Research Design 1

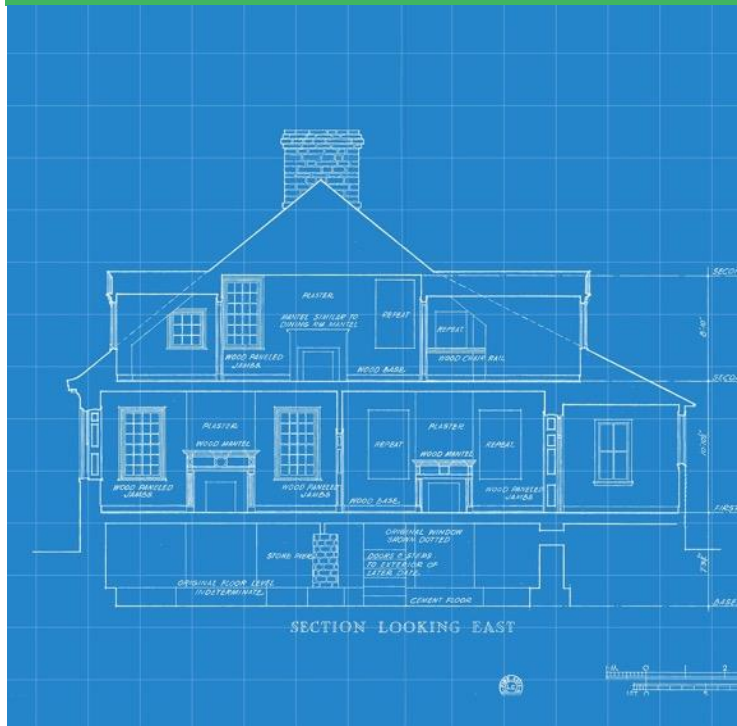
Quantitative Analysis
Week 2

Why study research design?

- Statistical methods are tools. They allow us to solve very specific problems and answer very specific questions.
  - For example: “what is the relationship between variable A and variable B”; or “what clusters of similar observations exist within data set X”.
- Having more of these tools in our toolbox – i.e. knowing more specific statistical methods – allows us to approach a wider range of problems.

Why study research design?

- But those questions are not research questions – they are statistics problems.
- We may be able to break a research question down into a set of statistics problems (this is called operationalisation), but the statistics methods themselves are not the research design.
- If statistics methods are the tools, research design is the blueprint. You don't start building a house by just picking up a toolbox and starting work; first you need detailed plans.

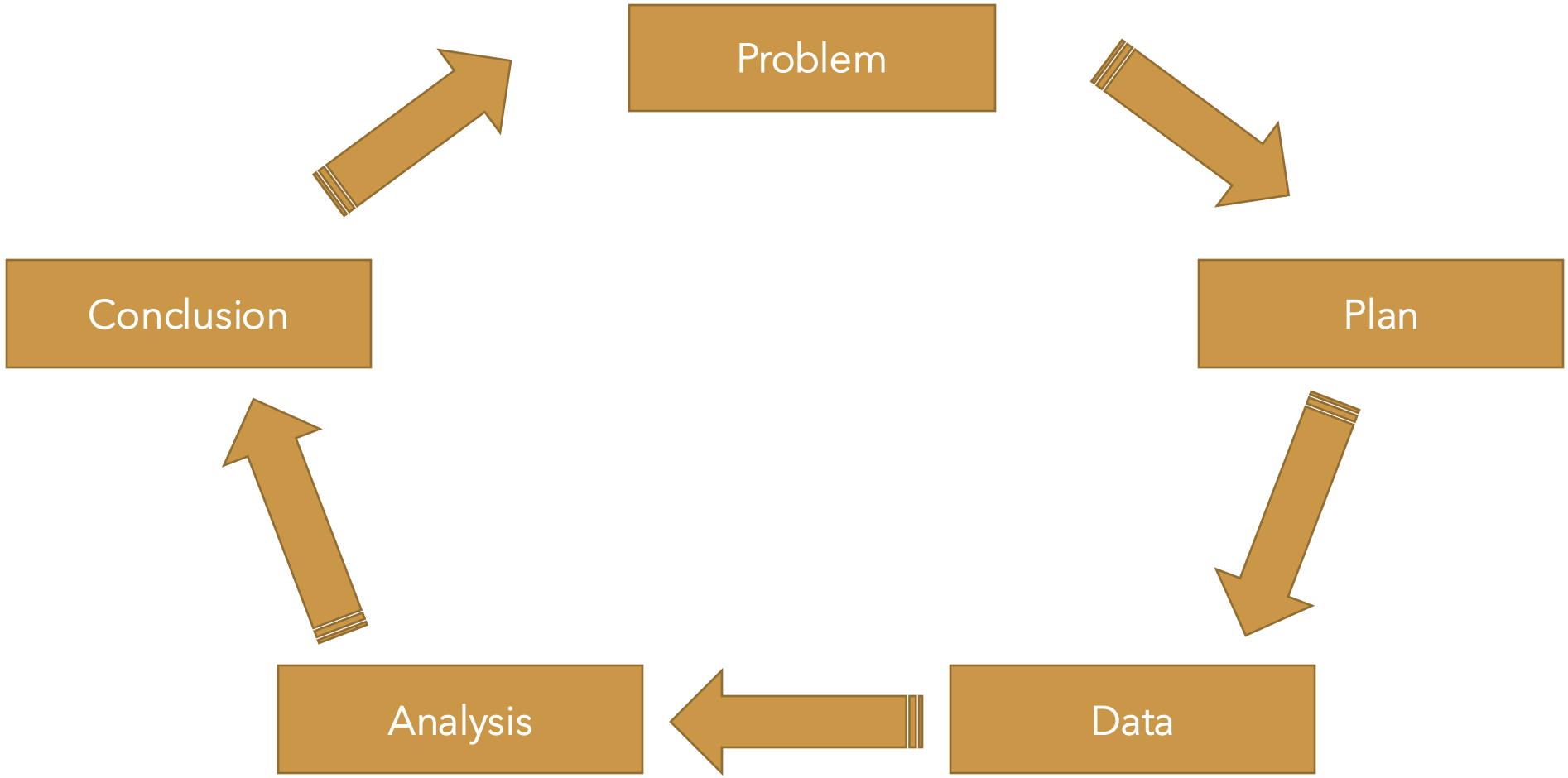


“When all you have is a hammer…

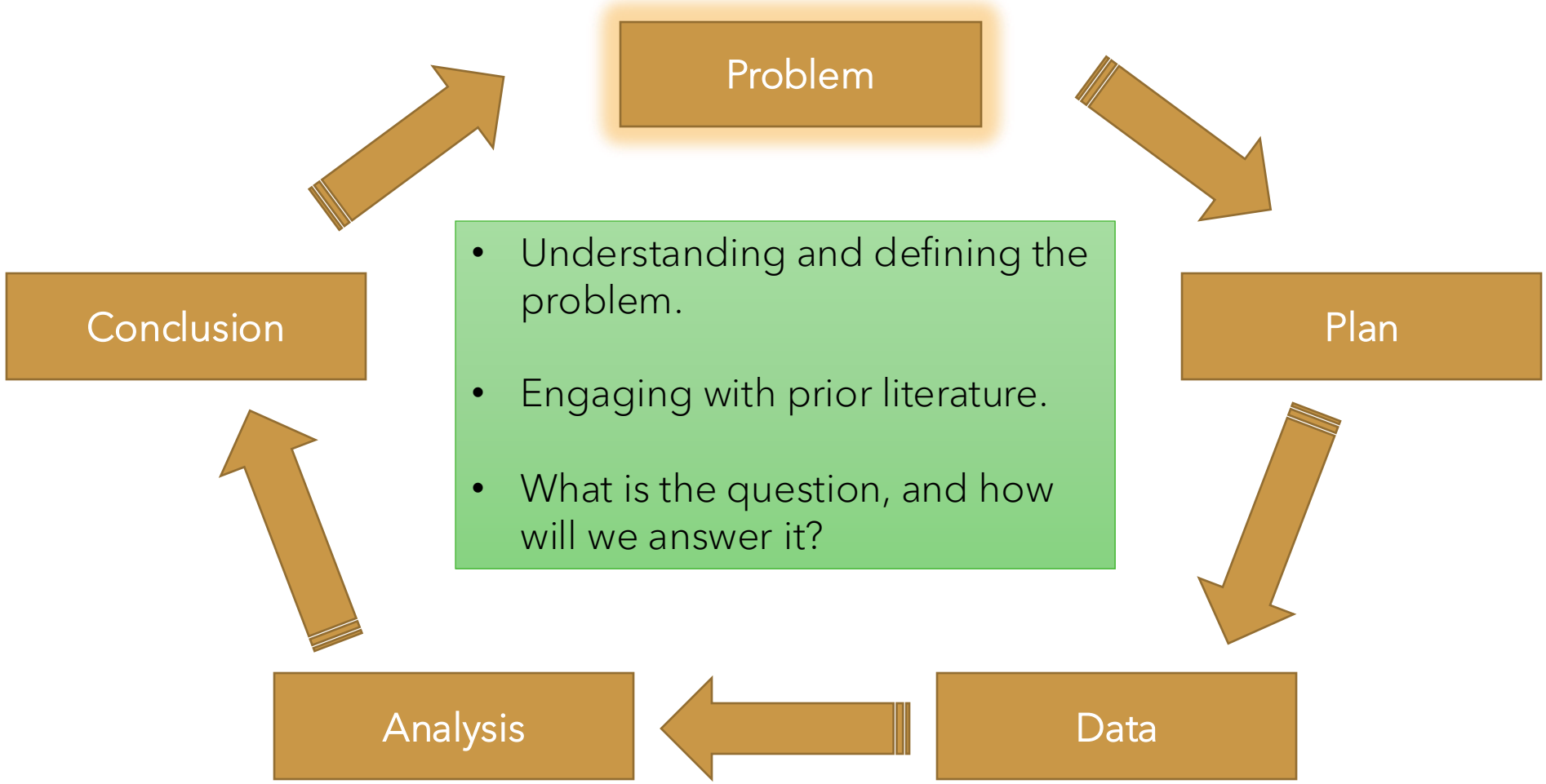
- …everything looks like a nail.”
- Using our knowledge of statistics as a starting point in research risks running into this problem: doing research based on which stats methods you prefer, instead of choosing methods based on what fits the research best.
- This is why it’s important to study research design alongside expanding our knowledge of statistical methods:

Good research design shows us where, when, and how to use our tools.

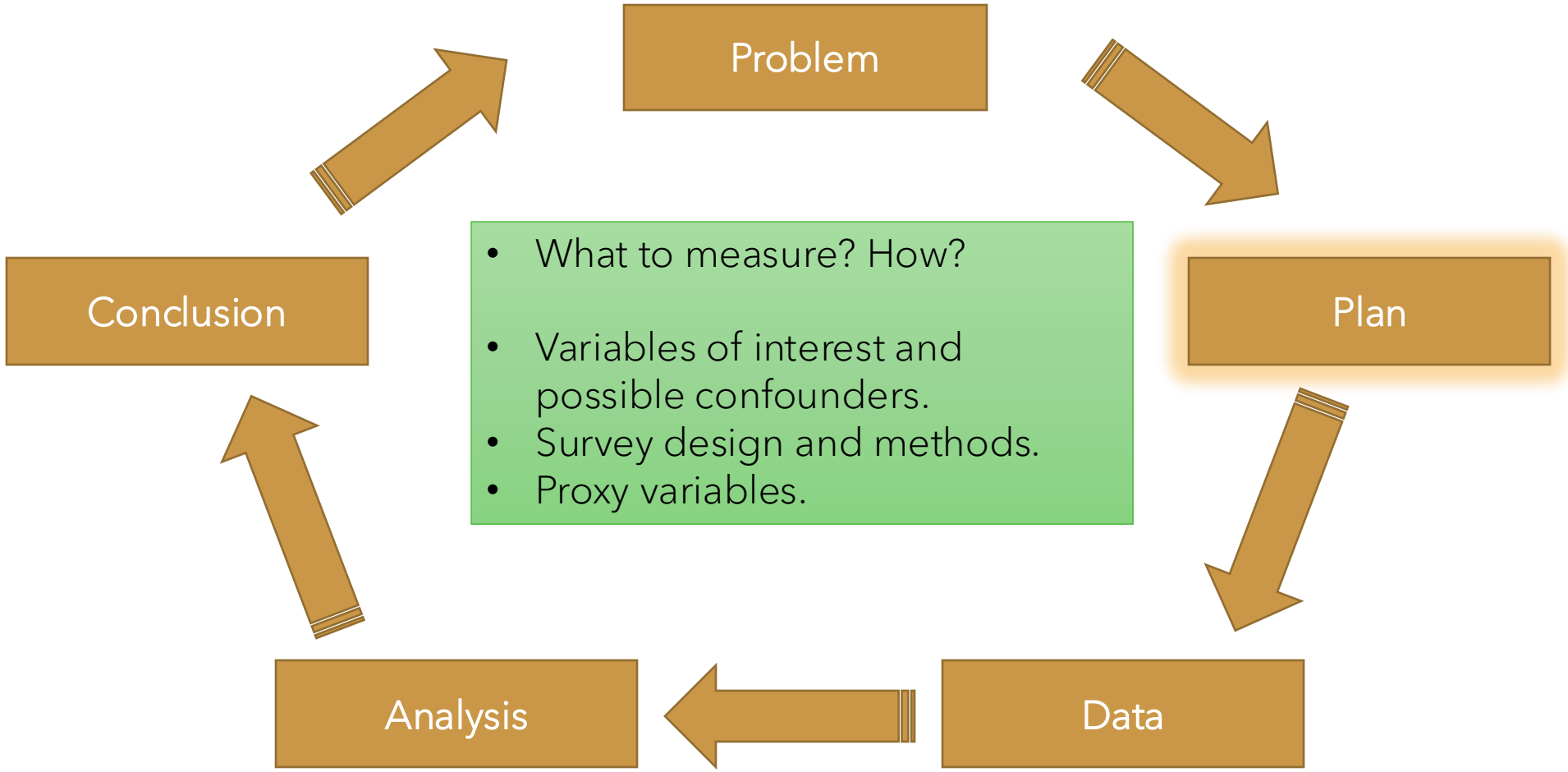
The PPDAC Cycle



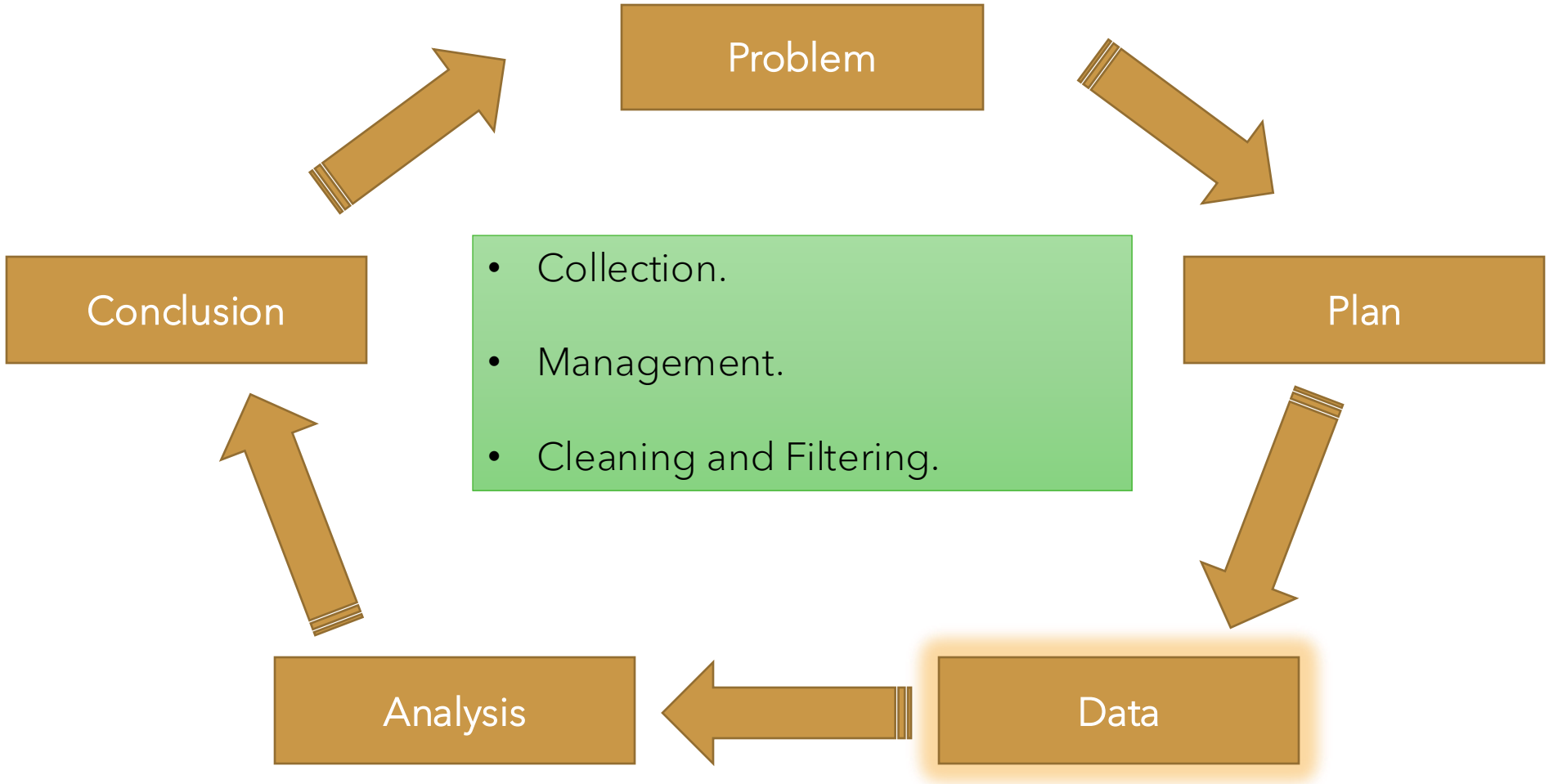
The PPDAC Cycle



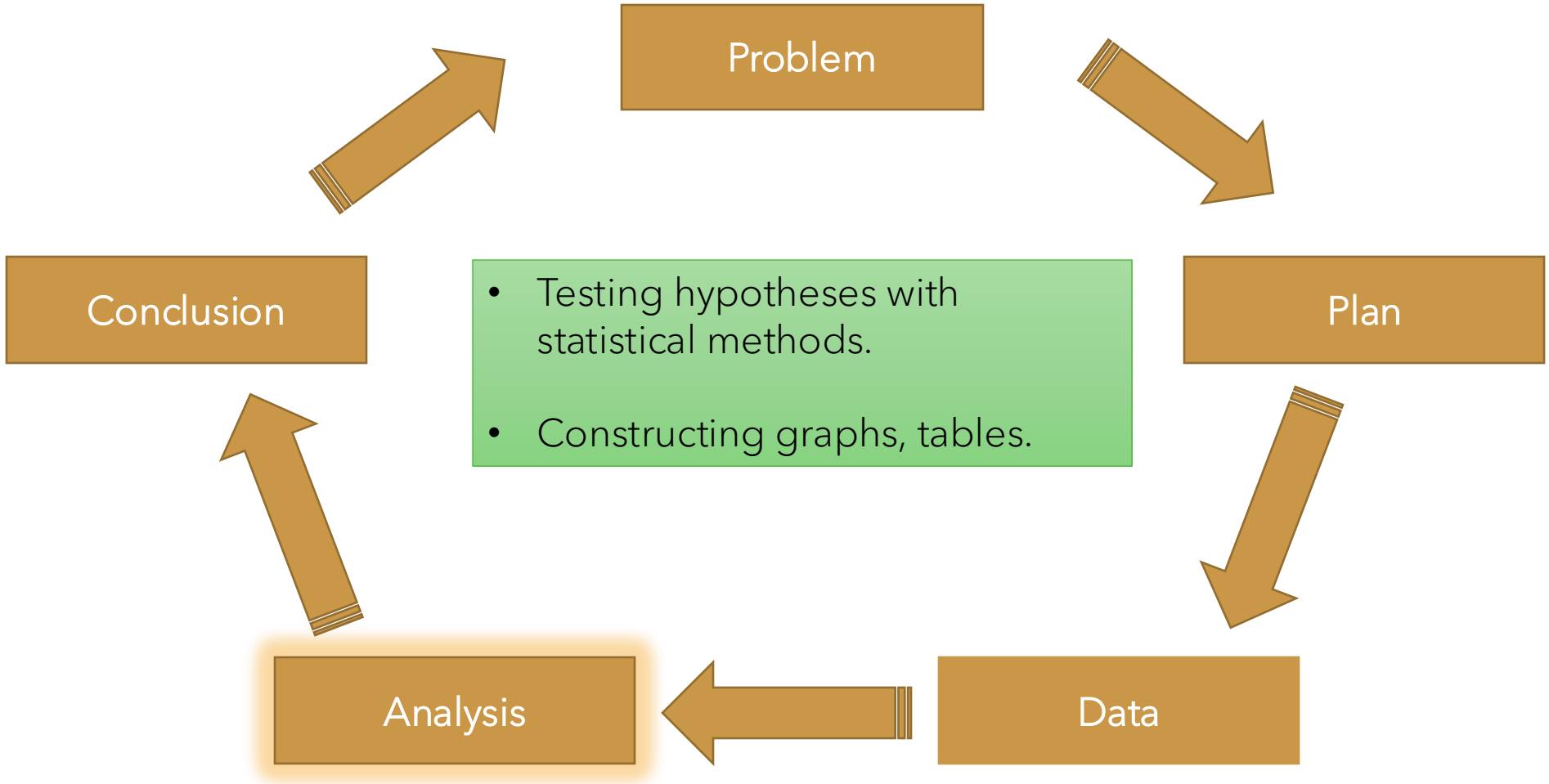
The PPDAC Cycle



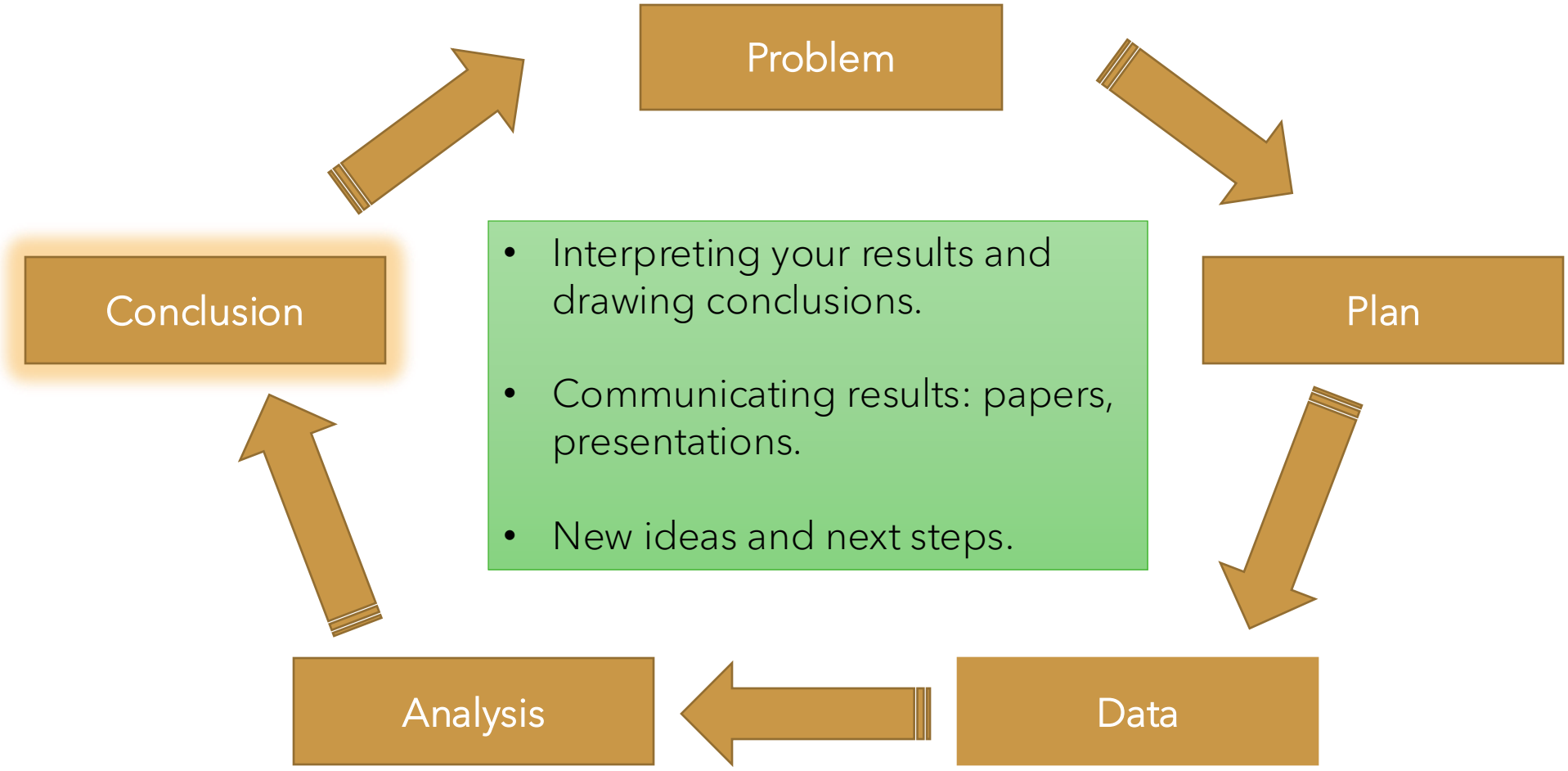
The PPDAC Cycle



The PPDAC Cycle



The PPDAC Cycle



Designing Social Inquiry (KKV) Chapter 1

- Core Question: What constitutes good social science?

  It should: *"arrive at valid inferences by the systematic use of well-established procedures of inquiry"*

- It should be scientific:

  - Inferential
  - Falsifiable
  - Uncertain
  - Replicable
  - Rules-Based
  - Public

What is “good social science”?

• Inferential

Good research uses limited observations to infer something broader and more general.

An obvious example: opinion polls only survey a tiny subset of the population, but through good analysis, they can infer knowledge about the overall population.

In other words, this is the ability to generalise your findings from a small data set to the wider world.

• Replicable

Good research could be carried out again by another researcher, following all the same steps, and they would get the same results.

This means all your research steps have to be properly and clearly described, and your inferences have to be robust – you shouldn’t make broad claims when in fact your results are only true for one specific data set!

What is "good social science"?

- Falsifiable

It must be possible to prove a research theory wrong. If no evidence can possibly exist to disprove an idea, then that idea is not scientific.

Religious claims like "all human beings have souls" are non-falsifiable.

Slippery claims like "most people secretly believe X, but they'll never admit it to you" are designed to be non-falsifiable.

- Rules-Based

Good research must follow the basic rules of inference – i.e. logic.

You can invent your own methodology – in fact, you're encouraged to! – but the entire process of the research must follow clear logical rules that others can follow.

Moreover, that logic must be explained – show your reason for taking each step in the research and analysis.

Good Social Science must be Uncertain

Good research recognises that it is uncertain and tries to be clear about the sources of uncertainty and the likelihood of being wrong.

This doesn't mean being humble – it means clearly acknowledging factors you couldn't measure or alternative hypotheses you couldn't check.

In quantitative research, it means fully reporting and interpreting probabilities.

- The key reason that scientific research must acknowledge uncertainty is down to complexity.
- One human being is very complex. A social system made up of a lot of human beings is exponentially complex.
- We can nonetheless make inferences about how those complex systems work – but we will never be able to control for every variable.
- The only certainty is uncertainty.

Good Social Science must be Public

This connects to the idea of being Replicable – everything about how you conduct your research must be public. You cannot have a secret method for doing some kind of analysis or collecting some kind of data.

No appeals to authority - "just trust me bro" is not a research methodology.

All scientific research must be open to criticism and challenge, no matter who does the research or what it's about.

- Scientific journals (including political science & economics) now often require that you publish both the **code** (in R, Python etc.) and the **data** that you used for statistical analysis.
- Some journals also encourage you to submit your **research plan** before you do the research – proving that you followed the plan properly.
- This is part of a broad push to make science more **open source**.

Research Types

Experiments and Observations

The Fundamental Problem of Causal Inference

- Generally speaking, the aim of research is to find out whether A caused B.
  - Did taking drug A cause side effect B?
  - Does economic condition A lead to election result B?
  - Does exposure to media A lead to change in social attitude B?
  - Etc.
- We call A the "treatment" and B the "outcome".

- Research is conducted by finding cases where the treatment variable is different and measuring change in the corresponding outcome variable.
- The Problem:

  For each case, only one treatment can be given, so only one outcome can be observed.

  We never get to see what was down the "road not taken" – what would really have happened in that case if we gave a different treatment.

  This is called the counterfactual.

Counterfactuals

Country A and Country B are next to each other. Country A adopts economic policy X, Country B does not. Later, unemployment in Country A drops; in Country B, it rises.

The simple conclusion is that policy X caused a drop in unemployment.

But we cannot measure what would have happened to unemployment in Country A if policy X had never been adopted – or what would have happened in Country B if policy X had been adopted.

- There could be many factors – pretreatment variables, or confounders – outside of policy X that influence unemployment rates in each country. We can never observe and control for all of those variables – the world is too complex.
- Moreover, the decision to adopt policy X was not random – Country A selected this policy, i.e. self-selected the treatment. This is selection bias.

The research "ideal"?
Randomised Controlled Trials

- One way to overcome these problems is to conduct an experiment called a Randomised Controlled Trial, or RCT.
- Researchers choose completely at random which study participants will receive the treatment. Some, called the control group, receive no treatment (or a placebo).
- By randomising the treatments, the pretreatment variables associated with each case should average out - so the only systematic difference is the treatment.

- RCTs have high internal validity – the causal effects are clear and easy to calculate.
- They're considered the "gold standard" in medical research.
- However, their external validity – i.e. how reliable they are for inferring a broader truth – can be low.
- This is because participants are usually self-selected – and may not be representative of the general public.

RCTs in Social Science

- In recent years there's been a surge of interest in doing RCT-style experiments in social science.
- Instead of a drug, the "treatment" in a social science RCT might be exposure to a certain kind of media, or a certain set of facts; or it might be implementing a certain policy only in certain villages and not in others.
- The simplest social science RCTs can be done in online surveys; the more complex require significant government support.

- It's important to consider the ethics of conducting a social science RCT.
- If the treatment you expose a participant to does change their attitudes, is that an ethical thing to do?
- How about if implementing a policy in some villages changes their electoral behaviour?
- Experiments require permission from the university's Institutional Review Board (IRB)

Observational Research

- Most social science research is observational – we observe things happening in the real world and use those measurements for our analysis.
- This means we have no control over the treatments – people, nations etc. may select their own treatments (self-selection bias), or treatments may be assigned naturally according to some systematic differences (confounding bias).

- The external validity of observational studies is great – it's real-world observational data. Robust conclusions from these studies can be generalised into inferences about the broader world.
- The challenge is improving the internal validity. Without control over treatment assignment, we must try to measure and account for as many relevant confounding variables and sources of bias as possible.

Choosing a Research Question

Assignment #1: Your Research Interests

Choosing a research question

- Most of the research process is guided by rules – there is room for innovation, but it must fall clearly within the rules of logic and inference.
- Choosing a research question is the exception – this is down to your own creativity and intuition.
- The only criteria are:
    1. Importance
    2. Contribution

- Importance is loosely defined - and subjective. You can make an argument for the importance of almost any research which explains something about human behaviour or society.
- Making a contribution means that you're adding something new to existing scientific knowledge in a field. You need to be familiar with existing knowledge about a topic before you can challenge it, or add something new to it.

Assignment: tell me about your research interest

- Your first assignment is simple: tell me what you're interested in researching.
- This can be as general or as specific as you like – maybe you already have a specific research topic in mind, or you might just have a broad idea of what area of politics, economics, or public policy you're interested in.
- You only need to write a brief paragraph – three or four sentences – explaining the area you're interested in, and what it is about that area that interests you.
- This assignment will be available on Moodle shortly after the class. Complete the assignment directly on Moodle – do not email it to me.
- Submission deadline is midnight next Tuesday (October 22nd).