# Some questions?

- Are humans rational decision makers?
- Does spanking hurt children?
- A bank hired 32 males and 61 females during an 8 year period for the same positions. The average male salary was \$818 more than the average female salary. Is their gender discrimination in the bank?
- Will Covaxin (or any other) vaccine for Covid-19 work?

# Question - 1: Are humans rational

- Questions have to be transformed into hypotheses with specific measurable outcomes and the factors that influence them.
- Consider the earlier list of questions and let us look at the first question.

  `Are humans rational?`.

  To find an answer:
  - Convert the question into a hypothesis: `'Humans are rational'`.
  - Devise a test of rationality.
  - Collect the responses from a set of experimental subjects.
  - Analyse the data to decide whether or not the hypothesis 'Humans are rational' can be maintained or has to be rejected.
  - We are not asserting that the hypothesis is true or false.

# Question - 2: coroporal punishment study

Consider the second question in the list.
`Does spanking/corporal punishment hurt children?`

- Clearly, 'hurt' refers to the psyche of the child over time and impact on behaviour and not physical pain.
- How can we measure hurt?
- In one study hurt was measured in terms of level of anti-social behaviour (ASB) as reported by mothers for children 6-9 years old. Data was collected over a short period of time for spanking and then two years later data was collected on ASB for the same children.
- Study conclusion: Spanking correlates positively with anti-social behaviour 2 years later.
- Study claimed: it controlled for economic status, gender, level of ASB at start, extent of emotional support and stimulation at home.

# Question - 3: Gender discrimination

Is there gender discrimination?

- Statistical tests (we study later) show that the data is consistent with the statement that there is gender discrimination.
- But is there gender discrimination? Cause-effect relationship?

  That is: Is the data that we see **due to** gender discrimination or there are other reasons for the skew in the data but it is also consistent with the possibility of gender discrimination.

# Possible gender discriminination: more data

1973 UC Berkeley admission data

| **Gender** | **No. of applicants** | **No. admitted** |
|---|---|---|
| Male | 8442 | 44% |
| Female | 4321 | 35% |

# More detailed data

Disaggregated data for 6 largest departments.

| | Male | | Female | |
|---|---|---|---|---|
| Dept. | No. applicants | % admitted | No. applicants | % admitted |
| 1 | 825 | 62% | 108 | 82% |
| 2 | 560 | 63% | 25 | 68% |
| 3 | 325 | 37% | 593 | 34% |
| 4 | 417 | 33% | 375 | 35% |
| 5 | 191 | 28% | 393 | 24% |
| 6 | 272 | 6% | 341 | 7% |

# Schematic of admissions

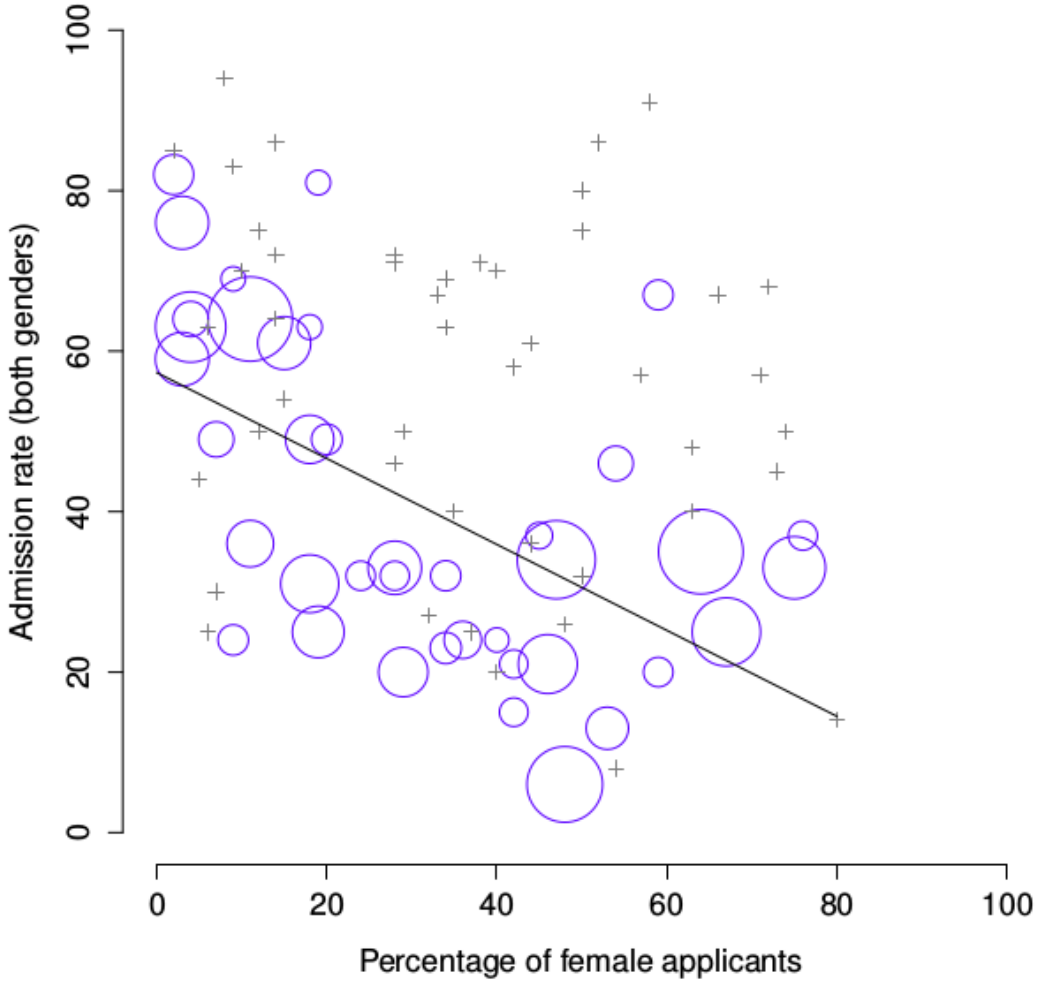


Figure: Schematic of admissions for 85 depts. with at least 1 female applicant. Circles had $> 40$ applicants, area proportional to number. $+$

# Conclusions from the previous example

- Data can be tricky and unless properly analysed the conclusions can be wrong.
- Superficial/ motivated/ ideologically coloured analysis of data is widespread. So, be very careful when accepting conclusions.
- Cause-effect relationships cannot be simply inferred from data alone. They require additional supporting evidence.
- All conclusions, theories are provisional and it is best to treat them as approximations to the 'truth' to a greater or lesser degree.

# Question - 4

- What is measured?

  Antibody concentration.
  The vaccine based approach to disease prevention pre-supposes the correctness of the current theory of the human immune system.
- There are many other preventive and curative prescriptions: e.g. hydorchloroquinine, blood plasma therapy, some homeopathic/ avurvedic remedies, practices like jal-neti, masks, lockdowns etc. Which are effective and which not?
- Most science knowledge builds new knowledge assuming the correctness of some previous knowledge. If the correctness of the assumed knowledge is in question we cannot draw reliable conclusions.

# Independent, dependent variables I

- Any study has independent or input variable(s) that the experimenter manipulates to see its effect on the output or dependent variable(s).
- The output or dependent parameter is the one that we believe depends on the input/independent parameter(s) and we are interested in studying this dependence.
- In some cases it may be just a diagnostic - that is the dependent variable is discrete and has just two values. For example, i) are humans rational?, ii) Am I infected with the Corona virus? But much more often we are interested in the existence of a systematic dependence between independent and dependent variables for theoretical and predictive reasons. In particular - is there a causal dependence and can we predict value(s) of the dependent variable if we know the values of the independent variable. **Can we build a model?**

# Independent, dependent variables II

- There are always other **known and unknown** parameters in every study that may affect the dependent variable(s) that we would like to avoid/ control - but may not be able to do so.