

Research Methodology

MTech (CSE) – Aug-Dec 2022

CSE 607

(Lecture#3 Sept 05,2022)

Dhiren Patel

Primary vs. secondary data

- Primary data is any original information that you collect for the purposes of answering your research question (e.g. through surveys, observations and experiments).
- Secondary data is information that has already been collected by other researchers (e.g. in a government census or previous scientific studies).
- If you are exploring a novel research question, you'll probably need to collect primary data. But if you want to synthesize existing knowledge, analyze historical trends, or identify patterns on a large scale, secondary data might be a better choice.

Descriptive vs. experimental data

- In descriptive research, you collect data about your study subject without intervening. The validity of your research will depend on your **sampling method**.
- In experimental research, you systematically intervene in a process and measure the outcome. The validity of your research will depend on your **experimental design**.

KAT

- Known Answer Tests
- For given input, known output
- Standards
- E.g. to test implementation of cryptographic algorithm
- E.g. Hash function, Encryption/Decryption
- E.g. Gauges (measuring instrument) – rigorous testing and calibration before product delivery

Research methods for analyzing data

Statistical analysis	Quantitative	To analyze data collected in a statistically valid manner (e.g. from experiments, surveys, and observations).
Meta-analysis	Quantitative	To statistically analyze the results of a large collection of studies.
Thematic analysis	Qualitative	To analyze data collected from interviews, focus groups or textual sources.
Content analysis	Either	To analyze large volumes of textual or visual data collected from surveys, literature reviews, or other sources. Can be quantitative (i.e. frequencies of words) or qualitative (i.e. meanings of words).

Cause and Effect

- To conduct an experiment, you need to be able to vary your **independent variable**, precisely measure your dependent variable, and control for **confounding variables**.
- If it's practically and ethically possible, this method is the best choice for answering questions about cause and effect.

Interesting ideas (2018)

- Personalized Education
- (study what you are interested in, self paced, flexibility to change course/university/teacher – online learning – which became large scale reality during COVID lockdowns and after)
- Personalized Medicines
- (E.g. speciality drugs (customized for a person – based on medical history, body conditions, existing parameters, situational parameters))

Research (@Stanford (Oct. 2014))

- Reverse aging (Monkey and Rats)
- Organ development for emergency transplant (Obama pigs - lever)
- Requires collaborations between Medical School and School of CSE (Data Analytics/Visualization/ML)

Research methods for collecting data

Method	Quantitative(1)/ Qualitative(2)	When to Use?
Experiment	Q1	To test cause-and-effect relationships.
Survey	Q1	To understand general characteristics (e.g. of a population)
Interview/focus group	Q2	To gain more in-depth understanding of a topic
Observation	Q1 and/or Q2	To understand how something occurs in its natural setting
Literature Review	Q1 and/or Q2	To situate your research in an existing body of work, or to evaluate trends within a research topic
Case Study	Q1 and/or Q2	To gain an in-depth understanding of a specific group or context, or when you don't have the resources for a large study.

Vaccine creation efforts - March 2020

- Ref: Micholas Dean Smith and Jeremy C. Smith;
Repurposing Therapeutics for COVID-19:
Supercomputer- Based Docking to the SARS-CoV-2 Viral
Spike Protein and Viral Spike Protein-Human ACE2
Interface (March 11, 2020), Journal ChemRxiv
- Research Paper from Oak Ridge National Laboratory
(Ref: The world's fastest supercomputer identifies 77
chemicals that could stop coronavirus from spreading -
(computing speed of 200 quadrillion calculations per
second - ran simulations of over 8,000 compounds,
found 77 that could be effective))

Researching Research

- Ack: Mischa Dohler (keynote at Valencia, Spain (2008), King's College London)

R versus D versus M

■ Research:

- *... is the process of going up alleys to see if they are blind*
- infinite number of problems but only finite resources
- challenge is to say no

■ Development:

- *... is too boring for research and never sufficiently fast for marketing*
- Murphy's law loves development as the devil lies in the detail
- challenge is to deliver

■ Market:

- *... is to make people buy things they don't actually need*
- marketing has much more in common with research than with development
- challenge is to predict