


Name:	Prashil Deepak Kadam
UID:	2021600031
Experiment No:	2
Aim:	To visualize data from a SocioEconomic dataset.
Dataset link:	<a href="https://github.com/its-me-abhishek/global-data-dashboard/blob/main/world-data-2023.csv">https://github.com/its-me-abhishek/global-data-dashboard/blob/main/world-data-2023.csv</a>
Results / Outputs:	<div> <p>Sum of Population by Official language</p>  <p>Sum of Minimum wage by Country</p>  <p>Sum of Population by Official language</p>  <p>Word map of Population and the official language spoken It is clearly visible that Hindi and Chinese are the most spoken languages</p> </div>

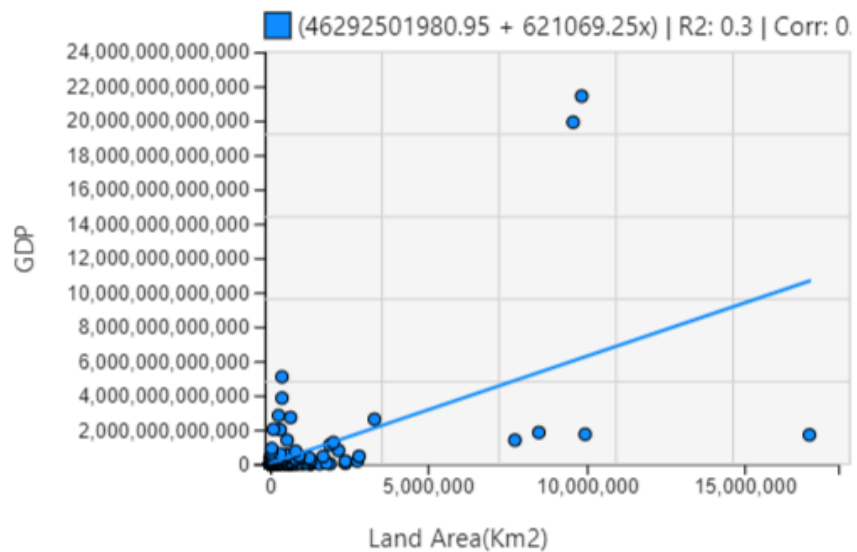
Sum of Minimum wage by Country



Word map of Minimum wage by country

Observing the word map it can be deciphered that Luxembourg, Australia, Monaco are one of the top countries with the highest sum of minimum wages.

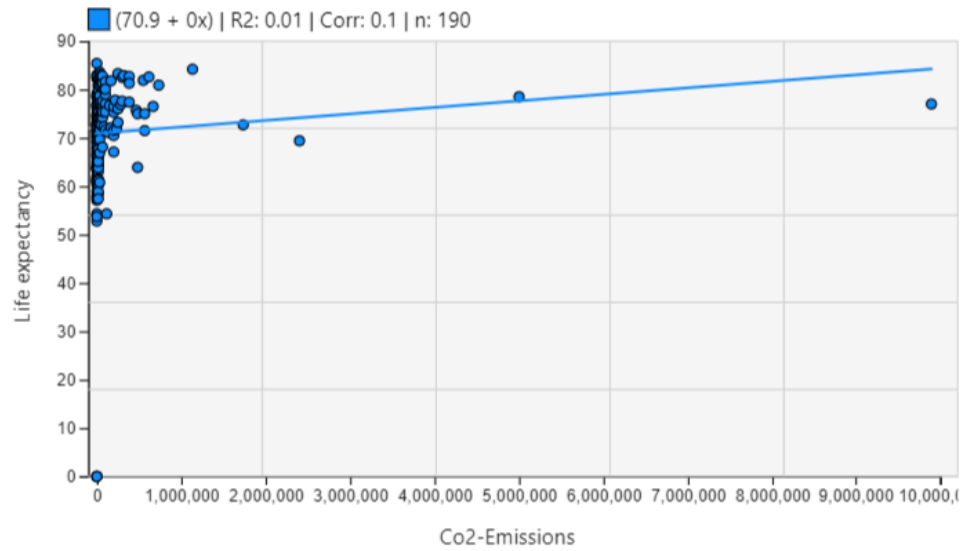
GDP and Land Area(Km2)



Linear Regression plot of GDP vs Land Area(km2)

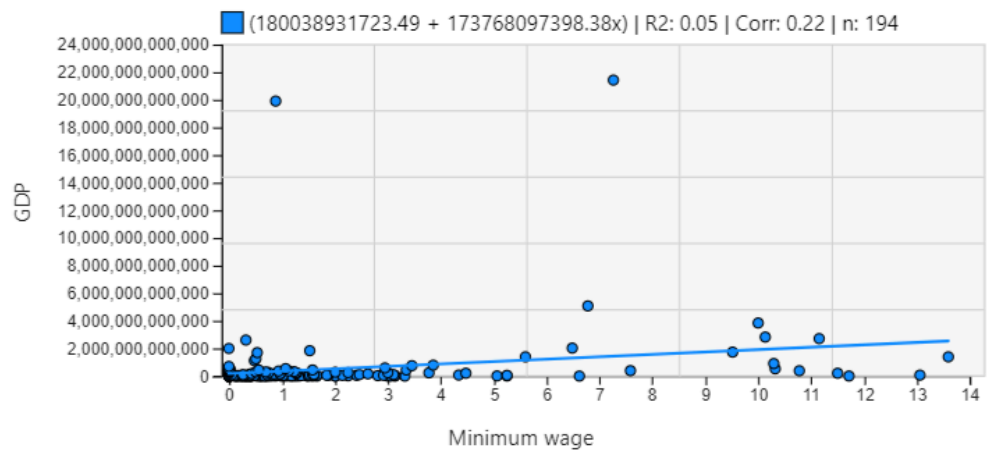
It can be seen that a line equation is formed and correspondingly the line is plotted.

Life expectancy and Co2-Emissions

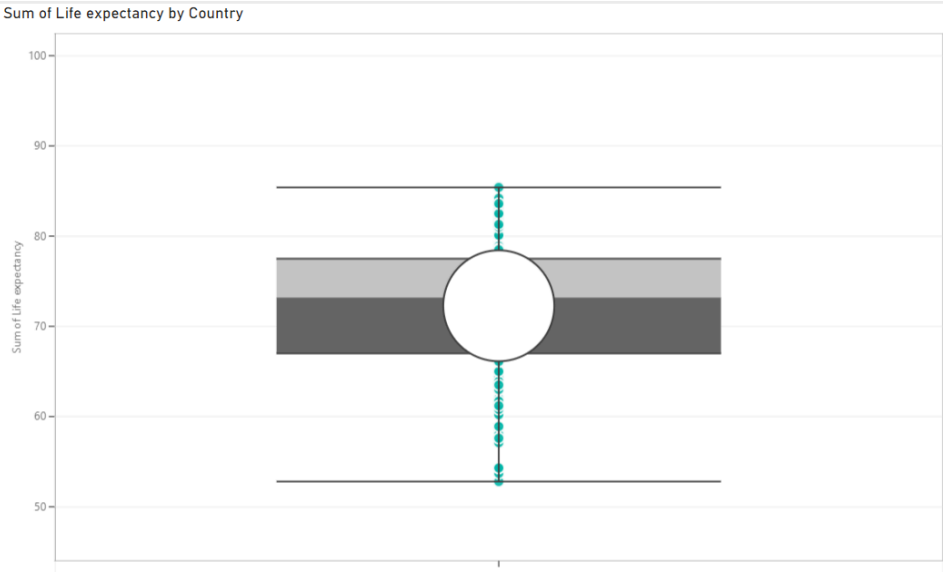


Linear regression plot of Life expectancy vs CO2 emissions

GDP and Minimum wage



Linear regression plot of Out of pocket health expenditure vs Physicians per thousand

	<div data-bbox="516 115 1453 682"><p>Sum of Life expectancy by Country</p></div> <p>Box and whiskers plot of Sum of Life expectancy vs Country</p>
<p>Conclusion</p>	<p><b>1. Sum of Population by Official Language:</b> The word cloud representing the sum of populations by official language highlights the dominance of languages like Chinese, Hindi, Standard (likely referring to Standard Mandarin), and English, indicating these languages are spoken by vast populations worldwide. Other languages such as Arabic, Portuguese, French, and Spanish also appear prominently, showcasing their widespread influence across multiple countries. This suggests a significant portion of the global population communicates using these major languages.</p> <p><b>2. GDP and Land Area (Km<sup>2</sup>):</b> The scatter plot examining the relationship between GDP and land area reveals a weak correlation, with an R<sup>2</sup> value of 0.3 and a correlation coefficient of near zero. This suggests that land area is not a strong determinant of GDP. While there is a slight positive trend indicating that larger land areas may correspond to higher GDPs, the wide scatter of data points shows significant variability, with many small countries achieving high GDPs despite their limited land size, indicating that factors beyond geography heavily influence economic output.</p> <p><b>3. Sum of Minimum Wage by Country:</b> The word cloud depicting the sum of minimum wages by country shows that nations like Monaco, Luxembourg, Australia, the United Kingdom, and the Netherlands have the highest aggregate minimum wages, reflecting higher standards of living and labor protections in these regions. The presence of other developed countries such as Germany, Ireland, and New Zealand in larger font sizes also aligns with their status as high-income nations with robust labor laws, suggesting these countries prioritize fair wages for their workers.</p> <p><b>5. Life Expectancy and CO2 Emissions:</b> The scatter plot shows a weak positive correlation (Corr: 0.1) between life expectancy and CO2 emissions. The R-squared value (R<sup>2</sup>: 0.01) is very low, indicating that CO2 emissions do not significantly explain the variation in life expectancy across the countries. This suggests that factors other than CO2 emissions play a more substantial role in determining life expectancy.</p>

**5. GDP and Minimum Wage:**

The plot reveals a weak positive correlation (Corr: 0.22) between GDP and minimum wage. The R-squared value ( $R^2$ : 0.05) is also low, suggesting that minimum wage levels have little explanatory power over the variation in GDP across different countries. There may be many other variables influencing GDP beyond minimum wage policies.

**6. Sum of Life Expectancy by Country:**

The box plot shows the distribution of life expectancy across different countries. The data appears to be clustered around a central value, with a few outliers. The majority of countries have a life expectancy around the median, with some countries significantly above or below this value.

In this experiment, different types of plots were used to understand the data better. Word maps helped identify the most important themes by showing which terms appeared most often. Linear regression plots were used to see if there was a relationship between two variables, like whether one could predict the other. Box and whiskers plots showed how the data was spread out, including the average and any outliers. Together, these plots gave a well-rounded view of the data, each offering its own unique insight.