Proportional Representation

The Principle of Proportional Ink

Introduction

- **What is Proportional Ink?**
  - Data values represented by graphical elements (e.g., bars, dots).
  - Need for consistency between graphical element size and data value.
- **Principle**: Sizes of shaded areas in a visualization should be proportional to the data values.

Common Visualization Errors

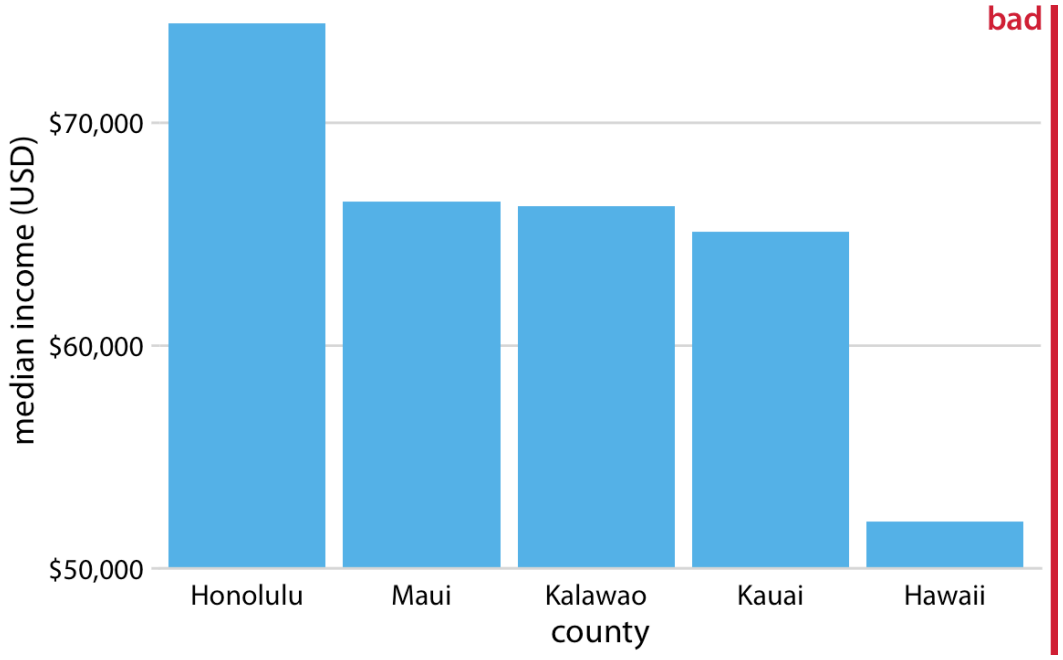
- **Internal Inconsistencies**
  - Bars starting at non-zero values can mislead perception of data.
  - Example: Bar lengths represent the extent from an arbitrary start, not the data value.

The Principle of Proportional Ink

- **Definition**: Sizes of graphical elements must be proportional to the data they represent.
- **Bergstrom and West's Concept** (2016).
- **Common Violations**: Misleading graphs, especially in media and finance.

Linear Axes and Misleading Visualizations

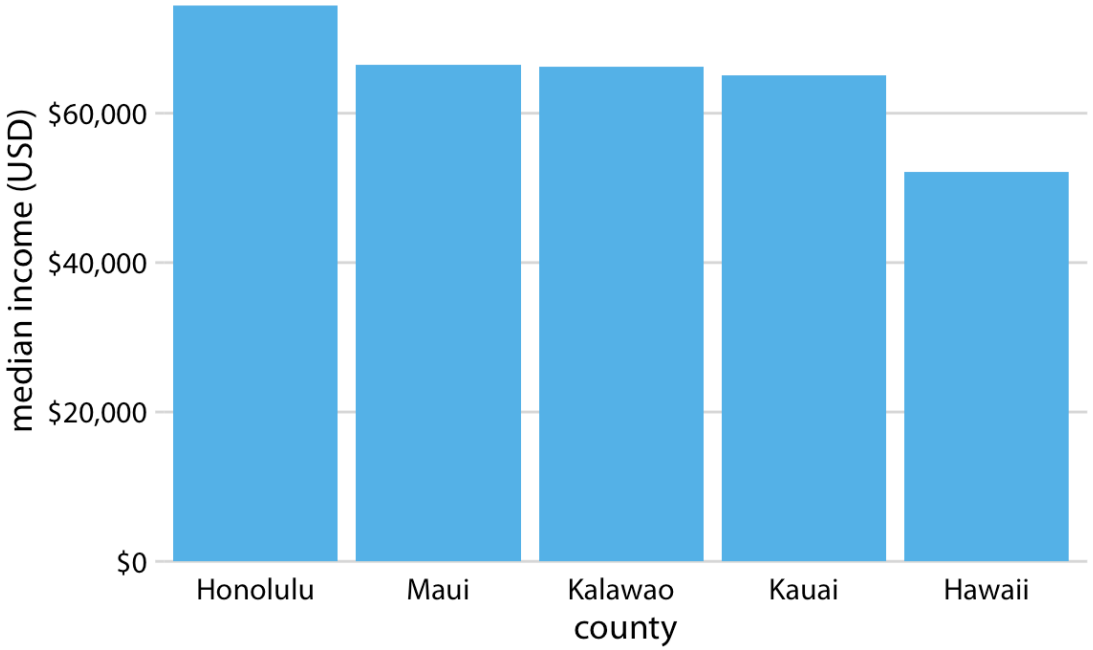
- **Example**: Median Income in Hawaii
  - Issue: Y-axis starts at $50,000.
  - Result: Exaggerates income differences.



Median income in five counties of Hawaii. The y-axis starts at $50,000, making income differences appear exaggerated.

Correct Visualization on Linear Axes

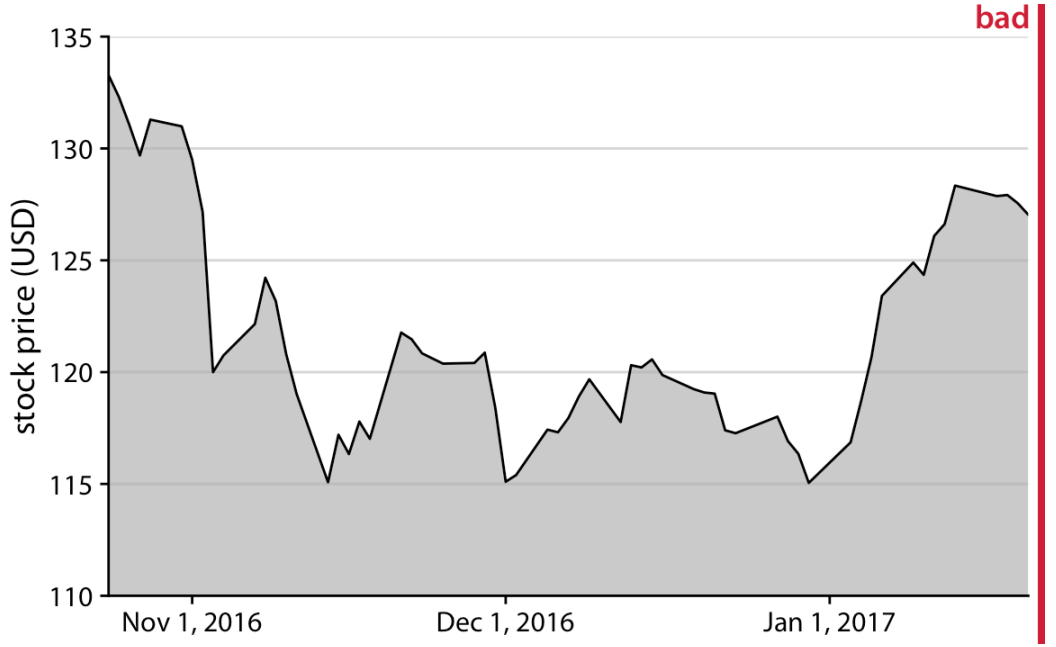
- **Example**: Median Income in Hawaii
  - Proper Y-axis starting at $0.
  - Accurate representation of relative magnitudes.

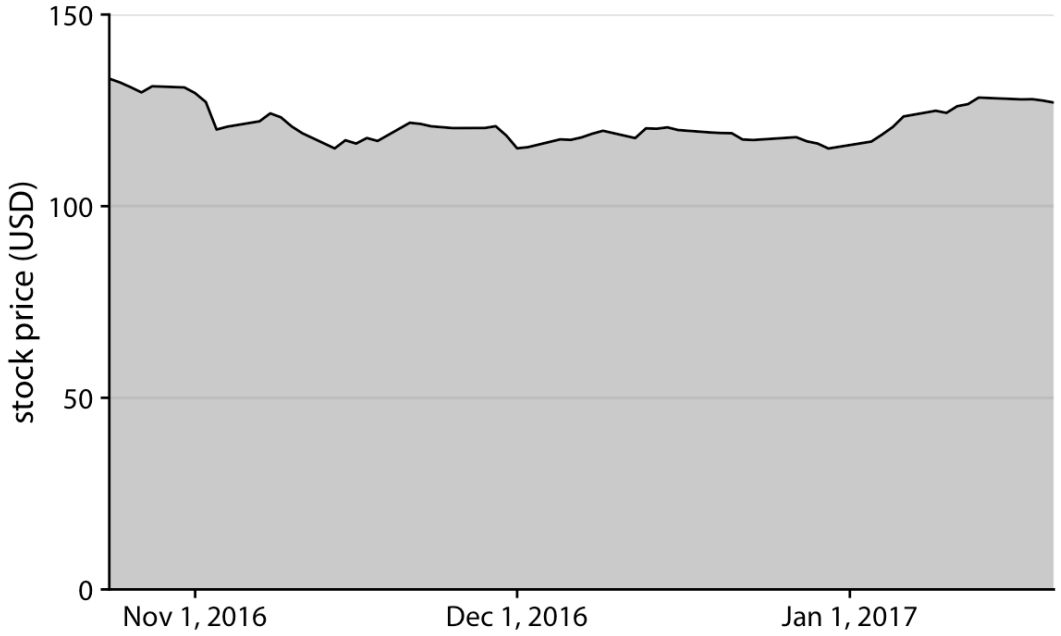


Median income in five counties of Hawaii with a y-axis starting at $0, accurately representing relative magnitudes.

Time Series and Stock Prices

- **Example**: Facebook Stock Price
  - Issue: Y-axis starting at $110, exaggerates price drop.
  - Corrected visualization with proper scale.

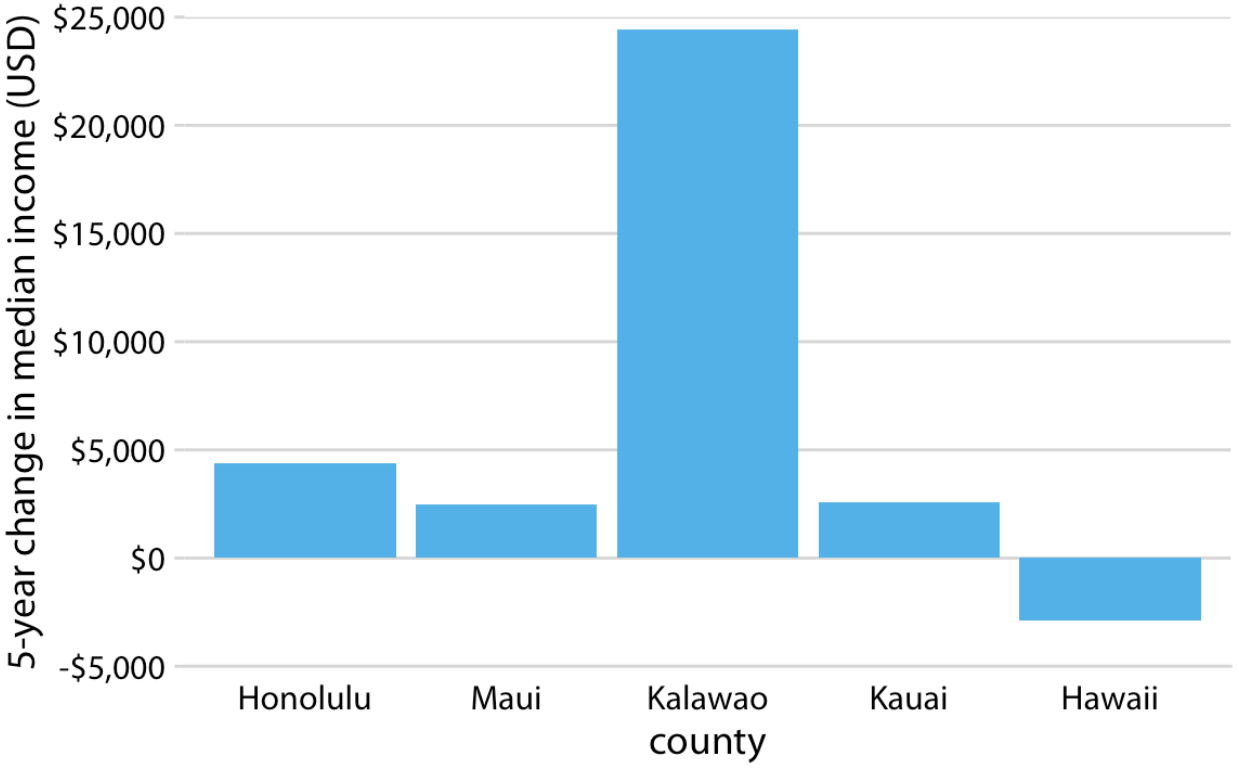




Stock price of Facebook (FB) from Oct. 22, 2016, to Jan. 21, 2017. The first figure exaggerates the price drop by starting the y-axis at $110, while the second, with a y-axis from $0 to $150, provides a more accurate representation.

Correct Use of Bars for Small Changes

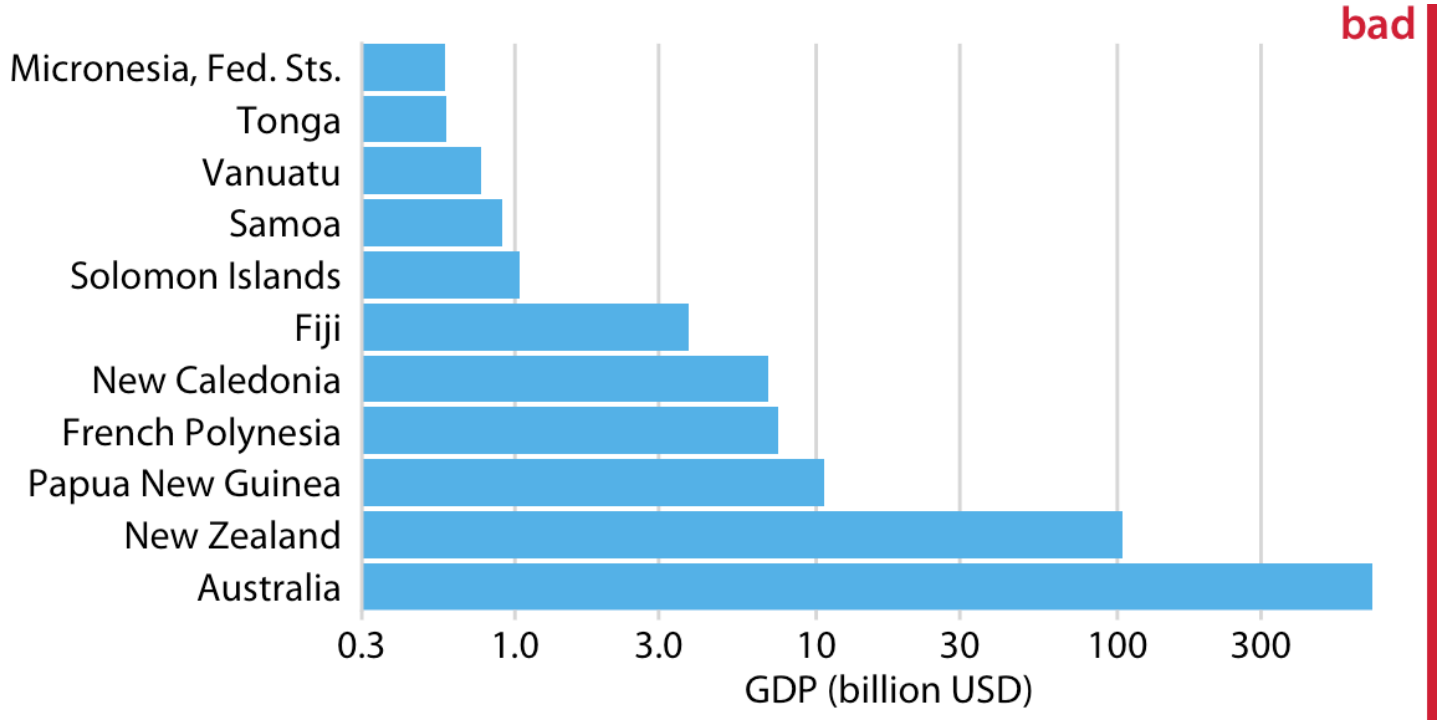
- **Visualizing Change**: Median Income Changes
  - Bars can show differences, as long as the starting point is clear (0 for amounts, baseline for changes).



Change in median income in Hawaiian counties (2010–2015).

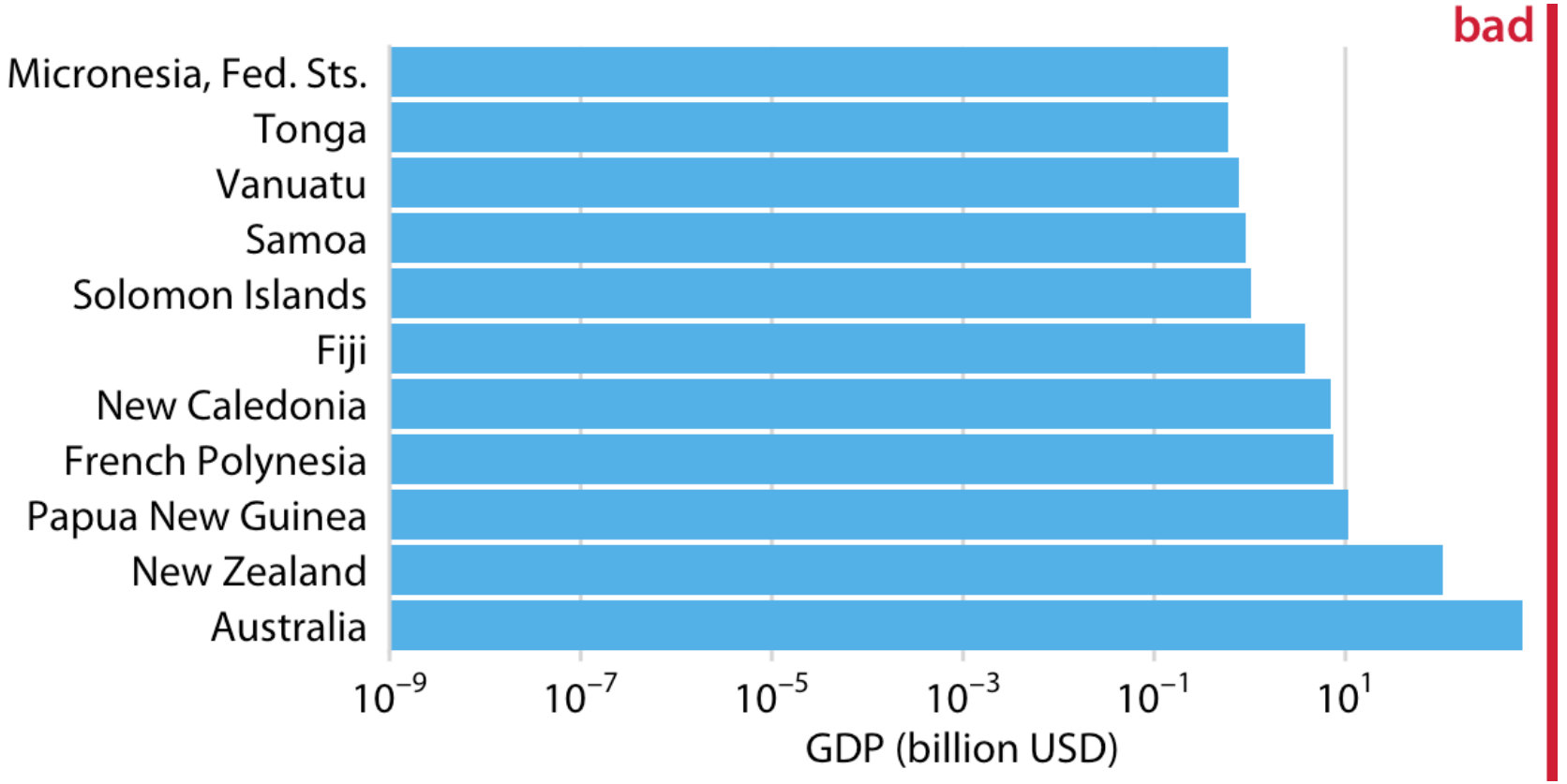
Logarithmic Axes

- **Log-Scale Visualizations**: Pros and Cons
  - Bars on log scales represent ratios, not amounts.
  - Example: GDP of countries



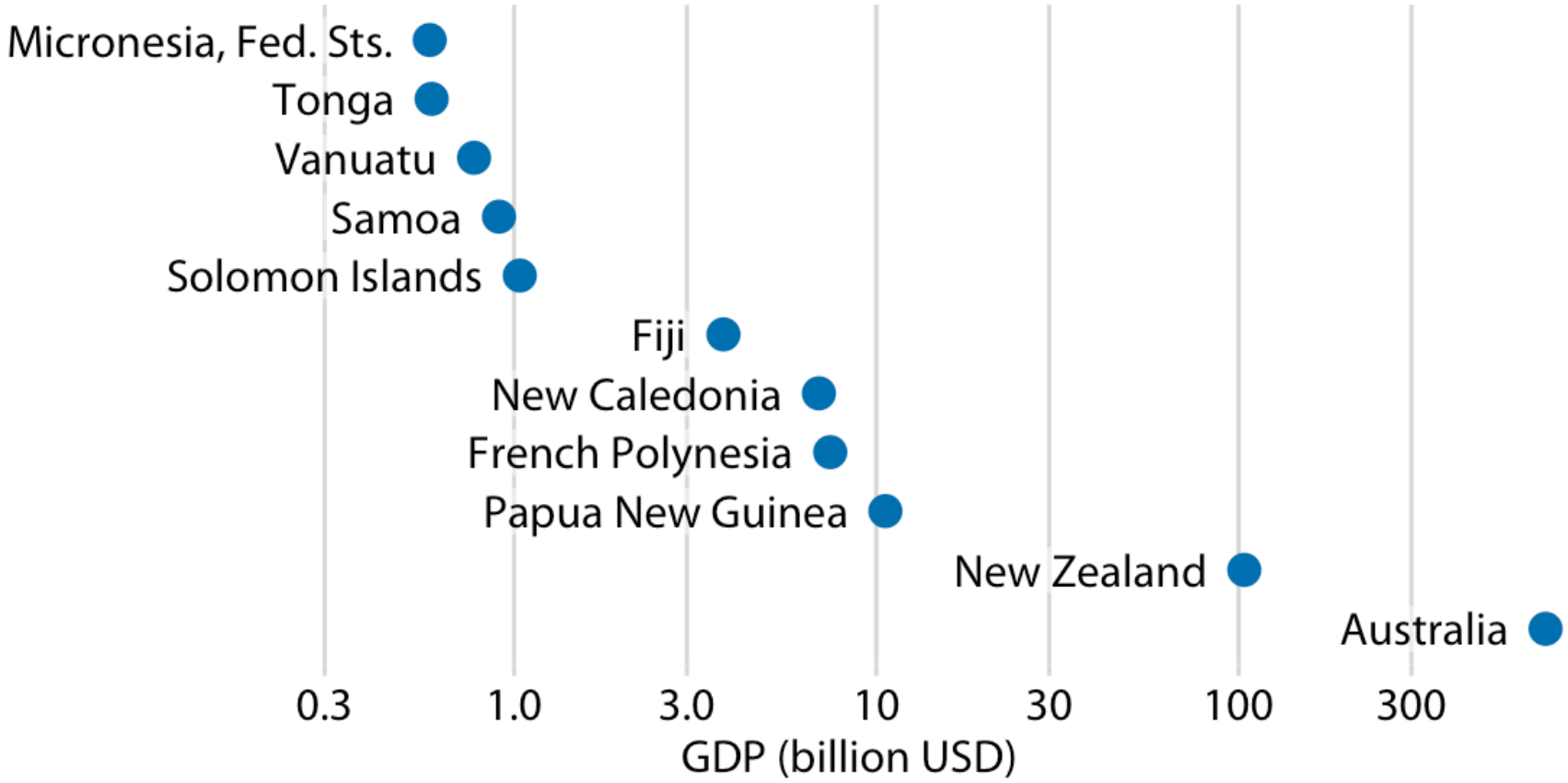
GDP of Oceania countries in 2007. Bar lengths are distorted as they start at 0.3 billion USD.

Bars on log scales (continued)



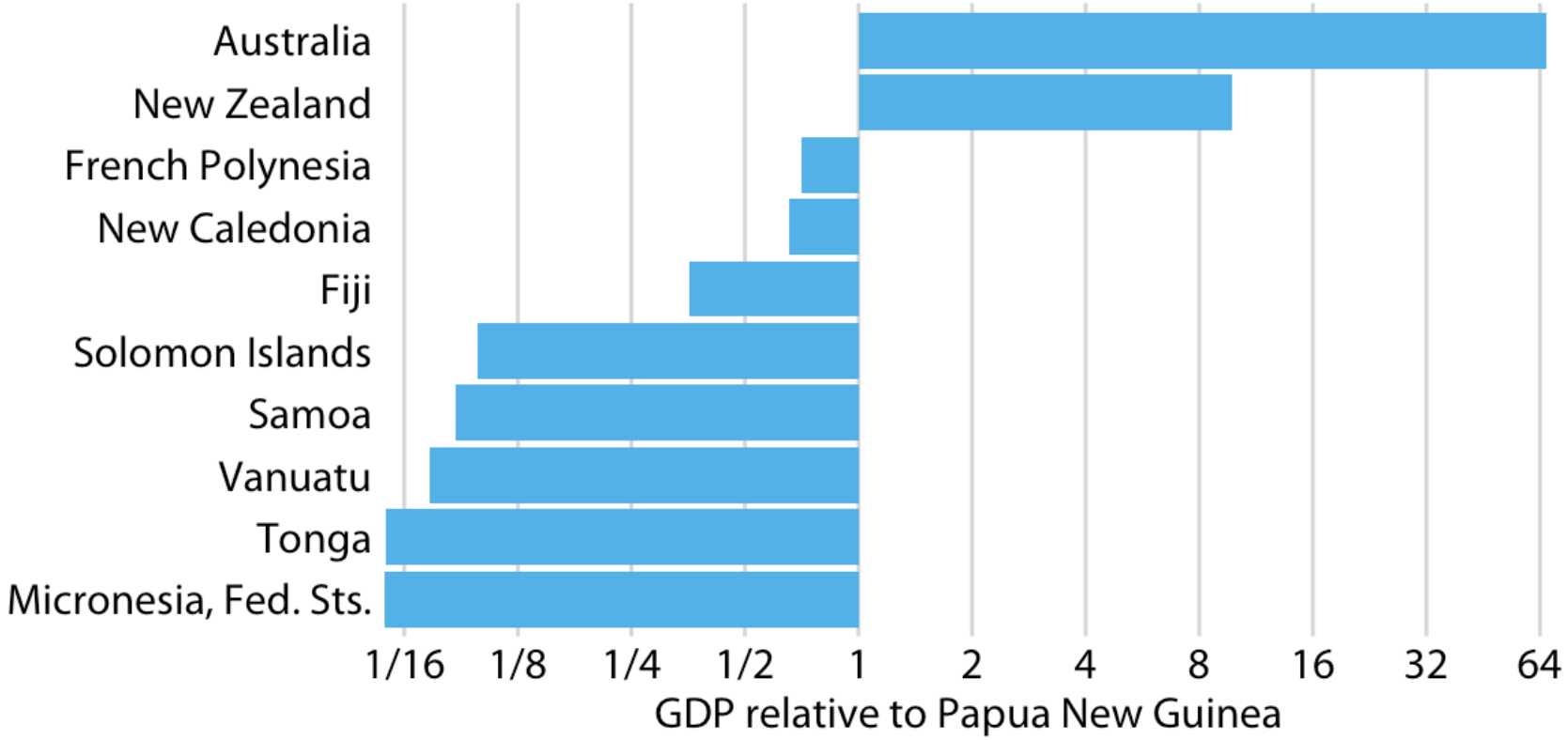
GDP of Oceania countries in 2007. Bars are distorted as they start at 10^{-9} billion USD.

Dots instead of bars



GDP in 2007 of countries in Oceania.

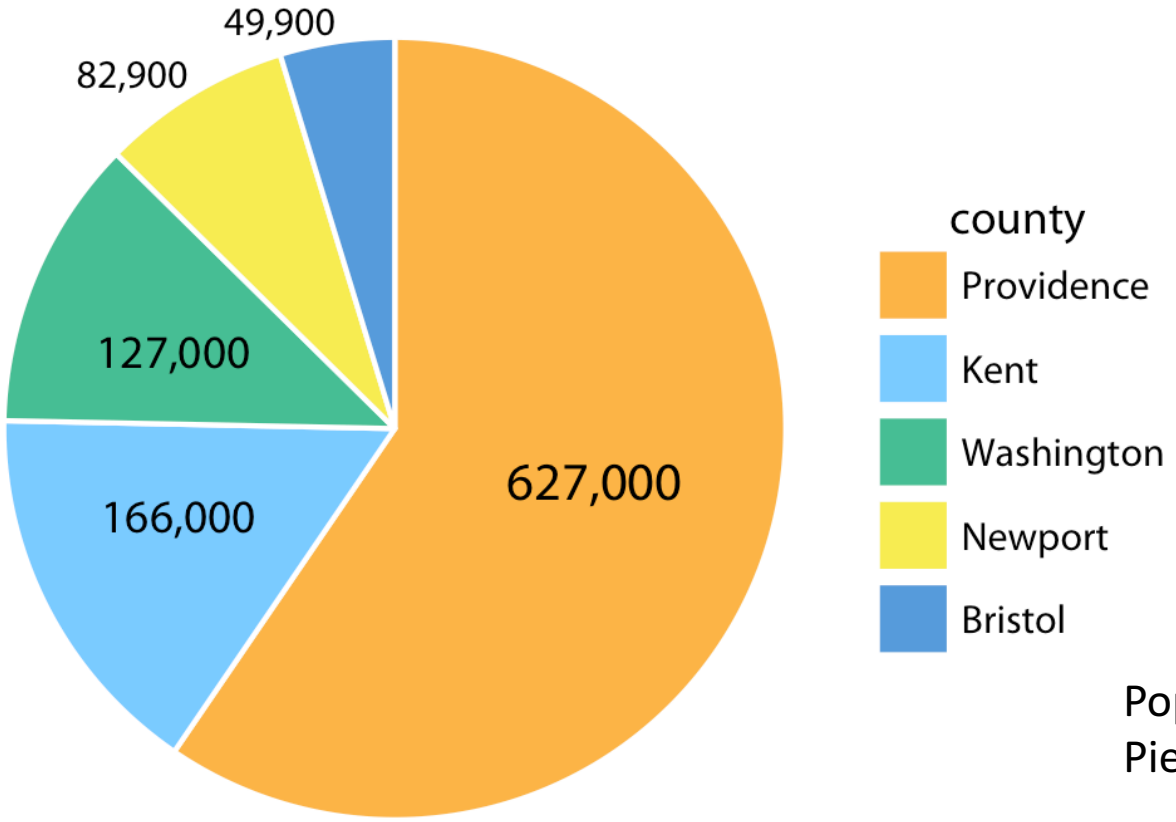
Bars on a log scale



GDP in 2007 of countries in Oceania, relative to the GDP of Papua New Guinea.

Direct Area Visualizations

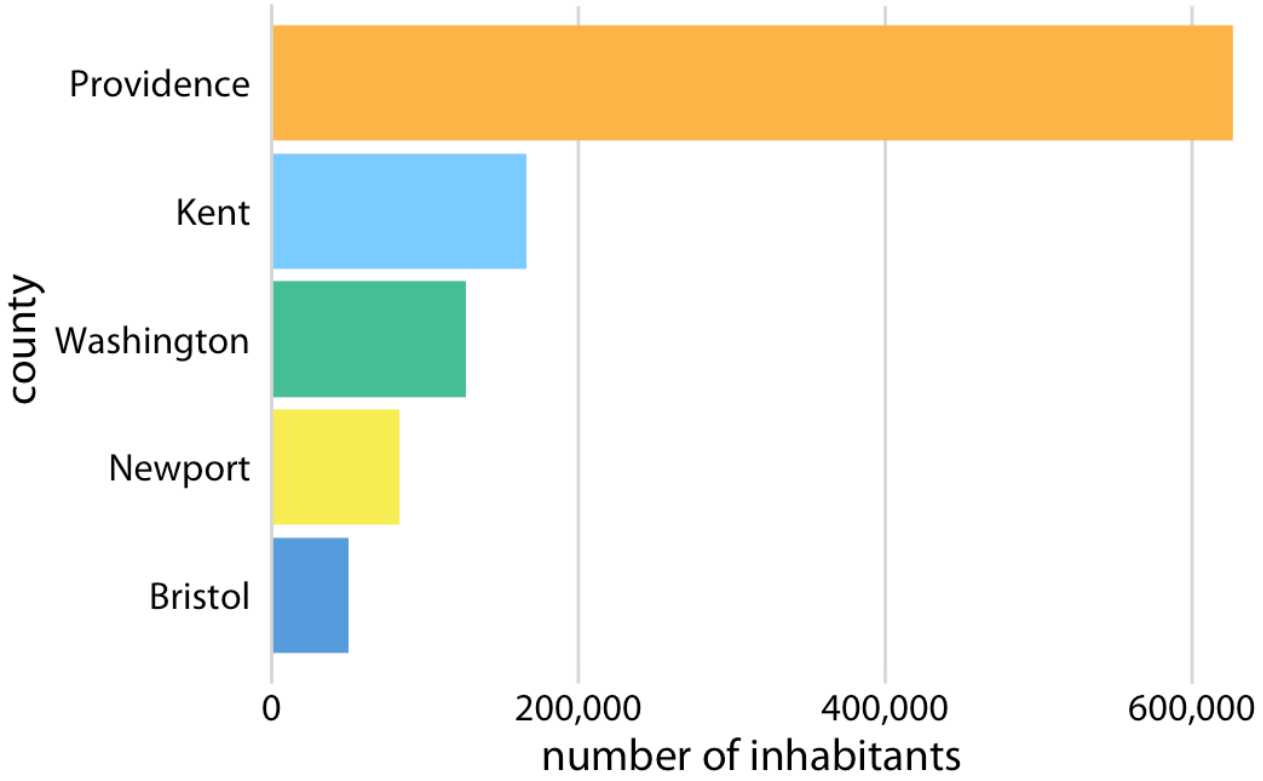
- **Pie Charts**: Area proportional to data values
  - Pie charts satisfy proportional ink but are less perceptible than bar charts.



Population distribution in Rhode Island counties.
Pie wedges are proportional to county populations.

Perception of Areas

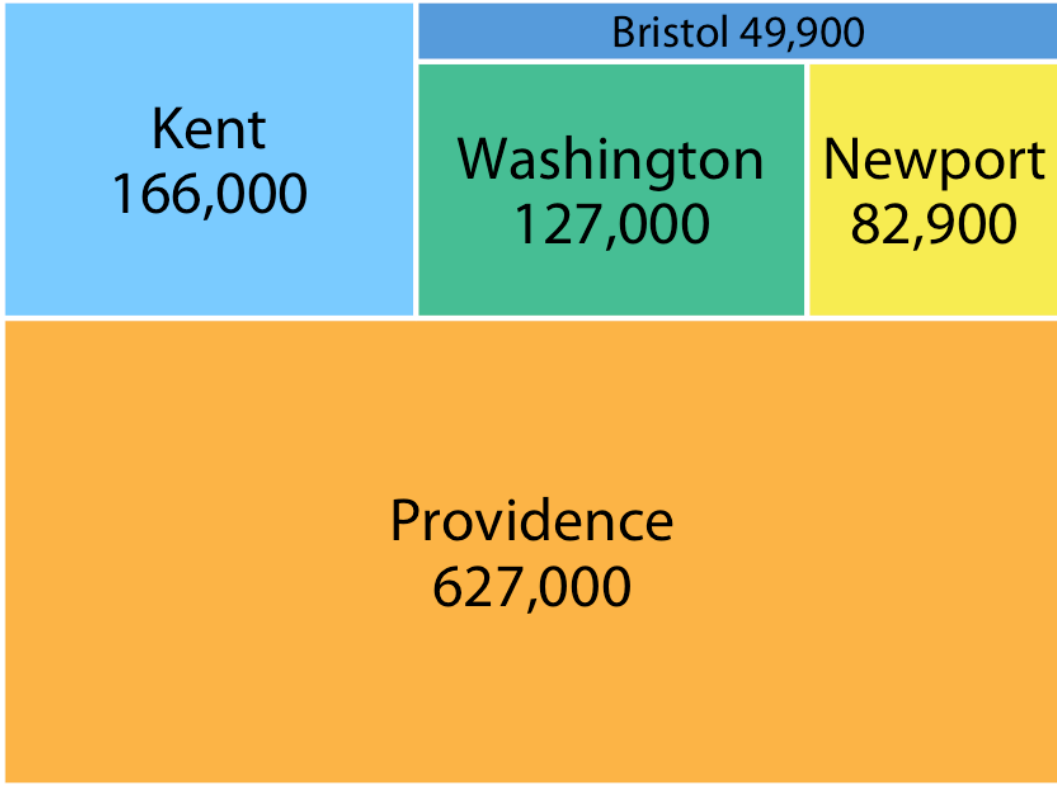
- **Human Perception**: Better at judging lengths than areas.
  - Example: Pie chart vs bar chart



Population in Rhode Island counties, shown as bars proportional to county populations.

Treemaps and Areas

- **Treemaps**: Proportional areas for data values.
  - Perception issues arise: Differences seem less pronounced compared to bar charts.



Population in Rhode Island counties, shown as a treemap with area proportional to county populations.