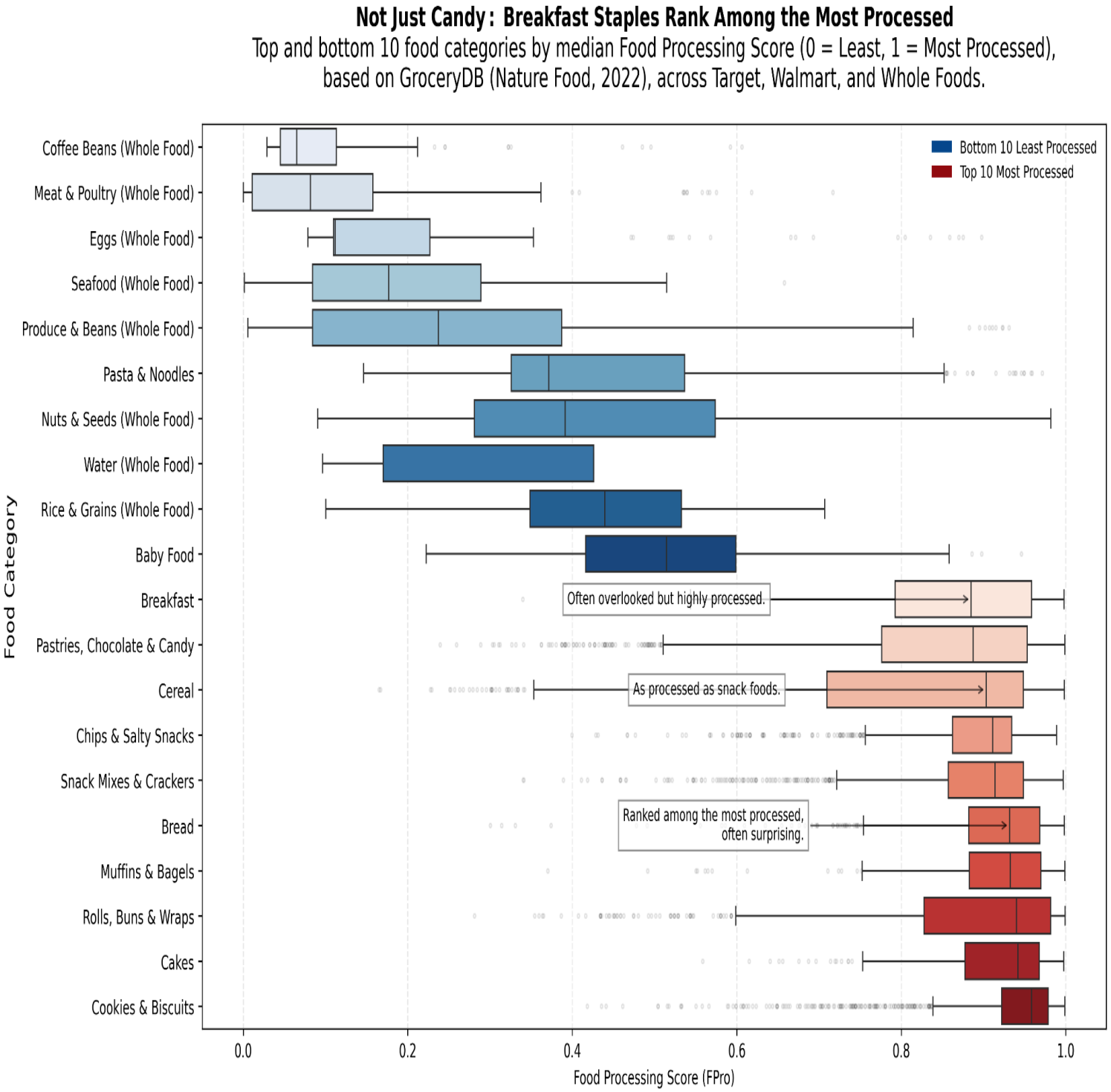
Visualization



Write-up

Many people assume that ultra-processed foods are limited to snacks, candy, or junk food, but I wanted to explore whether other everyday foods might also rank high in processing. My motivation for this project was to challenge that assumption and see if categories typically considered simple or healthy, such as cereal, bread, or breakfast, were among the most processed.

The final visualization shows the top 10 most processed and bottom 10 least processed food categories using the median Food Processing Score (FPro). The dataset comes from GroceryDB, published in Nature Food (2022), and includes items from major retailers like Target, Walmart, and Whole Foods. FPro scores range from 0 (least processed) to 1 (most processed). The categories are sorted vertically by median score. Unsurprisingly, the bottom 10 categories were almost entirely whole ingredients like coffee beans, meat, seafood, and produce. On the other hand, bakery and breakfast items dominated the top 10. Cookies and cakes ranked highest, but surprisingly, bread appeared to be the 5th most processed, cereal scored as high as salty snacks, and breakfast foods came in at 10th, just within the highly processed group.

I used a boxplot for the design to show each category's distribution, quartiles, and outliers of FPro scores. I limited the data to the top and bottom 10 categories to emphasize contrast and keep the story focused. The categories are sorted by median FPro, and the two groups are visually separated using color palettes. Blues are for the least processed, and reds are for the most processed. This created a clean visual divide and avoided the ambiguity that can come with a gradient scale. I also replaced acronyms like “wf” with full descriptors like “Whole Food” for clarity, made the outliers smaller and more transparent to minimize distraction in the readers, and added color intensity increase for both categories (the higher the processing, the darker the blue or the red). To guide the viewer, I added arrow annotations calling out the rank of bread, cereal, and breakfast foods, highlighting unexpected insights supporting the title’s message.

I started by plotting all food categories, but the result felt overwhelming and visually cluttered. After experimenting with several plot types (violin plot, bar with only median, etc.), I chose the boxplot to balance clarity and detail. I then refined the y-axis labels for readability and ranked the data to focus on the extremes. To improve contrast, I built two separate color palettes. I also spent time positioning the annotations so they would stay within the graph area and draw attention to the key points. Last, I tightened the layout, clarified the subtitle, and added a small legend to help the viewer interpret the color-coded ranks.