537160159af/download

Friendly, M., & Denis, D. (2005). The Early Origins and Development of the Scatterplot. *Journal of the History of the Behavioral Sciences*.

Gaižutis, M. D. (2010). Relation between GDP and life satisfaction in the European Union. *EKONOMIKA, 89*(1), 9-21.

Groebner, D., P.W.Shannon, & Fry, P. (n.d.). *Business Statistics.* Pearson.

Healey, J. F. (n.d.). *Statistics: A tool for social research.* Nelson Education.

Janes, J. (1999). "Descriptive statistics: where they sit and how they fall. *Library Hi Tech*, 405.

Miles, J., & Shevlin, M. (2001). *Applying Regression and Correlation: A guide for students and researchers.* SAGE.

Osborne, J. W., & Waters, E. (2002). Four Assumptions Of Multiple Regression That Researchers Should Always Test. *Practical Assessment: Research and Evaluation*, 1-6.

Proto, E. (2013). A Reassessment of the Relationship between GDP and Life Satisfaction. *PLOS one, 8*(11).