How to use ChatGPT to improve your research

Nikolai Czajkowski

```
M4 ← lm(Lykke~Venner+Mann+Venner:Mann, data=dt)
summary(M4)
```
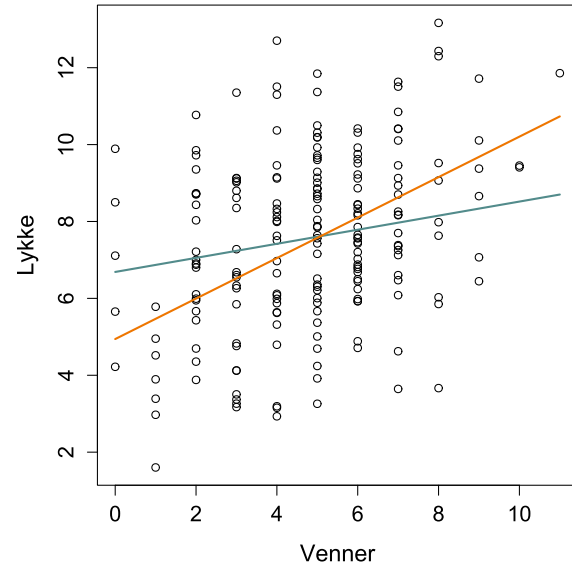
```
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.94162    0.52011   9.501  < 2e-16 ***
## Venner       0.52648    0.09633   5.465 1.39e-07 ***
## Mann         1.74562    0.72678   2.402   0.0172 *
## Venner:Mann -0.34342    0.13842  -2.481   0.0139 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.12 on 196 degrees of freedom
## Multiple R-squared:  0.1451,    Adjusted R-squared:  0.132
## F-statistic: 11.09 on 3 and 196 DF,  p-value: 9.373e-07
```


$$Lykke = b_0 + b_1 \cdot Venner_i + b_2 \cdot Mann_i + b_3 \cdot Venner_i \cdot Mann_i$$

$$Lykke_k = 4.94 + 0.53 \cdot Venner_i$$
$$Lykke_m = 4.94 + 0.53 \cdot Venner_i + 1.75 \cdot 1 + (-0.34) \cdot Venner_i \cdot 1$$
$$Lykke_m = (4.94 + 1.75) + (0.53 - 0.34) \cdot Venner_i$$

```
M4 ← lm(Happiness~Friends+Man+Friends:Man, data=dt)
summary(M4)
```

```
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.94162    0.52011   9.501  < 2e-16 ***
## Friends      0.52648    0.09633   5.465 1.39e-07 ***
## Man          1.74562    0.72678   2.402   0.0172 *
## Friends:Man -0.34342    0.13842  -2.481   0.0139 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.12 on 196 degrees of freedom
## Multiple R-squared:  0.1451,    Adjusted R-squared:  0.132
## F-statistic: 11.09 on 3 and 196 DF,  p-value: 9.373e-07
```
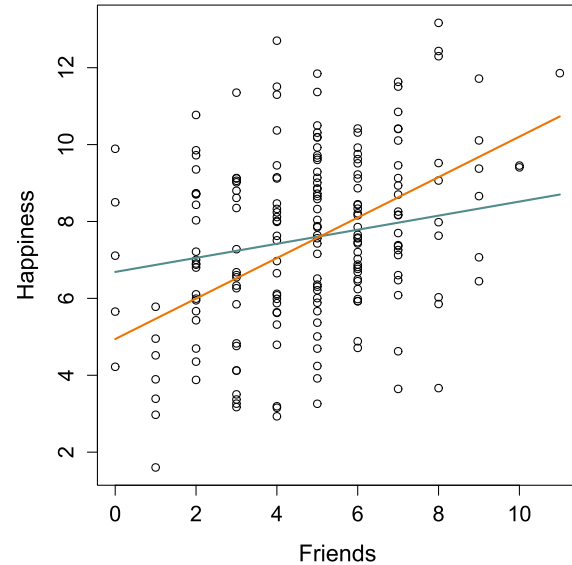

$$Happiness = b_0 + b_1 \cdot Friends_i + b_2 \cdot Man_i + b_3 \cdot Friends_i \cdot Man_i$$

$$Happiness_k = 4.94 + 0.53 \cdot Friends_i$$
$$Happiness_m = 4.94 + 0.53 \cdot Friends_i + 1.75 \cdot 1 + (-0.34) \cdot Friends_i \cdot 1$$
$$Happiness_m = (4.94 + 1.75) + (0.53 - 0.34) \cdot Friends_i$$

```
M4 ← lm(Inflation ~ GDP + Developed + GDP:Developed, data
summary(M4)
```

```
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)     2.309e+00  2.760e-01   8.365 1.11e-14 ***
## GDP             5.797e-05  2.792e-05   2.076   0.0392 *
## Developed       8.770e-01  3.949e-01   2.221   0.0275 *
## GDP:Developed  -2.523e-05  3.917e-05  -0.644   0.5202
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5643 on 196 degrees of freedom
## Multiple R-squared:  0.2665,    Adjusted R-squared:  0.2553
## F-statistic: 23.74 on 3 and 196 DF,  p-value: 3.775e-13
```


GDP's Effect on Inflation by Country Development
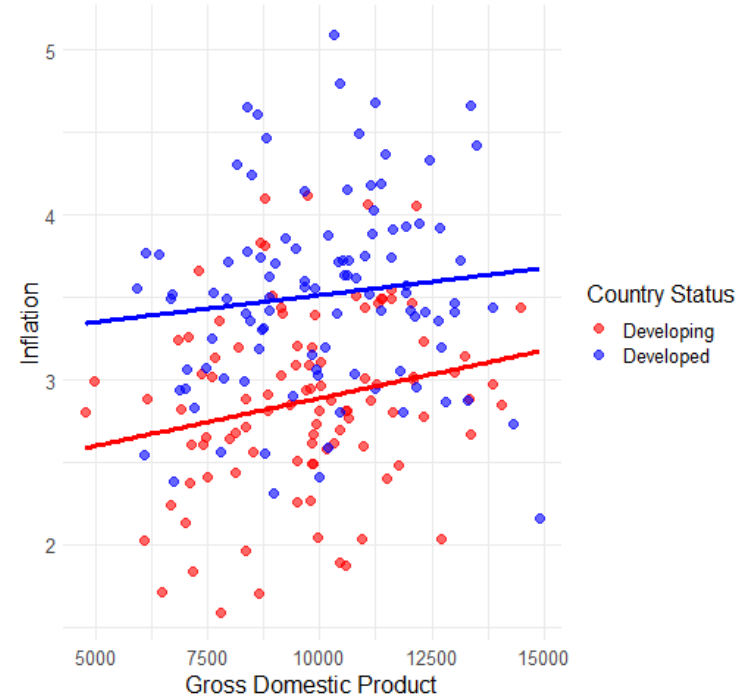
$$Inflation = b_0 + b_1 \cdot GDP_i + b_2 \cdot Developed_i + b_3 \cdot GDP_i \cdot Developed_i$$

$$Inflation_{devng} = 2.309 + 5.797 \times 10^{-5} \cdot GDP_i$$
$$Inflation_{dev} = 2.309 + 5.797 \times 10^{-5} \cdot GDP_i + 0.877 \cdot 1 + (-2.523 \times 10^{-5}) \cdot GDP_i \cdot 1$$
$$Inflation_{dev} = (2.309 + 0.877) + (5.797 \times 10^{-5} - 2.523 \times 10^{-5}) \cdot GDP_i$$

A smiling woman in a red dress on a beach with a lighthouse in the backgroud, birds flying through the sky, 85mm, cinematic lighting
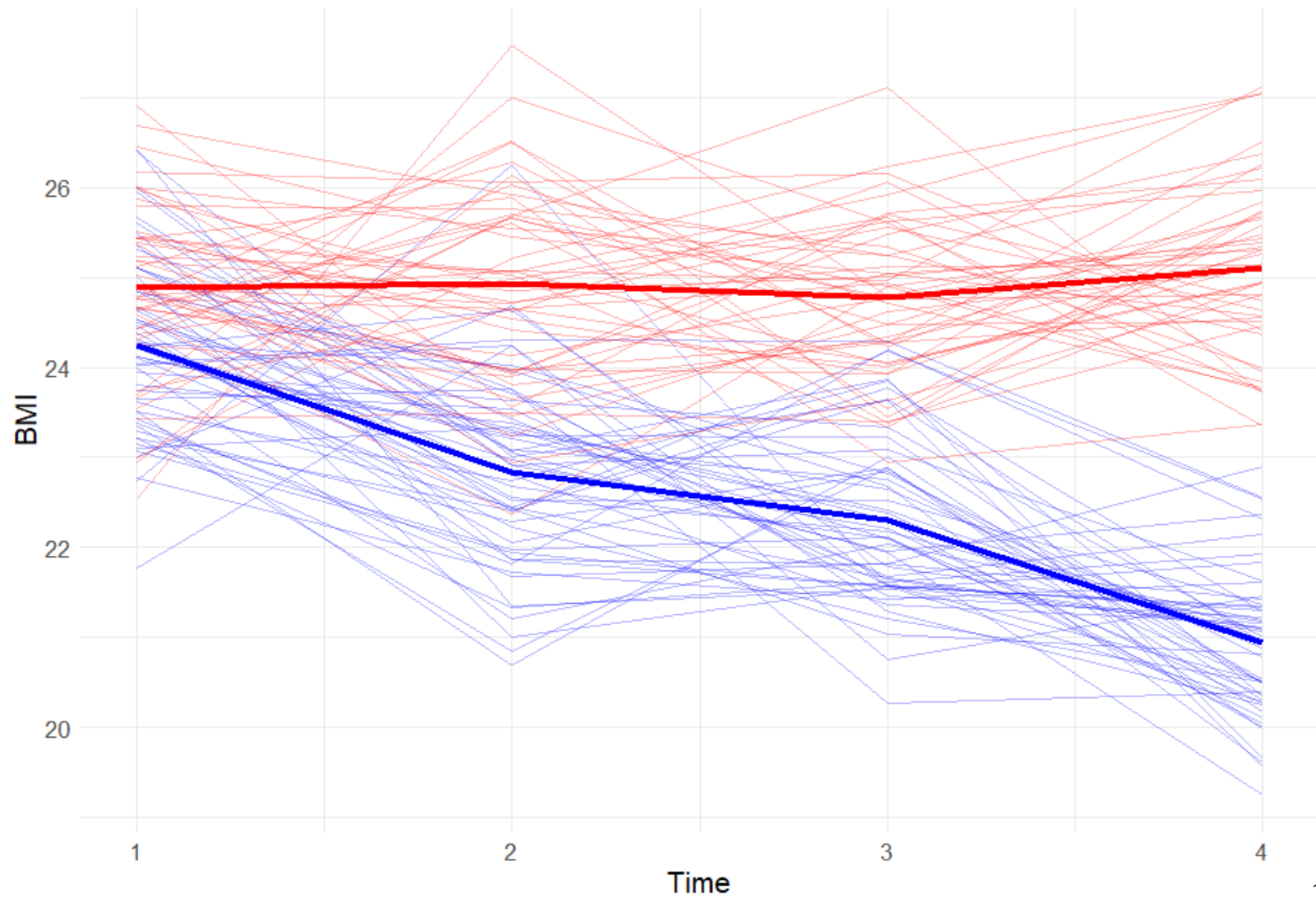
RESEARCH SUB-COMPONENTS?

1. Topic Selection
2. Literature Review
3. Formulating Research Questions
4. Choosing Research Design
5. Data Collection
6. Data Analysis
7. Interpreting Results
8. Writing and Publishing
9. Peer Review
10. Dissemination

https://chat.openai.com/share/6ec1456f-4f50-4ab6-8aac-c9b55061d789

- ChatGPT helpful for a broad set of analyses.
  - Might require some iterations, but we almost always get a working code.
  - Simulations useful for understanding and validation.

- Spectacularly more effective than google search.

- Modifying a prototype is far more effective than building from scratch.

- No longer locked in to a programming platform, language or library/package.

https://chat.openai.com/share/382e691b-f934-4557-a566-b8c5ffb0a10b
- Litterary review; Extracting and summarizing information from papers.
- Discussing content of specific paper

- Brainstorming course and lecture plans.
- Draft and improve lecture slides for target audience.
- Make exercises and interactive tutorials based on previous examples.
- Make individualized exams and exercises.
- Generate exam solutions.
- Make quizzes and indivualized practice exercises.

https://www.sv.uio.no/psi/personer/vit/nikolaic/div/sonia_nikolai_timeline1.html?vrtx=source

- Confounding occurs when an extraneous variable correlates with both the independent and dependent variables.
- It can lead to spurious associations and misinterpretation of causal relationships.
- In public health research, confounding can be particularly problematic due to the complex interplay of various health determinants.
- For example, consider a study investigating the effect of physical activity (PA) on cardiovascular disease (CVD) in Norway. Age might be a confounder, as it is associated with both PA and CVD.
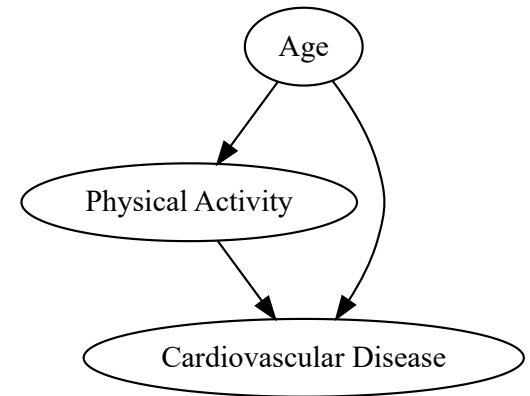
- Hallucinations
- Plagiarism
- Biases
- Data Privacy
- Lack of Understanding
- Reproducibility Concerns
- AI dependency and laziness

Immediate benefits
- More time to higher-level tasks, not tied to a specific platform
- Always a willing discussion partner on any part of the research process

Some novel challenges
- Takes time to adapt to a new way of interacting with computers (Unstructured, approximate, incomplete specifications)
- Not always obvious where it can assist, or how
- Many of our skills suddenly irrelevant

- ChatGPT has profoundly changed my workflow as a researcher.
- I believe it will be impossible to stay competitive without embracing these tools.