

# GDP Growth Graph and London Low-carbon Generator Map

Mengze Zhang

26 March 2023

## 1. Introduction

Analysis for this project uses the R statistical programming language (R Core Team 2022), as well as `tidyverse` (Wickham et al. 2019) programming package to select the relevant columns, rename the columns, and transform the data into a format that is ready to be visualized. The `WDI` (Arel-Bundock 2022) package provides access to a wide range of economic and development data across the world. The graph and map are created with `ggplot2` (Wickham 2016) and `ggmap` (Kahle and Wickham 2013).

## 2. Graph

Figure 1 shows the GDP growth of Brazil, France, India, and USA from 1961 to 2021.

In order to make this graph, I firstly searched for a specific indicator of GDP growth data using the `WDIsearch` function in the `WDI` (Arel-Bundock 2022) package. After acquiring the indicator, the `WDI` function was used to retrieve GDP growth data for the countries of Brazil, India, France, and the United States. The resulting data is then assigned to the `world_bank_data` object. Then, I renamed the column “NY.GDP.MKTP.KD.ZG” to “gdp\_growth” for easier readability and used the `select` function to only retain the columns “country”, “year”, and “gdp\_growth”. This processing step helps to clean and prepare the data for further analysis or visualization.

After preparing the dataset, I created a line graph with `ggplot2` (Wickham 2016). I set year as the x-axis and GDP growth as the y-axis to show the changes in GDP growth over the years.

Some of the adjustments I did to make the graph more effective include:

- Color the lines by country to differentiate between them. This allows the audience to quickly and easily identify the trends in each country.
- Set appropriate breaks for the x-axis and y-axis. I set the x-axis to display years in intervals of 10 using `scale_x_continuous(breaks = seq(1960, 2025, 10))`, and limited the y-axis to show values between -10% and 15% with intervals of 5% using `scale_y_continuous(limits = c(-10, 15), breaks = seq(-10, 15, 5))`. By setting breaks for the x-axis, it improves the readability and interpretability of the graph.
- Use a minimal theme to keep the focus on the data being represented and avoid distractions.
- Label the axes, color, and add a caption using `labs()` to provide context and information to the audience.

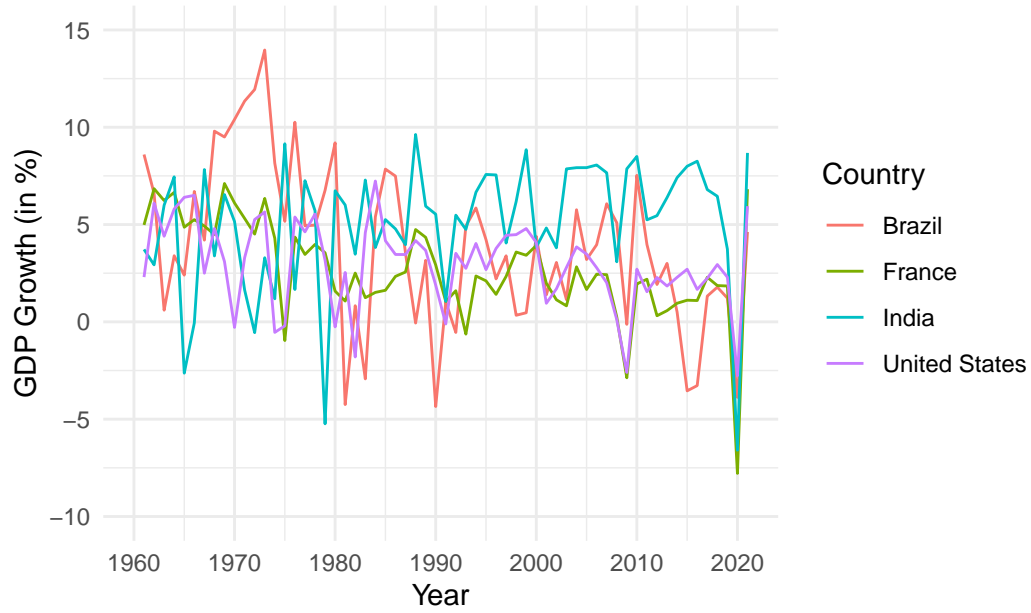
Other options to make the graph more effective may include:

- Use different shades of the same color to indicate different values. For example, in the case of Brazil, we could use scarlet for values over 5%, red for values between 0% and 5%, and pink for values below 0%. This makes it easier to distinguish between high and low values, and to quickly identify trends.
- Add a horizontal line at 0% GDP growth to show more clear whether GDP increased or decreased from the previous year.
- Add a vertical line and annotation to indicate significant events that may have impacted GDP growth. For example, in 2020 during the COVID-19 pandemic, GDP growth of all four countries dropped significantly.
- Use a custom color palette to improve the overall aesthetic of the graph.
- Use different line types, such as dashed or dotted lines, to differentiate between lines in a clear and concise way.

### 3. Map

Figure 2 shows the locations of low-carbon energy generators located at fire stations in London in 2011.

In order to generate this map, I firstly created a variable called `bbox` and set it equal to a vector of four values that define the spatial extent of London, UK. Then, I used the `get_stamenmap` function in the `ggmap` (Kahle and Wickham 2013) package to download a map of London based on the bounding box defined in `bbox` and store it in the `london_map` variable. After storing the map of London, I used the `readr::read_csv` function to read in and stores the data in a data frame called `low_carbon_generators` from the Greater London Authority (GLA) ( Greater London Authority 2022) about low-carbon energy generator locations in London. Then, I



Data source: World Bank.

Figure 1: GDP Growth in the Last 60 Years by Country

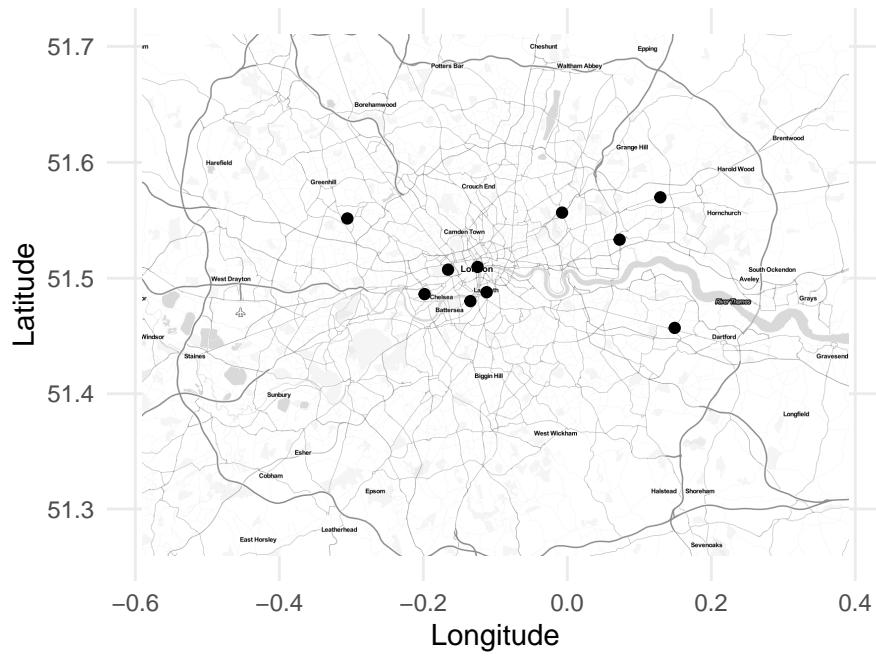


Figure 2: Low-carbon Generators on MPS buildings in London in 2011

selected only the `Longitude_WGS84` and `Latitude_WGS84` columns from the data frame and renamed them to `lon` and `lat` respectively. The result of these operations is stored in a new data frame called `reduced_table`.

At this point, the dataset and map are both ready and can be combined in the visualization. I used the `ggmap` function to overlay the `reduced_table` data points onto the `london_map` as points representing the low-carbon energy generators located at fire stations in London.

The quote “maps only show people who aren’t invisible to the makers” from Heather Krause means that maps only depict information and details that the people creating the map are aware of or choose to include. It suggests that maps are not a neutral or objective representation of reality, but rather a reflection of the views and beliefs of the people who make them, and there may be aspects of reality that are invisible and omitted or hidden from view. This quote highlights the importance of considering who is included and excluded in data visualizations and maps, and how this can impact the representation and interpretation of information.

The GLA is a strategic regional authority, with powers over transport, policing, economic development, and fire and emergency planning. Three functional bodies — Transport for London, the Mayor’s Office for Policing and Crime, and London Fire and Emergency Planning Authority — are responsible for delivery of services in these areas ( Greater London Authority 2022). The Metropolitan Police Service (MPS) collected and provided the dataset of low-carbon energy generation installations on MPS buildings in London in 2011. This dataset only includes information about low-carbon energy installations on MPS buildings, which may not accurately reflect the total amount of low-carbon energy generation in London. This may create a selection bias that leads to an underestimation of low carbon energy generation in the city. In addition, it may not be representative of energy generation in other buildings or across different regions in London, which could lead to incorrect conclusions about the distribution and extent of low-carbon energy generation in the city. Considering these potential biases, it is important to interpret the dataset with caution and to consider other sources of information and data when evaluating the state of low-carbon energy generation in London.

Similar insights could also be found in Chapter 3 of *Data Feminism: On Rational, Scientific, Objective Viewpoints from Mythical, Imaginary, Impossible Standpoints* (D’Ignazio and Klein 2020), where it argues that maps, like all data visualization tools, have a significant impact on how people understand and make decisions about the world. Maps have the power to reinforce and amplify existing social biases, leading to a distorted understanding of reality. For example, maps may only show certain types of data or information, in this case it might be the wealth distribution of different districts of London, while ignoring other important data, such as cultural landmarks or the lived experiences of marginalized communities.

In summary, the biases and ethics of mapping should be considered when creating and analyzing maps. This is because maps have the potential to reinforce existing social biases and distort our understanding of reality. It is important to ensure that the data used in mapping is accurate, diverse, and inclusive, and that the perspectives of marginalized communities are taken into account.

## References

- Arel-Bundock, Vincent. 2022. *WDI: World Development Indicators and Other World Bank Data*. <https://CRAN.R-project.org/package=WDI>.
- D'Ignazio, Catherine, and Lauren Klein. 2020. *Data Feminism*. <https://data-feminism.mitpress.mit.edu>.
- Greater London Authority. 2022. *London Datastore*. <https://data.london.gov.uk/dataset/>.
- Kahle, David, and Hadley Wickham. 2013. "Ggmap: Spatial Visualization with Ggplot2." *The R Journal* 5 (1): 144–61. <https://journal.r-project.org/archive/2013-1/kahle-wickham.pdf>.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolmund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.