

# COS30045 - Assignment 1 - Data Visualization

## Critique: Cost of Living

Ta Quang Tung  
Student ID: 104222196

### EXECUTIVE SUMMARY

This report critiques a number of data visualizations related to cost of living around the world, a pressing economic and societal issue. Through its analyses of three types of visualizations, namely line charts, tables, and choropleth maps, the report has found a number of recommendations for each type of chart. For line charts, the issue of overcrowdedness can be solved with interactivity or splitting them into smaller, one-line charts. For tables, the clever incorporation of color hue and size channels can increase its readability. For choropleth maps, diverging color schemes should be used over arbitrary color choices and interactivity can add important data that is otherwise impossible to display statically.

*Word count: approximately 2400 words.*

### Table of contents

Introduction .....	1
Visualization analysis .....	2
Changes in costs of basic items (line chart) .....	2
Expenditures on select categories in different countries (table-like graphic vs. table) .....	5
Cost of living in different countries (choropleth maps and color selection) .....	7
Conclusion .....	10
References .....	11

### Introduction

Cost of living is defined as the amount of money necessary to pay for basic expenses such as accommodation, sustenance, taxes, medical care at a certain place and time (*Cost of Living: Definition, How to Calculate, Index, and Example*, no date). It is of interest to people wanting to move to another country to live such as students, workers, or refugees. Cost of living is a pressing economic and societal issue as around the world it is becoming increasingly expensive to live. The cost of living is going up in many places such as Afganistan, Lebanon, Pakistan, Somalia, Venezuela (*What the cost of living crisis looks like around the world*, 2022), and even a developed economy as the UK has been going through a cost of living crisis since 2021, resulting in negative impacts on people's mental health such as children self-harming over financial pressures (Ambrose, 2022). In addition, about a quarter of the world's

population reports that they are experiencing financial difficulties (*People still feeling the pinch, fed up with 'shrinkflation' as 2023 wraps up*, 2023).

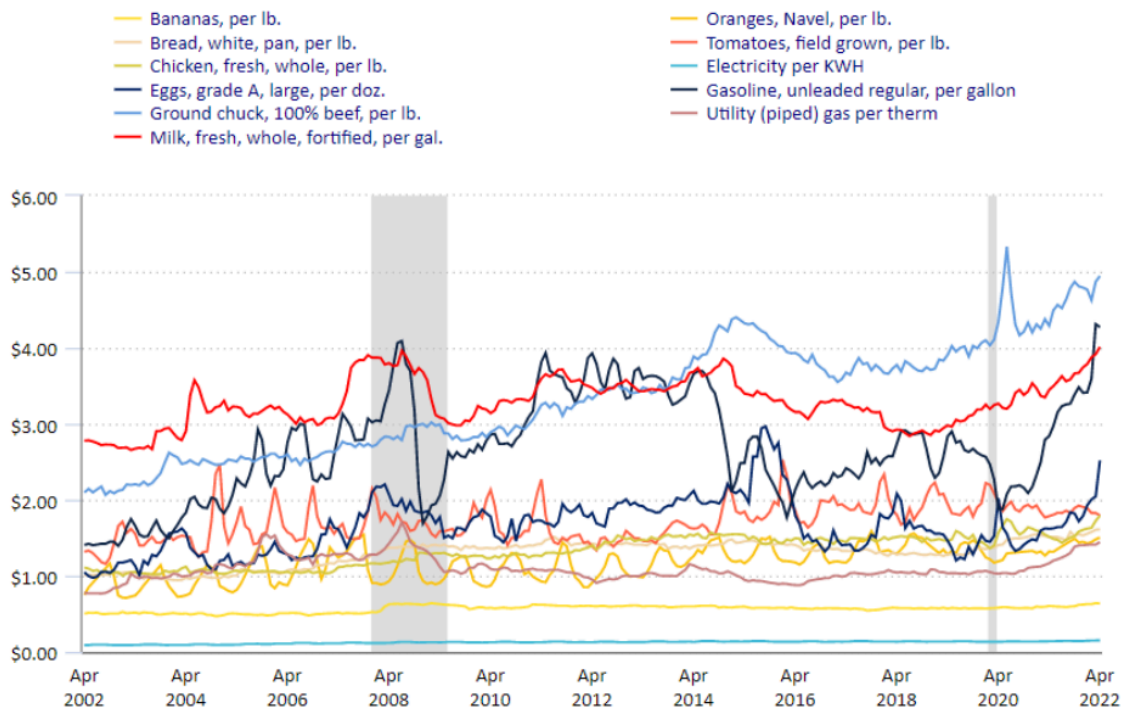
All of this makes cost of living a topic that is much discussed on the media. This discussion is accompanied with many data visualizations that attempt to put the issue into perspective. This report aims to assess the effectiveness of some of these visualizations and give recommendations as appropriate. This report is meant to be an analysis of existing data visualizations and not a discussion of the veracity of the data presented. It is aimed at designers who want to design more effective visualizations and avoid common design pitfalls. The first section discusses an overcrowded line chart and sees how interactivity can solve its issues. The second section looks at how presenting a table in a slightly different format can greatly enhance its readability. The final section compares two choropleth maps with an emphasis on color selection.

## **Visualization analysis**

### **Changes in costs of basic items (line chart)**

When discussing the rising cost of living around the world, an area of interest is the prices of basic food items and services. The line chart below, which has been obtained from a World Economic Forum article, presents the change in the average costs of select items between April 2002 and April 2022 in America. The dataset consists of three attributes: item category, average price, and time. Since the purpose of the chart in the article is to show how the average item prices have changed over time, a line chart is the appropriate type of visualization.

### Average price data (in U.S. dollars), selected items



As inflation is much too high, the average price of products is also increasing in the US. Image: US Bureau of Labor Statistics

Figure 1: The line chart shows how the average prices of different products have varied over 20 years. It originates from the US Bureau of Labor Statistics' website, but is featured on a World Economic Forum article.

Each item category in the chart is represented with a colored line. The costs, which are measured in US dollars, are treated by the chart as quantitative data and represented by the y-axis. The x-axis is a timescale between April 2002 and April 2022. Both the X and Y axes do not contain any scale distortions and the y-axis starts correctly at 0. The chart also encode information about recession in areas shaded in gray.

In the current static version found in the World Economic Forum article, the chart is extremely crowded and hard to read because there are too many lines. The majority of these lines are jagged and cross each other several times. In addition, some lines are almost similar in color (i.e. eggs vs. gasoline), making it very hard to tell them apart.

To make this chart more readable, it can be turned into an interactive chart that lets readers select the lines they want to see. This is actually what has been done for the original version of the chart, which can be found on the US Bureau of Labor Statistics' website. The interactive version solves the issue of overcrowded lines by allowing users to selectively add and remove lines. In addition, the user can

hover on the line that they want to inspect to highlight it and blur out the rest. This makes the lines much more readable and allows the user to easily compare lines of interest.

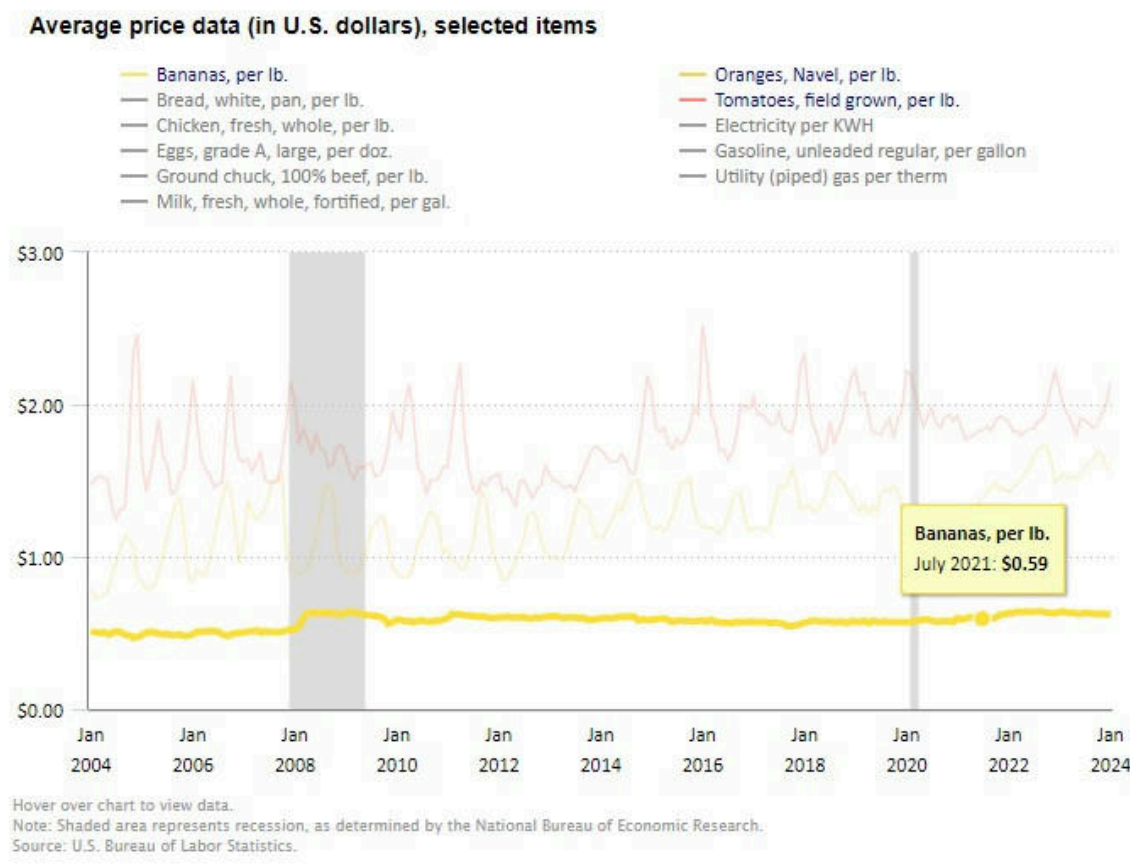


Figure 2: The original interactive line chart. Here only the lines for bananas, oranges, and tomatoes have been selected. Hovering on the banana line reveals the cost at a given point in time, highlights the line, and blurs out the other two lines. Source: US Bureau of Labor Statistics.

Making interactive chart is a good solution to line charts that must include many lines. That said, there are cases where it is impossible to present the readers with an interactive chart, for example in the printed media or on television. In these cases, a possible solution is to draw a line chart for each line. These charts must share the same X and Y scales and be placed side by side for easier comparison.

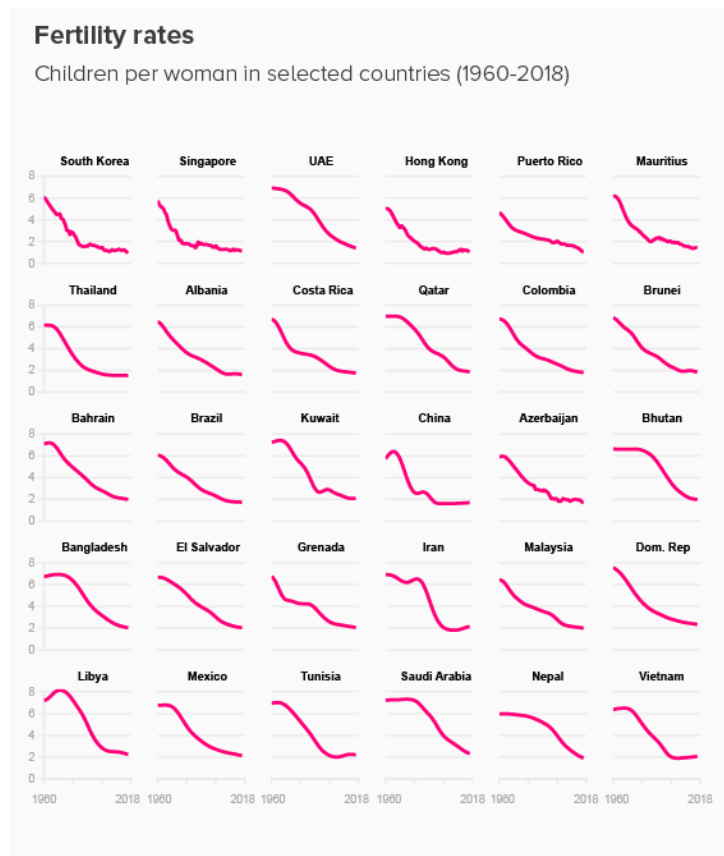


Figure 3: An example of how the crowded static line chart could be refactored. This example does not use the same dataset and is for demonstration only. Source: Add Two Digital.

## Expenditures on select categories in different countries (table-like graphic vs. table)

When it comes to deciding on which city to live in, the cost of living is an important factor for many people. The cost of living of a place encapsulates many aspects of life such as accommodation, food, recreation, etc. The following graphic from The Guardian compares the cost of living for expatriates in select cities across a range of categories. Three pieces of data are included: the city names, the expenditure categories, and the amount of spending for each category in a given city. This graphic greatly resembles a table in that the data is arranged in rows and columns. The rows denote the cities while the columns denote the expenditure categories (i.e. accommodation, cinema tickets, milk, etc.). The intersection of a row and column denotes the amount of money (in British pounds) needed for a particular category in a given city.

## Cost of living

Cost of living for ex-pats in selected cities, £

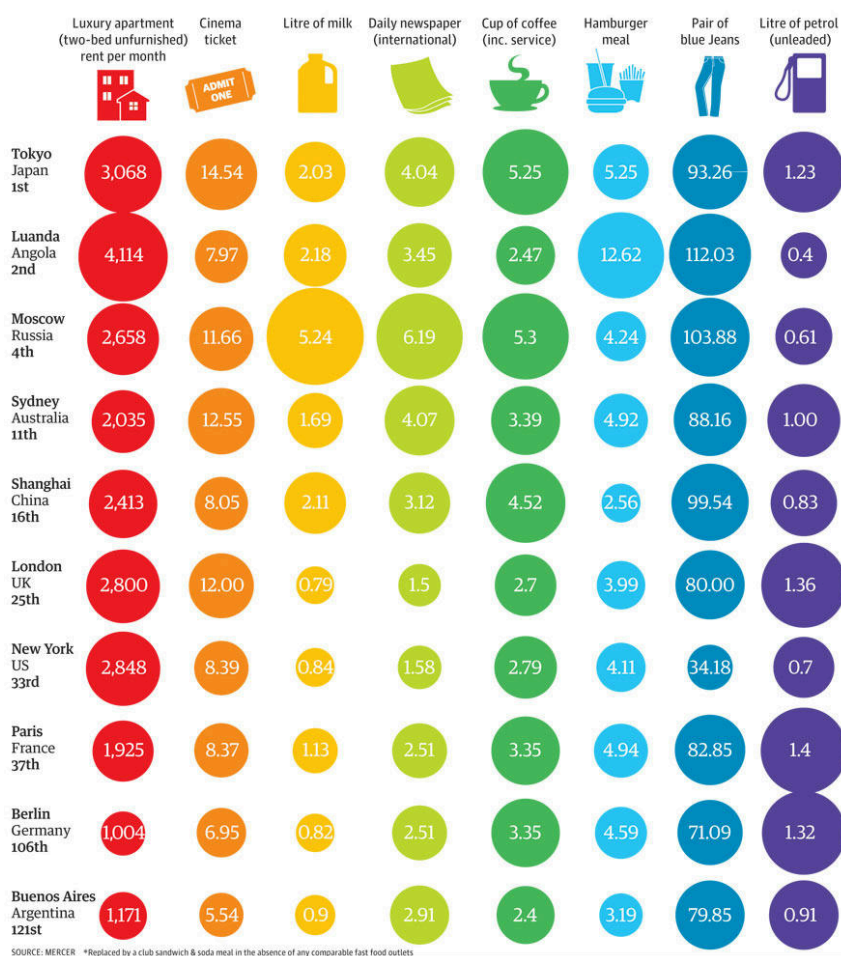


Figure 4: The graphic compares the cost of living among ten countries across a range of items. Source: The Guardian.

EXPENDITURE, £GBP	Rent of a luxury two bedroom unfurnished apartment (per month)	Cinema, international release, 1 seat	Milk, pasteurized whole milk, above 2.5% fat (1lt / 33.8 oz)	1 issue of international daily newspaper	1 cup of coffee, including service	Fast food hamburger meal	One liter of gasoline, unleaded 95 octane	1 pair of blue jeans
Tokyo	3,068.40	14.54	2.03	4.04	5.25	5.25	1.23	93.26
Luanda	4,114.24	7.97	2.18	3.45	2.47	12.62	0.40	112.03
Moscow	2,658.43	11.66	5.24	6.19	5.30	4.24	0.61	103.88
Sydney	2,035.15	12.55	1.69	4.07	3.39	4.92	1.00	88.16
Shanghai	2,413.04	8.05	2.11	3.12	4.52	2.56	0.83	99.54
London	2,800.00	12.00	0.79	1.50	2.70	3.99	1.36	80.00
New York	2,848.32	8.39	0.84	1.58	2.79	4.11	0.70	34.18
Paris	1,924.72	8.37	1.13	2.51	3.35	4.94	1.40	82.85
Berlin	1,004.20	6.95	0.82	2.51	3.35	4.59	1.32	71.09
Buenos Aires	1,170.98	5.54	0.90	2.91	2.40	3.19	0.91	79.85

Figure 5: The same data presented as an Excel table. Data source: Mercer, downloaded via The Guardian.

As can be seen from figures 4 and 5, the difference between The Guardian's graphic and the table is that the data cells of the graphic are represented with colored circles of different sizes. For these circles, two channels have been used: color hue and size. The color hue channel is used to signify the expenditure category (which is appropriate because it is categorical data) while the size channel indicates the amount of money needed for the category (which is quantitative data). It should be noted that the size of a circle is relative to other circles in the same column. As such, it does not make sense to compare two circles of different columns because they belong to two different spending categories.

A clear advantage that the graphic has over the table is that it is much easier to see which city is the most expensive for a given expenditure category. For instance, by simply scanning the first column, one can immediately see that the price of a luxury apartment is highest in Luanda (4,114 pounds). This is possible because the reader does not have to actually read the numbers. They can simply glance at the size of the circle to pick up this information. It is the opposite with the table, in which the reader must read the numbers carefully and be mindful of which number they encountered is the highest. The graphic is highly effective in communicating its information because it makes use of two highly effective channels - area for quantitative data and hue for categorical data (Munzner and Maguire, 2015) - instead of relying much on text. The text is present only for readers that want to inspect the data further.

### **Cost of living in different countries (choropleth maps and color selection)**

Many people are also interested in how the cost of living varies across countries around the world. The most common way of visualizing this data is with a colored map of countries in which the colors indicate how expensive the countries are. This section compares two of such choropleth maps that show the cost of living in different countries around the world. Choropleth maps are a type of map that colors, shades, or patterns geographical areas based on a data variable (*Choropleth Map - Learn about this chart and tools to create it*, no date) and visualizes how that variable changes across different areas. For the purpose of visualizing how the cost of living varies around the world, choropleth maps are ideal because they make it very easy to see which countries or regions are expensive and which ones are not.

The choropleth maps in question (denoted by A and B) take data from Numbeo, a crowd-sourced database containing cost of living data around the world. They color each country based on its cost of living index (COLI), a metric designed by Numbeo that expresses how expensive it is compared to New York City (which is given an index of 100) (*Overview of Cost of Living Section Indices on this Website*, no date). For example, Vietnam has a COLI of 30.8, meaning that it is around 70% cheaper than New York City.



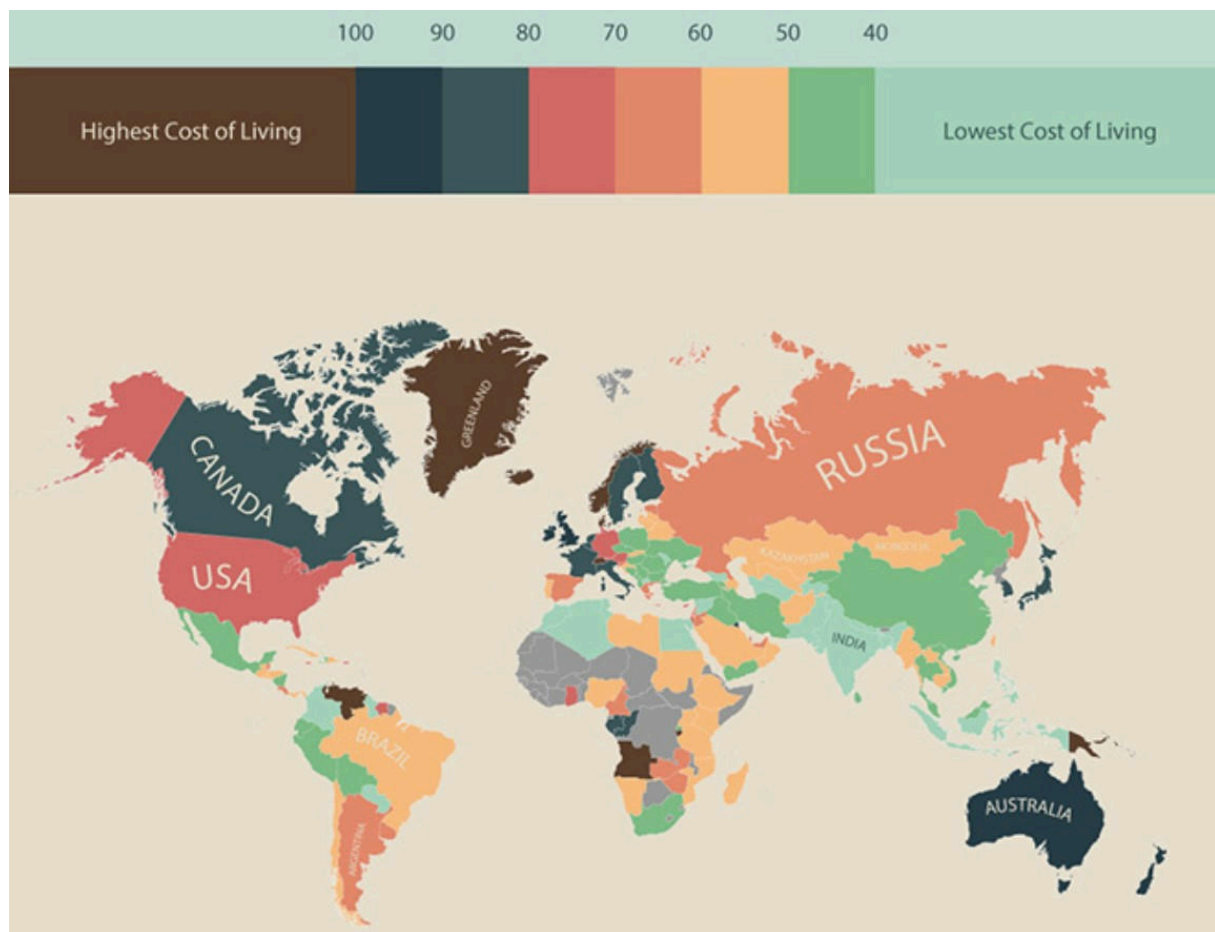


Figure 6: Choropleth map A. Source: MoveHub.

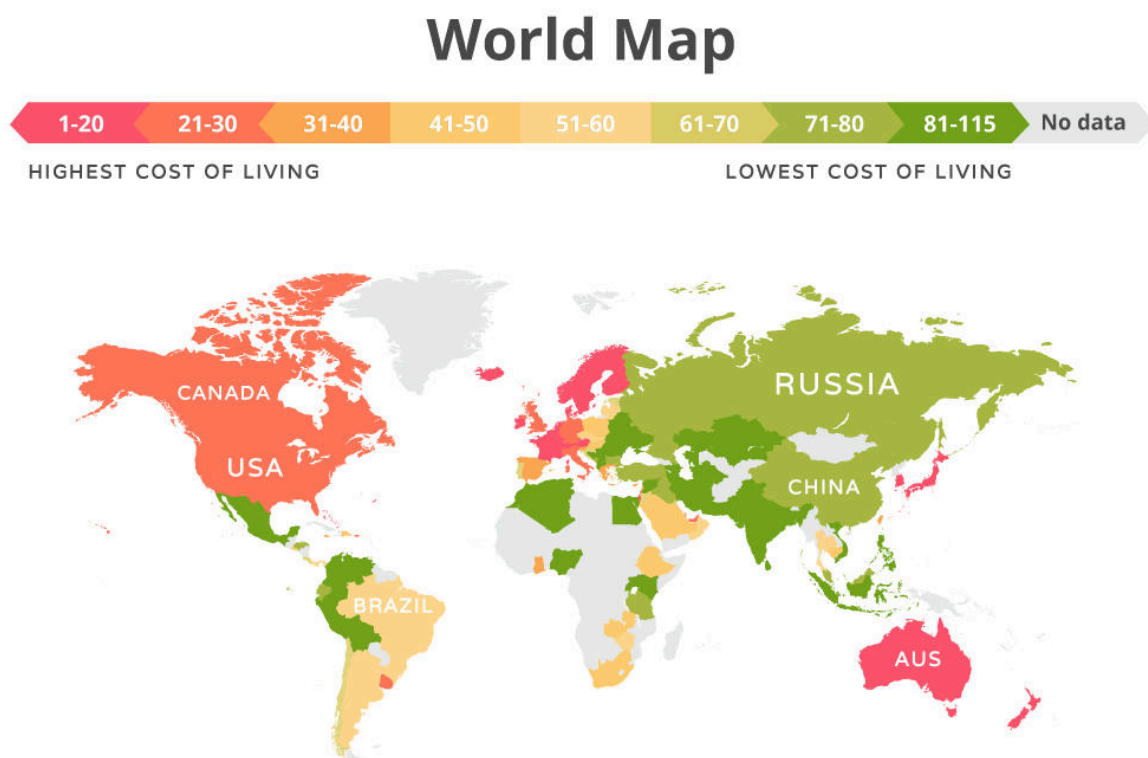


Figure 7: Choropleth map B. Source: MoveHub.



When thought of as a table, the dataset contains two columns: country and cost of living index. Both maps treat the country column as categorical data and encode it with spatial region. The cost of living index column is treated as ordinal data. For map A, the cost of living spectrum is divided into distinct ranges that go from high to low, and each range is assigned a color. Map B handles this data differently: it first sorts the countries in descending order of the COLI, then successively groups them into ranges such that the first range contains the top 20 countries, the second range contains the next 10 countries, etc. Similar to map A, each range is given a color.

Map A encodes the COLI data with color hue where countries in the higher COLI ranges are filled with dark brown, dark teal, or red while those in the lower COLI ranges are filled with colors ranging from orange to light green. This choice of color is the map's most glaring issue as the color progression does not make sense. From dark brown, it suddenly jumps to two dark teal colors, then to red, and finally to green near the end. The inclusion of teal in the higher COLI ranges is problematic because it is similar to green, which is the color of the lower COLI ranges. At a glance, readers might mistake dark-teal countries as being inexpensive while it is the opposite. A better color progression is to use red for the highest range, yellow for the middle range, green for the lowest, and interpolation for ranges in-between. If one is to consult the color wheel, which visualizes the relationship among different colors, they will find that this color progression is more natural.



Figure 8: The color wheel. Red naturally transitions into yellow, which transitions into green. Source: Britannica.

Another issue with map A is that it lacks an explanation for what the color gray means. Many parts of the map (i.e. Africa) have been colored gray without any indication as to why they are so. Readers

can only guess that gray represents unavailable data. A good visualization should be clear and obviate the need for readers to make assumptions.

Map B avoids the issues of map A by selecting a better color progression and clearly explaining what gray means. Map B uses a diverging color scheme, which pairs two sequential schemes that are based on different hues and that share a common color used for the midpoint (*Color Use Guidelines*, no date). The first sequential scheme goes from red (1-20) to bright yellow (51-60) while the second goes from bright yellow (51-60) to green (81-115). Naturally, red represents the most expensive countries while green represents the cheapest ones. Aside from being easier to read, diverging color schemes are also more effective in choropleth maps than other schemes, as pointed out by Brewer *et al.* (1997).

One issue with both maps is that it is hard to exactly compare countries. Readers can only tell whether a country is more expensive than another but not how much. One solution is to include the COLI directly inside each country's region, but this would clutter the map and be very hard to read if the map is small. A better solution is to make the map interactive and show a country's COLI when it is hovered on, as in the following map:

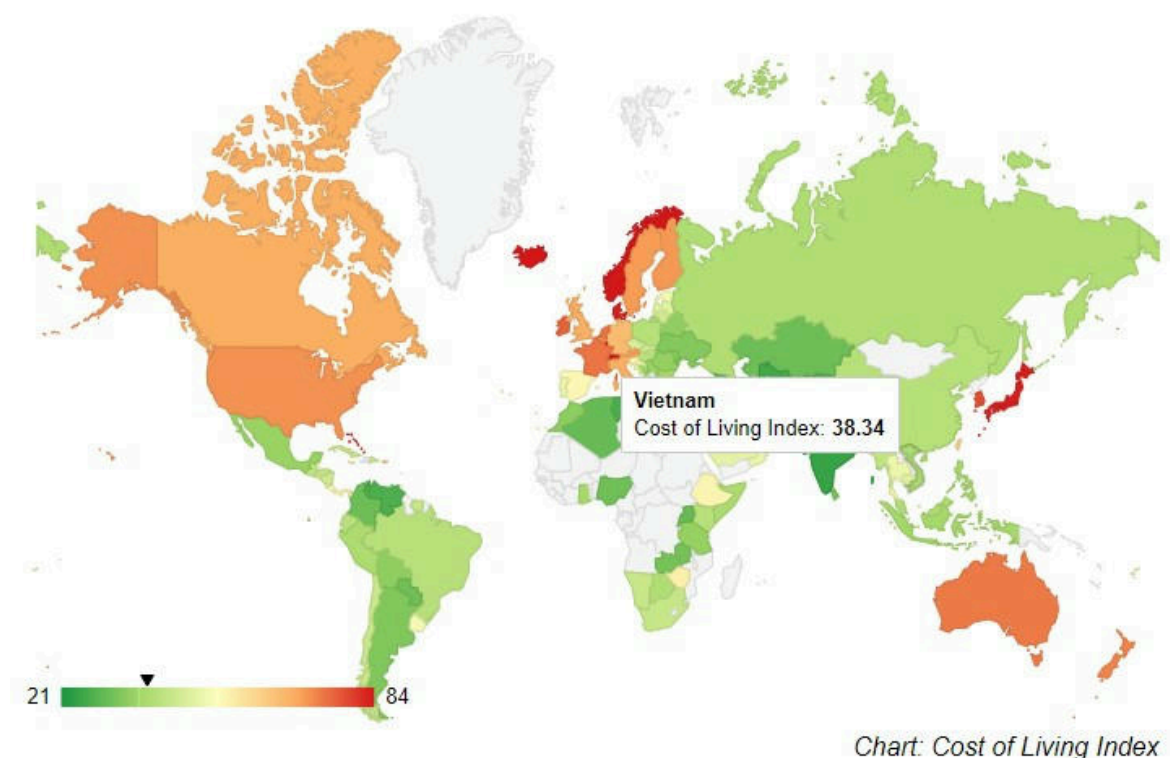


Figure 9: An interactive choropleth map which shows the COLI of Vietnam when the country is hovered on. Source: Numbeo.

## Conclusion

This report has analyzed a variety of data visualizations related to cost of living and reached a number of conclusions regarding how to design them effectively. The analysis of the line chart yielded the

conclusion that line charts with many data lines can be made interactive to selectively show lines of interest or split into a group of one-line charts placed side-by-side for better comparison. The analysis of the table-like graphic finds that incorporating the color hue and size channels into a table can boost its readability. Finally, the comparison between the two choropleth maps have concluded that the use of diverging color schemes in choropleth maps is more effective and that the maps can be made interactive to reveal data that is very hard to display statically.

## References

- Ambrose, T. (2022) "Children stressed and self-harming over UK cost of living crisis", *The Guardian* [Preprint]. Available at: <https://www.theguardian.com/society/2022/jun/18/children-stressed-and-self-harming-uk-cost-of-living-crisis-childhood-trust> (Accessed: February 18, 2024)
- Average price data (in U.S. dollars), selected items* (no date). Available at: <https://www.bls.gov/charts/consumer-price-index/consumer-price-index-average-price-data.htm> (Accessed: February 18, 2024)
- Brewer, C. A. et al. (1997) "Mapping Mortality: Evaluating Color Schemes for Choropleth Maps", *Annals of the Association of American Geographers*, 87(3), pp. 411–438
- Choropleth Map - Learn about this chart and tools to create it* (no date). Available at: <https://datavizcatalogue.com/methods/choropleth.html> (Accessed: February 18, 2024)
- Clissitt, C. (2020) *Cost of Living Around the World in 2020*. Available at: <https://www.movehub.com/blog/cost-of-living-worldwide/> (Accessed: February 18, 2024)
- Color Use Guidelines* (no date). Available at: [https://web.natur.cuni.cz/~langhamr/lectures/vtfg1/mapinfo\\_2/barvy/colors.html#diverging](https://web.natur.cuni.cz/~langhamr/lectures/vtfg1/mapinfo_2/barvy/colors.html#diverging) (Accessed: February 18, 2024)
- Colour wheel | Definition, Art, & Facts | Britannica* (2023). Available at: <https://www.britannica.com/science/color-wheel> (Accessed: February 18, 2024)
- Cost of Living Index by Country 2020* (no date). Available at: [https://www.numbeo.com/cost-of-living/rankings\\_by\\_country.jsp?title=2020](https://www.numbeo.com/cost-of-living/rankings_by_country.jsp?title=2020) (Accessed: February 18, 2024)
- Cost of Living: Definition, How to Calculate, Index, and Example* (no date). Available at: <https://www.investopedia.com/terms/c/cost-of-living.asp> (Accessed: February 18, 2024)
- Cost of living: This chart shows how the price of products has risen in the US* (2022). Available at: <https://www.weforum.org/agenda/2022/05/cost-of-living-price-rises/> (Accessed: February 18, 2024)
- Map Shows Cost of Living Around the World* (2015). Available at: <https://web.archive.org/web/20151210085404/https://www.movehub.com/blog/living-costs-world-map/> (Accessed: February 18, 2024)
- Munzner, T. and Maguire, E. (2015) *Visualization analysis & design*. Boca Raton, FL: CRC Press/Taylor, Francis Group (AK Peters Visualization series)

*Overview of Cost of Living Section Indices on this Website* (no date). Available at: [https://www.numbeo.com/cost-of-living/cpi\\_explained.jsp](https://www.numbeo.com/cost-of-living/cpi_explained.jsp) (Accessed: February 18, 2024)

*People still feeling the pinch, fed up with 'shrinkflation' as 2023 wraps up* (2023). Available at: <https://www.ipsos.com/en/people-still-feeling-pinch-fed-shrinkflation-2023-wraps> (Accessed: February 18, 2024)

*Rule 31: Line charts shouldn't have too many lines* (2022). Available at: <https://www.addtwodigital.com/add-two-blog/2022/5/4/rule-31-line-charts-shouldnt-have-too-many-lines> (Accessed: February 18, 2024)

*What the cost of living crisis looks like around the world* (2022). Available at: <https://www.rescue.org/article/what-cost-living-crisis-looks-around-world> (Accessed: February 18, 2024)

*Which is the world's most expensive city? Cost of living survey 2012 | News | theguardian.com* (no date). Available at: <https://www.theguardian.com/news/datablog/2012/jun/12/city-cost-of-living-2012-tokyo> (Accessed: February 18, 2024)