

Data Visualisation Design — Workbook

This workbook turns the course into a working method you apply to one real dataset. Choose a dataset now, something with at least one category column and one number column, and carry it through every exercise so you finish with a portfolio-ready data story rather than abstract notes. Each section maps to one course module, moving from perception through chart choice, colour and clutter, to a finished dashboard and narrative.

How People Read a Chart

Ground every later decision in perception: rank your encodings, decide what should pop, and choose the dataset you will carry through the course.

Exercise: Choose Your One Dataset

Pick a single real dataset you genuinely care about and will use for the entire course. It needs at least one category column and one numeric column; a time column is a bonus. Public sources like Our World in Data, your own business numbers, or a sports table all work. Open it and answer the prompts before doing anything visual.

- What is the dataset, where did it come from, and what does each row represent?

- Which columns are categories, which are numbers, and is there a time or ordered column?

- What is the single most important comparison a reader should be able to make from it?

- What is one question about this data you personally want answered?

Worksheet: Encoding Accuracy Audit

Take any chart you find in the wild this week, a news graphic, a report, a dashboard, and audit it against the Cleveland-McGill hierarchy. The goal is to start seeing charts as encodings ranked by accuracy, not as pictures.

Chart source and what it is trying to show

What variable is encoded with the highest-accuracy channel (position or length)?

What variable is stuck on a weak channel (area, angle, or colour) that needs precision?

Is the most important comparison on a high-accuracy channel, yes or no?

One change that would move the key comparison onto a more accurate encoding

Exercise: Find the Pop

For your own dataset, write the single sentence the chart must say, then decide which preattentive attribute will carry it. The point is to prove you can name what should be loud before you style anything.

- In one sentence, what is the message your chart must communicate?

- Which single data point, bar, or series proves that sentence?

- Which preattentive attribute will make it pop (a single accent colour, size, or position), and how will everything else stay muted?

- If you cannot name what should pop, what is the missing analysis you need to do first?

Choosing the Right Chart

Move from perception to practice: pick the chart family from the question, build the right workhorse chart, and stress-test it for honesty.

Worksheet: Question-First Chart Selector

Before choosing a chart, name the question. Fill this in for your dataset so the chart is the answer to a stated question rather than a gallery pick.

The question as a full sentence

Chart family it implies (comparison, trend, distribution, relationship, composition, geospatial)

Number of variables and number of categories to show

Chart type chosen within the family (e.g. sorted horizontal bars, single line, small multiples)

Confirm the key comparison sits on position or length (yes/no)

Ten-second test: can a stranger answer the question from this chart? (yes/no plus note)

Exercise: Build Three, Keep One

Make the same data three different ways, for example a pie, a sorted bar chart, and a dot plot, then judge which lets a reader answer your question fastest and most accurately. This makes the accuracy hierarchy real in your own hands.

- Which three chart types did you build for the same data?

- Which one answered your question fastest and most accurately, and why?

- Did sorting the bars by value change how quickly you could read the ranking?

- Which version are you keeping, and what specifically made the other two weaker?

Worksheet: Misleading-Chart Stress Test

Run your chosen chart through the common ways charts distort data and record the verdict for each. Treat this as the honesty contract: the visual proportions must match the data proportions.

Bar value axis starts at zero (yes / n/a)

No dual y-axis inviting a false correlation (confirmed / fixed how)

Time range is the full fair span, not cherry-picked (confirmed)

Areas and bubbles sized by area, not radius; no 3D (confirmed)

Using rates or per-capita where raw counts would mislead (yes / n/a)

Checked whether splitting by a group reverses the story, Simpson's paradox (result)

Checklist: Honest Chart Checklist

- Bars baselined at zero; lines may use a non-zero axis to reveal variation
- Bars sorted by value for ranking, or kept in natural order for time and ordered scales
- A line connects points only where the space between them is meaningful
- Single clear axis rather than a dual-axis chart
- Correlation is not presented as causation; trend lines only drawn when a relationship is real
- Denominator and time range stated plainly so the comparison is fair

Colour, Type, and Clutter

Make the chart clear and accessible: match the palette type to the data, prove it survives colour blindness and greyscale, and strip everything that is not data.

Worksheet: Palette Decision Sheet

Choose colour deliberately by first classifying the data, then picking the matching palette type from a tested source. Fill this in before colouring anything.

Data type: ordered, ordered with a midpoint, or unordered categories

Palette type chosen: sequential, diverging, or categorical

Source palette used (ColorBrewer scheme, Viridis, Okabe-Ito, or brand palette)

Number of categorical colours used (cap around seven)

Which two or three categories are highlighted, and which are muted to grey

How chart colour ties to the brand palette without harming meaning or access

Exercise: Greyscale and Colour-Blind Test

Take your chart and prove it works for everyone. Simulate it under colour-vision deficiencies and convert it to greyscale, then fix any failure with a redundant cue. This is the single most important accessibility step.

- When you simulated red-green colour blindness (or used Viz Palette), did any categories become indistinguishable?
- When you converted the chart to greyscale, could you still read every series?

-
- What redundant cue did you add so colour is never the only signal (direct labels, line style, markers, patterns)?
 - Did any meaningful element rely on red versus green, and what colour-blind-safe pair did you switch to?
-

Exercise: Strip It Down

Apply data-ink discipline to your chart: remove non-data ink until the data dominates, then replace the legend with direct labels and rewrite the title as a message. Compare a before and after screenshot.

- What decoration did you remove (border, gridlines, shadows, gradients, extra decimals)?
 - Did replacing the legend with direct labels remove the back-and-forth eye movement?
 - What was your old title, and what is the new message-style title that states the finding?
 - Looking at before and after, what specifically makes the after easier to read?
-

Checklist: Accessibility and Clarity Checklist

- Palette type matches the data type (sequential, diverging, or categorical)
- No meaning carried by red versus green alone; a redundant cue is present
- Chart still readable when converted to greyscale
- Text and key marks meet WCAG contrast (4.5 to 1 body, 3 to 1 large and graphical)
- Legend replaced with direct labels wherever they fit
- Title states the takeaway; one annotation highlights the key insight

Dashboards and Data Stories

Assemble finished work: tidy the data, build in a real tool, lay out a dashboard by priority, and frame a single honest data story.

Worksheet: Tidy-Data and Tool Plan

Cleaning beats styling: get the data into tidy shape and pick the right tool before you build. Fill this in so the chart step takes minutes, not hours.

Is the data tidy: one variable per column, one observation per row, one value per cell? (yes / what to fix)

Cleanup needed (remove totals rows, unpivot wide columns, fix text dates, standardise units)

Tool chosen and why (Datawrapper for fast trustworthy charts, Flourish for story, Tableau Public for interactive dashboard)

Chart type to build, from your question-first selector

Source line text and any method or denominator note to display

Exercise: Publish One Chart End to End

Take your tidy data into Datawrapper or Flourish and produce a finished, publishable chart: correct chart type, zero-baselined bars, colour-blind-safe palette, direct labels, message title, and a source line. Export both an embed and a static image.

- Which tool did you use, and what chart type did you publish?
- Which course rules did you apply in the refine step (sort, zero baseline, palette, direct labels)?

-
- What is your final message-style title and subtitle with units and date range?
 - Where will this chart live, and did you export both a responsive embed and a static version?
-

Worksheet: Dashboard Layout Plan

Design the dashboard as one answer, not a wall of charts. Decide the audience and priority before placing anything, following overview first, then details on demand.

Audience and the three to five questions the dashboard must answer

The single most important metric or chart that goes top-left

Supporting charts and how they are grouped (which belong together)

Consistent rules across panels: one colour meaning, one number format, one date format

Which charts share axis scales because they are meant to be compared

How detail is provided (filters, tooltips, drill-down) instead of cramming the screen

Exercise: Frame the Data Story

Turn your best chart into an explanatory data story built around one message, framed honestly, and guided by annotation. This is the capstone that proves the full method on your dataset.

- What is the single takeaway, written as one sentence and used as the title?
 - What did you mute or remove so only the supporting data is loud?
 - What annotation did you add to name the insight (a note on the key moment, a reference line, a shaded period)?
 - How did you keep the framing honest (full fair range, stated source and denominator, no distorted encoding)?
-

Your Action Plan

1. Choose one real dataset and write down what each row means and the key comparison it should support
2. Audit a chart in the wild against the Cleveland-McGill hierarchy and name what should pop in your own chart
3. Write your question as a sentence and let it pick the chart family and type
4. Build the same data three ways and keep the version that answers the question fastest and most accurately
5. Run your chart through the misleading-chart stress test, including a Simpson's paradox split check
6. Classify your data and choose a matching, tested palette; cap categorical colours and mute the long tail
7. Prove accessibility with a colour-blind simulation and a greyscale test, adding a redundant cue where needed

8. Strip non-data ink, replace the legend with direct labels, and rewrite the title as a message
9. Tidy the data, build a finished chart in Datawrapper or Flourish, and export an embed plus a static image
10. Lay out a dashboard by priority and frame one honest, annotated data story as your portfolio capstone

