**1. Bar plots**

Bar plots are a close relative of box plots.

**2. When should you use a bar plot?**

They are usually used when you want counts or percentages of a categorical variable. Though less common, it is possible to calculate a different number for each category. An important constraint is that the value zero should be important in some way, since the bars extend to zero.

**3. bros**

In 1988, the pop band Bros asked the question "When will I, will I be famous". Let's use data science to find out.

**4. ESPN 100 most famous athletes from 2017**

Let's take a look at a dataset of the world's most famous athletes from 2017, as judged by ESPN, a US TV channel. The athletes were ranked according to things like how many social media followers they have, how much money they make from endorsing products, and internet search popularity.

1. 1 https://data.world/wesmagee/2017-espn-world-fame-100

**5. Bar plot of counts by country**

Here's a bar plot of the number of athletes on the list from each country. The categories, that is the countries, are on the y-axis, and the x-axis shows the counts.

**6. Vertical bars**

It's possible to swap the axes to show categories on the x-axis and counts on the y-axis. That makes the countries harder to read because you have to tilt your head to read the vertical writing, so horizontal bars are preferable here.

**7. Sorting by count**

Usually, you'll want to sort the bars by the count. This makes it easier to see that, for example, Spain and India are tied for fifth place, with four athletes each.

**8. Children's fruit and veg consumption**

Here's a dataset from the Health Survey for England in 2018. The survey asks many health-related questions, and this particular dataset focuses on a question about how many portions of fruit and vegetables children eat per day. We have two categorical variables; the number of portions eaten and the year. The metric is a percentage rather than a count.

1. 1 https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/2018/health-survey-for-england-2018-data-tables

**9. Stacking bars**

Since the percentages of children for each year always add up to 100%, it's helpful to stack the bars on top of each other. In 2001, for example, you can see that the bottom two blocks reach 25%, meaning that 25% of children ate at least four pieces of fruit and veg per day in that year. In 2003, the UK government started a campaign to encourage people to eat five portions of fruit and veggies per day. Look at the bottom blocks, and notice that the percentage of children eating five portions increased each year from 2003 to 2006, and stayed roughly constant until 2014. Similarly, the pale blocks at the top of the plot show that the percentage of children eating zero portions per day decreased from 2003 to 2006 then stayed constant. It looks like the campaign was a success.

**10. Bar plots vs. box plots**

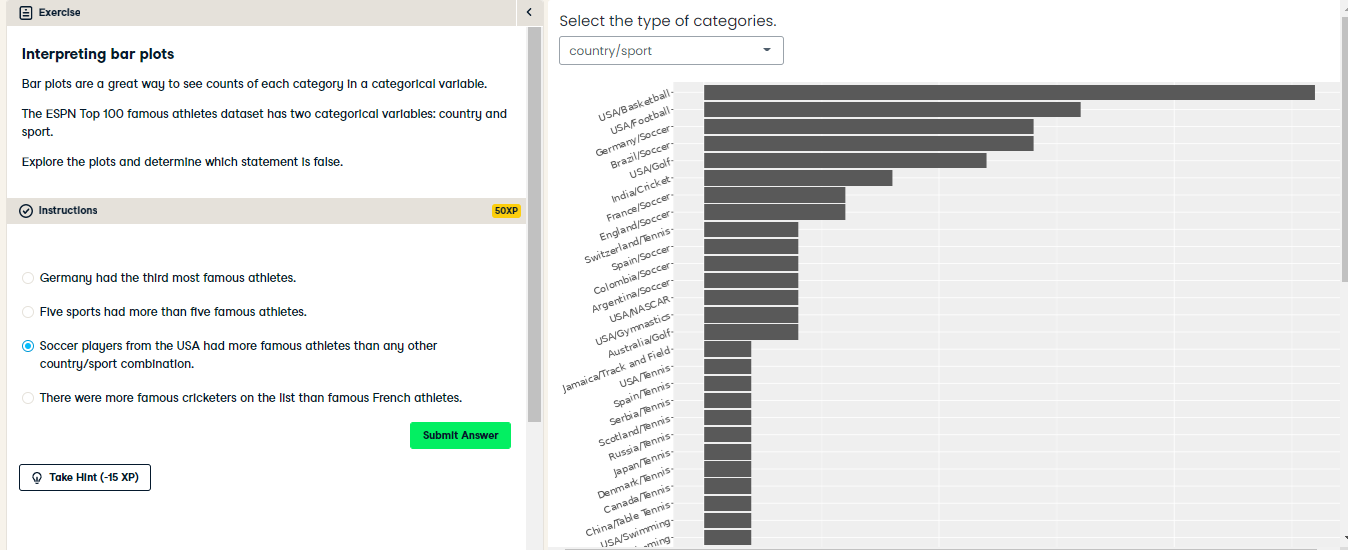
Let's consider the relationship between box plots and bar plots. Here are the box plots of the age that English and British monarchs started ruling, split by royal house. A bar plot of counts by house has a similar form: the categories are on the y-axis, and the x-axis is numeric. The difference is that the box plot is designed to answer questions about the spread of a variable, and the bar plot is designed to answer questions about a single metric relative to zero, in this case count.

**11. Other metrics than counts**

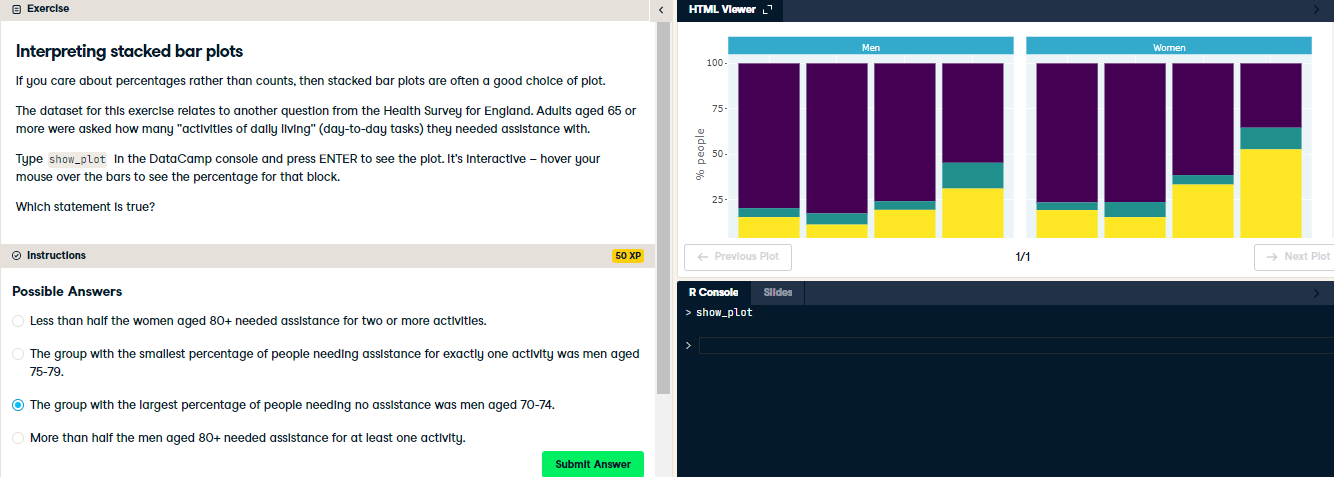
As mentioned earlier, count is not the only metric you could show on a bar plot. Here, the mean age at the start of rule is shown instead. It's a perfectly valid plot, but since it only shows one value per bar, it feels less exciting than the box plot on the left.

**12. Let's practice!**

Time to begin bar plotting.



Badass bar plot interpretation! Although USA had the most famous athletes by country, and soccer had the most famous athletes by sport, there were no famous soccer players from the USA. By looking at both categories at once, you saw that the most common categories for famous athletes were basketball and football players in the USA, and soccer players in Germany and Brazil.



Stunning stacked bar plot interpretation! For percentages of categorical variables, stacked bar plots are an excellent choice of visualization.