SMI606: Week 1
What is quantitative social science?

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

Sign In

Learning Objectives

What will I learn? | How does this week fit into my course?

By the end of this week you will:

- Know where to find relevant information for participating in the module, including where to locate module resources and how to prepare for each week's class time.
- Be able to differentiate between quantitative social research approaches and qualitative research approaches and the way they shape our view of the social/human world.
- Be able to describe some of the reasons why quantitative research can be useful, while being cautious around the pitfalls of its methods and methodology.
- Have installed **R** and Rstudio and be able to interact with the console, run scripts, check help documentation, distinguish between environmental object types, and read in data.

Learning Objectives

What will I learn? | How does this week fit into my course?

- Help you schedule your independent learning for the rest of the module to fit within the demands of the programme.
- Give you an indication of when quantitative research projects (or components of projects) may be useful, and see how the content of SMI606 fits together with SMI607 Principles of Social Research I and Principles of Research Design II.
- Introduce you to **R**, a flexible open-source statistical programming language that can be used for quantitative data analysis but also for presenting, reporting, and other purposes.

Who we are:

Calum Webb

c.j.webb@sheffield.ac.uk

Research on poverty and the child welfare system in the UK using secondary survey and administrative data in **R**. My most recent project has been about how investment in services for children improves their welfare and safety.

Eleanor Bale

Eleanor's PhD focuses on the impact of migration on segregation and the effect of social frontiers on social and economic factors such as educational outcomes, mental health, and crime. Eleanor is interested in research about inequality, social mobility especially focusing on children's outcomes.


FULLY EXPECTING ~~for~~ HATE THIS CLASS!
STATS
CODE
@allison_horst

Module design, timetable, and support

- Introductions to important statistical principles and methods in quantitative social science.
    - Intuition *over* equations.

Module design, timetable, and support

- Introductions to important statistical principles and methods in quantitative social science.
    - Intuition *over* equations.
- 'Hands-on' experience putting these principles into practice in **R** using real-world data.

Module design, timetable, and support

- Introductions to important statistical principles and methods in quantitative social science.
    - Intuition *over* equations.
- 'Hands-on' experience putting these principles into practice in **R** using real-world data.
- Assessment on two small quantitative research projects where the topic is decided by you.

Module design, timetable, and support

Time & Location: 12:00-14:50 in The Wave Computer Room 2a. *You can bring your lunch.*

Consultation & Feedback Hours: 10:00-11:00 Tuesdays (Week 1-12) 15:00-16:00 Thursdays (Week 1-12)
Can be booked online or face-to-face. Book here.

Maths and Statistics Help (MASH)

The Maths and Statistics Help (MASH) service on Glossop Road provide 1:1 and group support on learning statistics in **R**. You can book a tutorial by visiting their webpages.

R for Ultra Beginners Workshops will usually be running through the first few weeks of the semester!

Responsibilities and expectations

As a rough guideline you **should expect to devote around 10 hours of study time per week to this module** during a typical 15-week semester.

Responsibilities and expectations

As a rough guideline you **should expect to devote around 10 hours of study time per week to this module** during a typical 15-week semester.

It is up to you to identify your learning needs and use the resources provided to meet them. I am happy to help you with this process.

Responsibilities and expectations

As a rough guideline you **should expect to devote around 10 hours of study time per week to this module** during a typical 15-week semester.

It is up to you to identify your learning needs and use the resources provided to meet them. I am happy to help you with this process.

To ensure your independent learning is productive, you should **regularly think about what parts of the module are challenging you**.

Responsibilities and expectations

As a rough guideline you **should expect to devote around 10 hours of study time per week to this module** during a typical 15-week semester.

It is up to you to identify your learning needs and use the resources provided to meet them. I am happy to help you with this process.

To ensure your independent learning is productive, you should **regularly think about what parts of the module are challenging you**.

But as a general rule you should follow the reading list and **work through ~1 chapter of the textbook per week**, as well as **finish any practical work started in the class**.

Responsibilities and expectations

As a rough guideline you **should expect to devote around 10 hours of study time per week to this module** during a typical 15-week semester.

It is up to you to identify your learning needs and use the resources provided to meet them. I am happy to help you with this process.

To ensure your independent learning is productive, you should **regularly think about what parts of the module are challenging you**.

But as a general rule you should follow the reading list and **work through ~1 chapter of the textbook per week**, as well as **finish any practical work started in the class**.

And lastly, remember to **ask for help if you need it!** There is no such thing as a stupid question.

What will we learn?

Week	Topic	Required Reading
1	What is quantitative social science?	Powell, T. C. (2019). '*Can Quantitative Research Solve Social Problems?*' & Mehmetoglu & Mittner (2022) *Chapter 1, 2 & 3, Applied Statistics Using R*
2	Types of quantification	Mehmetoglu & Mittner (2022) *Chapter 4, Applied Statistics Using R*
3	Relationships between variables	Mehmetoglu & Mittner (2022) *Chapter 5, Applied Statistics Using R*
4	Inference	Mehmetoglu & Mittner (2022) *Chapter 6, Applied Statistics Using R*
5	Causality	Goldthorpe, J. H. '*Causation, Statistics and Sociology*'
6	Bivariate Linear Regression	Mehmetoglu & Mittner (2022) *Chapter 7, Applied Statistics Using R*
7	Reading Week	Catch up reading week (**Ready for assessment 1!**)
8	Multiple Linear Regression	Mehmetoglu & Mittner (2022) *Chapter 8 & 9, Applied Statistics Using R*
9	Logistic Regression	Mehmetoglu & Mittner (2022) *Chapter 11, Applied Statistics Using R*
10	Cluster Analysis	UC Business Analytics **R** Programming Guide 'k-means Cluster Analysis' (Available free online) '*Hierarchical Cluster Analysis*' (Available free online)
11	Spatial Analysis	Imai, K. Chapter 5.3 '*Quantitative Social Science: An introduction* (Available online through the library) (**Ready for assessment 2!**)

Why R?

- R has a **steep learning curve**.
- R, mostly, **does not use menus**.
- R can have **unhelpful error messages**.
- R requires **learning how to code at the same time as learning methods and statistical theory**.

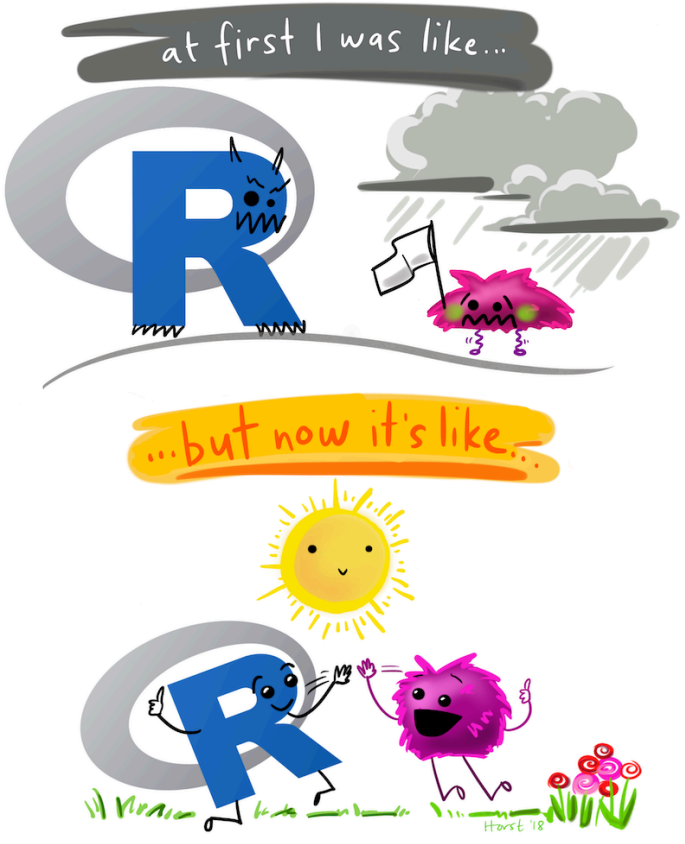


Illustration by Allison Horst

Why R?

- R has a **steep learning curve**.
- R, mostly, **does not use menus**.
- R can have **unhelpful error messages**.
- R requires **learning how to code at the same time as learning methods and statistical theory**.

- R is **free**.
- R is **flexible**: there are dozens of different ways to do the same thing. No way to use a certain method? You can even program it yourself.
- R promotes **reproducibility** through the use of script.
- R has a **great community** and a lot of **free resources**.

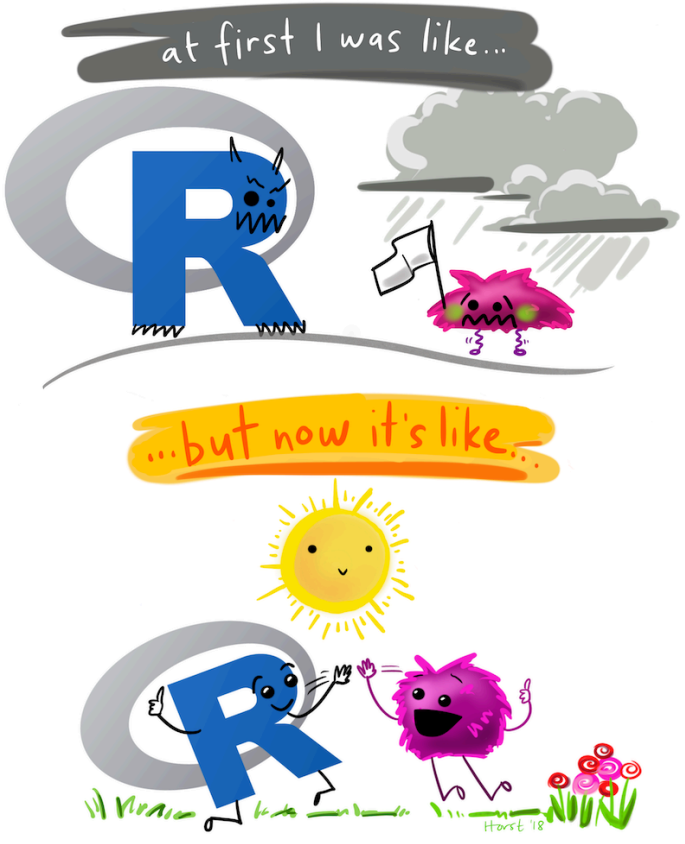


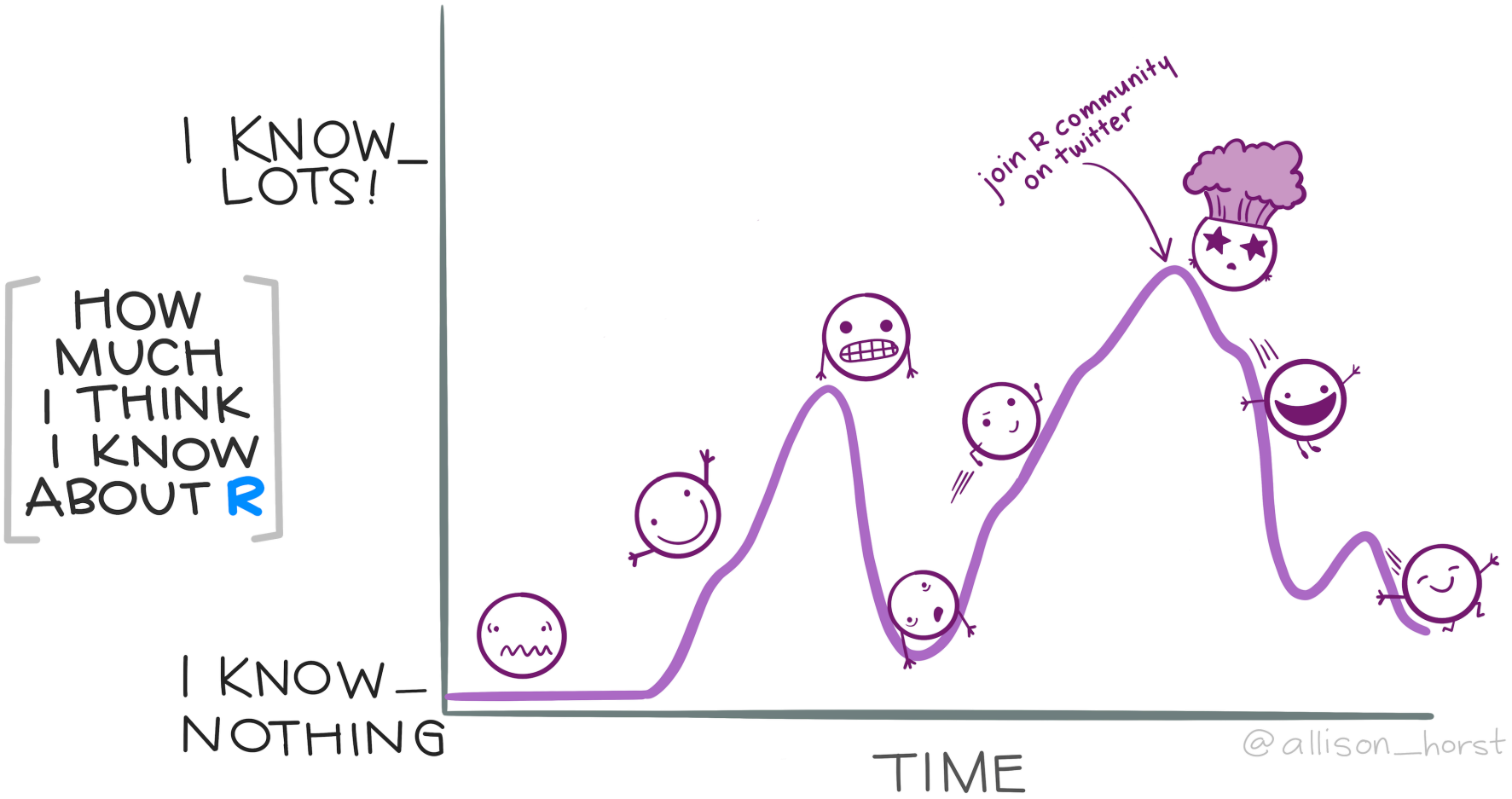
Illustration by Allison Horst

Base R

```
data$log_income <- log(data$income)
summary(lm(data = subset(data, age > 15), formula = log_income ~ age))
```

Tidyverse

```
data %>%
  filter(age > 15) %>%
  mutate(log_income = log(income)) %>%
  lm(., formula = log_income ~ age) %>%
  summary(.)
```


I KNOW LOTS!
HOW MUCH I THINK I KNOW ABOUT R
I KNOW NOTHING
join R community on twitter
TIME
@allison_horst

How to use the reading/resource list?

- Probably at least 50% of module conveners effort goes towards putting together a good quality reading & resource list. But not all students use it!

How to use the reading/resource list?

- Probably at least 50% of module conveners effort goes towards putting together a good quality reading & resource list. But not all students use it!

The reading and resource list for this module includes:

- Detailed information about the topic each week as well as **strongly suggested/required preparatory reading** as well as additional or alternative reading.
- Contains extra resources for:
  - Learning **R**
  - Open data to practice on or use for your assessment

Additional Resources & Workshops from the University

Library

Research Skills and Critical Thinking (RSCT) workshops:

- Discovering Information: Getting Started
- Discovering Information: A masterclass

Online resources:

- How to read a journal article
- Active reading for understanding
- Tutorials A-Z

301 Skills Centre

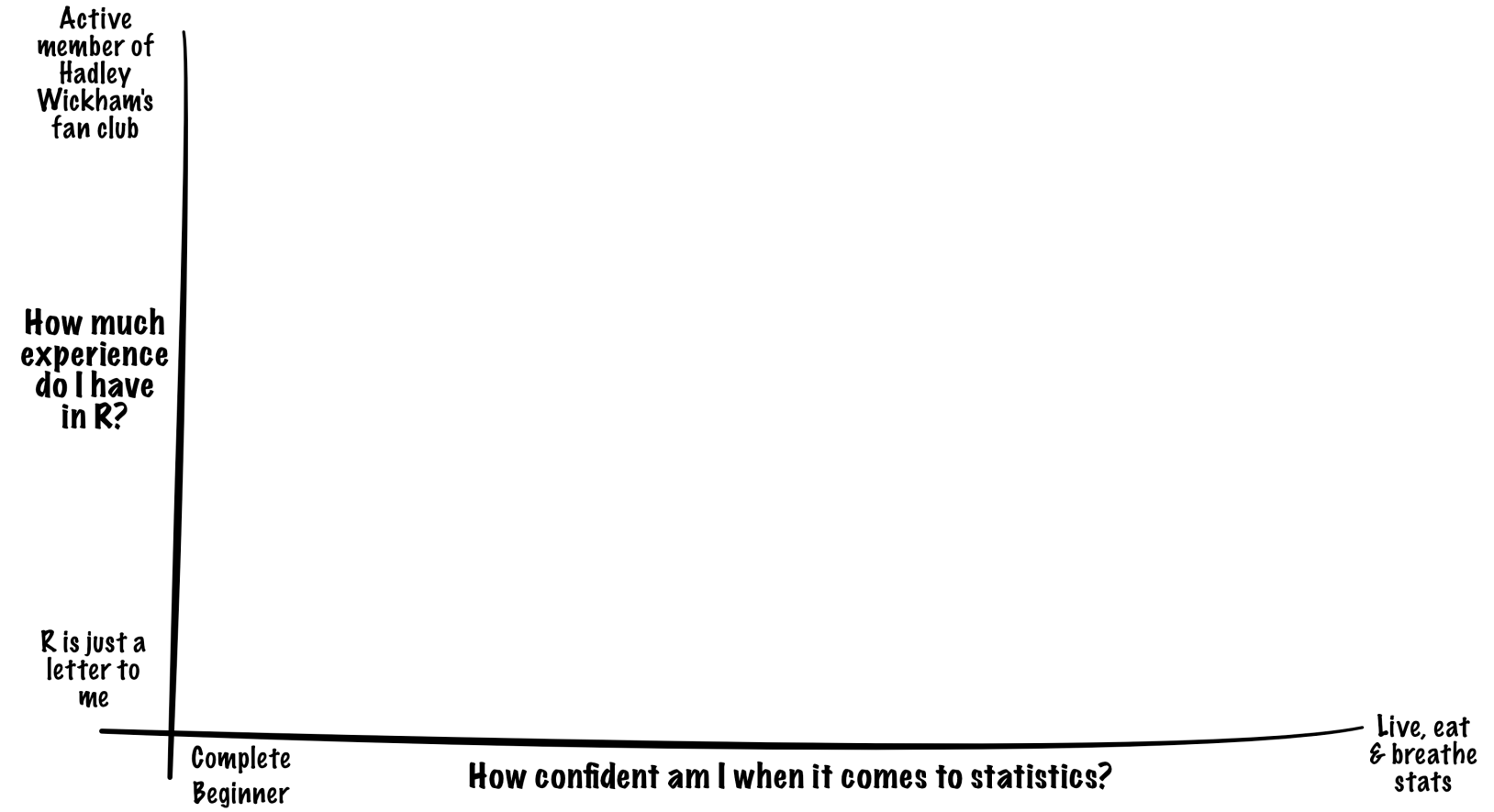
- Academic skills refresher
- Reading and note-taking at university
- Managing your time and avoiding distractions
- Study Skills 1-to-1s

M.A.S.H.

Statistics resources:

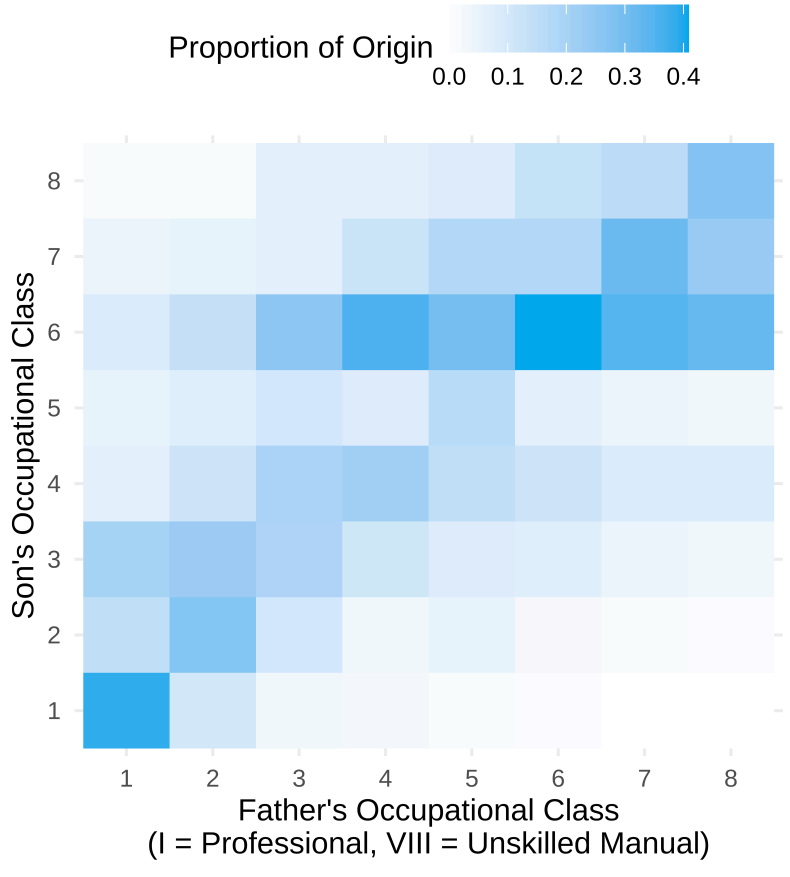
- First steps in R
- R/Excel/SPSS for Ultra-Beginners
- 1-to-1 Support

Exercise: Who are you?



What is quantitative social science?

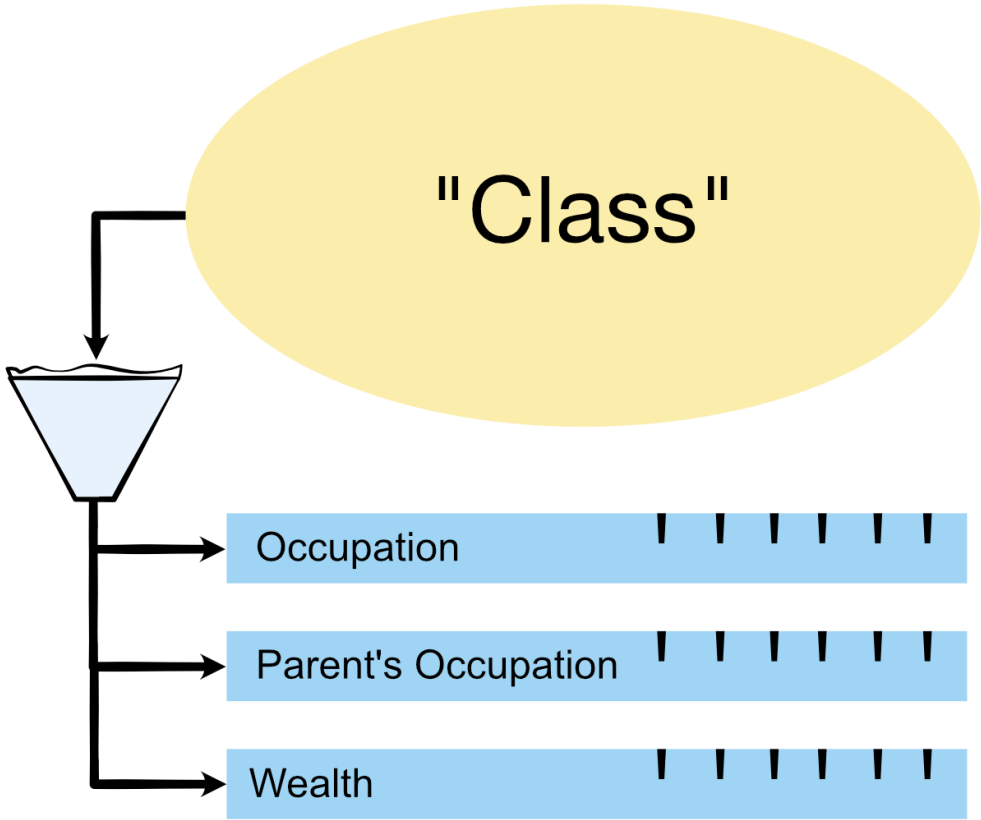
- The identification of **social patterns** through the quantification of data.
- **Operationalising** a concept so that it can be measured (on a numeric scale or codified with finite classifications).
- Checking the **validity** and **reliability** of that concept.
- Designing studies with **samples** that allow us to make **generalisations** about populations of interest, and/or claims about **causality**.
- **Collecting data** that can be organised, **explored** and **analysed** using a wide range of statistical methods.
- Make use of standardised, "**objective**" tools for collecting data, doing analysis, and answering research questions.



Data from Goodman, 1979 via Glass, 1954. Social Mobility in Britain. Data from 1949.

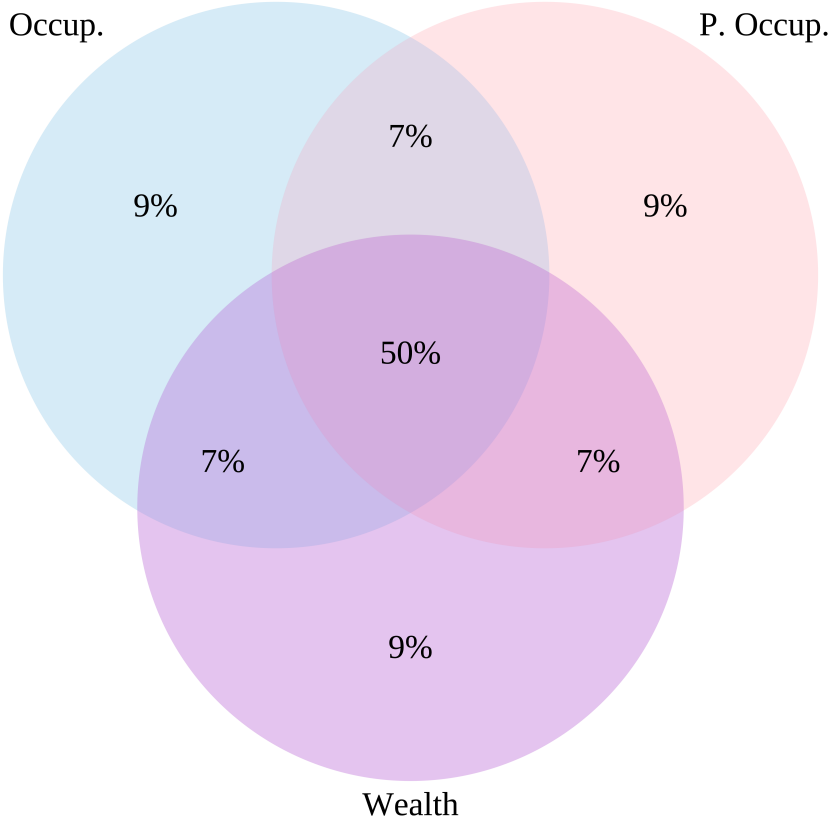
What is quantitative social science?

- The identification of **social patterns** through the quantification of data.
- **Operationalising** a concept so that it can be measured (on a numeric scale or codified with finite classifications).
- Checking the **validity** and **reliability** of that concept.
- Designing studies with **samples** that allow us to make **generalisations** about populations of interest, and/or claims about **causality**.
- **Collecting data** that can be organised, **explored** and **analysed** using a wide range of statistical methods.
- Make use of standardised, "**objective**" tools for collecting data, doing analysis, and answering research questions.



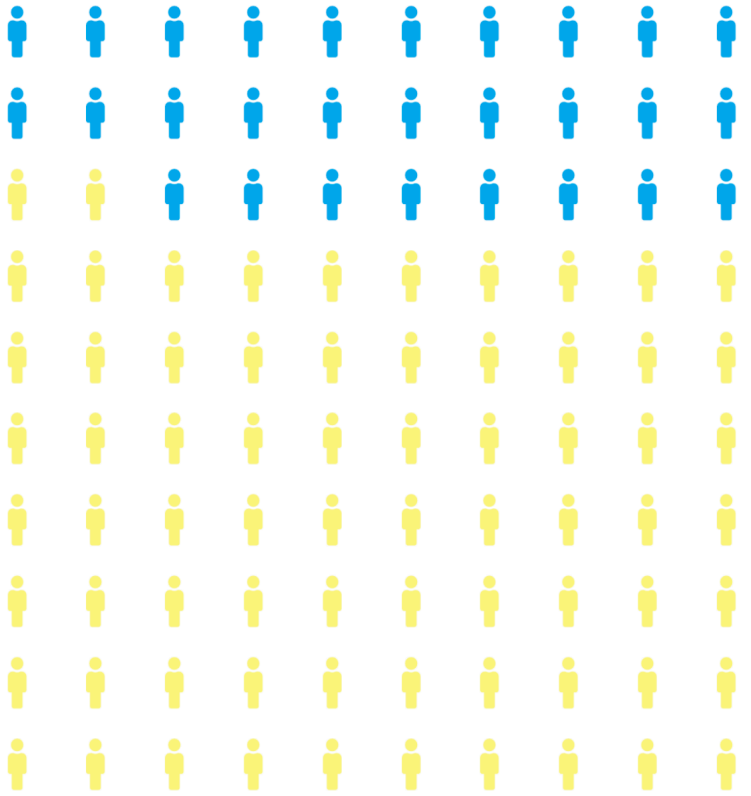
What is quantitative social science?

- The identification of **social patterns** through the quantification of data.
- **Operationalising** a concept so that it can be measured (on a numeric scale or codified with finite classifications).
- Checking the **validity** and **reliability** of that concept.
- Designing studies with **samples** that allow us to make **generalisations** about populations of interest, and/or claims about **causality**.
- **Collecting data** that can be organised, **explored** and **analysed** using a wide range of statistical methods.
- Make use of standardised, "**objective**" tools for collecting data, doing analysis, and answering research questions.



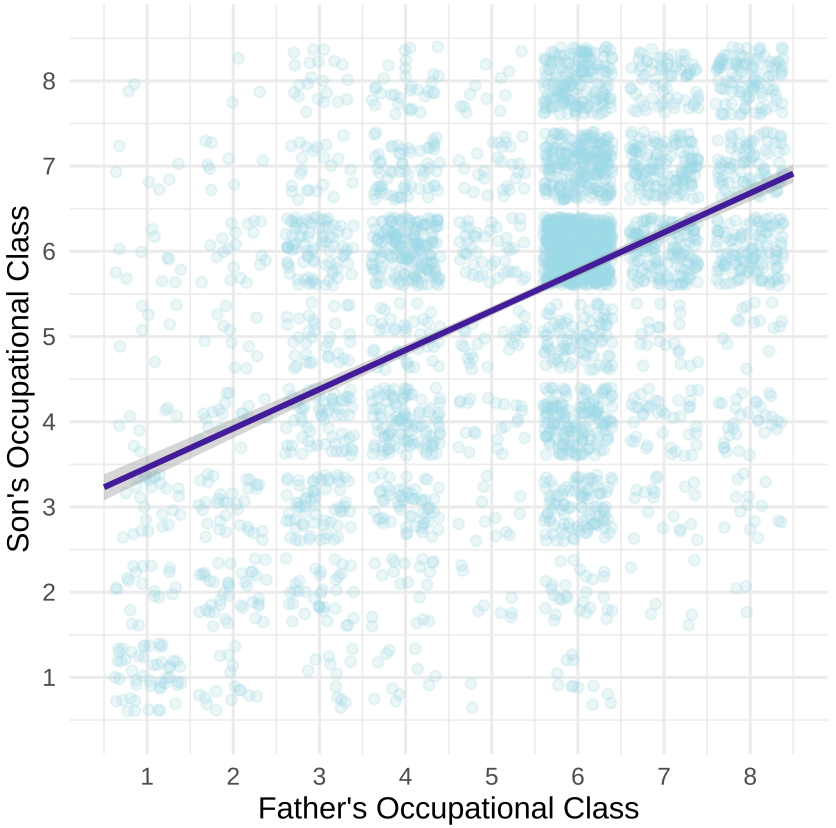
What is quantitative social science?

- The identification of **social patterns** through the quantification of data.
- **Operationalising** a concept so that it can be measured (on a numeric scale or codified with finite classifications).
- Checking the **validity** and **reliability** of that concept.
- Designing studies with **samples** that allow us to make **generalisations** about populations of interest, and/or claims about **causality**.
- **Collecting data** that can be organised, **explored** and **analysed** using a wide range of statistical methods.
- Make use of standardised, "**objective**" tools for collecting data, doing analysis, and answering research questions.



What is quantitative social science?

- The identification of **social patterns** through the quantification of data.
- **Operationalising** a concept so that it can be measured (on a numeric scale or codified with finite classifications).
- Checking the **validity** and **reliability** of that concept.
- Designing studies with **samples** that allow us to make **generalisations** about populations of interest, and/or claims about **causality**.
- **Collecting data** that can be organised, **explored** and **analysed** using a wide range of statistical methods.
- Make use of standardised, "**objective**" tools for collecting data, doing analysis, and answering research questions.



$$\hat{Destination} = 3 + 0.46 \cdot Origin$$

What is quantitative social science?

- The identification of **social patterns** through the quantification of data.
- **Operationalising** a concept so that it can be measured (on a numeric scale or codified with finite classifications).
- Checking the **validity** and **reliability** of that concept.
- Designing studies with **samples** that allow us to make **generalisations** about populations of interest, and/or claims about **causality**.
- **Collecting data** that can be organised, **explored** and **analysed** using a wide range of statistical methods.
- Make use of standardised, "**objective**" tools for collecting data, doing analysis, and answering research questions.

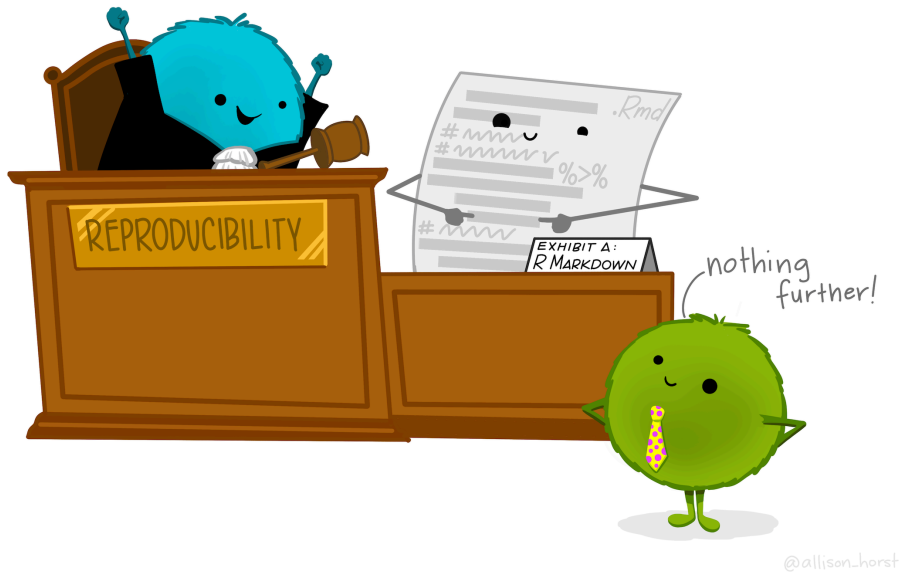


Illustration by Allison Horst

Positivism

The way we research something reflects (implicitly or explicitly) **how we understand the human world**. You will explore this in much more detail in SMI607 Principles of Social Research I.

- Quantitative research has a long association with Positivism: the **idea that we can apply the same scientific principles we use to study the natural world to understand the human/social world**.

Positivism

The way we research something reflects (implicitly or explicitly) **how we understand the human world**. You will explore this in much more detail in SMI607 Principles of Social Research I.

- Quantitative research has a long association with Positivism: the **idea that we can apply the same scientific principles we use to study the natural world to understand the human/social world**.
- "**Qualitative and quantitative research methods are now regarded as forming different, but equally vital, aspects of the social science research endeavour.**" Walter, 2006: 25. While quantitative research can be useful for establishing consistent **social patterns** and challenging conventional wisdom, it is limited in developing **social meaning**.

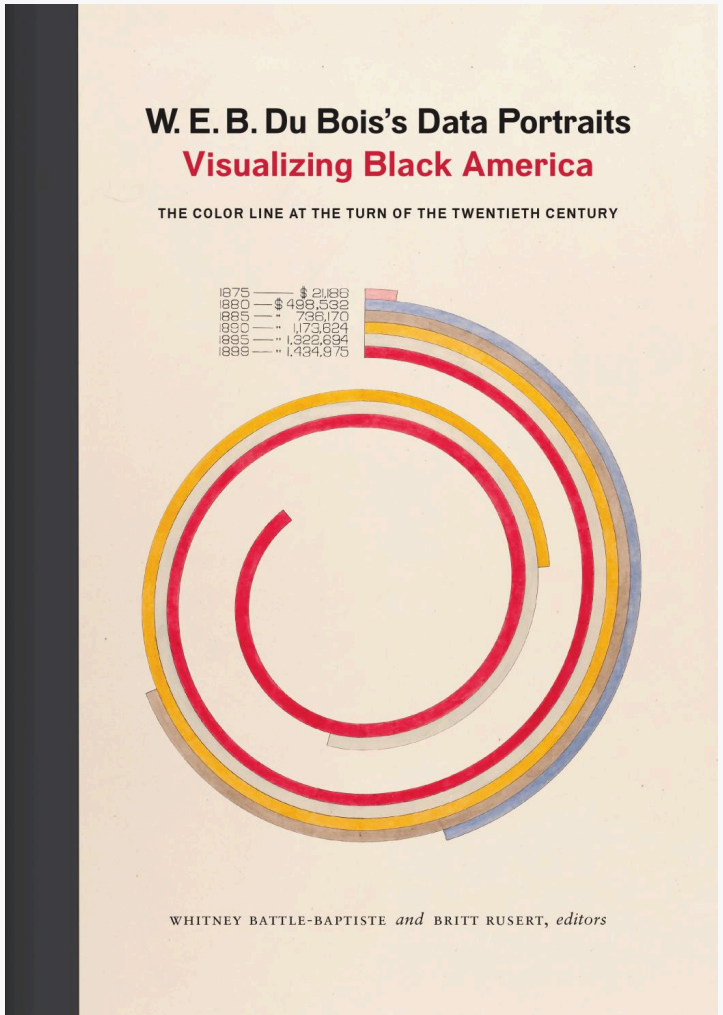
Beyond Positivism: Pragmatism, social problems, and politics

Can quantitative social research solve social problems?

- Quantitative research as **activism** and **counter-storytelling** epitomised by **W.E.B. Du Bois**.

> **"The twisted logic ran if the black man was inferior he was not oppressed-his place in society was appropriate to his meager talent and intellect. Dr. Du Bois recognized that the keystone in the arch of oppression was the myth of inferiority and he dedicated his brilliant talents to demolish it."**

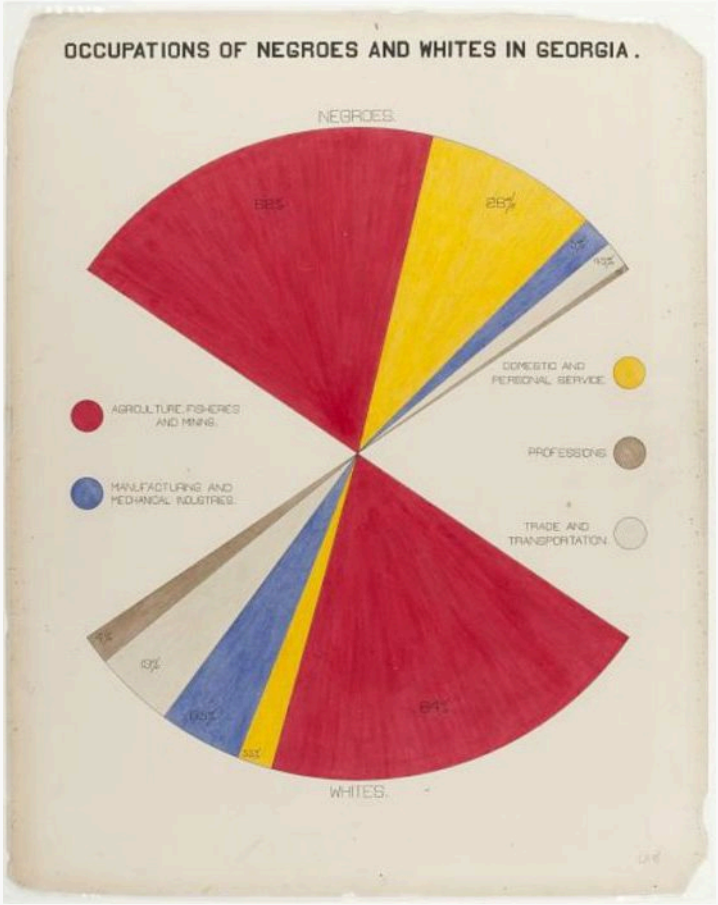
Dr. Martin Luther King Jr., cited in Morris, n.d.



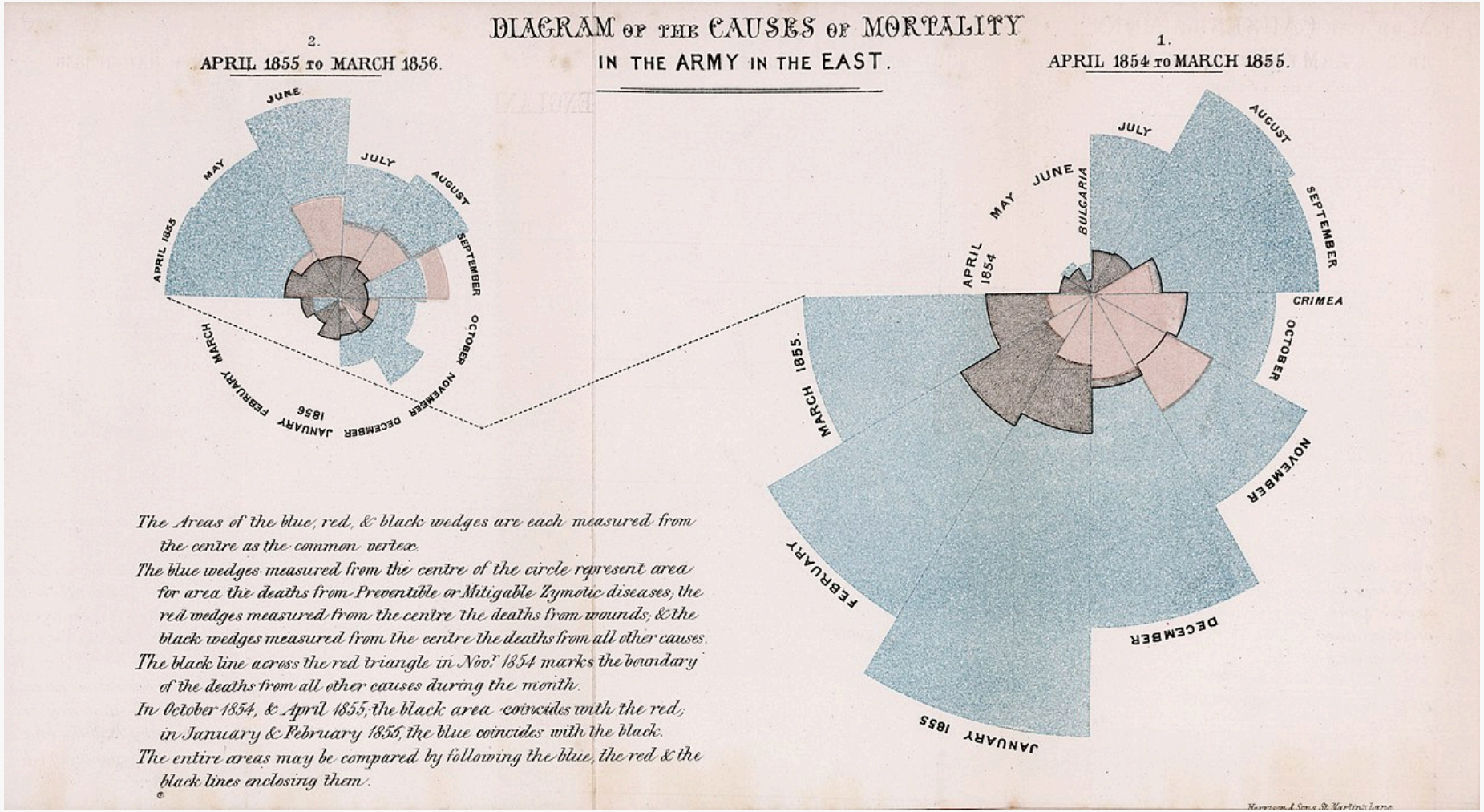
Beyond Positivism: Pragmatism, social problems, and politics

Du Bois engaged in the continuous systemic collection, analysis, interpretation and dissemination of public health-related data in people of various racial groups [between 1896 and 1897] ... Du Bois' analysis and interpretation of racism led to his summative pronouncement that **systemic inequality was associated with increasing health risks and high mortality rates** ... When he connected the systemic inequality of racism to education, criminal justice, poverty, and employment, Du Bois demonstrated the importance of a multidisciplinary collaborative lens of social epidemiological research. ... Du Bois' recommended interventions and strategies to confront the determinants of health and determinants of death in the 7th Ward would have been well posed for policy development aimed at combating racism and eliminating health disparities.

Jones-Eversley & Dean. (2018). *After 121 Years, It's Time to Recognize W.E.B. Du Bois as a Founding Father of Social Epidemiology*



One of Du Bois' Data Portraits presented at the 1900 Paris Exhibition.



Florence Nightingale's famous 1850s data visualiation showing the impact of the Sanitary Commission on reducing the deaths of front line soldiers in the Crimean War.

"Whenever I am infuriated, I revenge myself with a new diagram."

Florence Nightingale (August 1857), cited in Gupta, 2020

"It’s worth noting, before getting started, that this material is hard. If you find yourself confused at any point, you are normal. Any sense of confusion you feel is just your brain correctly calibrating to the subject matter. Over time, confusion is replaced by comprehension..."

Richard McElreath. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan. 2nd Ed.*. Boca Raton, FL: CRC Press. p.193

Value Objectivity

- Can we ever be free of all bias and political motivation in our research, as historical Positivism promoted (**dispassionate social inquiry**)?
- Would we even enjoy or want to do research if we were? **What would the point be?**

Methodological Objectivity

- What we *can* do is ensure the methods we use, our data, and the steps we went through are **transparent, replicable, and well-communicated**.
- What do we lose by accepting and making our values explicit but retaining our methodological objectivity? What do we gain? (e.g. see: Indigenous Methodologies in Walter, 2006 and Andersen & Walter, 2013).

Bifurcation of objectivity: Jenkins, 2018; Powell, 2020.

References

- **Battle-Baptiste, W., & Rusert, B. (Eds.). (2018).** WEB Du Bois's data portraits: Visualizing black America. Chronicle Books.
- **Goodman, L. A. (1979).** Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association*, 74(367), 537-552.
- **Gupta, S. (2020).** Florence Nightingale understood the power of visualizing science. *ScienceNews.*
- **Jenkins, R. (2018).** Foundations of sociology: Towards a better understanding of the human world. *Macmillan International Higher Education*.
- **Jones-Eversley, S. D., & Dean, L. T. (2018).** After 121 years, it's time to recognize WEB Du Bois as a founding father of social epidemiology. *The Journal of Negro Education*, 87(3), 230-245.
- **Morris, A. (n.d.)** What's in a Name? W. E. B. Du Bois vs. W.E.B. DeBois. *UC Press Blog*.
- **Powell, T. C. (2020).** Can quantitative research solve social problems? Pragmatism and the ethics of social research. *Journal of Business Ethics*, 167(1), 41-48.
- **Walter, M. (2006).** The nature of social science research. *Social research methods: An Australian perspective*, 1-28.
- **Walter, M., & Andersen, C. (2016).** Indigenous statistics: A quantitative research methodology. *Routledge.*

Getting set up in R

- Go to the **SMI606 Blackboard page -> Learning Materials -> Week 1**, and follow the tutorial for downloading and installing `R` and Rstudio.
- Use the **visual tour of Rstudio sheet** to familiarise yourself with the software.
- Follow the **"Creating a New R Project and Script" tutorial** to create a new project called "week-1-exercise"

Week 1 `R` Exercise

- Copy the files in the "**Week 1 R Exercise**" folder in R into your new project directory.
- Open `week-1.R` in Rstudio and follow the tutorial written into the script to take your first steps programming in `R`!
- *You may not have time to finish this entire tutorial before the end of this session, but should complete it during your independent study time.*
- *In your own time, complete the challenge at the end and post about any errors in `R` you experienced on the class discussion board (link on Blackboard)..*