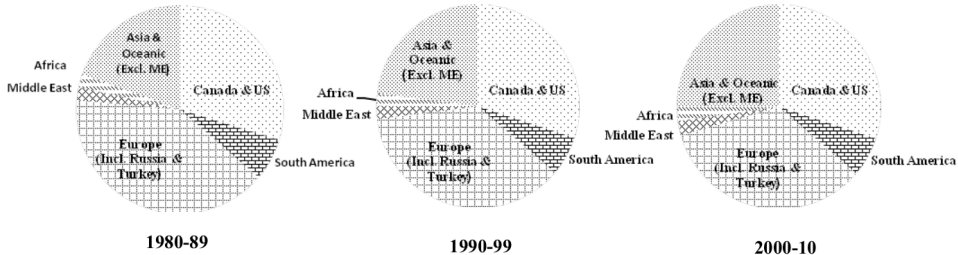




Principles of Statistical Graphics

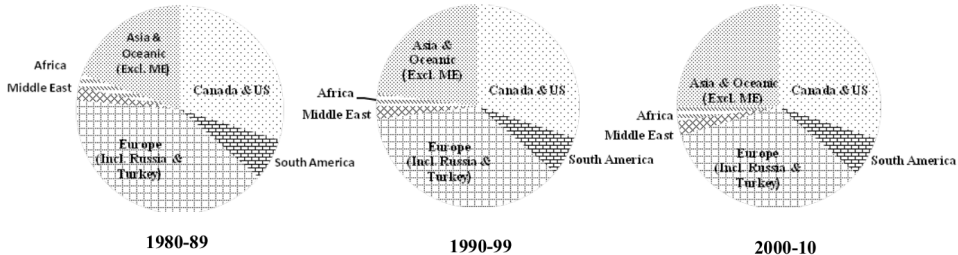
What's wrong with this?

Annual GDP



What's wrong with this?

Annual GDP



How would you graph the same data?

Cleveland's paradigm

A graph *encodes* quantitative and categorical information using symbols, geometry and color. Graphical perception is the *visual decoding* of the encoded information.

- The graph may be beautiful but a failure: the visual decoding has to work.
- To make graphs work we need to understand graphical perception: what the eye sees and can decode well. visual perception.

Cleveland's paradigm

A graph *encodes* quantitative and categorical information using symbols, geometry and color. Graphical perception is the *visual decoding* of the encoded information.

- The graph may be beautiful but a failure: the visual decoding has to work.
- To make graphs work we need to understand graphical perception: what the eye sees and can decode well. visual perception.

Cleveland's paradigm

A graph *encodes* quantitative and categorical information using symbols, geometry and color. Graphical perception is the *visual decoding* of the encoded information.

- The graph may be beautiful but a failure: the visual decoding has to work.
- To make graphs work we need to understand graphical perception: what the eye sees and can decode well. visual perception.

Cleveland's paradigm

A specification and ordering of *elementary graphical-perception tasks*.

Ten properties of graphs:

- Angle
- Area
- Colour hue
- Colour saturation
- Density (amount of black)
- Length (distance)
- Position along a common scale
- Position along a divergent scale
- Position along a non-linear scale
- Volume

Cleveland's paradigm

A specification and ordering of *elementary graphical-perception tasks*.

Ten properties of graphs:

- Angle
- Area
- Colour hue
- Colour saturation
- Density (amount of black)
- Length (distance)
- Position along a common scale
- Position along identical, non-aligned scales
- Slope
- Volume

Cleveland's paradigm

A specification and ordering of *elementary graphical-perception tasks*.

Ten properties of graphs:

- Angle
- Area
- Colour hue
- Colour saturation
- Density (amount of black)
- Length (distance)
- Position along a common scale
- Position along identical, non-aligned scales
- Slope
- Volume

Cleveland's paradigm

A specification and ordering of *elementary graphical-perception tasks*.

Ten properties of graphs:

- Angle
- Area
- Colour hue
- Colour saturation
- Density (amount of black)
- Length (distance)
- Position along a common scale
- Position along identical, non-aligned scales
- Slope
- Volume

Cleveland's paradigm

A specification and ordering of *elementary graphical-perception tasks*.

Ten properties of graphs:

- Angle
- Area
- Colour hue
- Colour saturation
- Density (amount of black)
- Length (distance)
- Position along a common scale
- Position along identical, non-aligned scales
- Slope
- Volume

Cleveland's paradigm

A specification and ordering of *elementary graphical-perception tasks*.

Ten properties of graphs:

- Angle
- Area
- Colour hue
- Colour saturation
- Density (amount of black)
- Length (distance)
- Position along a common scale
- Position along identical, non-aligned scales
- Slope
- Volume

Cleveland's paradigm

A specification and ordering of *elementary graphical-perception tasks*.

Ten properties of graphs:

- Angle
- Area
- Colour hue
- Colour saturation
- Density (amount of black)
- Length (distance)
- Position along a common scale
- Position along identical, non-aligned scales
- Slope
- Volume

Cleveland's paradigm

A specification and ordering of *elementary graphical-perception tasks*.

Ten properties of graphs:

- Angle
- Area
- Colour hue
- Colour saturation
- Density (amount of black)
- Length (distance)
- Position along a common scale
- Position along identical, non-aligned scales
- Slope
- Volume

Cleveland's paradigm

A specification and ordering of *elementary graphical-perception tasks*.

Ten properties of graphs:

- Angle
- Area
- Colour hue
- Colour saturation
- Density (amount of black)
- Length (distance)
- Position along a common scale
- Position along identical, non-aligned scales
- Slope
- Volume

Cleveland's paradigm

A specification and ordering of *elementary graphical-perception tasks*.

Ten properties of graphs:

- Angle
- Area
- Colour hue
- Colour saturation
- Density (amount of black)
- Length (distance)
- Position along a common scale
- Position along identical, non-aligned scales
- Slope
- Volume

Cleveland's paradigm

A specification and ordering of *elementary graphical-perception tasks*.

Ten properties of graphs:

- Angle
- Area
- Colour hue
- Colour saturation
- Density (amount of black)
- Length (distance)
- Position along a common scale
- Position along identical, non-aligned scales
- Slope
- Volume

Cleveland's paradigm

Cleveland's order of accuracy:

- 1 Position along a common scale
- 2 Position along identical, non-aligned scales
- 3 Length
- 4 Angle and slope
- 5 Area
- 6 Volume
- 7 Colour hue, colour saturation, density

Encode data on a graph so that the visual decoding involves tasks as high as possible in the ordering.

Cleveland's paradigm

Cleveland's order of accuracy:

- 1 Position along a common scale
- 2 Position along identical, non-aligned scales
- 3 Length
- 4 Angle and slope
- 5 Area
- 6 Volume
- 7 Colour hue, colour saturation, density

Encode data on a graph so that the visual decoding involves tasks as high as possible in the ordering.

Cleveland's paradigm

Cleveland's order of accuracy:

- 1 Position along a common scale
- 2 Position along identical, non-aligned scales
- 3 **Length**
- 4 Angle and slope
- 5 Area
- 6 Volume
- 7 Colour hue, colour saturation, density

Encode data on a graph so that the visual decoding involves tasks as high as possible in the ordering.

Cleveland's paradigm

Cleveland's order of accuracy:

- 1 Position along a common scale
- 2 Position along identical, non-aligned scales
- 3 Length
- 4 Angle and slope
- 5 Area
- 6 Volume
- 7 Colour hue, colour saturation, density

Encode data on a graph so that the visual decoding involves tasks as high as possible in the ordering.

Cleveland's paradigm

Cleveland's order of accuracy:

- 1 Position along a common scale
- 2 Position along identical, non-aligned scales
- 3 Length
- 4 Angle and slope
- 5 Area
- 6 Volume
- 7 Colour hue, colour saturation, density

Encode data on a graph so that the visual decoding involves tasks as high as possible in the ordering.

Cleveland's paradigm

Cleveland's order of accuracy:

- 1 Position along a common scale
- 2 Position along identical, non-aligned scales
- 3 Length
- 4 Angle and slope
- 5 Area
- 6 **Volume**
- 7 Colour hue, colour saturation, density

Encode data on a graph so that the visual decoding involves tasks as high as possible in the ordering.

Cleveland's paradigm

Cleveland's order of accuracy:

- 1 Position along a common scale
- 2 Position along identical, non-aligned scales
- 3 Length
- 4 Angle and slope
- 5 Area
- 6 Volume
- 7 Colour hue, colour saturation, density

Encode data on a graph so that the visual decoding involves tasks as high as possible in the ordering.

Cleveland's paradigm

Cleveland's order of accuracy:

- 1 Position along a common scale
- 2 Position along identical, non-aligned scales
- 3 Length
- 4 Angle and slope
- 5 Area
- 6 Volume
- 7 Colour hue, colour saturation, density

Encode data on a graph so that the visual decoding involves tasks as high as possible in the ordering.

Cleveland's paradigm

Cleveland's order of accuracy:

- 1 Position along a common scale
- 2 Position along identical, non-aligned scales
- 3 Length
- 4 Angle and slope
- 5 Area
- 6 Volume
- 7 Colour hue, colour saturation, density

Encode data on a graph so that the visual decoding involves tasks as high as possible in the ordering.

Cleveland's paradigm

Corollaries

- 1 Pie charts perform poorly because they rely on comparing angles rather than distances.
- 2 Time series should be plotted as lines with time on the horizontal axis.
- 3 Avoid representing data using volumes.
- 4 If a key point is represented by a changing slope, consider plotting the rate of change itself rather than the original data.
- 5 Think of simplifications which enhance the detection of the basic properties of the data.
- 6 Think of how the distance between related representations of data affects their interpretation.

Cleveland's paradigm

Corollaries

- 1 Pie charts perform poorly because they rely on comparing angles rather than distances.
- 2 Time series should be plotted as lines with time on the horizontal axis.
- 3 Avoid representing data using volumes.
- 4 If a key point is represented by a changing slope, consider plotting the rate of change itself rather than the original data.
- 5 Think of simplifications which enhance the detection of the basic properties of the data.
- 6 Think of how the distance between related representations of data affects their interpretation.

Cleveland's paradigm

Corollaries

- 1 Pie charts perform poorly because they rely on comparing angles rather than distances.
- 2 Time series should be plotted as lines with time on the horizontal axis.
- 3 **Avoid representing data using volumes.**
- 4 If a key point is represented by a changing slope, consider plotting the rate of change itself rather than the original data.
- 5 Think of simplifications which enhance the detection of the basic properties of the data.
- 6 Think of how the distance between related representations of data affects their interpretation.

Cleveland's paradigm

Corollaries

- 1 Pie charts perform poorly because they rely on comparing angles rather than distances.
- 2 Time series should be plotted as lines with time on the horizontal axis.
- 3 Avoid representing data using volumes.
- 4 If a key point is represented by a changing slope, consider plotting the rate of change itself rather than the original data.
- 5 Think of simplifications which enhance the detection of the basic properties of the data.
- 6 Think of how the distance between related representations of data affects their interpretation.

Cleveland's paradigm

Corollaries

- 1 Pie charts perform poorly because they rely on comparing angles rather than distances.
- 2 Time series should be plotted as lines with time on the horizontal axis.
- 3 Avoid representing data using volumes.
- 4 If a key point is represented by a changing slope, consider plotting the rate of change itself rather than the original data.
- 5 Think of simplifications which enhance the detection of the basic properties of the data.
- 6 Think of how the distance between related representations of data affects their interpretation.

Cleveland's paradigm

Corollaries

- 1 Pie charts perform poorly because they rely on comparing angles rather than distances.
- 2 Time series should be plotted as lines with time on the horizontal axis.
- 3 Avoid representing data using volumes.
- 4 If a key point is represented by a changing slope, consider plotting the rate of change itself rather than the original data.
- 5 Think of simplifications which enhance the detection of the basic properties of the data.
- 6 Think of how the distance between related representations of data affects their interpretation.

Case study 1

Contribution to product sales by region

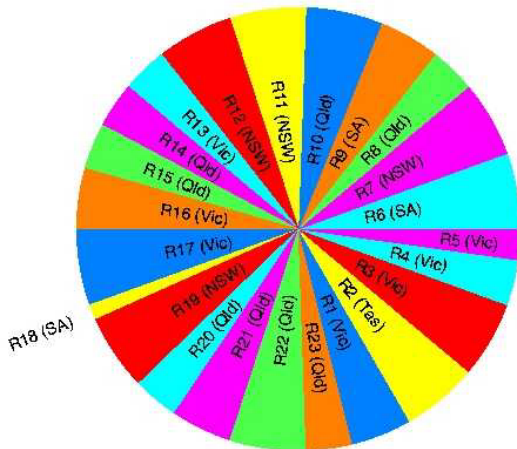


Figure 2. Pie chart, showing the relative contributions to the profits of an enterprise from various Divisions around Australia.

Case study 1

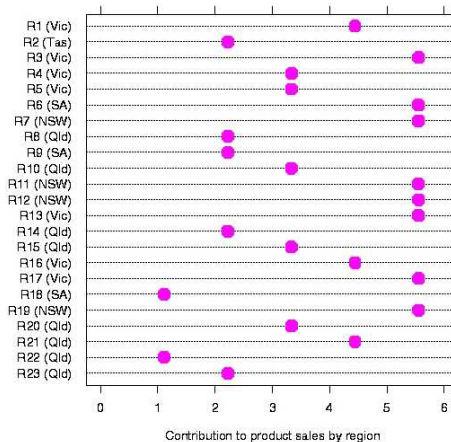


Figure 4. A dotplot of the data used in Figure 2, showing the relative contributions to enterprise profits from its various Divisions around Australia. The discrete nature of the data is immediately evident.

Case study 1

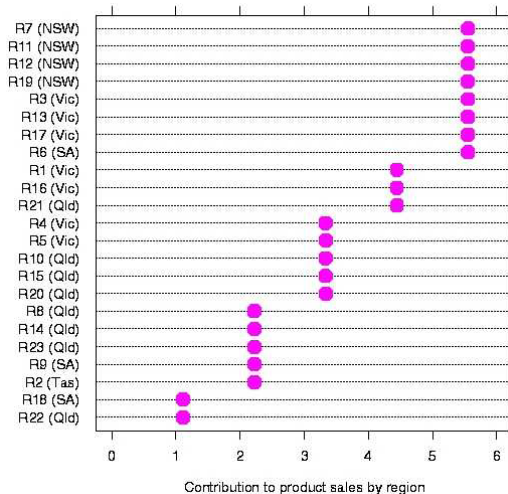


Figure 5. The display from Figure 4 has been modified, so that the data plot from largest to smallest. A further pattern emerges: different States tend to contribute differently to enterprise profits.

Case study 1

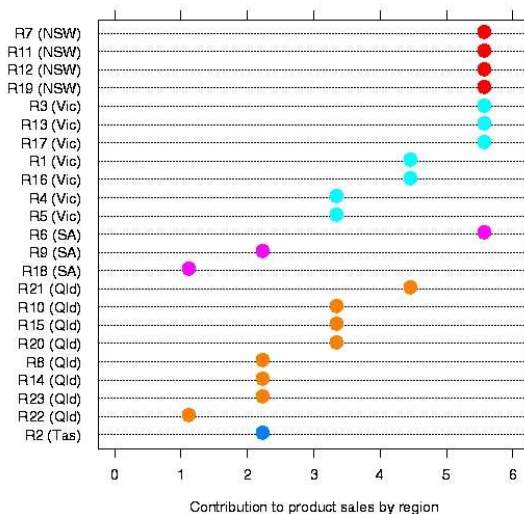


Figure 6. The contributions to group profit by different regions are plotted by State. The clear differences between States are evident.

Case study 2

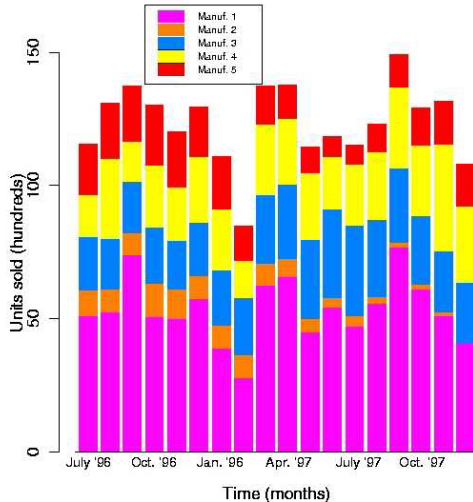


Figure 3. Divided bar chart, showing monthly sales of different brands of whitegoods over an 18-month period.

Case study 2

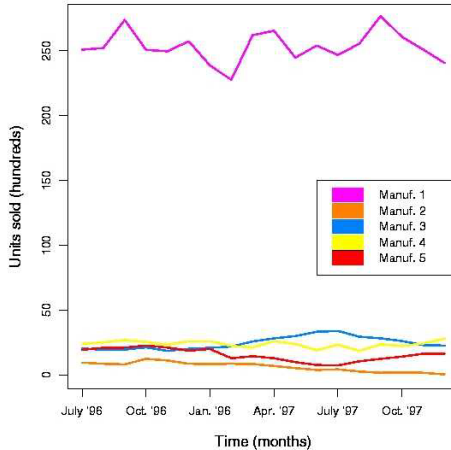


Figure 7. The monthly sales data from Figure 3 have been replotted so that sales patterns for each manufacturer can be seen without distortion. However, the curve for the dominant manufacturer is compressing patterns in the other curves.

Case study 2

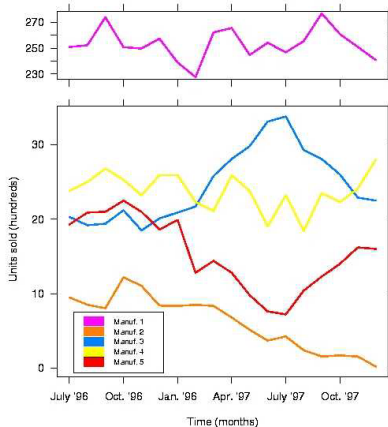


Figure 8. The monthly sales data from Figure 7 have been replotted so that the dominant curve is displayed separately, with a false origin, and the other curves that are measured on a much smaller scale can then be plotted using a better aspect ratio. This reveals more information about individual and comparative trends in the curves.

Case study 2

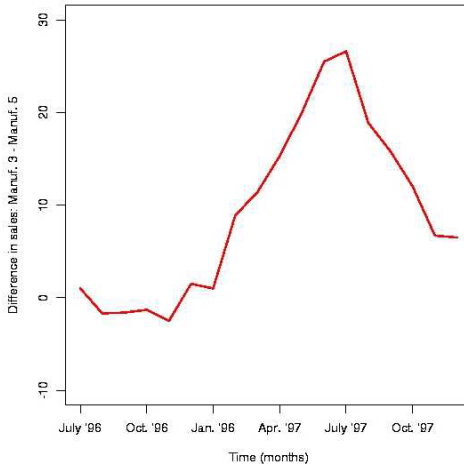


Figure 9. This graph shows the difference between sales of Manufacturers 5 and 3. Over the period January – July 1997 there was a marked increase in sales in favour of Manufacture 5; after July, this advantage declined steadily to the end of the year.

Twenty rules for good graphics

- 1 Use vector graphics such as eps or pdf.
- 2 Use readable fonts.
- 3 Avoid cluttered legends.
- 4 If you must use a legend, move it inside the plot.
- 5 No dark shaded backgrounds.
- 6 Avoid dark, dominating grid lines.
- 7 Keep the axis limits sensible.
- 8 Do not forget to specify units.
- 9 Tick intervals should be at nice round numbers.
- 10 Axes should be properly labelled.

Twenty rules for good graphics

- 1 Use vector graphics such as eps or pdf.
- 2 Use readable fonts.
- 3 Avoid cluttered legends.
- 4 If you must use a legend, move it inside the plot.
- 5 No dark shaded backgrounds.
- 6 Avoid dark, dominating grid lines.
- 7 Keep the axis limits sensible.
- 8 Do not forget to specify units.
- 9 Tick intervals should be at nice round numbers.
- 10 Axes should be properly labelled.

Twenty rules for good graphics

- 1 Use vector graphics such as eps or pdf.
- 2 Use readable fonts.
- 3 **Avoid cluttered legends.**
- 4 If you must use a legend, move it inside the plot.
- 5 No dark shaded backgrounds.
- 6 Avoid dark, dominating grid lines.
- 7 Keep the axis limits sensible.
- 8 Do not forget to specify units.
- 9 Tick intervals should be at nice round numbers.
- 10 Axes should be properly labelled.

Twenty rules for good graphics

- 1 Use vector graphics such as eps or pdf.
- 2 Use readable fonts.
- 3 Avoid cluttered legends.
- 4 If you must use a legend, move it inside the plot.
- 5 No dark shaded backgrounds.
- 6 Avoid dark, dominating grid lines.
- 7 Keep the axis limits sensible.
- 8 Do not forget to specify units.
- 9 Tick intervals should be at nice round numbers.
- 10 Axes should be properly labelled.

Twenty rules for good graphics

- 1 Use vector graphics such as eps or pdf.
- 2 Use readable fonts.
- 3 Avoid cluttered legends.
- 4 If you must use a legend, move it inside the plot.
- 5 **No dark shaded backgrounds.**
- 6 Avoid dark, dominating grid lines.
- 7 Keep the axis limits sensible.
- 8 Do not forget to specify units.
- 9 Tick intervals should be at nice round numbers.
- 10 Axes should be properly labelled.

Twenty rules for good graphics

- 1 Use vector graphics such as eps or pdf.
- 2 Use readable fonts.
- 3 Avoid cluttered legends.
- 4 If you must use a legend, move it inside the plot.
- 5 No dark shaded backgrounds.
- 6 Avoid dark, dominating grid lines.
- 7 Keep the axis limits sensible.
- 8 Do not forget to specify units.
- 9 Tick intervals should be at nice round numbers.
- 10 Axes should be properly labelled.

Twenty rules for good graphics

- 1 Use vector graphics such as eps or pdf.
- 2 Use readable fonts.
- 3 Avoid cluttered legends.
- 4 If you must use a legend, move it inside the plot.
- 5 No dark shaded backgrounds.
- 6 Avoid dark, dominating grid lines.
- 7 **Keep the axis limits sensible.**
- 8 Do not forget to specify units.
- 9 Tick intervals should be at nice round numbers.
- 10 Axes should be properly labelled.

Twenty rules for good graphics

- 1 Use vector graphics such as eps or pdf.
- 2 Use readable fonts.
- 3 Avoid cluttered legends.
- 4 If you must use a legend, move it inside the plot.
- 5 No dark shaded backgrounds.
- 6 Avoid dark, dominating grid lines.
- 7 Keep the axis limits sensible.
- 8 Do not forget to specify units.
- 9 Tick intervals should be at nice round numbers.
- 10 Axes should be properly labelled.

Twenty rules for good graphics

- 1 Use vector graphics such as eps or pdf.
- 2 Use readable fonts.
- 3 Avoid cluttered legends.
- 4 If you must use a legend, move it inside the plot.
- 5 No dark shaded backgrounds.
- 6 Avoid dark, dominating grid lines.
- 7 Keep the axis limits sensible.
- 8 Do not forget to specify units.
- 9 Tick intervals should be at nice round numbers.
- 10 Axes should be properly labelled.

Twenty rules for good graphics

- 1 Use vector graphics such as eps or pdf.
- 2 Use readable fonts.
- 3 Avoid cluttered legends.
- 4 If you must use a legend, move it inside the plot.
- 5 No dark shaded backgrounds.
- 6 Avoid dark, dominating grid lines.
- 7 Keep the axis limits sensible.
- 8 Do not forget to specify units.
- 9 Tick intervals should be at nice round numbers.
- 10 Axes should be properly labelled.

Twenty rules

- 11 Use linewidths big enough to read.
- 12 Avoid overlapping text on plotting characters or lines.
- 13 Follow Tufte's principles by removing chart junk and keeping a high data-ink ratio.
- 14 Plots should be self-explanatory.
- 15 Use a sensible aspect ratio.
- 16 Use points not lines if element order is not relevant.
- 17 Use lines if element order is relevant (e.g., time).
- 18 Use common scales for comparison (or a single plot).
- 19 Avoid pie-charts. Especially 3d pie-charts. Especially 3d pie-charts with exploding wedges.
- 20 Never use fill patterns such as cross-hatching.

Twenty rules

- 11 Use linewidths big enough to read.
- 12 Avoid overlapping text on plotting characters or lines.
- 13 Follow Tufte's principles by removing chart junk and keeping a high data-ink ratio.
- 14 Plots should be self-explanatory.
- 15 Use a sensible aspect ratio.
- 16 Use points not lines if element order is not relevant.
- 17 Use lines if element order is relevant (e.g., time).
- 18 Use common scales for comparison (or a single plot).
- 19 Avoid pie-charts. Especially 3d pie-charts. Especially 3d pie-charts with exploding wedges.
- 20 Never use fill patterns such as cross-hatching.

Twenty rules

- 11 Use linewidths big enough to read.
- 12 Avoid overlapping text on plotting characters or lines.
- 13 Follow Tufte's principles by removing chart junk and keeping a high data-ink ratio.
- 14 Plots should be self-explanatory.
- 15 Use a sensible aspect ratio.
- 16 Use points not lines if element order is not relevant.
- 17 Use lines if element order is relevant (e.g., time).
- 18 Use common scales for comparison (or a single plot).
- 19 Avoid pie-charts. Especially 3d pie-charts. Especially 3d pie-charts with exploding wedges.
- 20 Never use fill patterns such as cross-hatching.

Twenty rules

- 11 Use linewidths big enough to read.
- 12 Avoid overlapping text on plotting characters or lines.
- 13 Follow Tufte's principles by removing chart junk and keeping a high data-ink ratio.
- 14 **Plots should be self-explanatory.**
- 15 Use a sensible aspect ratio.
- 16 Use points not lines if element order is not relevant.
- 17 Use lines if element order is relevant (e.g., time).
- 18 Use common scales for comparison (or a single plot).
- 19 Avoid pie-charts. Especially 3d pie-charts. Especially 3d pie-charts with exploding wedges.
- 20 Never use fill patterns such as cross-hatching.

Twenty rules

- 11 Use linewidths big enough to read.
- 12 Avoid overlapping text on plotting characters or lines.
- 13 Follow Tufte's principles by removing chart junk and keeping a high data-ink ratio.
- 14 Plots should be self-explanatory.
- 15 Use a sensible aspect ratio.
- 16 Use points not lines if element order is not relevant.
- 17 Use lines if element order is relevant (e.g., time).
- 18 Use common scales for comparison (or a single plot).
- 19 Avoid pie-charts. Especially 3d pie-charts.
Especially 3d pie-charts with exploding wedges.
- 20 Never use fill patterns such as cross-hatching.

Twenty rules

- 11 Use linewidths big enough to read.
- 12 Avoid overlapping text on plotting characters or lines.
- 13 Follow Tufte's principles by removing chart junk and keeping a high data-ink ratio.
- 14 Plots should be self-explanatory.
- 15 Use a sensible aspect ratio.
- 16 Use points not lines if element order is not relevant.
- 17 Use lines if element order is relevant (e.g., time).
- 18 Use common scales for comparison (or a single plot).
- 19 Avoid pie-charts. Especially 3d pie-charts.
Especially 3d pie-charts with exploding wedges.
- 20 Never use fill patterns such as cross-hatching.

Twenty rules

- 11 Use linewidths big enough to read.
- 12 Avoid overlapping text on plotting characters or lines.
- 13 Follow Tufte's principles by removing chart junk and keeping a high data-ink ratio.
- 14 Plots should be self-explanatory.
- 15 Use a sensible aspect ratio.
- 16 Use points not lines if element order is not relevant.
- 17 Use lines if element order is relevant (e.g., time).
- 18 Use common scales for comparison (or a single plot).
- 19 Avoid pie-charts. Especially 3d pie-charts. Especially 3d pie-charts with exploding wedges.
- 20 Never use fill patterns such as cross-hatching.

Twenty rules

- 11 Use linewidths big enough to read.
- 12 Avoid overlapping text on plotting characters or lines.
- 13 Follow Tufte's principles by removing chart junk and keeping a high data-ink ratio.
- 14 Plots should be self-explanatory.
- 15 Use a sensible aspect ratio.
- 16 Use points not lines if element order is not relevant.
- 17 Use lines if element order is relevant (e.g., time).
- 18 Use common scales for comparison (or a single plot).
- 19 Avoid pie-charts. Especially 3d pie-charts.
Especially 3d pie-charts with exploding wedges.
- 20 Never use fill patterns such as cross-hatching.

Twenty rules

- 11 Use linewidths big enough to read.
- 12 Avoid overlapping text on plotting characters or lines.
- 13 Follow Tufte's principles by removing chart junk and keeping a high data-ink ratio.
- 14 Plots should be self-explanatory.
- 15 Use a sensible aspect ratio.
- 16 Use points not lines if element order is not relevant.
- 17 Use lines if element order is relevant (e.g., time).
- 18 Use common scales for comparison (or a single plot).
- 19 Avoid pie-charts. Especially 3d pie-charts.
Especially 3d pie-charts with exploding wedges.
- 20 Never use fill patterns such as cross-hatching.

Twenty rules

- 11 Use linewidths big enough to read.
- 12 Avoid overlapping text on plotting characters or lines.
- 13 Follow Tufte's principles by removing chart junk and keeping a high data-ink ratio.
- 14 Plots should be self-explanatory.
- 15 Use a sensible aspect ratio.
- 16 Use points not lines if element order is not relevant.
- 17 Use lines if element order is relevant (e.g., time).
- 18 Use common scales for comparison (or a single plot).
- 19 Avoid pie-charts. Especially 3d pie-charts. Especially 3d pie-charts with exploding wedges.
- 20 Never use fill patterns such as cross-hatching.

References

- Cleveland, W.S. (1985) *The elements of graphing data*, Wadsworth.
- Cleveland, W.S. (1993) *Visualizing data*, Hobart Press
- Tufte, E.R. (1983) *The visual display of quantitative information*, Graphics Press.
- Tufte, E.R. (1990) *Envisioning information*, Graphics Press.

References

- Cleveland, W.S. (1985) *The elements of graphing data*, Wadsworth.
- Cleveland, W.S. (1993) *Visualizing data*, Hobart Press
- Tufte, E.R. (1983) *The visual display of quantitative information*, Graphics Press.
- Tufte, E.R. (1990) *Envisioning information*, Graphics Press.

References

- Cleveland, W.S. (1985) *The elements of graphing data*, Wadsworth.
- Cleveland, W.S. (1993) *Visualizing data*, Hobart Press
- Tufte, E.R. (1983) *The visual display of quantitative information*, Graphics Press.
- Tufte, E.R. (1990) *Envisioning information*, Graphics Press.

References

- Cleveland, W.S. (1985) *The elements of graphing data*, Wadsworth.
- Cleveland, W.S. (1993) *Visualizing data*, Hobart Press
- Tufte, E.R. (1983) *The visual display of quantitative information*, Graphics Press.
- Tufte, E.R. (1990) *Envisioning information*, Graphics Press.