

Ethical Data Science Analysis of Amazon E-Commerce Data

Data Visualization and Ethical Storytelling

Introduction

Visual analytics is a dynamic contributor to the data science process, and converts raw statistics into shapes, trends and stories that technical and less technical stakeholders easily comprehend. When buying a product in the Amazon product dataset, visualization allows us to see the interaction of factors that may drive consumer behavior: price, ratings, and reviews. But visualization is not only technical work, it is also ethically responsible. The implications of the design of charts, the stories they support, and the stories they suppress all have implications for fairness, transparency, and justice within the digital economy.

Here shall explore how the data visualization method can be applied to the data set and critically consider the ethical and social justice connotations.

1. Price Distribution Product-wise

A box or histogram of product price distribution shows that prices are distributed across price ranges centered mostly at 100-400 dollars, with a few products costing above 1000 dollars. Price clustering and absurd outliers are well explained with this kind of visualization.

Ethical Implications:

- Showing this sort of visualization to consumers might further develop the belief that a budget phone is somehow a lesser-valuable possession merely due to its prevalence at the bottom of the price bracket.
- To business enterprises, pointing out differences in prices that are not in context makes predatory pricing strategies seem valid, including the artificial inflation of original prices to make discounts look large.

- From a social justice perspective, this picture also brings to mind economic disparities: income or access to technology are divided by another factor: economic status. A visualisation that sweeps this context under the carpet is prone to elitist narratives.

2. Scatter Plot Rating AND Price

Pricing correlates to ratings by using a scatter diagram where it will be determined that highly rated products are likely to be priced expensively. Practically, there is frequently no significant correlation with such plots, with budget phones even getting as good ratings as premium brands.

Ethical Implications:

- This is opposed to stereotypical consumer attitudes of competing based on price: the higher the product's price, the better it is. An ethical visualisation highlights this conclusion, stimulating more informed, fairer buying behaviour.
- But when companies only provide parts of this visualization (e.g. it only presents high-end clusters), it ends up as a manipulation tool instead of an empowerment tool.
- Data scientists are thus expected to defend the ability to achieve bias-free representation practice through visual storytelling.

3. Reviews Bar Chart vs. Ratings

A bar chart with the values of the number of reviews per rating category indicates that most ratings are concentrated in the 4-5 stars range, with a lower amount of the product rating of less than three stars. This skew implies an inflation of ratings everywhere.

Ethical Implications:

- This understanding poses the exactness of reviews. If businesses influence reviews through incentives, the visualization becomes not the consumer truth but the corporate steam.
- To the readers, it could build false confidence in some products due to misleading review distributions.
- A socially just practice would be marking visualizations with caution regarding possible biases to make people aware that the perceived patterns cannot always be impartial facts.

4. Brand-Level Visualization

The chart in the form of a bar gives us the slightest idea of how certain manufacturers conquer the consumer perception regarding average ratings. Some cheap brands may report just as high satisfaction as high-end brands such as Apple or Samsung may report a high average.

Ethical Implications:

- Unless the visualization is cautiously weaponized, it risks negatively stigmatizing smaller or less well-known brands and further existing market disparities.
- Ethical storytelling could mean placing the dominance of specific brands within a specific context (e.g., marketing budgets, visibility in search results, cultural trust) instead of the context of it being the difference of pure consumer choice.

Ethical Storytelling Visualization

Visualisation is not a simple aesthetics it is a narration with a moral load. Every design decision, such as axis scaling, color schemes, etc structure interpretation. For example:

- A logarithmic price scale may dilute inequalities and economic differences to reduce the importance of differences.

- Excluding outliers may cleanse the state of luxury consumerism by removing the evidence of those rarer and ultra-premium products available.
- Implicit associative elements, such as color schemes, can be treated as cues of bias (e.g., low-quality brands portrayed in red, and premium product lines, in green), which fortifies stereotypes.

To engage in ethical visualization, the data scientists should:

1. Bias reporting within the data (e.g. lack of products, product manipulation).
2. To avoid illusive scales or selective processing of patterns.
3. Contextualize so that the viewers do not forget that figures and imagery depict social phenomena, not laws.
4. Focus on center equity or findings and insert these facts so that they tend to result in fairness, as in demonstrating that budget items frequently perform similarly to high-quality ones.

Conclusion

In the Amazon dataset, visualization may be used as a two-edged sword; on the one hand, it can help consumers understand the situation clearly, or on the other hand, it can mislead consumers by discriminating against the information being framed. Ethical storytelling should move beyond the technical aspect and consider the following: What are we asserting? Who gains from these visualizations, and who is hurt? Data visualization can be a weapon of the social justice, instead of exploitation, and to achieve this, the principles of fairness, transparency and accountability should be embraced.