Worldwide Overproduction of Cars

Jingquan Chen, John Wang, Keng Li Chia, Jourdan Lim, Perpetua Raj, Joey Tan
{2101268, 2101925, 2102718, 2102516, 2101771, 2001885}@sit.singporetech.edu.sg
Computing Science

Introduction

CrudeOilPeak[1], a website that tracks global oil production and consumption as well as its impact on the economy, states that over the past decade, global car production has surged by 52%, predominantly driven by China's growth. In contrast, oil supplies have only increased by 16%[2]. This disparity highlights a significant challenge: the demand for cars is outpacing the available oil supply.
To further illustrate this issue, they have created a visualization that compares the growth in car production by country/region against the growth in global liquid supplies (Figure 1). We found the visualization to be severely lacking and messy, and believe that it can be improved to be more visually appealing and easier to digest.

Previous Visualization

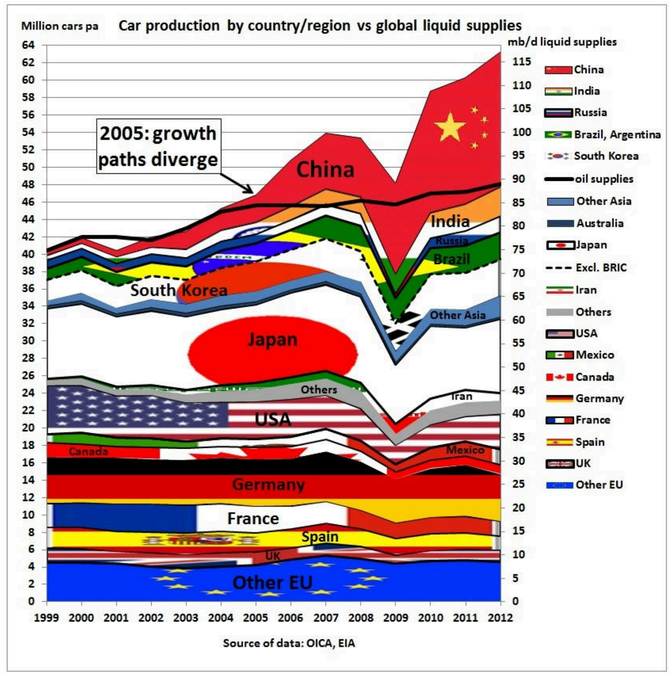

Figure 1: Car production by country/region vs global liquid supplies

Strengths

1. High amount of information shown on the graph. Users can pinpoint and gather data specific to each country.
2. Area filling with country flags make it possible to identity country in question without referencing the plot legends.
3.

Suggested Improvements

1. *Avoid having two y-axis* as it can be confusing and misleading. Two different graphs could be created instead.
2. *Too much visual clutter* can make it difficult to interpret the data. Stick to a consistent color scheme and remove unnecessary elements.
3. *Do not add flags to the graph* as they do not add value.
4. *The growth paths divergence* can be better emphasized.

Implementation

Data

- Car production data by country/region were obtained from the International Organization of Motor Vehicle Manufacturers (OICA) for the year 2017[3].
- Similarly, the oil production data were obtained from the U.S. Energy Information Administration[4]. However, the data was not bundled together initially, so we had to manually combine and summarize the data from the different sources.

Software

We used the Quarto publication framework and the R programming language, along with the following third-party packages:

- *readxl* for data import
- *tidyverse* for data transformation, including *ggplot2* for visualization based on the grammar of graphics
- *knitr* for dynamic document generation

Improved Visualization

Further Suggestions for Interactivity

Because our visualization was intended for a poster, we did not implement any interactive features, including the infotip. However, if the data are visualized in an HTML document, interactive features can be achieved using the R packages such as *plotly*. In that case, we recommend that the tile does not change its fill color. In contrast, the original visualization changes the fill color of the activated tile to light blue (see **?@fig-infotip_color_change**), which can be misinterpreted as a change in incidence. Instead, we suggest highlighting the activated tile by thickening its border.

Conclusion

We successfully implemented all suggested improvements for the non-interactive visualization. By labeling every state and choosing a colorblind-friendly palette, the revised plot is more accessible. The logarithmic color scale makes the decrease in incidence after the introduction of the vaccine less striking but enables readers to detect patterns in the low-incidence range more easily.

[1]https://crudeoilpeak.info
[2]https://crudeoilpeak.info/world-car-production-grows-3-times-faster-than-global-oil-supplies
[3]https://www.oica.net/category/production-statistics/
[4]https://www.eia.gov/international/data/world#/?