Computational social science

An Introduction
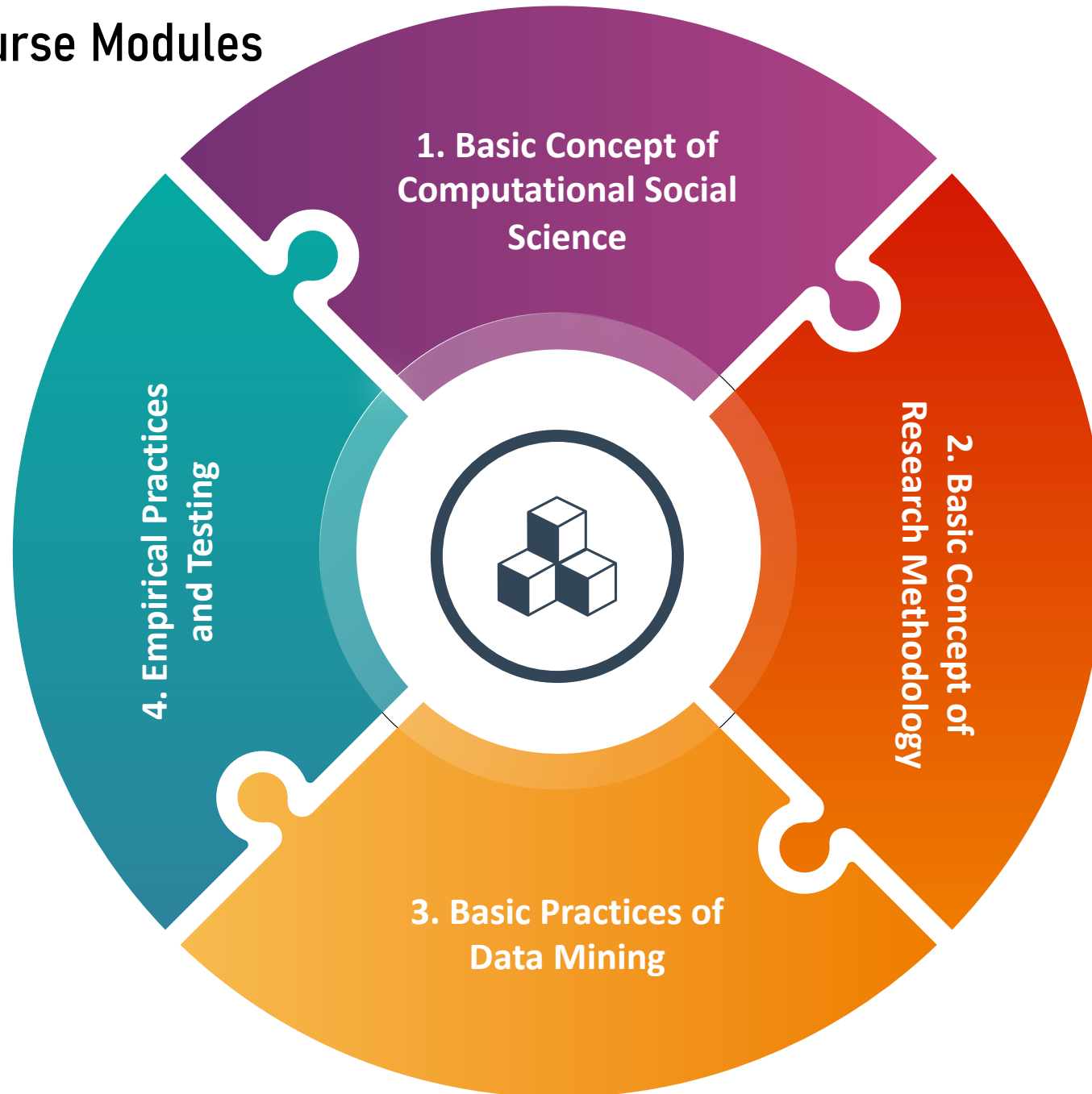
Data Mining Social Science

Week 01

Orientation

Basic Concept of Course Modules

1. Basic Concept of Computational Social Science

2. Basic Concept of Research Methodology

3. Basic Practices of Data Mining

4. Empirical Practices and Testing
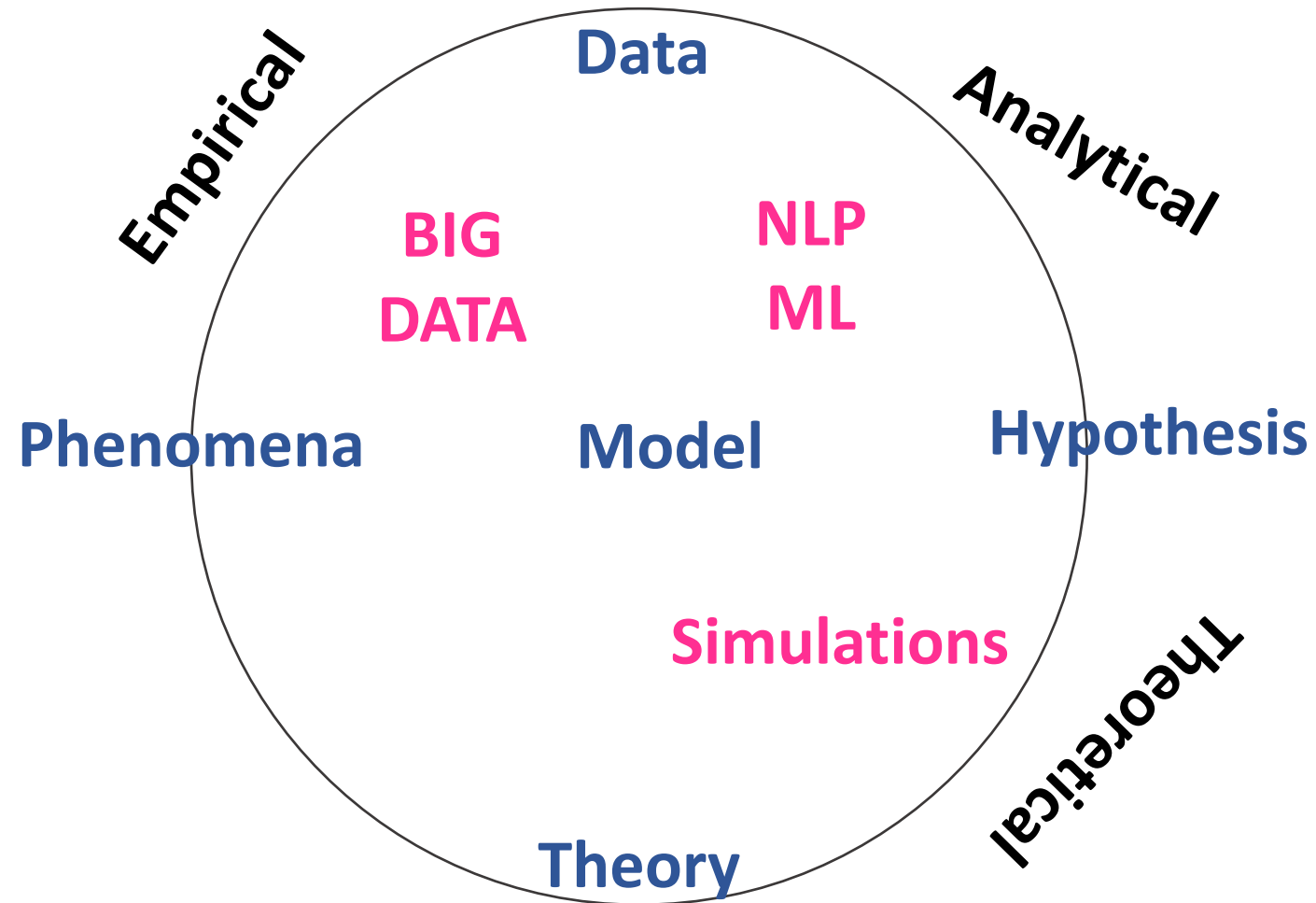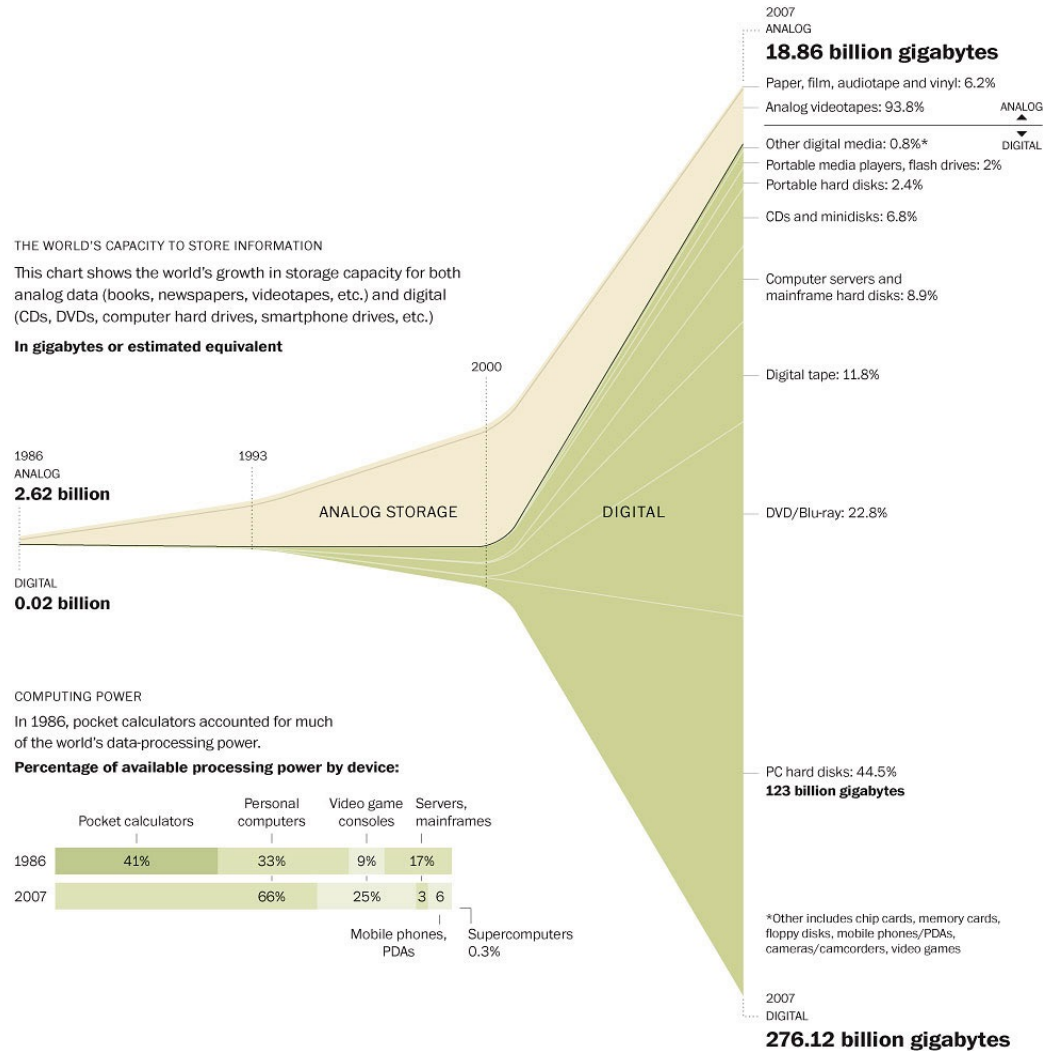
Computational Social Science

- Computational social science is an interdisciplinary field that includes economics, political science, public policy, psychology, and sociology, among others.

- The discipline employs data to grasp and address social and behavioural issues.

- For instance, researchers investigate racial discrimination in the labour market, evaluate the impact of the news media on the asset market, and anticipate government stability based on social media usage.

- A data-driven methodology could be utilised in law, journalism, linguistics, and literature.

- Since social scientists are concerned with a vast array of real-world concerns, the outcomes of their studies have the potential to have a significant impact on society, government policies, and business practices.
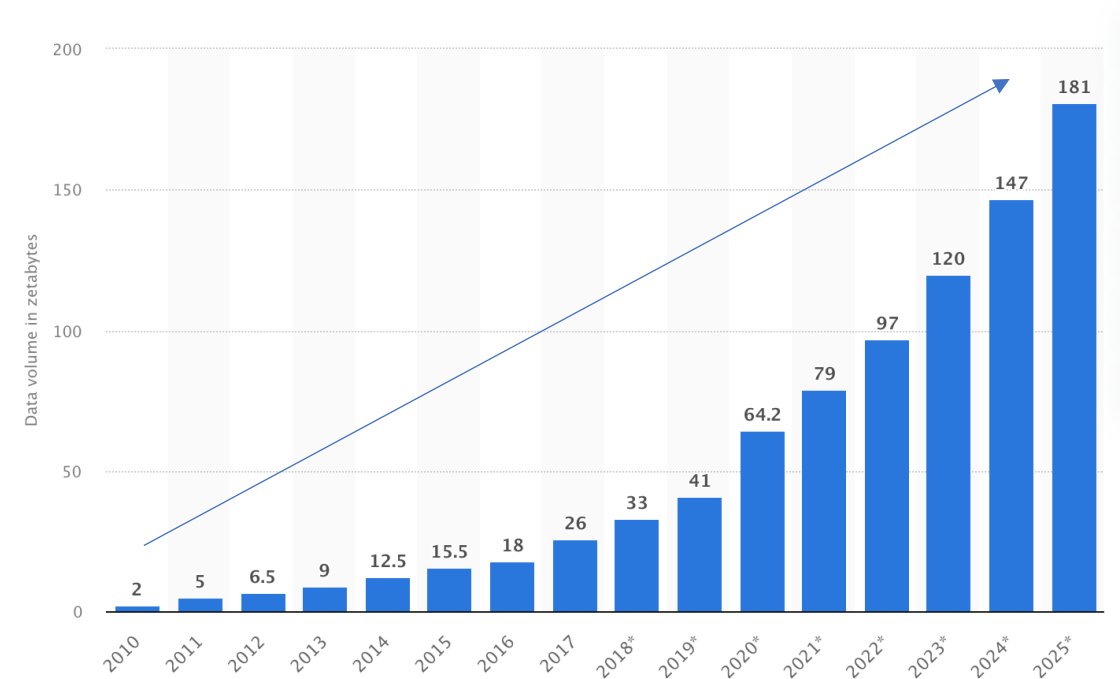
Computational Social Science

Computational Social Science

THE WORLD'S CAPACITY TO STORE INFORMATION

This chart shows the world's growth in storage capacity for both analog data (books, newspapers, videotapes, etc.) and digital (CDs, DVDs, computer hard drives, smartphone drives, etc.)

In gigabytes or estimated equivalent

1986 ANALOG **2.62 billion**

DIGITAL **0.02 billion**

ANALOG STORAGE

DIGITAL

2000

1993

COMPUTING POWER

In 1986, pocket calculators accounted for much of the world's data-processing power.

Percentage of available processing power by device:

	Pocket calculators	Personal computers	Video game consoles	Servers, mainframes
1986	41%	33%	9%	17%
2007		66%	25%	3 6

Mobile phones, PDAs

Supercomputers 0.3%

2007 ANALOG **18.86 billion gigabytes**

Paper, film, audiotape and vinyl: 6.2%
Analog videotapes: 93.8%   ANALOG ▲ ▼ DIGITAL

Other digital media: 0.8%*
Portable media players, flash drives: 2%
Portable hard disks: 2.4%

CDs and minidisks: 6.8%

Computer servers and mainframe hard disks: 8.9%

Digital tape: 11.8%

DVD/Blu-ray: 22.8%

PC hard disks: 44.5% **123 billion gigabytes**

*Other includes chip cards, memory cards, floppy disks, mobile phones/PDAs, cameras/camcorders, video games

2007 DIGITAL **276.12 billion gigabytes**

Volume of data created in worldwide from 2010 to 2020 with forecasts from 2021 to 2025

Data volume in zetabytes

Year	Value
2010	2
2011	5
2012	6.5
2013	9
2014	12.5
2015	15.5
2016	18
2017	26
2018*	33
2019*	41
2020*	64.2
2021*	79
2022*	97
2023*	120
2024*	147
2025*	181

Mining Social Science

Qualitative, Quantitative and Mining

Method	Data	Methodology
Qualitative Research	Text (Qualitative data)	Qualitative analysis
Quantitative Research	Numerical value (Quantitative data)	Quantitative analysis
Text Mining	Text (Quantitative data)	Quantitative analysis
Data Mining	Numerical value (Quantitative data)	Quantitative analysis

- Text mining is a quantitative research that uses qualitative forms of text with quantitative analysis techniques like statistics and multivariate analysis.

- Text mining is a method that strengthens the dependability, validity, and evidence of qualitative research as well as its accountability.

Qualitative Data Analysis in a Digital World

Qualitative analysis, at its very core, can be condensed to a close and repeated review of data, categorising, interpreting, and writing. In other words, qualitative analysis is based on secondary data review.

It focuses on using logical reasoning to determine what caused what or how important something is based on different accounts of observed facts and how other researchers judged those facts.

- Why do we do data mining with qualitative data?

- As the data source, the text is probably one of the most studied in social sciences, especially in sociology and political science.

- Data mining is a new way to investigate things that go beyond surveys for social science research.

- Text mining can confirm, supplement and contradict quantitative analysis results.

Quantitative Data Analysis in a Digital World

"Most qualitative researchers reject the positive aspects of regularity, predictability, and causal analysis via variable-centred research as inapplicable to the meaningful, human, social world."
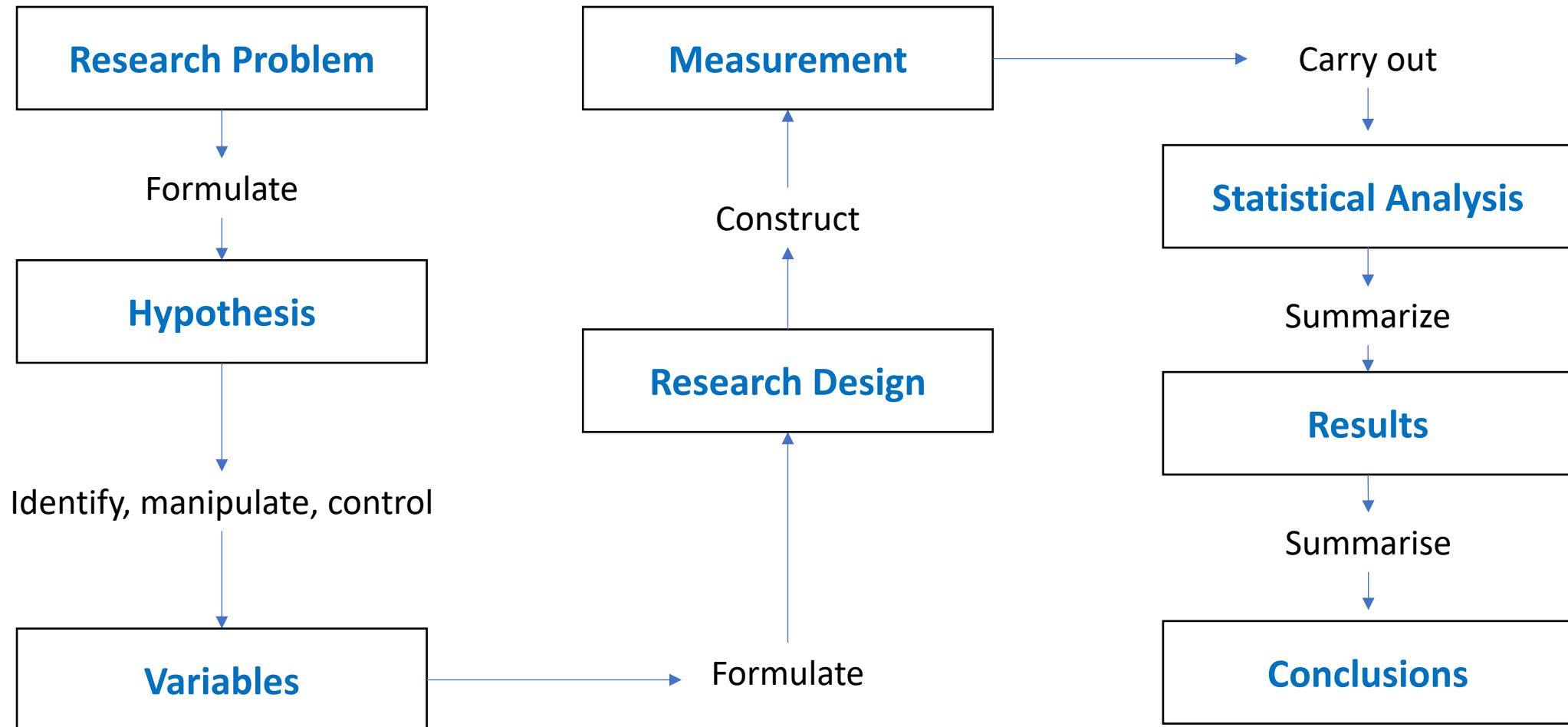
Quantitative analysis, which looks at the construction of cause-and-effect relationships by putting observations, events, and their results into numbers, is used to determine how things work.

- Why do we do data mining with quantitative data?

- Quantitative means that we are working with numbers, primarily data, and social science is the vast field of research we are studying, such as politics, economics, and perhaps law.

- Quantitative social science is the study of social science through the collection and analysis of numbers.

Scientific Research

- A systematic, controlled, empirical, and critical investigation is the researcher's key to investigating social phenomena.

- The primary goal of scientific inquiry is to uncover and construct a corpus of organised information.

- Methodical and empirical research creates hypotheses, conceptions, generalisations, and principles.

- Variables with cause-and-effect relationships can be predicted and managed using research.

Importance of research steps

Lecture Schedule

The course will cover popular text data collection, processing, and analysis methods. Approaches are supervised and unsupervised.

The training will cover topics such as:

- Web scraping methods use

- Text cleaning and preprocessing

- Dictionary methods

- Machine learning approaches

- Topic modelling

- Scaling with applied statistics

Lost Key

"I'm searching for my keys."

Emergence

Emergence – How Stupid Things Become Smart Together

- https://www.youtube.com/watch?v=16W7c0mb-rE

Next Week

Please bring your laptop.

Week 2: Hello World and Computational Social Science

This section overviews the rest of the course's ideas in computer science and R.

- Data as Qualitative and Quantitative
- What is computational social science?
- An introduction to R and RStudio and the most common commands and functions.
- Hello World
- Arithmetic Operations, Objects, Vectors, Functions, Data Files, Saving Objects and Packages
- Programming and Exercises