Why your data viz needs alt text

Readers who are blind or visually impaired can miss out on important content if it's in the form of a plot or diagrams. Nathan Green talks to Mine Dogucu, lecturer in statistics at UCL and co-author of an open access book, on how straightforward it has become to ensure nobody misses out on your data-based story.

When we tell stories through data, how do we know if we are reaching everyone we need to?

Alternative (alt) text (also called alt tags and alt descriptions) is the written copy that appears in place of an image on a webpage or other HTML document. It is meant to describe the appearance or function of an image on a page. Alt text should convey the purpose of an image in relation to the content of a document and, if written well, helps screen- reading tools describe images aloud to blind or visually impaired readers to aid their understanding. It also allows search engines to better crawl and rank your website, and displays on the page if the image fails to load.

In statistics and data science, images often consist of plots and diagrams. Academic publishers can take advantage of alt text to enable web accessibility to all by describing their images to those who are unable to see them clearly, if at all. Routine use of alt text will directly benefit the reader, as well as widen the readership of the work to benefit the author and science more generally.

In this interview, statistics lecturer Mine Dogucu talks about the importance of providing alt text in academic publishing, how she sought to make her first book, Bayes Rules! An Introduction to Applied Bayesian Modeling, released in 2022, accessible and inclusive, and what the process taught her and her co- authors about accessibility. At the Bayes Rules! team's request, tools in R such as R Markdown and Quarto now support alternative text.

Bayes Rules! is an online book, right?

That is absolutely right. It's freely available at bayesrulesbook.com and a hard copy is also available from CRC Press. When we first started writing the book, before even reaching out to publishers for proposals, our first criterion was making the book freely available online. Thankfully, our editor agreed to it.

Can you explain to me how you incorporate accessibility into your work?

We just talked about having an open-access book, for instance. So the book website is actually visited by hundreds of readers, sometimes even thousands, every day from many different countries. If we just had a physical book, I am not sure it would have been so widely read around the world. One aspect of accessibility is lowering costs, which can be a big barrier in education. That is why I try to make my teaching materials, slides, handouts, etc. open-access as much as possible. When we were writing the book, we held ourselves accountable to learn more about accessibility and focused on visual accessibility. For instance, we wanted to make sure the plots would be accessible to colour-blind people. Also something exciting happened for the R community at this time. We wanted to provide descriptions of plots for blind and visually impaired readers. We were writing the book using the Bookdown package which is R Markdown- based, but at the time R Markdown did not support alternative text. We requested this feature from the R Markdown team and now R Markdown – as well as Quarto, which is the new-generation R Markdown – supports alternative text. So we were able to write our book with alternative texts for figures, which was important for us.

How does that practically work? You write the description and then record an audio which is attached to the book?

Alt text is actually just text, not an audio file. If provided, it is available on the back end of a web page. Users who rely on screen readers can hear the text of the book being read out. However, if there is a figure without an alt text then they would hear something like "figure2.png". When we provide a detailed description of the figure in the R Markdown document as alt text, the screen reader would actually read the alt text. I should also note that the alt text feature is not specific to R Markdown or statistics per se. We should be utilising them on Twitter or even in PowerPoint presentations, for instance.

Is there a format that you can use to describe a certain type of plot, for example? Has it been standardised?

I would not say it has been standardised, but there are recommended ways of doing it. First of all, there is an automatic way to do this. The BrailleR package in R supports this with the VI() function. For instance, for a bar plot, the VI() function would return text with the labels on the x- and y-axes, the height of the bars, and so on. However, any automated tool comes with limitations. For instance, one cannot utilise this function for verbalising scatterplots. Providing a manually written alt text is a better option. There are two resources I frequently rely on. One is Amy Cesal's post on Medium on writing alternate text for visualisations (bit. ly/3iWKPK7). The other is a talk by Silvia Canelón and Elizabeth Hare given at the CSV Conference (bit.ly/3BySbKq). Overall the recommendations include writing the chart type and variables as well as their units, for example: "Bar chart showing revenue in billions of dollars made by meat substitutes where there has been a 66% increase between 2012 and 2022". Most importantly, the bigger message of the plot needs to be included in the alt text.

Could I use that in journal papers?

Great question! We absolutely should, but the journals make it very difficult to do so. Once I wrote a short piece for a professional organisation's magazine where I was actually talking partially about alt text. When the editor approached me, I had told them that I would only write the piece under the condition that they are able to include my alt text in the online version of the publication, to which they agreed. This publication has issues dating back to the 1950s hosted online. Unfortunately, my piece was not published with alt text and I had to insist on it. Eventually, it was published by itself in HTML as a separate page. I have not checked recently but it might still be the only piece with alt text in this publication. The only organisation I have encountered that supports alt text is American Computing Machinery, at least in my limited experience. They also have an active accessibility group. Also, it is worth noting that the useR! conference provides guidelines on accessible presentations. Some of the organisers of useR! 2021 recently published on how to host an inclusive conference (bit.ly/3VSsAEr). I hope more organisations will consider accessibility. Without accessibility, we are deliberately excluding people from the statistics community.

So if anyone wants to do it in Bookdown, now they can?

Yes, they can do this in any package from the R Markdown family such as Bookdown, Blogdown or Distill. It is also possible to use it in Quarto. This experience made me think about why it took so many years for the R community to have this feature. I blame statistics educators because I am one! We don't know enough about accessibility to teach it because we were not taught about it. Then our students design tools that are not accessible. So I wanted to break this cycle and teach about accessibility. Luckily I was able to get grant support from the Teach Access network. Through this work, I collaborated with one of the most prominent experts, JooYoung Seo, to develop curricular

materials on teaching accessibility as part of the data science curriculum. There is still a lot of work that needs to be done, but we really need to make sure the next generation of statisticians is more educated than we were on this topic.