IDS 701 - Solving Real Problems with Data Science
Spring 2025
Peter de Guzman

Feedback Exercise - Tiger Exploratory Question Report

Summarization:

This project aims to address the climate impact of the last-mile logistics involved in the delivery of goods purchased through e-commerce platforms. In order to address this problem, the authors aim to understand how large of a problem are carbon emissions in the industry, what are the most energy-intensive stages of the e-commerce supply chain, how do these stages contribute to the overall environmental impact, and how long does it take for a package to be delivered. Using data from a last-mile delivery dataset from Shanghai, the authors conduct descriptive analyses of the distribution of delivery times and the proportion of deliveries completed within 30 minutes by region of Shanghai. They find that last-mile delivery is the most carbon-intensive part of the e-commerce supply chain. They also found that the overall supply chain is complex and energy-intensive, and that deliveries in some dense metros like Shanghai are possible in short amounts of time. They interpret these findings as evidence that last-mile deliveries are a key part of the way that e-commerce contributes to negative environmental impacts.

Strategy and Effectiveness:
The report provides content on the e-commerce industry, its environmental impact, and its particular inefficiencies. However, I did not feel that the question statement was very clear. The three sections do not flow coherently in their current order. I think that the second and first sections are redundant, and could be combined. An alternative structure could be to first lay out the most energy-intensive stages of the e-commerce supply chain, before going more specifically into the carbon footprint impacts and then how the end user experiences the supply chain. The third point feels disconnected from the environmental focus of the first two points. I understand the authors were trying to tie together the last-mile delivery part of the supply chain to the overall environmental impact of the e-commerce supply chain, but this connection was not explicitly made between these sections or in a conclusion statement.

Successes:

The structure of the report was clear and easy to read. The data analysis was generally easy to follow, although the region for the 30 minute benchmark was not explained. The subheaders in the second question section made it easy to follow the distinction between the different stages of the e-commerce supply chain for viewers who are unfamiliar with the industry.

Areas for Improvement:

I think the plots could be improved by including the source of the data in a note below the plot so that each plot can stand alone as a source of information. I also think it would have been beneficial to sort the region IDs numerically, or to sort them by the highest number of deliveries. As the region IDs are two-digit codes, they are meaningless to someone unfamiliar with the region. I believe that some of them could be omitted to make the plots less busy.

The plot in the appendix appears to have the bars sorted from the highest proportion of deliveries to the lowest. The first plot could be changed to be sorted in the same manner (but instead by the highest number of deliveries). Since the second plot is displaying a proportion, I think the y-axis should be standardized to be out of 1 or 100 (to represent 100%).

This report could be strengthened by a conclusion section that syntheses the findings of the analysis and the three individual sections into 1-2 paragraphs. These synthesized key takeaways should also be present in the executive summary.