The data provided contains information relating to *Gross Domestic Product* and *life expectancy (years*). This task involves the investigation and analysis of a variety of questions using the data. Students will analyse associations between *life expectancy (years*) and *Gross Domestic Product (GDP)*. GDP provides an economic snapshot of a country, it is used to estimate the size of an economy and growth rate giving a measure of the overall wealth of the country.

Use the data from the table below to investigate the relationship between *life expectancy (years*) and *Gross Domestic Product (GDP)* for these 20 countries.

|  |  |  |
| --- | --- | --- |
| **Country** | **Gross Domestic Product ($)** | **Life Expectancy (Years)** |
| Afghanistan | 1559 | 67 |
| Argentina | 19502 | 76 |
| Bahrain | 22943 | 77 |
| Bangladesh | 3402 | 71 |
| Belarus | 18136 | 74 |
| Belgium | 25301 | 83 |
| Botswana | 14190 | 72 |
| Bulgaria | 16373 | 74 |
| Cambodia | 3086 | 69 |
| Chile | 21589 | 80 |
| China | 10351 | 69 |
| Colombia | 9782 | 69 |
| Congo | 8644 | 68 |
| Costa Rica | 8075 | 68 |
| Croatia | 6937 | 67 |
| Cuba | 6368 | 67 |
| Denmark | 5799 | 67 |
| Djibouti | 5230 | 66 |
| Dominica | 4661 | 66 |
| Dominican Republic | 4092 | 66 |

*Use the entire list provided to investigate the following tasks.*

**Task:** Is there a relationship between a country’s between *life expectancy (years*) and *Gross Domestic Product*?

***Examine the data provided and answer the following questions in the development of your overall report.   
Show all working or briefly explain the steps used in finding your answers.***

**Your investigation may include the following:**

1. Construct a scatterplot of between *life expectancy (years*) and *Gross Domestic Product* for the given data.
2. Perform a regression analysis that enables *life expectancy* for a country to be predicted from the *Gross Domestic Product* of the country. Write a brief report outlining your findings, including relevant regression statistics.
3. Looking at the data for a select country (Afghanistan). Using the least squares regression line predict the *life expectancy* for Afghanistan from the *Gross Domestic Product.*  Compare this answer with the actual and hence calculate the residual value.
4. Complete a residual analysis, including a residual plot based on your regression line and comment on whether the linear model is suitable.
5. Based on your analysis of the residual plot, apply two appropriate transformations and explain whether they are effective.
   1. Using your most suitable least squares regression line predict the *life expectancy* (years) for a country with a *Gross Domestic Product* of $10 000
   2. Using the most suitable least squares regression line predict a country’s *Gross Domestic Product* for a country with a *life expectancy* of 90 years.
   3. Comment on the reliability of these predictions.