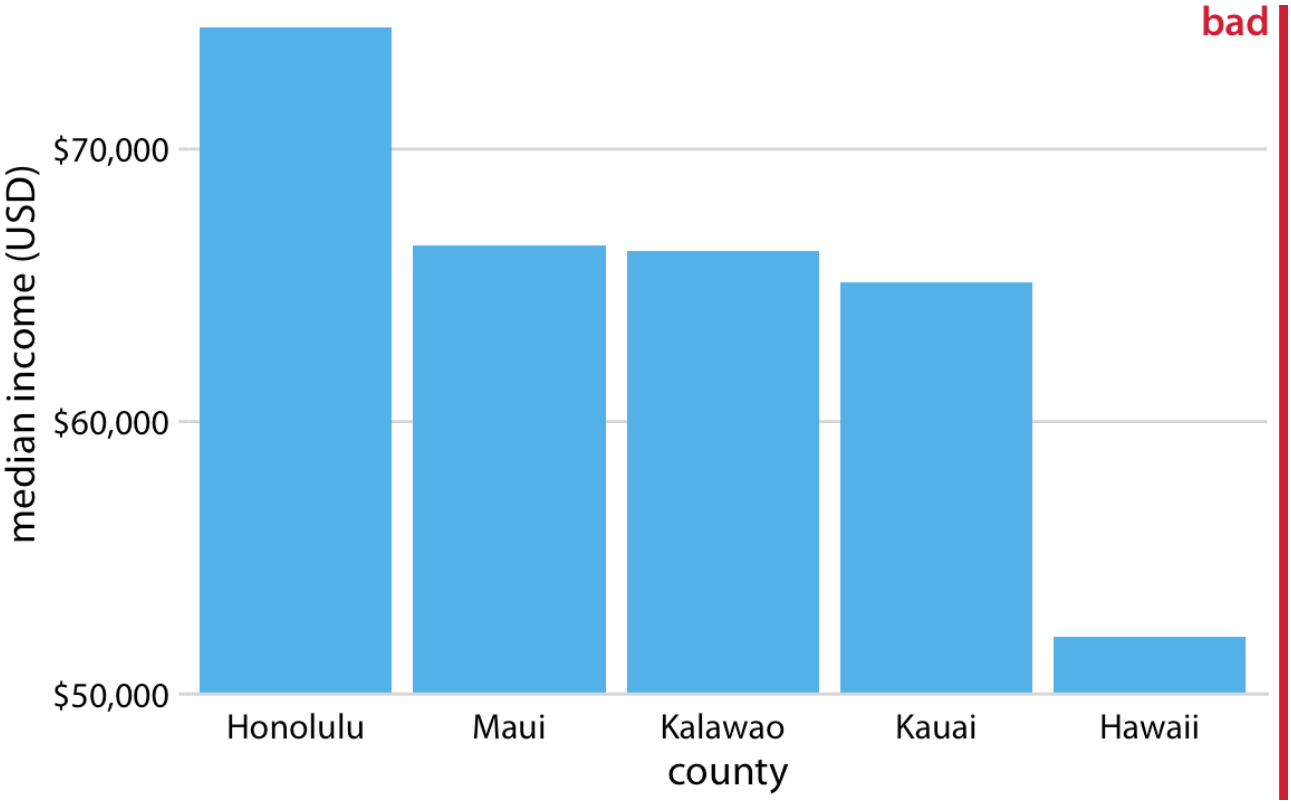
The principle of proportional ink -

The sizes of shaded areas in a visualisation need to be proportional to the data values they represent.

In short the the graphical content(shape, size, colour, etc.) we plot on graph must represent the data value they should not be inconsistent.

Visualization along linear axes -

Starting with example, figure shown below, shows the median income in the five counties. This figure suggest that the county of Hawaii is incredibly poor while the county of Honolulu is much richer than other counties. It is quite misleading, because all bars begin at $50,000 median income. In the case of bar plot the bar height is the key quantity we perceive when looking at this figure, not the location of the bar endpoint relative to the y axis.



But when we look at the figure below median income is comparable to other counties. This figure doesn't tell the story that the Hawaii is incredibly poor and Honolulu is much richer than other this is because the bars are starting from 0.

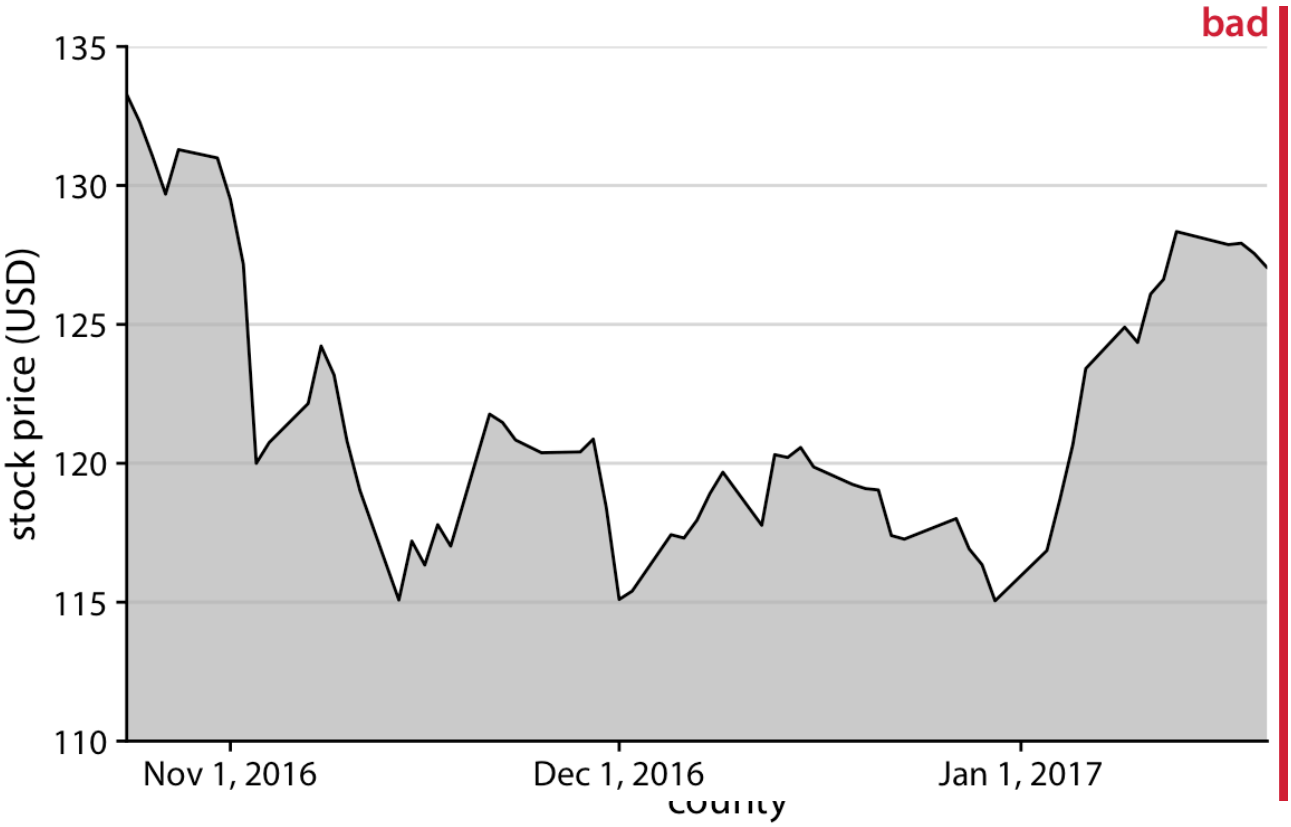
Note - Bars on a linear scale should always start at 0.

Similar visualization problems frequently arises in the visualisation of time series, such as those of stock prices. Let's take an example of Facebook stock price massive collapse occurred on Nov. 1, 2016. In reality, the price decline was moderate relative to the total price of the stock. Shaded area underneath the curve becomes problematic, because shading emphasises the distance of y axis for a particular value of x, thus it creates the visual impression that the height of the shaded area at given given day represents the stock price of that day instead, it only represents the difference in stock price from the baseline which is $110.

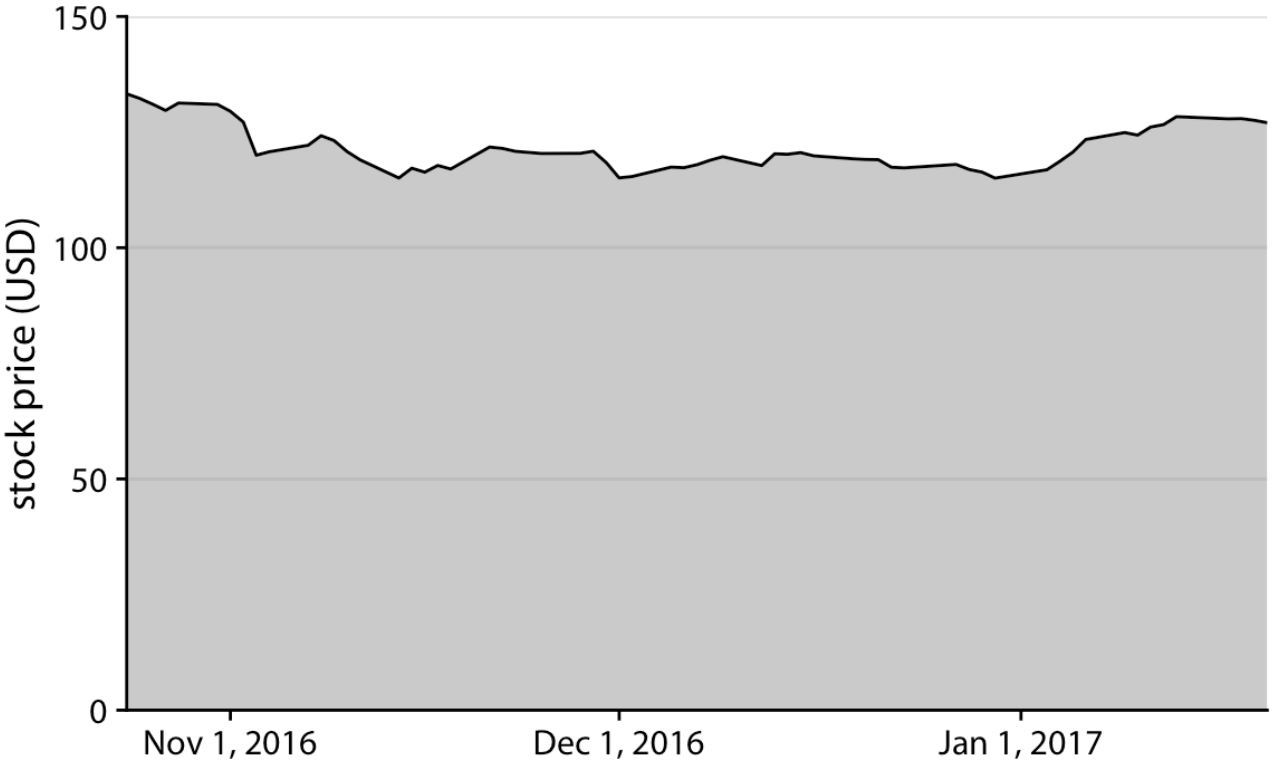
Figure: Stock price of Facebook (FB) from Oct. 22, 2016 to Jan. 21, 2017. By showing the stock price on a *y* scale from $0 to $150, this figure more accurately relays the magnitude of the FB price drop around Nov. 1, 2016.

Figure: Loss in Facebook (FB) stock price relative to the price of Oct. 22, 2016. Between Nov. 1, 2016 and Jan. 1, 2017, the price remained approximately $15 lower than it was at its high point on Oct. 22, 2016. But then the price started to recover in Jan. 2017.

Visualizations along logarithmic axes -

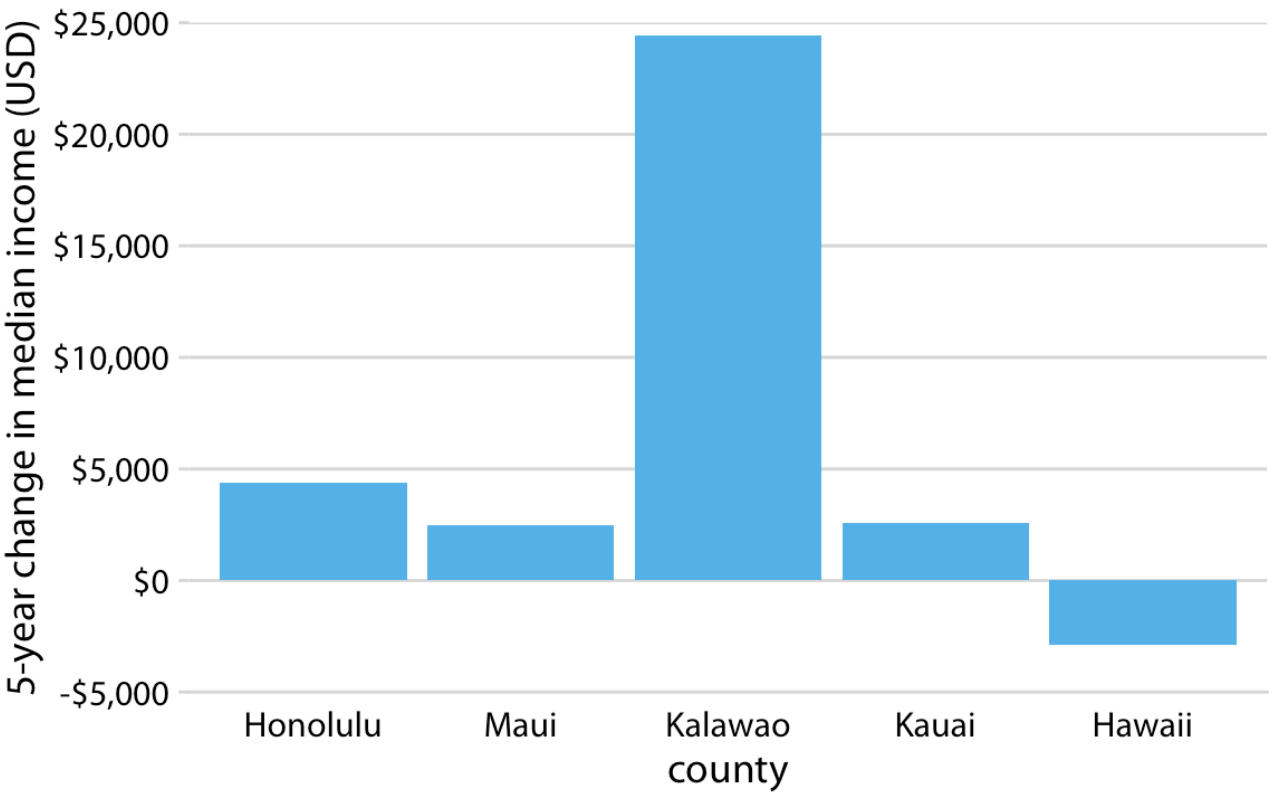


If we wants to visualize ratios instead of amounts then log scale are perfect.

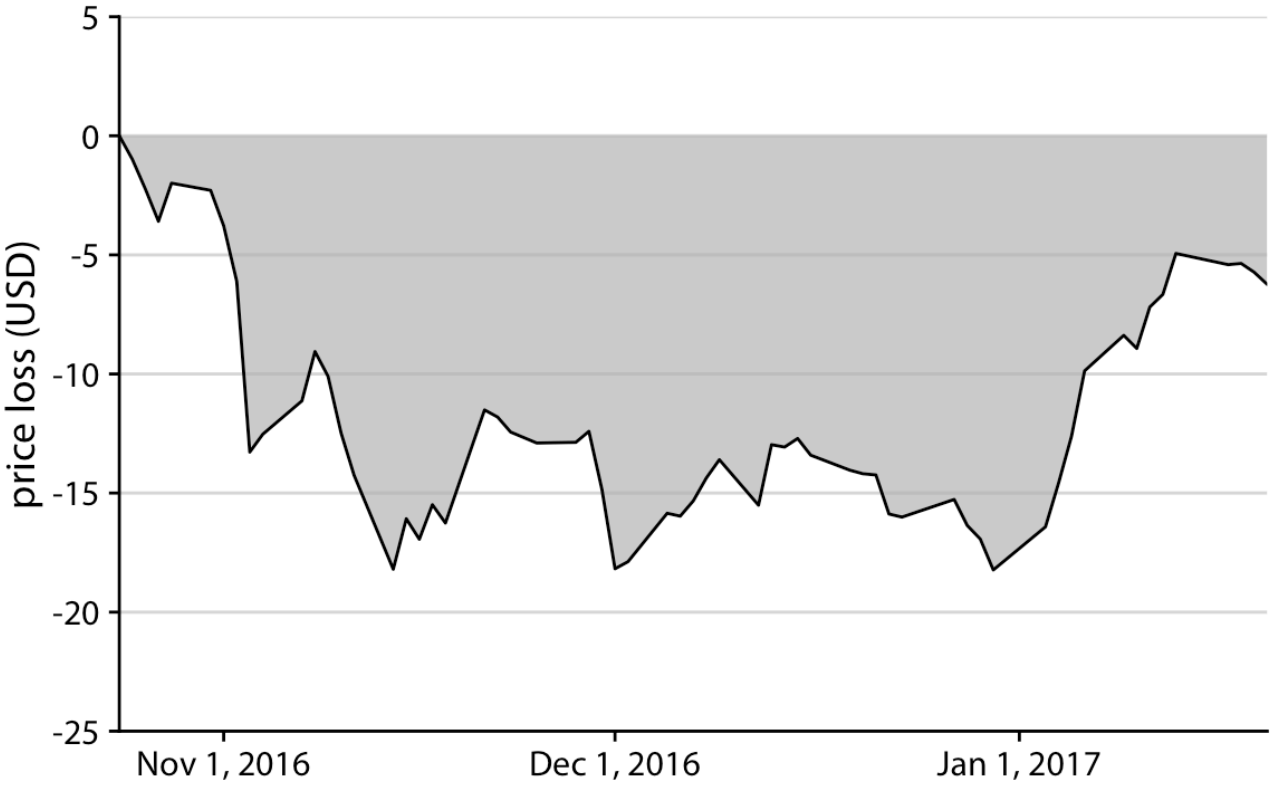


Remark - the bars or trends that doesn't start from zero is like we are drawing the zoom in version of that(that are starts from zero) graph, we can see this in above examples.

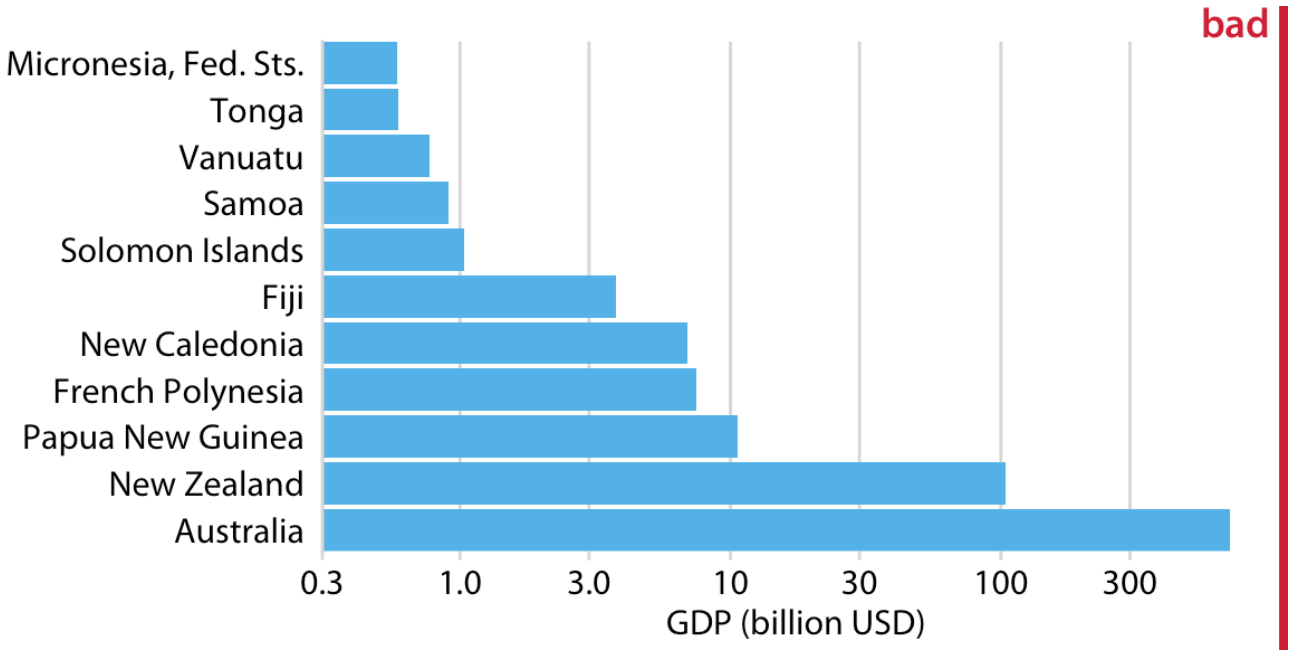
To represent the small changes over time or difference between conditions, since we always have to draw the whole bar or area starting from 0. Let's take above example of median in come in five counties, since it is not prefect case. For all counties except Kalawao, this change amounts to less than $5000(Kalawao is an unusual county, with fewer than 100 inhabitants, and it can experience large swings in median income from a small number of people moving into or out of the county.) And for Hawaii county the change is negative.



Similar we can daw the change in Facebook stock price over time as the difference from its temporary high point Oct. 22, 2016.



When we are visualising data along a linear scale(eg. amount in a bar plot shows the value along a single axes either x or y), the area of bars, rectangles, or other shapes are automatically proportional to data values but it is not true for logarithmic scale because data values are not linear. That's why bar graphs are falls under 'bad' category but on the other side area of each bars is proportional to the logarithmic of the data value, and thus bar graphs satisfy the principle of proportional ink in log transformed coordinates.



We already discussed that bars height must always start from zero otherwise it leads to data manipulation. The above figure is also a 'bad' because it is not starting from zero. Think again, a bar should start from 0 but when we transformed it to log it becomes infinite(log(0) = undefined). For that, unwillingly we have to start scale from an arbitrary number, so the problem is always there.