Wrapping up and assessment preparation.

SMI105: Week 11

Dr. Calum Webb
Sheffield Methods Institute, the University of Sheffield
c.j.webb@sheffield.ac.uk

Sign in

Learning outcomes

What will I learn?

By the end of this week you will know:

- How the skills you have learned relate to the current skills shortage in the U.K. and globally, and how you can record your skills on MySkills — I'd recommend doing this now, as you might have forgotten by third year!
- What the second and final assessment for SMI105 requires you to do.
- **How** to go about critiquing a data visualisation and thinking of how you would improve it **using the skills we have developed on this module and the module materials**
- Some top tips for your assessment: how to give yourself the best chance of succeeding.

What skills do you have after finishing SMI105 that you didn't have before?

Analysis skills

- How to visualise data and interpret data visualisations
- How to choose the right kind of visualisation to analyse your data
- How to summarise and aggregate data
- How to create new variables from data in order to improve analysis

Communication skills

- How to apply best graphic design principles to clearly communicate messages using data visualisation

- How to make data visualisations accessible (e.g. contrast, colour choice, fonts)
- How to use data storytelling principles such as active titles to consider your audience and clearly communicate the story behind the data visualisation

Analysis skills

- How to visualise data and interpret data visualisations
- How to choose the right kind of visualisation to analyse your data
- How to summarise and aggregate data
- How to create new variables from data in order to improve analysis

Communication skills

- How to apply best graphic design principles to clearly communicate messages using data visualisation
- How to make data visualisations accessible (e.g. contrast, colour choice, fonts)
- How to use data storytelling principles such as active titles to consider your audience and clearly communicate the story behind the data visualisation

Data processing skills

- How to work with various data formats within **R**
- How to join and work with relational data
- How to scrape data from the web
- How to tidy data and prepare it for analysis and visualisation

Critical thinking skills

- How to critically assess the design of data visualisations
- How to critically assess the impact of data visualisation on various groups and audiences, and in wider society

IT skills

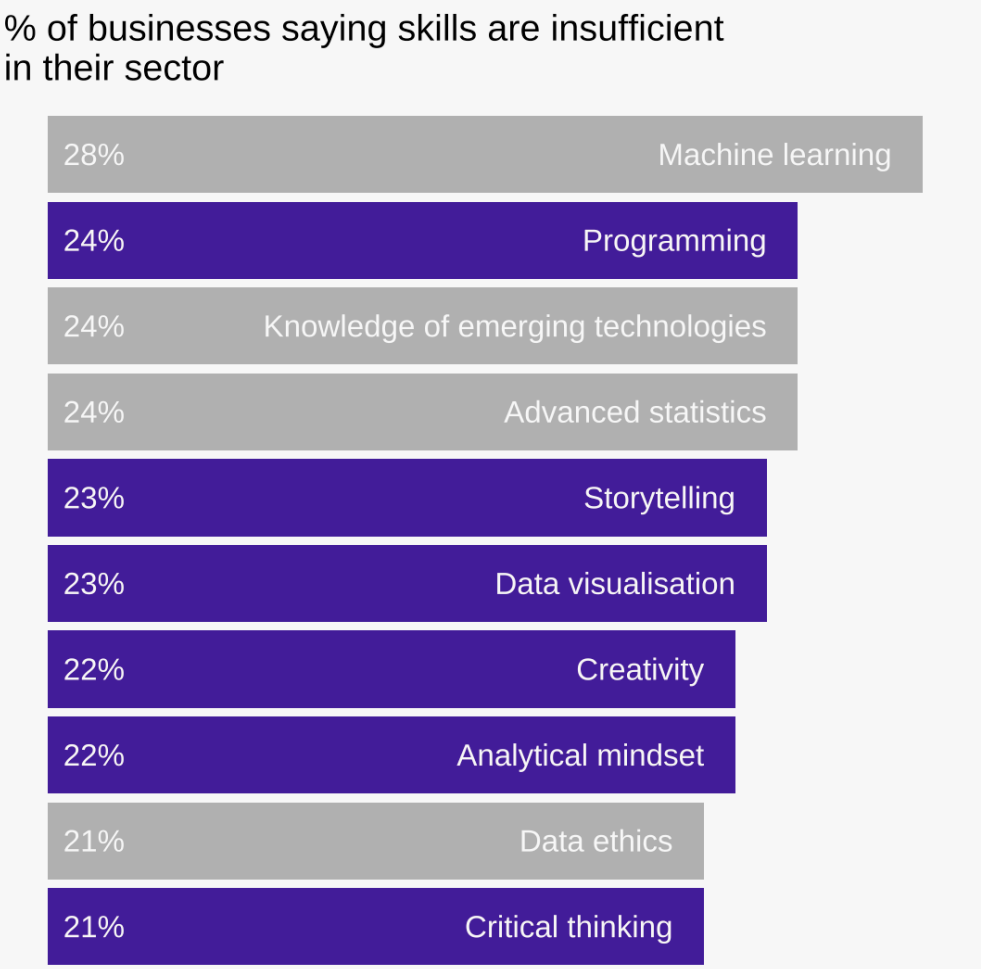
- How to manage a data visualisation project from start to finish using **R** projects
- Basic programming skills in the **R** statistical programming language
- Knowledge of optimal data formats for data visualisation (tidy data)
- How to export data visualisations to be suitable for publication (different technical export formats, e.g. svg, png)
- How to use generative AI within data tidying workflows

Will these skills help you get a job?

Your skills will be in demand.

The research found that there is significant demand for data skills with UK companies recruiting for 178,000 to 234,000 roles requiring hard data skills. Almost half of businesses (48%) are recruiting for roles that require hard data skills but under half (46%) have struggled to recruit for these roles over the last 2 years. **The supply of graduates with specialist data skills from universities is limited.**

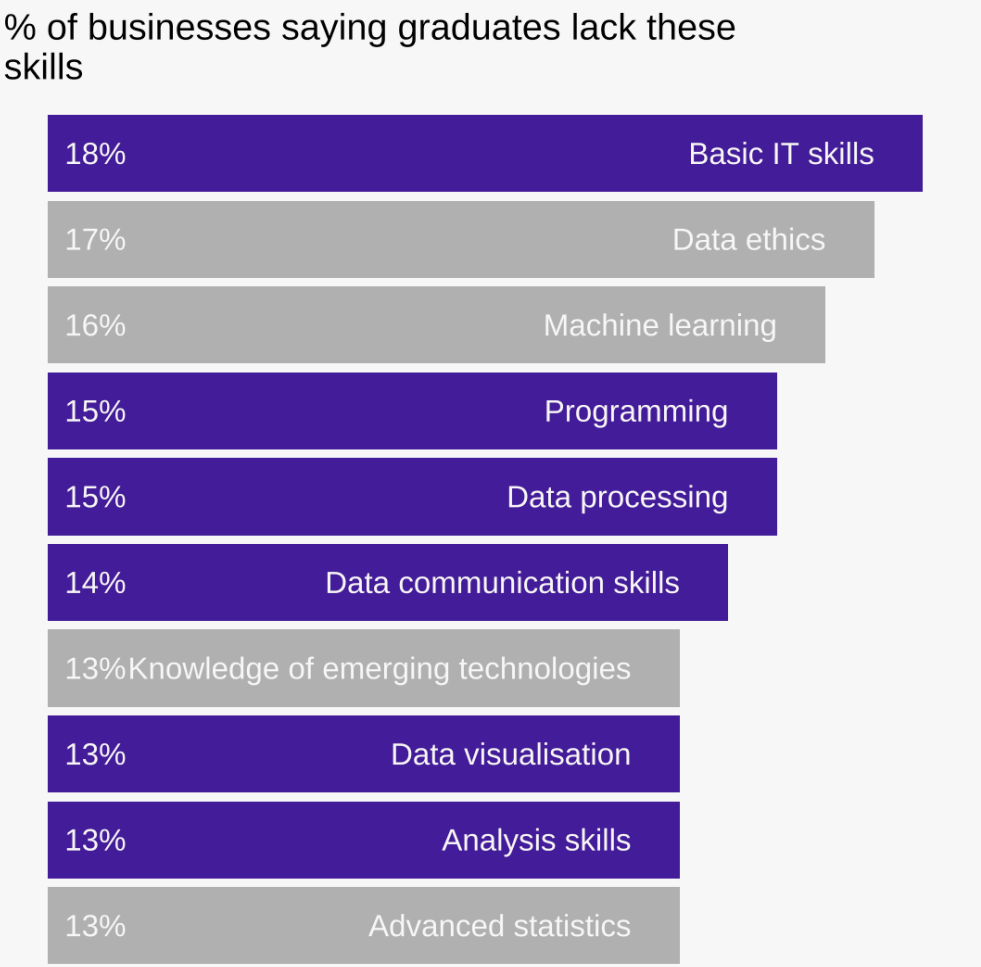
When recruiting for data roles, businesses were most often looking for a “data analyst” (12% of businesses surveyed were recruiting for this role), especially at junior or entry level.



Your skills will be in demand.

The research found that there is significant demand for data skills with UK companies recruiting for 178,000 to 234,000 roles requiring hard data skills. Almost half of businesses (48%) are recruiting for roles that require hard data skills but under half (46%) have struggled to recruit for these roles over the last 2 years. **The supply of graduates with specialist data skills from universities is limited.**

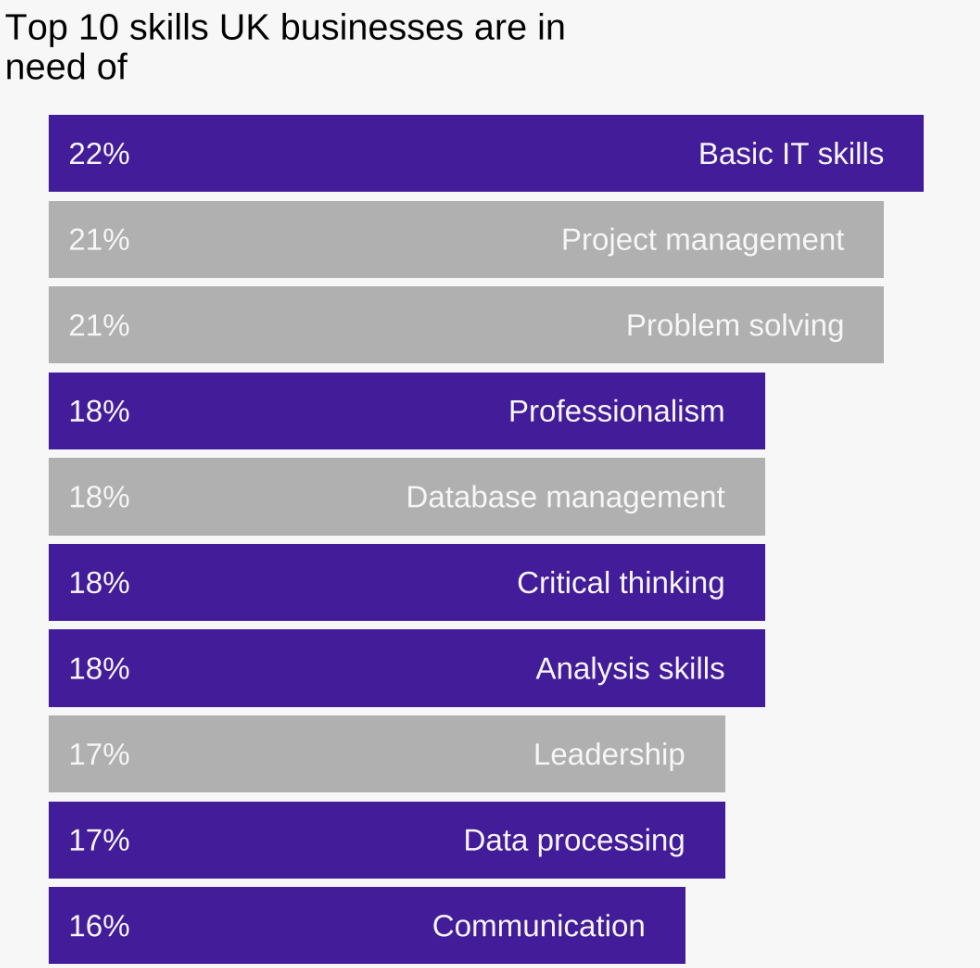
When recruiting for data roles, businesses were most often looking for a “data analyst” (12% of businesses surveyed were recruiting for this role), especially at junior or entry level.



Your skills will be in demand.

The research found that there is significant demand for data skills with UK companies recruiting for 178,000 to 234,000 roles requiring hard data skills. Almost half of businesses (48%) are recruiting for roles that require hard data skills but under half (46%) have struggled to recruit for these roles over the last 2 years. **The supply of graduates with specialist data skills from universities is limited.**

When recruiting for data roles, businesses were most often looking for a “data analyst” (12% of businesses surveyed were recruiting for this role), especially at junior or entry level.



“There [are] not many people qualified to work with big data”

Quantifying the UK Data Skills Gap (2021)

“Less people are training for this type of job and not enough qualified ones are available”

Quantifying the UK Data Skills Gap (2021)

“Our market is quite niche, and finding good data analysts who can put things into context has been difficult.”

Quantifying the UK Data Skills Gap (2021)

"There are simply too few people with the right skillset available."

Quantifying the UK Data Skills Gap (2021)

Gaps exist across all soft skills, but most notably for ‘communication’, ‘adaptability’ and 'critical thinking'

Quantifying the UK Data Skills Gap (2021)

Final Assessment

‘In the second and final assessment you’ll be asked to reflect on a published article that contains a data visualisation, to assess the data visualisation’s strengths and weaknesses, and consider how it might be improved. Then, remaining with the same topic, you’ll be asked to tell a relevant story by creating your own data visualisation(s), and then explain the decisions you made about which data and which type of visualisation you thought were most appropriate for the audience.’

1. Select a data visualisation on a specific topic

Pick one of the five options on the Blackboard page for Assessment 2.

1. Select a data visualisation on a specific topic

Pick one of the five options on the Blackboard page for Assessment 2.

2. Critique the presentation of data you have chosen

Write a critique of the presentation of data in the web page you choose (up to 500 words).

1. Select a data visualisation on a specific topic

Pick one of the five options on the Blackboard page for Assessment 2.

2. Critique the presentation of data you have chosen

Write a critique of the presentation of data in the web page you choose (up to 500 words).

3. Visualise data to answer a question on a related topic

Using either the same data as the page you've chosen, or different data, describe how the data in question can be used to answer a question on a related topic, including whichever graphs and maps you find appropriate. This problem should be related to the discussion in the page you've chosen, and your presentation must include at least one data visualisation you've created yourself using R (and include the code you used). Include an explanation of how your presentation of this data improves on that of the page you critiqued above (500-750 words).

Final assessment guidance: How to use what we have learned to critique and improve on a data visualisation

1. **Is the data visualisation clearly interpretable?** (Week 2). Is there a title? Does it tell you what the visualisation shows? Does the data support what the title claims? Are axes clearly labelled? Are axes readable? Does the visualisation have a title? Does the visualisation provide a source? Does the visualisation contain a key/legend or are all colours, shapes, shading, etc. used without a clear explanation of what they mean? Have scales been transformed or otherwise changed

in any way, and what impact does this have on the interpretation? Key

Final assessment guidance: How to use what we have learned to critique and improve on a data visualisation

1. **Is the data visualisation clearly interpretable?** (Week 2). Is there a title? Does it tell you what the visualisation shows? Does the data support what the title claims? Are axes clearly labelled? Are axes readable? Does the visualisation have a title? Does the visualisation provide a source? Does the visualisation contain a key/legend or are all colours, shapes, shading, etc. used without a clear explanation of what they mean? Have scales been transformed or otherwise changed in any way, and what impact does this have on the interpretation? Key

2. **Does the data visualisation effectively tell a story?** (Week 3). Is there a story underlying the data visualisation? Is that story clearly communicated in the design choices made around the data visualisation? Are colours, shapes, titles, used to effectively focus the audience on the data story? Is the visualisation and story clearly focused on a target audience? How could the data visualisation be changed to target a different audience? Does the data visualisation use a specific style of narrative plot? How would the use of a narrative plot improve the audience's engagement with the data visualisation? Is there a clear "character" in the data visualisation? Key reading: Feigenbaum & Alamalhodaei, 2020

Final assessment guidance: How to use what we have learned to critique and improve on a data visualisation

3. **Does the data visualisation use an appropriate type of chart?** (Week 4). Does the data visualisation use an appropriate type of chart for the underlying data (continuous, categorical, ordinal)? Would a different type of visualisation do a better job of illustrating certain things about the data, such as how it varies? Would a different type of visualisation be better suited for telling the story (e.g. Categorical, Hierarchical, Relational, Temporal, Spatial)? Key reading: Kirk, A. (2024)

Final assessment guidance: How to use what we have learned to critique and improve on a data visualisation

3. **Does the data visualisation use an appropriate type of chart?** (Week 4). Does the data visualisation use an appropriate type of chart for the underlying data (continuous, categorical, ordinal)? Would a different type of visualisation do a better job of illustrating certain things about the data, such as how it varies? Would a different type of visualisation be better suited for telling the story (e.g. Categorical, Hierarchical, Relational, Temporal, Spatial)? Key reading: Kirk, A. (2024)

4. **Does the data visualisation follow good information graphic design practice?** (Week 5). Is the data visualisation free from chart-junk? Is the visualisation elegant and clear, letting the data speak for itself? Is the visualisation free of misrepresentation? Does the data visualisation make appropriate use of colour and text? Do colours complement one another, and are colour palettes chosen to complement the data story? Is the layout of the data visualisation able to guide the reader's eye comfortably? Is the data visualisation accessible, with appropriate contrast levels, text size, and colour-blind friendly colour use? Key reading: Tufte, 2001, Muth, 2022, 2023

Final assessment guidance: How to use what we have learned to critique and improve on a data visualisation

5. **Has the data visualisation been published in a way that is most appropriate for its platform?** (Week 7). Does the data visualisation "work" on the web or in print? If not, why not and how could it be changed to work better? Would it work on social media (i.e. is it available in the right aspect ratios)? How might it be adjusted to work better for some platforms? Is the data visualisation at a high enough resolution? Is it in an appropriate format (vector versus raster)? Key reading: Chang, W. (2024), Chapter 14

Final assessment guidance: How to use what we have learned to critique and improve on a data visualisation

5. **Has the data visualisation been published in a way that is most appropriate for its platform?** (Week 7). Does the data visualisation "work" on the web or in print? If not, why not and how could it be changed to work better? Would it work on social media (i.e. is it available in the right aspect ratios)? How might it be adjusted to work better for some platforms? Is the data visualisation at a high enough resolution? Is it in an appropriate format (vector versus raster)? Key reading: Chang, W. (2024), Chapter 14

6. **Does your data visualisation relate to space: is it a map or could it be a map?** (Week 8). Does the data in the data visualisation refer to space in any way? Is the data presented as a map? If yes, is it appropriate to visualise this data as a map or not (i.e. does it suffer from the "land doesn't vote" problem)? If yes, would the map be better presented as a cartogram than a choropleth? If the data is not a map, would it benefit from being a map? In what ways? Is the colour scheme in the map appropriate, or does it overstate differences? Key reading: Healy, K. (2018) Chapter 7; Ericson, (2011)

Final assessment guidance: How to use what we have learned to critique and improve on a data visualisation

7. **What impact does this data visualisation have on society?** (Week 10). How could this data visualisation be changed in *content*, *form*, or *process* in order to engage with feminist principles? How have identities or opinions been simplified into binaries or limited groups and could the data visualisation be recreated with more complex categorisations (or none)? How could multiple perspectives on the matter be presented to promote pluralism? How can the data visualisation interrogate power and oppression? How could data be recontextualised? How can the process of data visualisation involve greater participation of "data subjects"? How could the form of the data visualisation be changed to create affect and embodiment? How could the data visualisation better represent the voices of those it relates to, to push back on impersonalisation? Key reading: D'Ignazio & Klein, 2016

Look at the example student assessments (Example 1 - Spotify Article and Example 2 - FT article) that are attached to the assessment brief to get an idea of how to complete the assessment.

In your first assessment, you were being tested on being able to create specific data visualisations that were appropriate for the type of data. In *this* assessment, you will be marked *more heavily* based on your ability to use data storytelling, graphic design, presentational, and critical thinking skills.

Don't try to critique *everything* from the list of options I just went through — you will not have enough words to do so. Pick three or four things *maximum* to critique about the original data visualisation and to use to inform your improved data visualisation.

Remember that you can critique *both* "bad" things about the visualisation in the article *and* good things. Critique is not just about weaknesses, but also strengths. What did they do *well* in their visualisation?

All of the materials that you need to succeed in this assessment *have been provided in the module*. Look in the slides and worksheets from each week that you have been given in order to try and solve a problem *first*, and only consult Google, etc. afterwards.

Workshop this week

This week we will use our workshop time in order to work through some of the assessment options, so you can make a decision about which article you choose to critique and how you might go about it. A very good goal for this is to decide on an article in advance, and then focus in the workshop on how to get the data into **R** and tidied, ready for you to work on your data visualisation. That way we can help with any coding problems.

That's all!

This is our last lecture together, but I'll see you again in the workshop and you can still meet with me in my office hours. A reminder that there will be an additional drop-in session next week where you can get support with some of the trickier parts of your assessment, such as finding, reading in, and tidying data related to the topic.

If you don't take advantage of the workshop this week and the drop-in next week, you might find that you will have to figure a lot of these difficult things out on your own! While I always do my best to help everyone, there will be limited opportunities for support with the assessment over the holiday period — you have to meet me halfway and use the opportunities for support that are available.