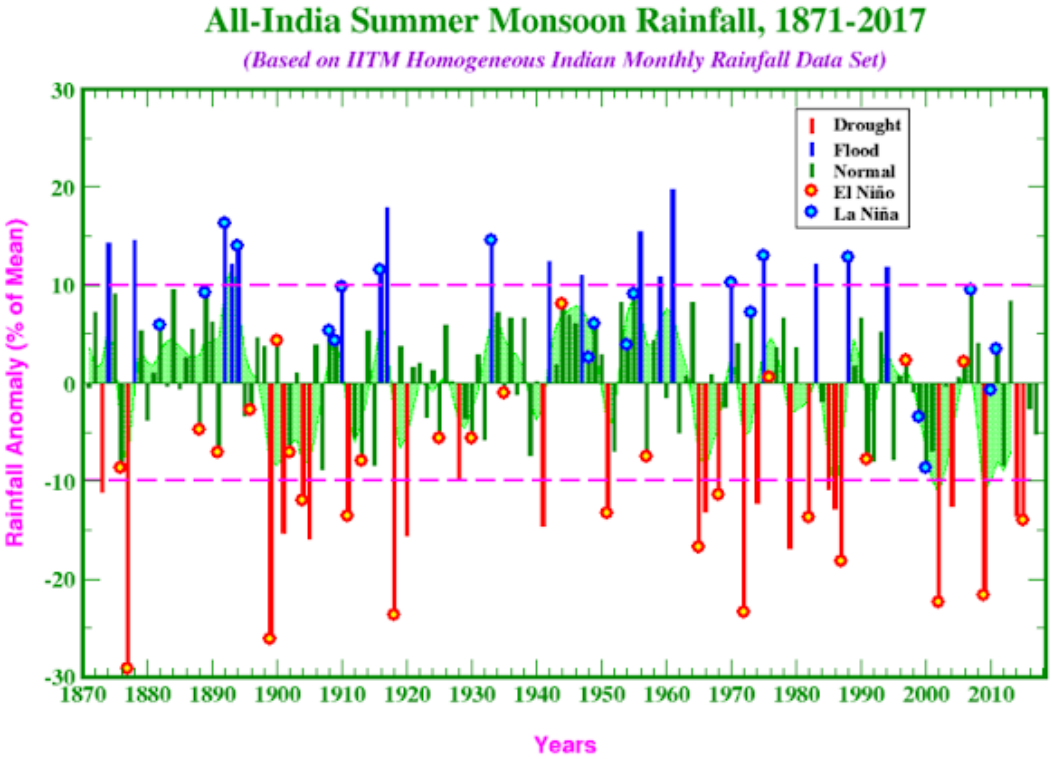
Component II: Data Visualization

Part 1: Graphic Inquisition

Figure 1.

Graphic for Inquisition



Note. Graphic obtained from Macromyths blog authored by Srinivas Thiruvadanthai (2018). It is unclear whether Thiruvadanthai is the author of the graph or if he is only referencing it.

Discussion

Figure 1 was obtained from a blog by Thiruvadanthai (2018), and it depicts the rainfall in India from 1871–2017. It also seems to aim for the inclusion of other relevant weather phenomena (El Niño/La Niña, drought, flood, and average rainfall). The figure is difficult to understand and read, and including so many variables does not serve its objective.

Gestalt Principles and Visual Structure. The figure is not characterized by simplicity since multiple elements like lines, area, data points, and colors coexist. Other Gestalt Principles are acceptable such as figure-ground differentiation and symmetry through a balance of elements.

Decoding and Operations. The figure uses the most effective elements for decoding, such as position in a scale and length, but then makes heavy use of the least effective: color—a lot of color. Anchoring through a grid is omitted, which helps keep some simplicity to it, but apart from that, the figure has too many superimposed elements that make its processing difficult in the end. Superimposing was a poor choice for this data.

Chartjunk and Data-Ink Ratio. There is some chartjunk and excess ink for the data presented in this figure: the drought/flood lines could be unnecessary and seem invasive. Color inclusion is indiscriminate and plain chartjunk. Axes are duplicated in the opposed borers of the plot, and minor breaks are unnecessary. There is a green area that is unclear which variable it represents.

Graphical Data Integrity and Lie Factor. The graphical integrity is not too bad, as there is no evidence of severe disproportionality to deceive the reader.

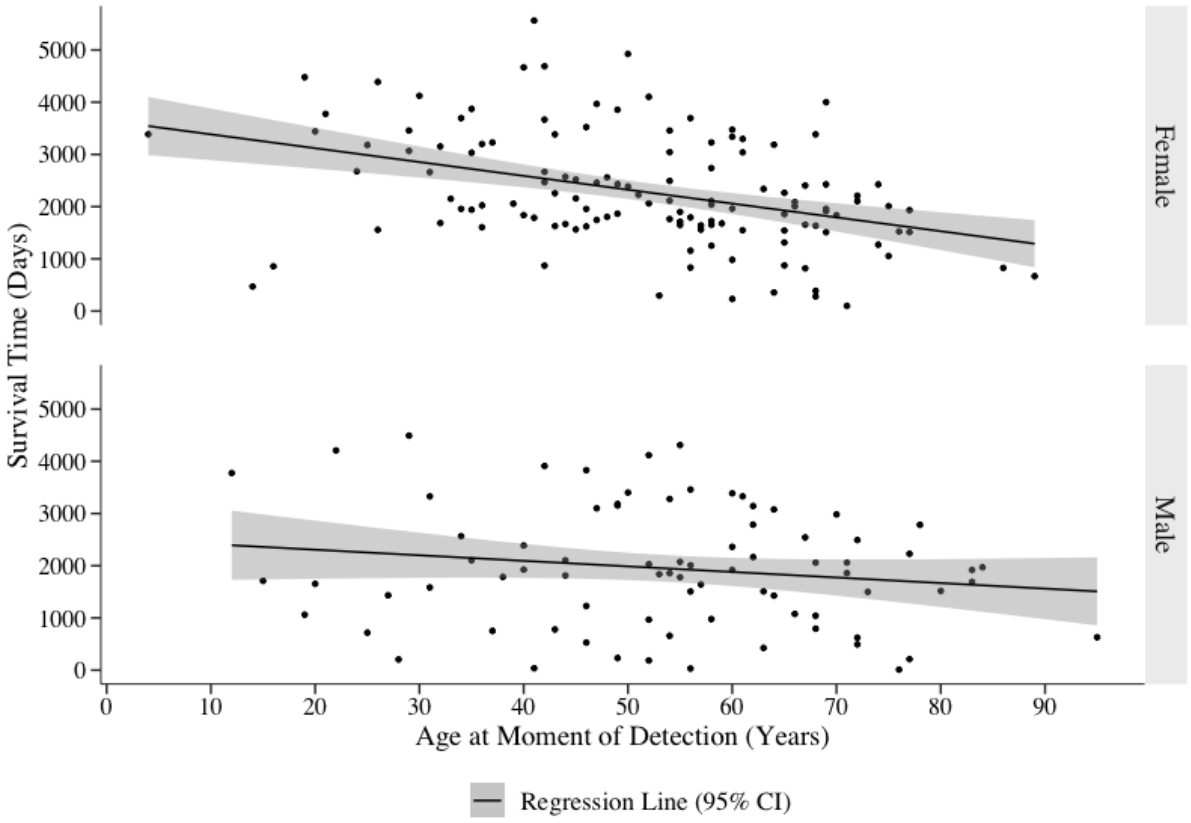
Annotation and Stand-Alone Readability. The figure can be a stand-alone read, but this is at the expense of a lot of text and information in the plot itself, so this item could be considered as failed. The annotation does not follow the APA style complicating its readability.

Part 2: Graphic Design

Designed Figure

Figure 2.

Scatter Plot Representing the Relationship Between the Age at Melanoma Detection and Survival Time



Note. Data obtained from the "Melanoma" dataset which is included in MASS R package (Venables & Ripley, 2022).

Discussion

Figure 2 is a scatter plot based on the "Melanoma" dataset from the MASS package, containing data on 205 patients with malignant melanoma in Denmark. The chosen variables were patient age in years (continuous), days of survival (continuous), and gender (binary). With the figure, I wanted to account for the potential relationship between age and survival time and compare this among gender. I opted for a scatter plot since both variables were continuous, and the data set had a relatively significant number of data points.

Gestalt Principles and Visual Structure. The designed figure complies with the Gestalt Principles. The plot is simple and clear. Both facets are symmetrical and can be read with continuity. Data points and lines are simple and similar across the figure. Proximity and similarity between the panels help to understand that they come from the same sample.

Decoding and Operations. To ease the decoding of the figure, both panels use common scales and share one axis. Color was omitted, and saturation was used only when strictly necessary. Regarding operations, grid lines were omitted as the objective was to present a general idea of the relationship rather than locate every data point. Scanning is eased by presenting consistent breaks in both survival time scales. To compare by gender, projection was chosen instead of superimposing to ease the graphic's processing.

Chartjunk and Data-Ink Ratio. The graph contains only strictly necessary information. Grid was omitted as it would not add information. No chartjunk elements are present.

Graphical Data Integrity and Lie Factor. Labels are consistent and meaningful, and the effect size of the data and graph are consistent and not exaggerated in any direction. The aspect ratio of the axes helps preserve the data's integrity.

Annotation and Stand-Alone Readability. The figure is readable by itself and is self-explanatory as it contains all the information for its understanding. It is annotated and formatted according to APA reference style.

Annex: Reference List

Thiruvadanthai, S. (2018, December 15). *Myths About India's 1991 BOP Crisis*. Macromyths. https://macromyths.blogspot.com

Venables, W. N., & Ripley, B. D. (2022). *Modern Applied Statistics with S*. Springer. https://www.stats.ox.ac.uk/pub/MASS4/