BASIC STATISTICS: FROM ZERO TO ANOVA

Dr.Jacopo Lanzone
Centro Regionale Epilessia
Neurologia - Servizio di Neurofisiologia Clinica
Direttore: Professor M. Filippi
IRCCS Ospedale San Raffaele

,

LICE
Lega Italiana Contro l'Epilessia

ILAE YOUNG EPILEPSY SECTION

Crea un'immagine per la mia presentazione

Ringraziamento al mio selezionatore

Ideazione di giornata rilassante

Scelta di un outfit fotogenico

Hai raggiunto il limite del piano Free per GPT-4o.
Le nuove risposte useranno un altro modello fino a quando il limite non verrà reimpostato dopo le ore 14:28.

Passa a Plus

Ciao devo effettua

STATISTICAL THINKING

Performing the task is easy, and in the future the interaction with software will be mostly AI mediated.

This shifts the focus from a “HOW” problem to a “WHY” problem.

Are we asking the right questions?
Are we envisioning data correctly ?
Are we thinking rationally?

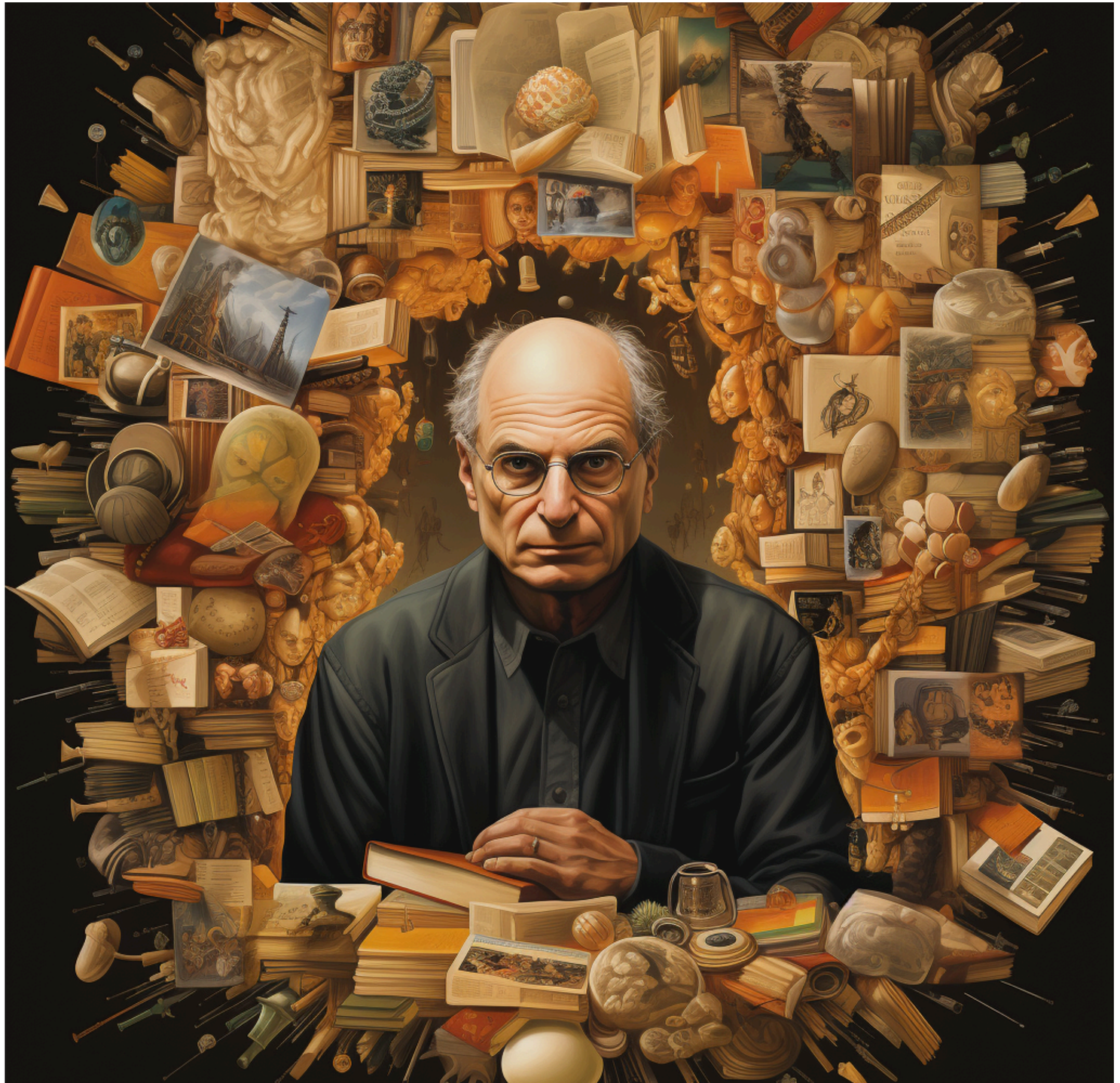
And mostly: what does this even mean/imply?

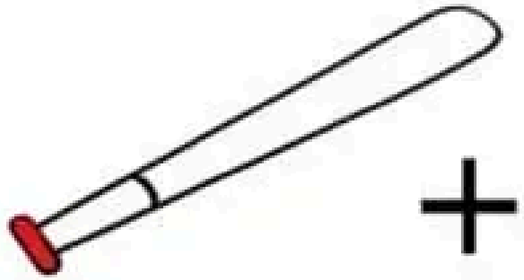
THINKING, FAST AND SLOW

DANIEL KAHNEMAN

WINNER OF THE NOBEL PRIZE IN ECONOMICS

"[A] masterpiece . . . This is one of the greatest and most engaging collections of insights into the human mind I have read." —WILLIAM EASTERLY, *Financial Times*



If  +  = $1.10

and the bat cost $1.00 more than the ball.

How much does the ball cost?

System 1

Fast, intuitive and emotional

System 2

Slow, conscious and effortful

A Higher Caseload Translates to Greater Reliance on System 1

SYSTEM 1

- Fast/automatic/easy
- Performs familiar or practiced routines
- Fine for small talk
- Undemanding
- Can perform while tired, sick or stressed
- Impressions/intuitions/feelings
- Susceptible to errors

SYSTEM 2

- Slow/effortful/hard
- Necessary for novel decisions or routines
- Useful for harder questions
- Tiring/draining
- Impaired by fatigue, illness or stress
- Logic/analysis/reflection
- Can override errors through careful thought

Memory

Social

Learning

Belief

Money

Politics

Fundamental Attribution Error

We judge others on their personality or fundamental character, but we judge ourselves on the situation.

Sally is late to class; she's lazy. You're late to class; it was a bad morning.

Self-Serving Bias

Our failures are situational, but our successes are our responsibility.

You won that award due to hard work rather than help or luck. Meanwhile, you failed a test because you hadn't gotten enough sleep.

In-Group Favoritism

We favor people who are in our in-group as opposed to an out-group.

Francis is in your church, so you like Francis more than Sally.

Bandwagon Effect

Ideas, fads, and beliefs grow as more people adopt them.

Sally believes fidget spinners help her children. Francis does, too.

Groupthink

Due to a desire for conformity and harmony in the group, we make irrational decisions, often to minimize conflict.

Sally wants to go get ice cream. Francis wants to shop for T-shirts. You suggest getting T-shirts with pictures of ice cream on them.

Halo Effect

If you see a person as having a positive trait, that positive impression will spill over into their other traits. (This also works for negative traits.)

"Taylor could never be mean; she's so cute!"

Moral Luck

Better moral standing happens due to a positive outcome; worse moral standing happens due to a negative outcome.

"X culture won X war because they were morally superior to the losers."

False Consensus

We believe more people agree with us than is actually the case.

"Everybody thinks that!"

Curse of Knowledge

Once we know something, we assume everyone else knows it, too.

Alice is a teacher and struggles to understand the perspective of her new students.

Spotlight Effect

We overestimate how much people are paying attention to our behavior and appearance.

Sally is worried everyone's going to notice how lame her ice cream T-shirt is.

Availability Heuristic

We rely on immediate examples that come to mind while making judgments.

When trying to decide on which store to visit, you choose the one you most recently saw an ad for.

Defensive Attribution

As a witness who secretly fears being vulnerable to a serious mishap, we will blame the victim less and attacker more if we relate to the victim.

Sally sat too long at a green light because she was playing with her phone. She got rear-ended. Greg, who is known to text and drive, got out and yelled at the person who smacked into her.

Just-World Hypothesis

We tend to believe the world is just; therefore, we assume acts of injustice are deserved.

"Sally's purse was stolen because she was mean to Francis about their T-shirt and had bad karma."

Naïve Realism

We believe that we observe objective reality and that other people are irrational, uninformed, or biased.

"I see the world as it really is — other people are dumb."

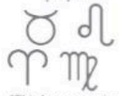
Naïve Cynicism

We believe that we observe objective reality and that other people have a higher egocentric bias than they actually do in their intentions/actions.

"The only reason this person is doing something nice is to get something out of me."

Forer Effect (aka Barnum Effect)

We easily attribute our personalities to vague statements, even if they can apply to a wide range of people.



"This horoscope is so accurate!"

Dunning-Kruger Effect

The less you know, the more confident you are. The more you know, the less confident you are.

Francis confidently assures the group that there's no kelp in ice cream. They do not work in the dairy industry.

Anchoring

We rely heavily on the first piece of information introduced when making decisions.

"That's 50% off? It must be a great deal."

Automation Bias

We rely on automated systems, sometimes trusting too much in the automated correction of actually correct decisions.

Your phone auto-corrects "its" to "it's," so you assume it's right.

Google Effect (aka Digital Amnesia)

We tend to forget information that's easily looked up in search engines.

"What was the name of that actor in that funny movie? I've looked it up like eight times…."

Reactance

We do the opposite of what we're told, especially when we perceive threats to personal freedoms.

One of Alice's students refuses to do his homework, even though both she and his parents tell him to.

Confirmation Bias

We tend to find and remember information that confirms our perceptions.

You can confirm a conspiracy theory based on scant evidence while ignoring contrary evidence.

Backfire Effect

Disproving evidence sometimes has the unwarranted effect of confirming our beliefs.

The evidence that disproves your conspiracy theory was probably faked by the government.

Third-Person Effect

We believe that others are more affected by mass media consumption than we ourselves are.

"You've clearly been brainwashed by the media!"

Belief Bias

We judge an argument's strength not by how strongly it supports the conclusion but how plausible the conclusion is in our own minds.

Sally mentions her supporting theory about your conspiracy theory, which you adopt wholeheartedly despite the fact that she has very little evidence for it.

Availability Cascade

Tied to our need for social acceptance, collective beliefs gain more plausibility through public repetition.

A story about razor blades appearing in candy eventually led to many people no longer offering homemade treats on Halloween in America.

Declinism

We tend to romanticize the past and view the future negatively, believing that societies/institutions are by and large in decline.

"In my day, kids had more respect!"

Status Quo Bias

We tend to prefer things to stay the same; changes from the baseline are considered to be a loss.

Even though an app's terms of service invade Sally's privacy, she'd rather not switch to another app.

Sunk Cost Fallacy (aka Escalation of Commitment)

We invest more in things that have cost us something rather than altering our investments, even if we face negative outcomes.

"In for a penny, in for a pound!"

Gambler's Fallacy

We think future possibilities are affected by past events.

Alice has lost nine coin tosses in a row, so she's sure to win the next one!

Zero-Risk Bias

We prefer to reduce small risks to zero, even if we can reduce more risk overall with another option.

"You should probably buy the warranty."

Framing Effect

We often draw different conclusions from the same information depending on how it's presented.

Alice hears that her favorite candidate is "killing it" with a 45% approval rating. Sally hears that the candidate is "disappointing the country" with a 45% rating. They have wildly different interpretations of the same statistic.

Stereotyping

We adopt generalized beliefs that members of a group will have certain characteristics, despite not having information about the individual.

"That guy with the fancy mustache is a hipster. He probably has a vinyl collection."

Outgroup Homogeneity Bias

We perceive out-group members as homogeneous and our own in-groups as more diverse.

Alice is not a gamer, but she believes "all gamers are the same."

Authority Bias

We trust and are more often influenced by the opinions of authority figures.

"My teacher told me this was fine."

Placebo Effect*

If we believe a treatment will work, it often will have a small physiological effect.

Alice was given a placebo for her pain, and her pain decreased.

Survivorship Bias

We tend to focus on those things that survived a process and overlook ones that failed.

Greg tells Alice her purse business is going to be great because a successful fashion company had the same strategy. (But 10 other failed companies also had the same strategy.)

Tachypsychia

Our perceptions of time shift depending on trauma, drug use, and physical exertion.

"When the car almost hit me, time slowed down …."

Law of Triviality (aka "Bike-Shedding")

We give disproportionate weight to trivial issues, often while avoiding more complex issues.

Rather than figuring out how to help the homeless, a local city government spends a lot of time discussing putting in a bike path and bike sheds.

Zeigarnik Effect

We remember incomplete tasks more than completed ones.

Greg feels guilty for never getting anything done, until he sees all of the tasks he's checked off on his task list.

IKEA Effect

We place higher value on things we partially created ourselves.

"Don't you love this pot I spent $20 on? I painted it myself!"

Ben Franklin Effect

We like doing favors; we are more likely to do another favor for someone if we've already done a favor for them than if we had received a favor from that person.

Greg loaned Francis a pen. When Francis asked to borrow $5, Greg did it readily.

Bystander Effect*

The more other people are around, the less likely we are to help a victim.

In a crowd of students, no one called 911 when someone got hurt in a fight.

Suggestibility

We, especially children, sometimes mistake ideas suggested by a questioner for memories.

"So did you fall off the couch before or after your mom hit you?"

False Memory

We mistake imagination for real memories.

Greg is certain Sally said a really funny joke about pineapples, when that joke actually came from a TV show.

Cryptomnesia

We mistake real memories for imagination.

Greg thinks he visited a graveyard, but he's pretty sure he just had a spooky dream.

Clustering Illusion

We find patterns and "clusters" in random data.

"That cloud looks like your cat, Alice!"

Pessimism Bias

We sometimes overestimate the likelihood of bad outcomes.

"Nothing will ever get better."

Optimism Bias

We sometimes are over-optimistic about good outcomes.

"It's going to turn out great!"

Blind Spot Bias

We don't think we have bias, and we see it in others more than ourselves.

"I am not biased!"

Memory

Social

Learning

Belief

Money

Politics

Fundamental Attribution Error

We judge others on their personality or fundamental character, but we judge ourselves on the situation.

Sally is late to class; she's lazy. You're late to class; it was a bad morning.

Self-Serving Bias

Our failures are situational, but our successes are our responsibility.

You won that award due to hard work rather than help or luck. Meanwhile, you failed a test because you hadn't gotten enough sleep.

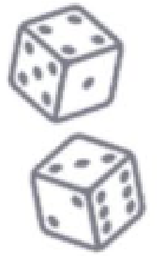
Confirmation Bias

We tend to find and remember information that confirms our perceptions.

You can confirm a conspiracy theory based on scant evidence while ignoring contrary evidence.

Gambler's Fallacy

We think future possibilities are affected by past events.



Alice has lost nine coin tosses in a row, so she's sure to win the next one!

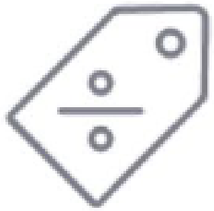
Availability Heuristic

We rely on immediate examples that come to mind while making judgments.

When trying to decide on which store to visit, you choose the one you most recently saw an ad for.

Anchoring

We rely heavily on the first piece of information introduced when making decisions.



"That's 50% off? It must be a great deal."

Placebo Effect*

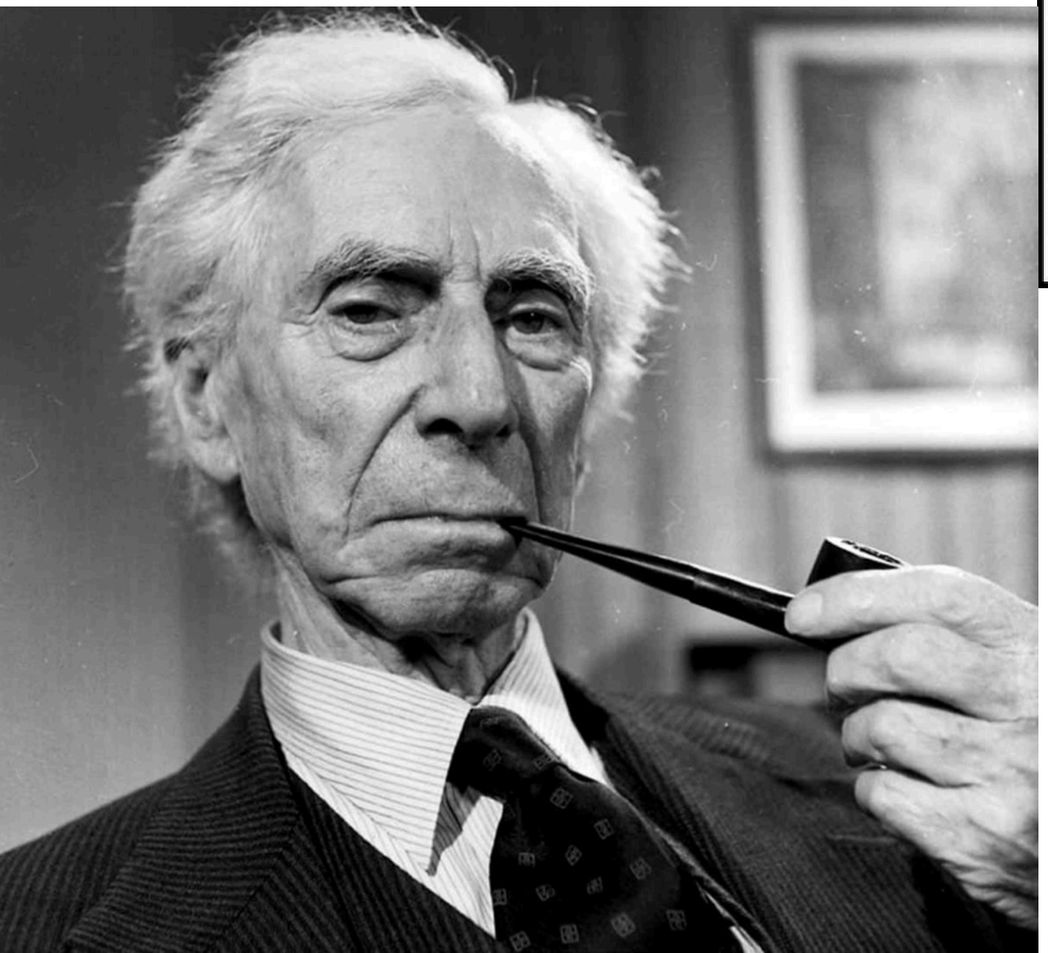
If we believe a treatment will work, it often will have a small physiological effect.

Alice was given a placebo for her pain, and her pain decreased.

Bystander Effect*

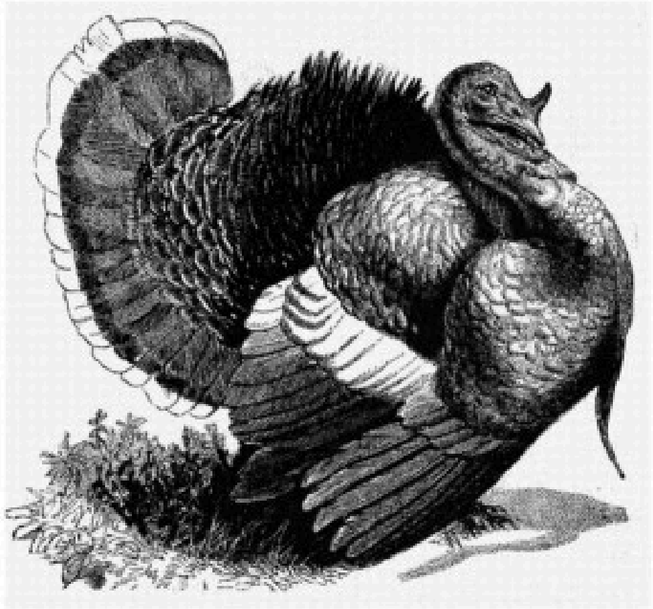
The more other people are around, the less likely we are to help a victim.

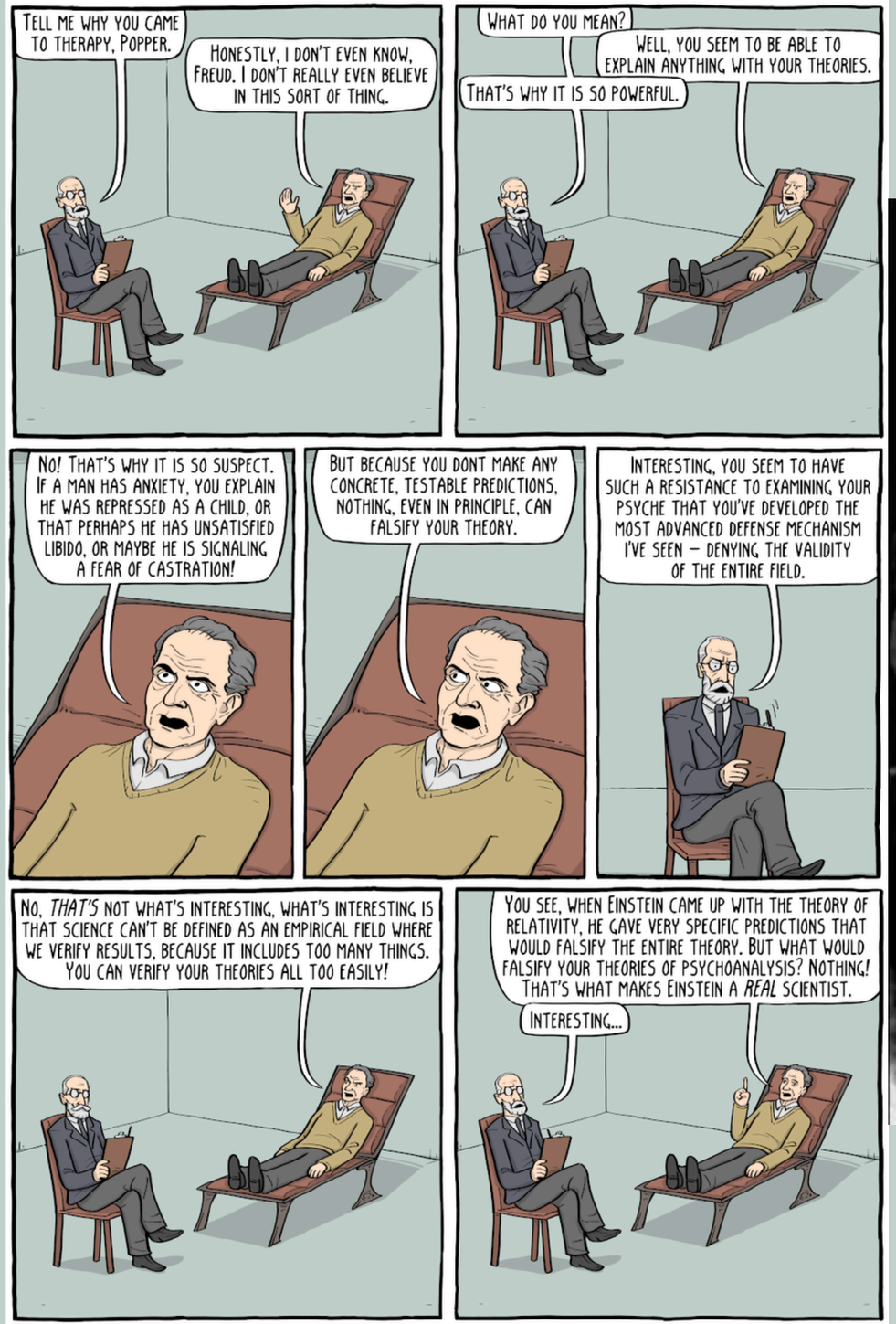
In a crowd of students, no one called 911 when someone got hurt in a fight.

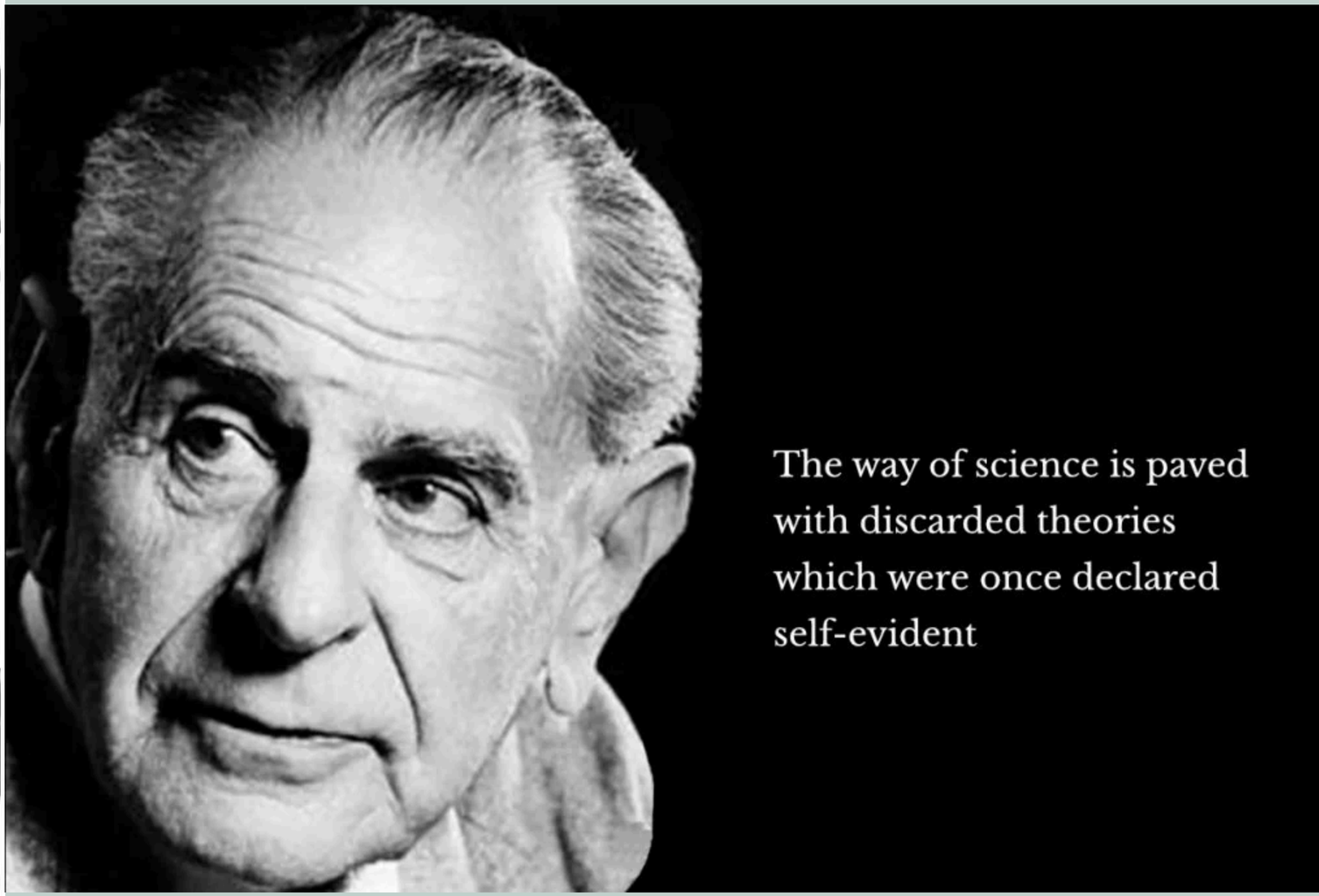


RUSSELL'S INDUCTIVIST TURKEY

The turkey found that, on his first morning at the turkey farm, he was fed at 9am. Being a good inductivist turkey, he did not jump to conclusions. He waited until he collected a large number of observations that he was fed at 9am. and made these observations under a wide range of circumstances, on Wednesdays, on Thursdays, on cold days, on warm days. Each day he added another observation statement to his list. Finally he was satisfied that he had collected a number of observation statements to inductively infer that 'I am always fed at 9am'. However on the morning of Christmas eve he was not fed but instead had his throat cut, in order to be cooked and served at the dining table.

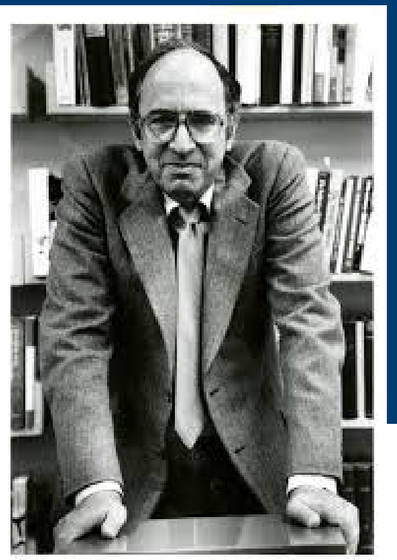


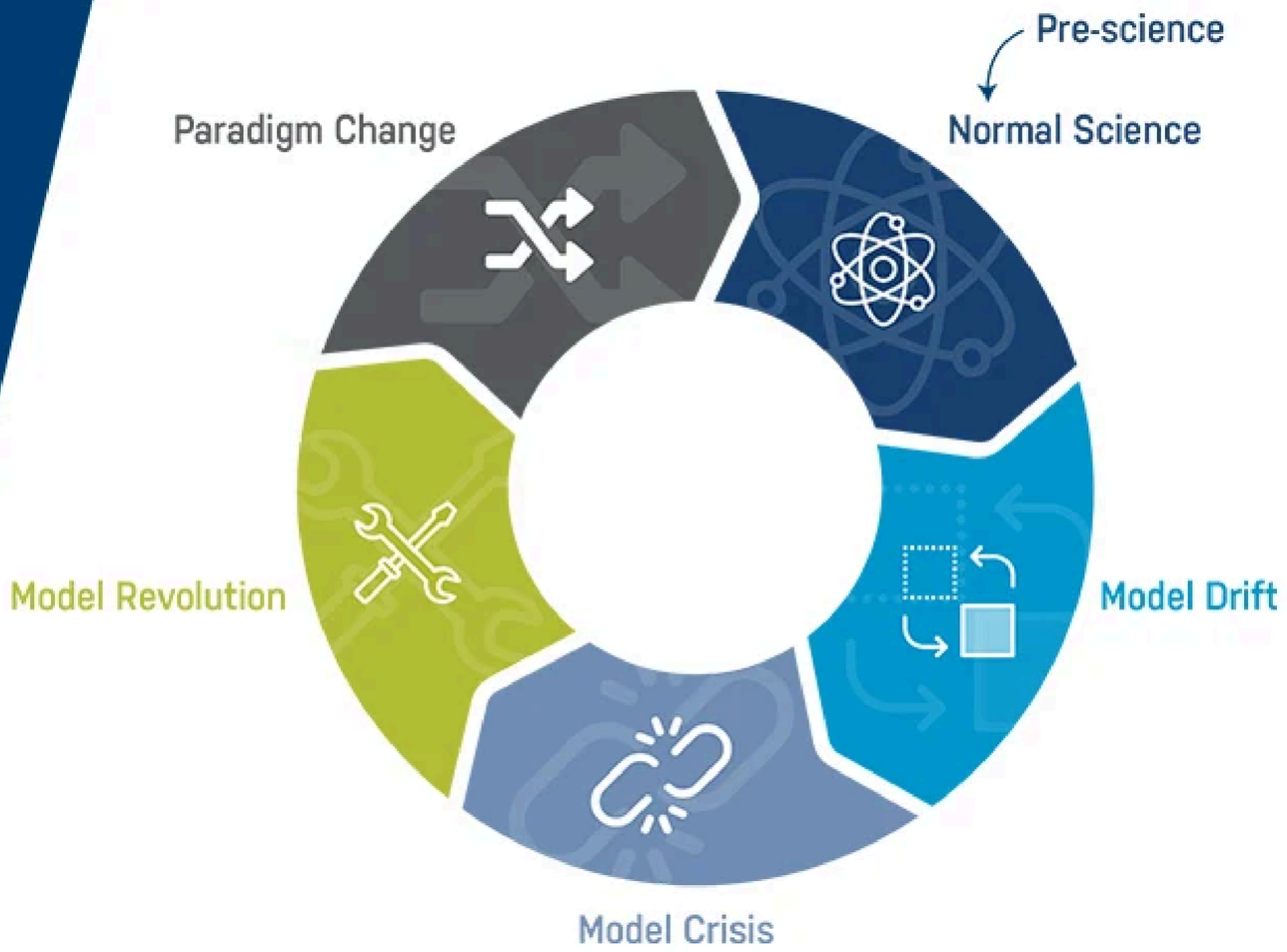

TELL ME WHY YOU CAME TO THERAPY, POPPER.
HONESTLY, I DON'T EVEN KNOW, FREUD. I DON'T REALLY EVEN BELIEVE IN THIS SORT OF THING.
WHAT DO YOU MEAN?
WELL, YOU SEEM TO BE ABLE TO EXPLAIN ANYTHING WITH YOUR THEORIES.
THAT'S WHY IT IS SO POWERFUL.
NO! THAT'S WHY IT IS SO SUSPECT. IF A MAN HAS ANXIETY, YOU EXPLAIN HE WAS REPRESSED AS A CHILD, OR THAT PERHAPS HE HAS UNSATISFIED LIBIDO, OR MAYBE HE IS SIGNALING A FEAR OF CASTRATION!
BUT BECAUSE YOU DONT MAKE ANY CONCRETE, TESTABLE PREDICTIONS, NOTHING, EVEN IN PRINCIPLE, CAN FALSIFY YOUR THEORY.
INTERESTING, YOU SEEM TO HAVE SUCH A RESISTANCE TO EXAMINING YOUR PSYCHE THAT YOU'VE DEVELOPED THE MOST ADVANCED DEFENSE MECHANISM I'VE SEEN — DENYING THE VALIDITY OF THE ENTIRE FIELD.
NO, *THAT'S* NOT WHAT'S INTERESTING, WHAT'S INTERESTING IS THAT SCIENCE CAN'T BE DEFINED AS AN EMPIRICAL FIELD WHERE WE VERIFY RESULTS, BECAUSE IT INCLUDES TOO MANY THINGS. YOU CAN VERIFY YOUR THEORIES ALL TOO EASILY!
YOU SEE, WHEN EINSTEIN CAME UP WITH THE THEORY OF RELATIVITY, HE GAVE VERY SPECIFIC PREDICTIONS THAT WOULD FALSIFY THE ENTIRE THEORY. BUT WHAT WOULD FALSIFY YOUR THEORIES OF PSYCHOANALYSIS? NOTHING! THAT'S WHAT MAKES EINSTEIN A *REAL* SCIENTIST.
INTERESTING...


The way of science is paved with discarded theories which were once declared self-evident

Karl Popper

The Kuhn Cycle





Kuhn[1] and Lakatos[2]

Immature Science

Immature Science: No prevailing school of thought, Various disparate theories, Competition

Revolution

Old Theory: well established, many followers, politically powerful, well understood, many anomalies
New Theory: few followers, untested, new concepts/techniques, accounts for anomalies, asks new questions

Normal Science

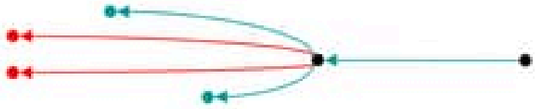
Normal Science: Stability, Determination of significant facts, Matching facts with theories, Articulation of theories (refinement and extension), "puzzle -solving" neither tests nor confirms its theories.
Driven by a paradigm: Commonly held set of beliefs, procedures, techniques. Agreement upon questions of import, upon what counts as a solution, and upon standards of evaluation. Hard core assumptions distinguished from auxiliary hypotheses.

Crisis

Crisis: Weight of accumulated anomalies, No agreement on how anomalies are to be dealt with, doubts arise. Hard core assumptions challenged.

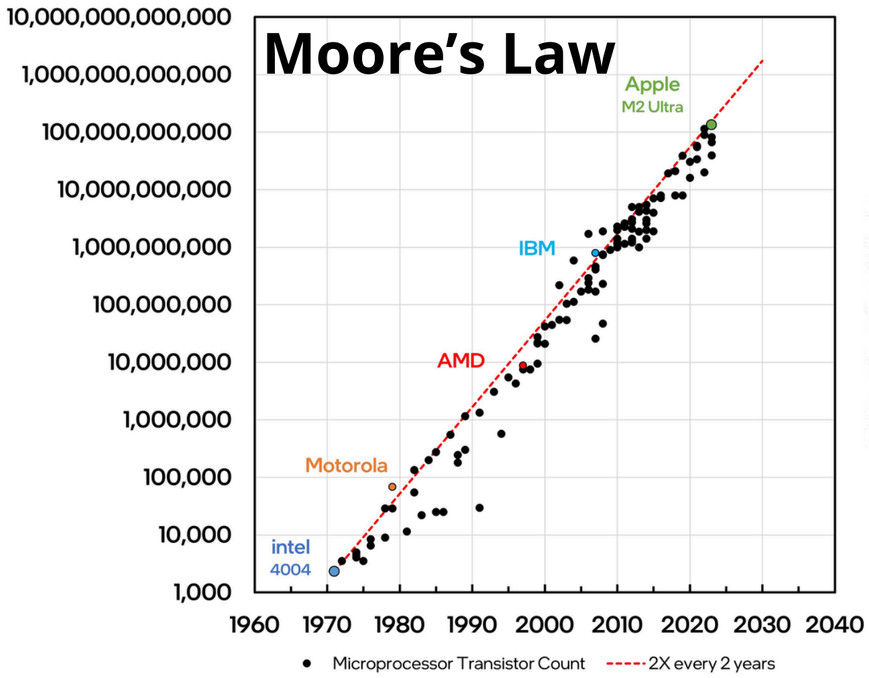
Anomalies

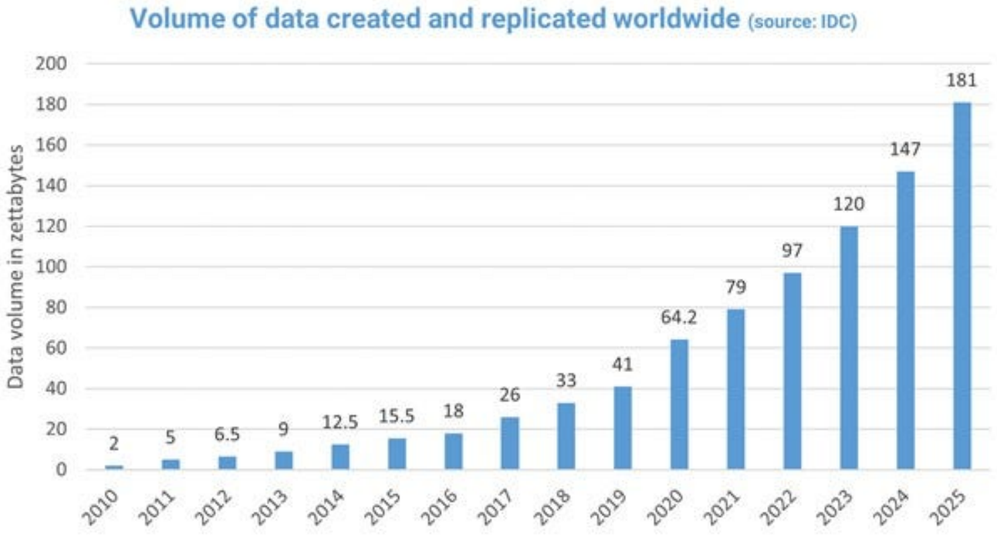
Anomalies: Not all expectations are borne out, Some anomalies lead to further discoveries, Some simply ignored.
Troublesome anomalies: Challenge key theoretical concepts, Resist solutions, Inhibit application of theory



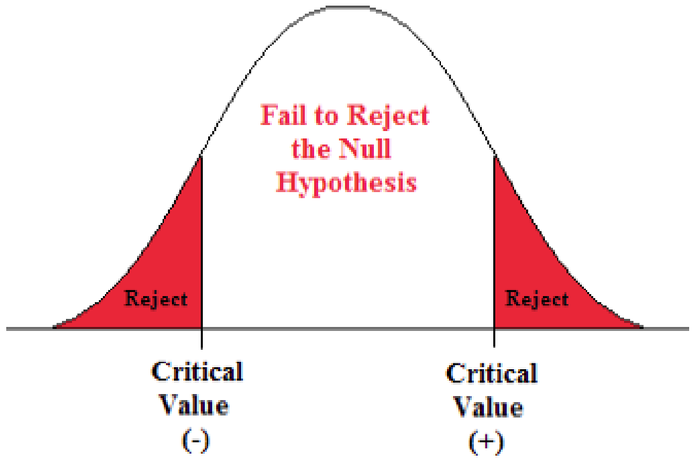
Alternative research programs: in addressing anomalies, some programs are generative of new facts, and some degenerative (i.e. post-rationalising others' facts but not generating new ones). Determining which is which takes time…

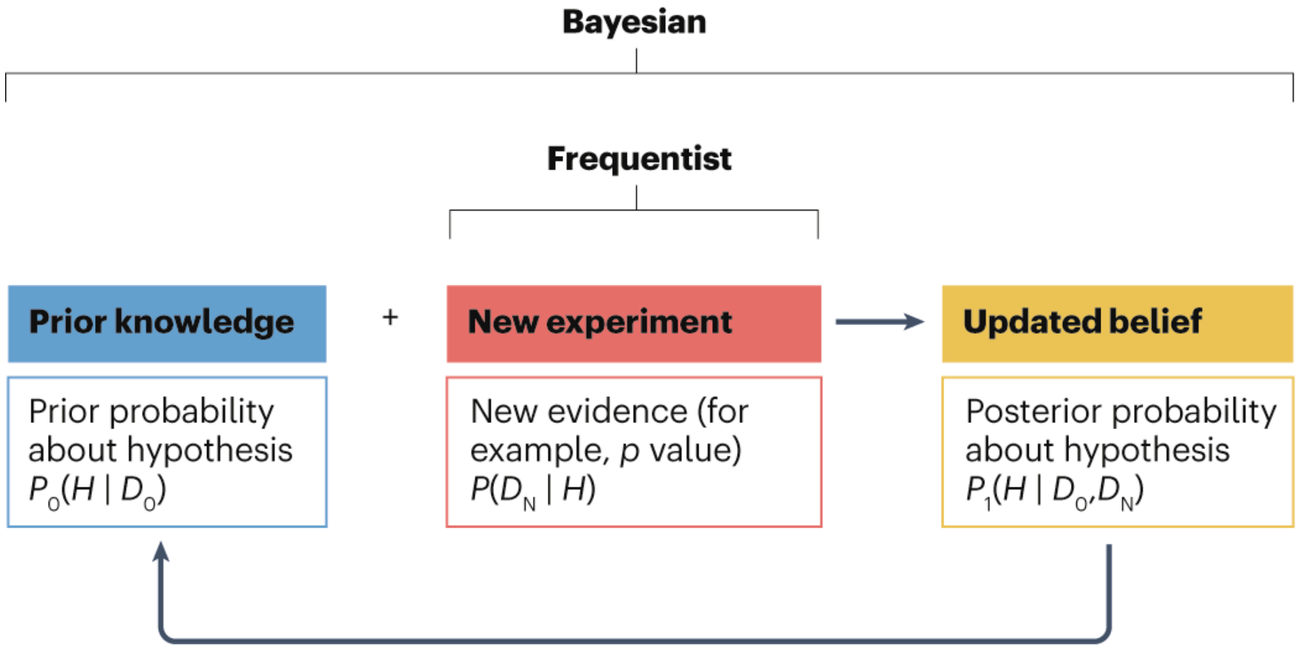
1 - Kuhn, T. S. (1962). The Structure of Scientific Revolutions, University of Chicago.
2 - Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. Criticism and the Growth of Knowledge. I. Lakatos and A. Musgrave. Cambridge, Cambridge University Press: 91-196.

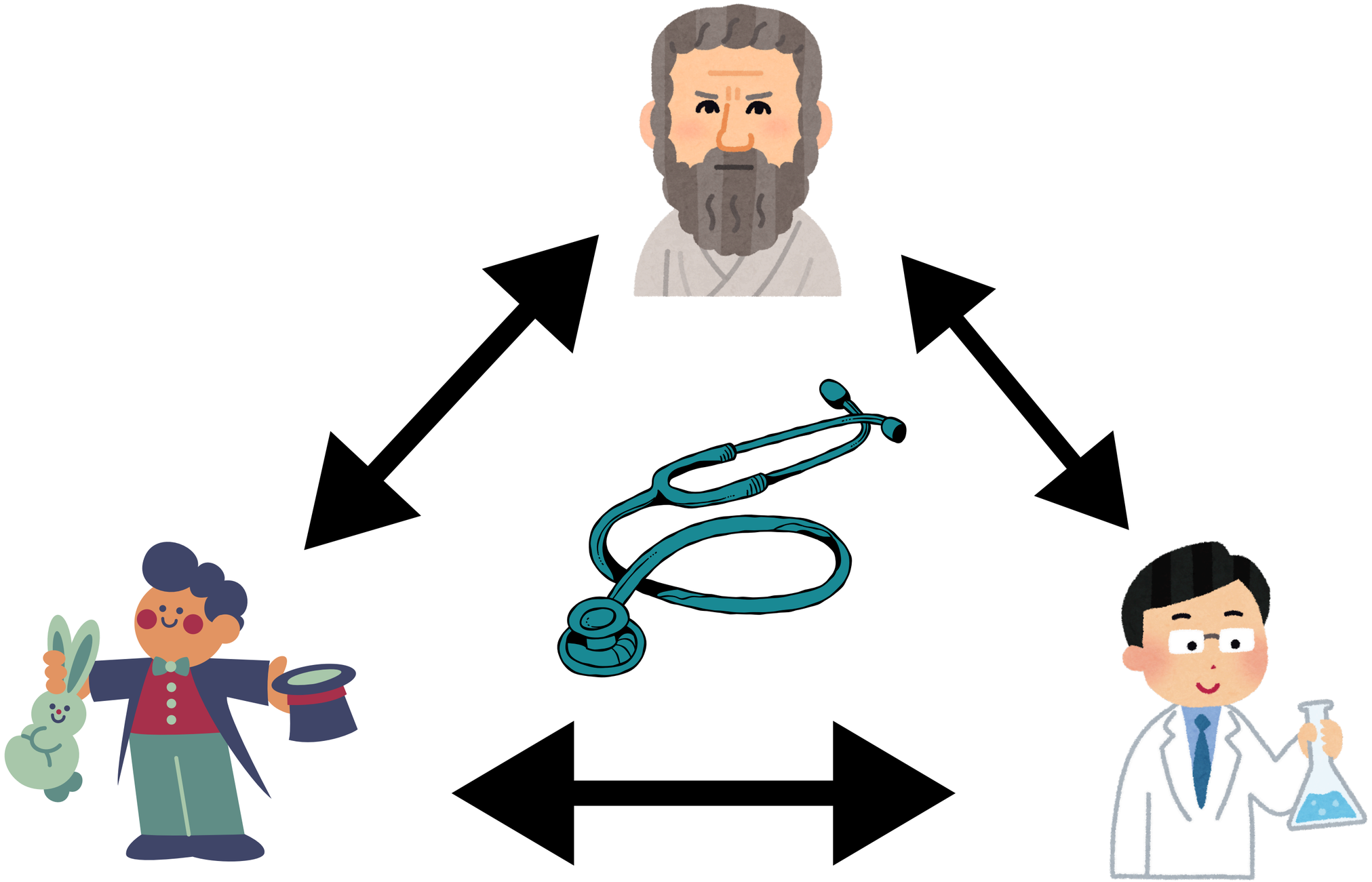

Moore's Law
Apple
M2 Ultra
IBM
AMD
Motorola
intel
4004
10,000,000,000,000
1,000,000,000,000
100,000,000,000
10,000,000,000
1,000,000,000
100,000,000
10,000,000
1,000,000
100,000
10,000
1,000
1960 1970 1980 1990 2000 2010 2020 2030 2040
Microprocessor Transistor Count
2X every 2 years


Volume of data created and replicated worldwide (source: IDC)
Data volume in zettabytes
200
180
160
140
120
100
80
60
40
20
0
2 5 6.5 9 12.5 15.5 18 26 33 41 64.2 79 97 120 147 181
2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025

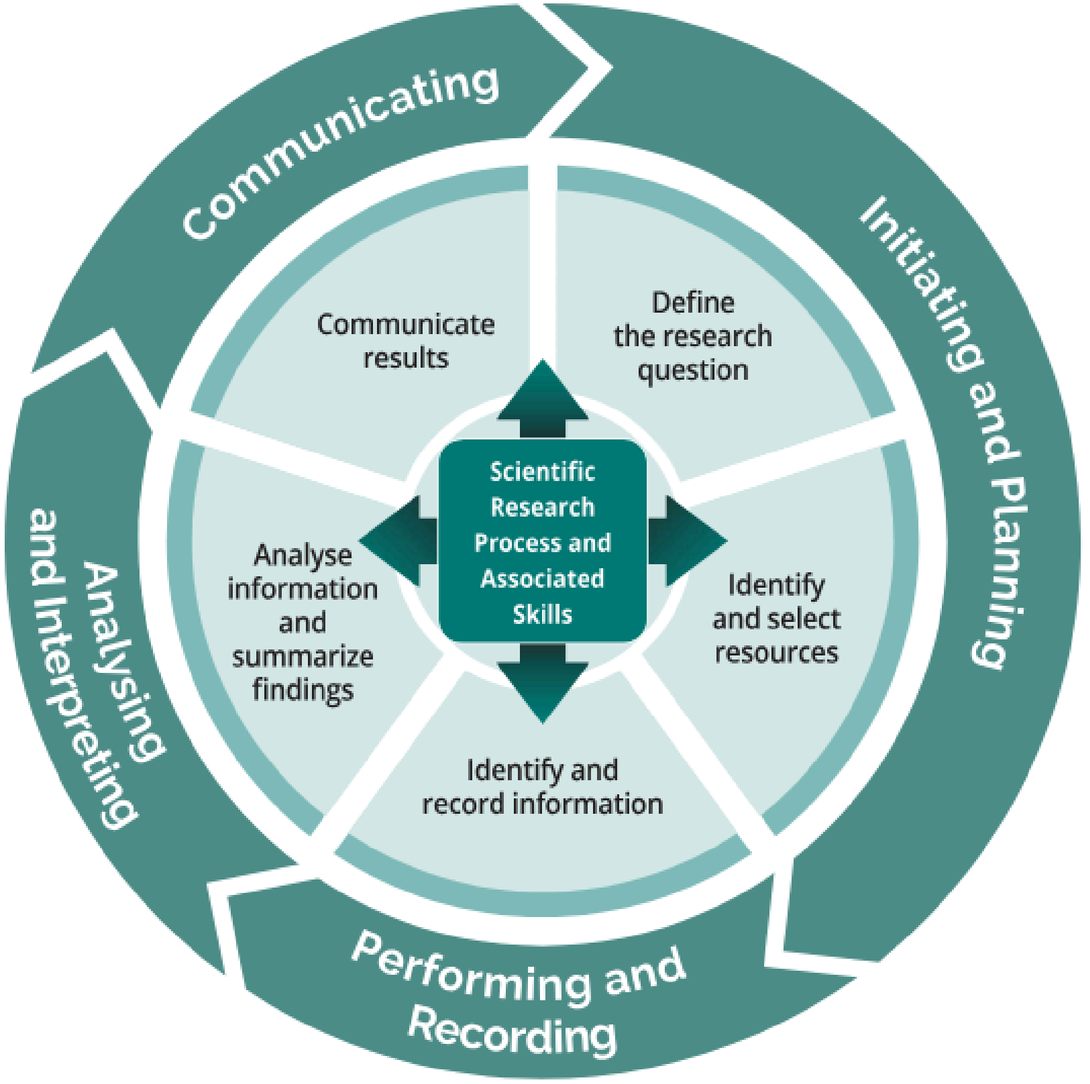
ChatGPT


Fail to Reject the Null Hypothesis
Reject
Reject
Critical Value (-)
Critical Value (+)


Bayesian
Frequentist
Prior knowledge
+
New experiment
Updated belief
Prior probability about hypothesis $P_0(H \mid D_0)$
New evidence (for example, p value) $P(D_N \mid H)$
Posterior probability about hypothesis $P_1(H \mid D_0, D_N)$



Science	Pseudoscience
Willingness to change with new evidence	Fixed ideas
Ruthless peer review	No peer review
Takes account of all new discoveries	Selects only favourable discoveries
Invites criticism	Sees criticism as conspiracy
Verifiable results	Non-repeatable results
Limits claims of usefulness	Claims of widespread usefulness
Accurate measurement	“Ball-park” measurement



A Rough Guide to

TYPES OF SCIENTIFIC EVIDENCE

Being able to evaluate the evidence behind a claim is important, but scientific evidence comes in a variety of forms. Here, the different types of scientific evidence are ranked and described, particularly those relevant to health and medicinal claims.

INCREASING STRENGTH OF EVIDENCE

ANECDOTAL & EXPERT OPINIONS

Anecdotal evidence is a person's own personal experience or view, not necessarily representative of typical experiences. An expert's stand-alone opinion, or that given in a written news article, are both considered weak forms of evidence without scientific studies to back them up.

ANIMAL & CELL STUDIES *(experimental)*

Animal research can be useful, and can predict effects also seen in humans. However, observed effects can also differ, so subsequent human trials are required before a particular effect can be said to be seen in humans. Tests on isolated cells can also produce different results to those in the body.

CASE REPORTS & CASE SERIES *(observational)*

A case report is a written record on a particular subject. Though low on the hierarchy of evidence, they can aid detection of new diseases, or side effects of treatments. A case series is similar, but tracks multiple subjects. Both types of study cannot prove causation, only correlation.

CASE-CONTROL STUDIES *(observational)*

Case control studies are retrospective, involving two groups of subjects, one with a particular condition or symptom, and one without. They then track back to determine an attribute or exposure that could have caused this. Again, these studies show correlation, but it is hard to prove causation.

COHORT STUDIES *(observational)*

A cohort study is similar to a case-control study. It involves selection of a group of people sharing a certain characteristic or treatment (e.g. exposure to a chemical), and compares them over time to a group of people who do not have this characteristic or treatment, noting any difference in outcome.

RANDOMISED CONTROLLED TRIALS *(experimental)*

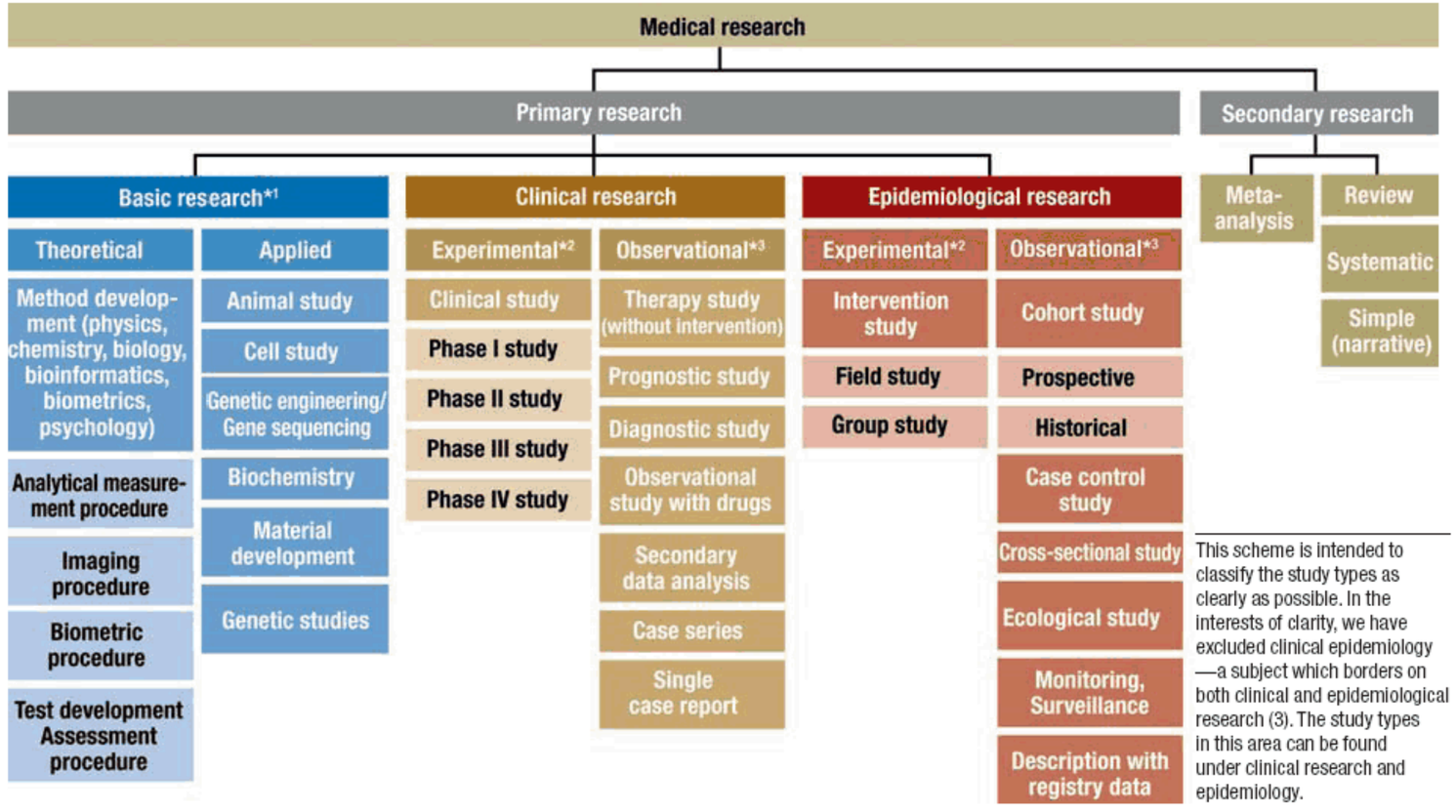
Subjects are randomly assigned to a test group, which receives the treatment, or a control group, which commonly receives a placebo. In 'blind' trials, participants do not know which group they are in; in 'double blind' trials, the experimenters do not know either. Blinding trials helps remove bias.

SYSTEMATIC REVIEW

Systematic reviews draw on multiple randomised controlled trials to draw their conclusions, and also take into consideration the quality of the studies included. Reviews can help mitigate bias in individual studies and give us a more complete picture, making them the best form of evidence.

Note that in certain cases, some of these types of evidence may not be possible to procure, for ethical or other reasons.

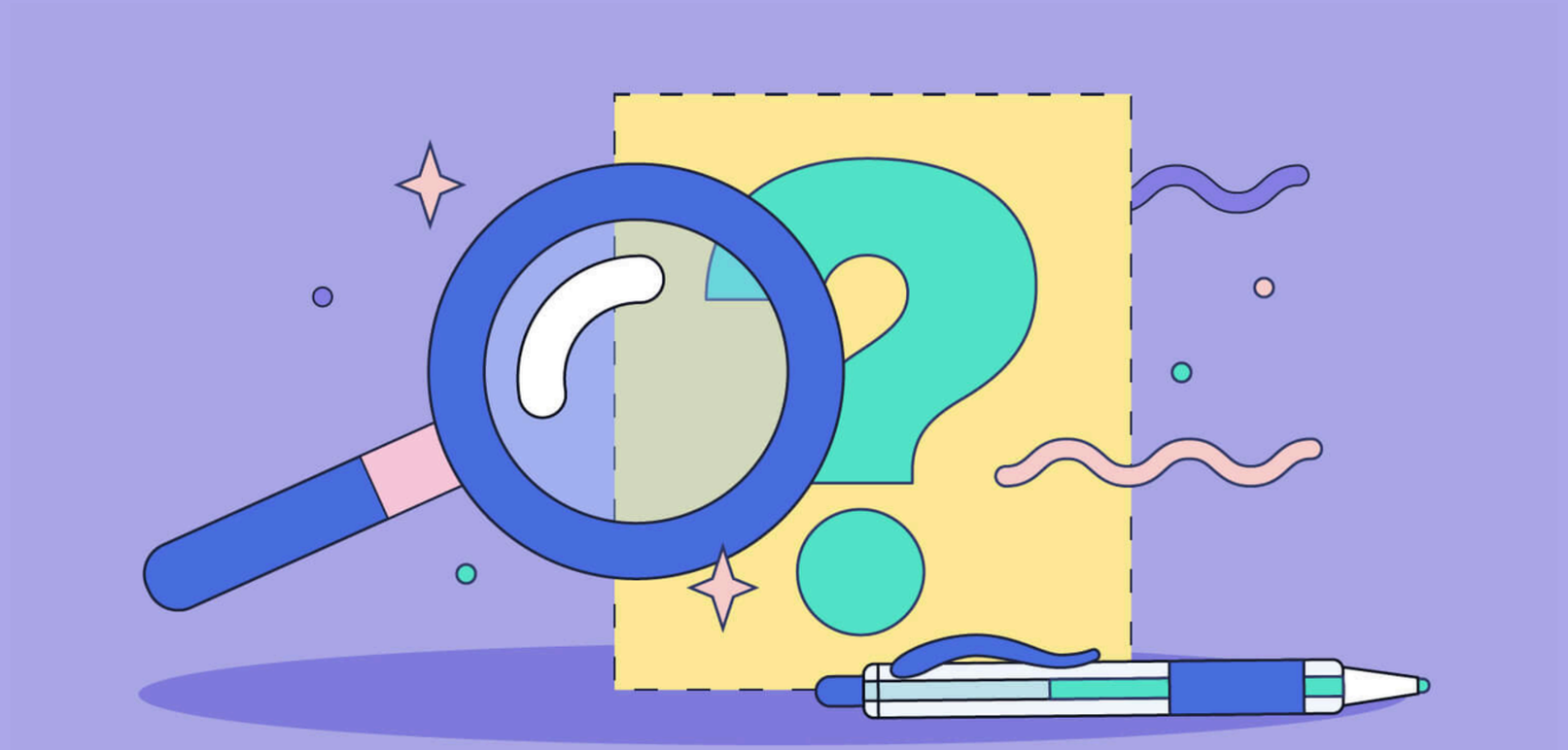
FIGURE 1



This scheme is intended to classify the study types as clearly as possible. In the interests of clarity, we have excluded clinical epidemiology —a subject which borders on both clinical and epidemiological research (3). The study types in this area can be found under clinical research and epidemiology.

Classification of different study types

*1, sometimes known as experimental research; *2, analogous term: interventional; *3, analogous term: noninterventional or nonexperimental



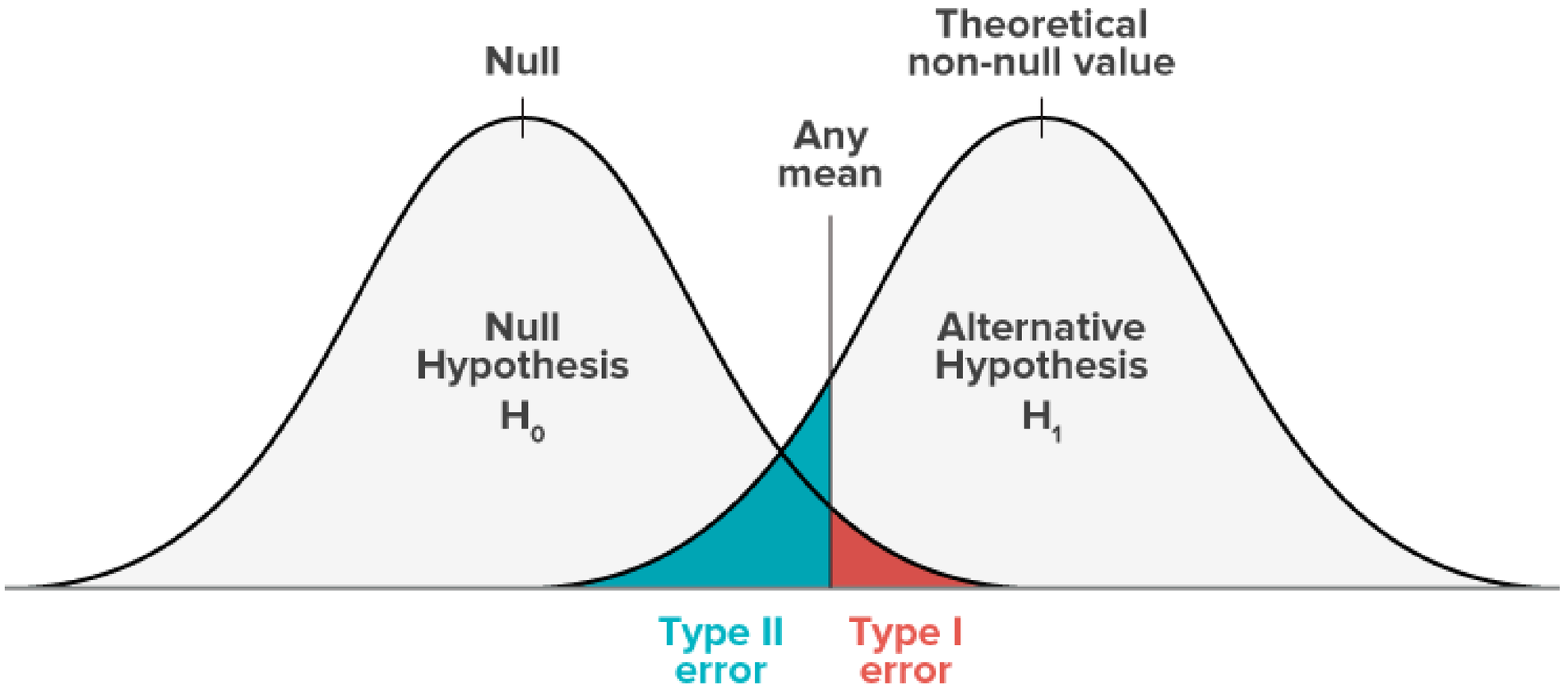
Build our hypothesys

H0 (null hypothesis) and H1 (alternative hypothesis) approach
Be **bold** make choices and commit

Hypothesis Testing and Comparing Two Proportions

- Hypothesis Testing: Deciding whether your data shows a "real" effect, or could have happened by chance
- Hypothesis testing is used to decide between two possibilities:
  - The Research Hypothesis
  - The Null Hypothesis

A good hypothesis is one that can be easily proved wrong if this is the case
Hypothesis are never "just true", but rather they hold true until falsified


Null
Theoretical
non-null value
Any
mean
Null
Hypothesis
H_0
Alternative
Hypothesis
H_1
Type II
error
Type I
error

Type I Error (False Positive)

Definition: A Type I error occurs when the null hypothesis is incorrectly rejected when it is actually true. In other words, it's concluding that there is an effect or difference when, in reality, there isn't one.

Probability Notation: The probability of making a Type I error is denoted by α (alpha), which is also known as the significance level of the test.

Common Significance Levels: Researchers often choose α = 0.05 (5%), meaning there's a 5% risk of rejecting the null

Type II Error (False Negative)

Definition: A Type II error happens when the null hypothesis is incorrectly accepted (i.e., failing to reject it) when the alternative hypothesis is true. Essentially, it's not detecting an effect or difference that actually exists.

Probability Notation: The probability of making a Type II error is denoted by β (beta).
Power of a Test: The power of a statistical test is defined as 1 - β. It represents the probability of correctly rejecting the null hypothesis when the alternative hypothesis is true.

Example Continued: If the new drug truly has a higher recovery rate than the existing drug (H_1 is true), but the study fails to detect this difference and concludes that there's no significant improvement, a Type II error has occurred.

The traditional alpha level of p<0.05p ical research due to several limitations and

- **Arbitrariness of the Threshol** en as arbitrary, originating from historical practices rather than from rigorou oes not reflect the variability and context of different r tringent standards.
- **Dichotomization of Results:** Sett nt" based on p<0.05p < 0.05p<0.05 can lead to oversimplification. This dicho ls to convey the strength of the evidence, pote the threshold.
- **Reproducibility Crisis:** Relying hea reproducibility issues in science. Studies that achieve significance by a narrow reflect sample-specific noise rather than a
- **Overemphasis on Statistical vs. Practical Significance:** Achieving p<0.05p < 0.05p<0.05 does not imply that an effect is practica ficant results that
- **Encoura** acking or selective report ance threshold.
- **Misint** lity that the null hypothes ll hypothesis were

Alternatives porting confidence intervals alongside effect sizes to provide a fuller picture of the data and its practical implications

storytelling WITH data®

SIMPLE TEXT

Just because you have numbers doesn't mean you need a graph!

TABLE

OFTEN THERE ARE MORE EFFICIENT WAYS

Avoid using tables in live presentations because people stop listening & start reading

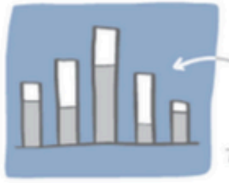
BAR CHARTS

Great for categorical data

Easy for our eyes - comparing heights to a consistent baseline

Rule: Must have a zero baseline. No exceptions!

VERTICAL

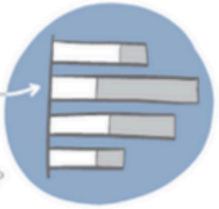


HORIZONTAL

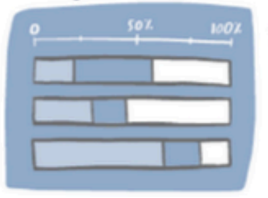
Good when category names are long

STACKED

OFTEN MISUSED... EASIER TO COMPARE TOTAL & FIRST SERIES, BUT SEGMENTS UP THE STACK DON'T LINE UP



100% STACKED



TWO BASELINES FOR COMPARISON

HEAT MAP

EYES CAN EASILY PICK OUT BIG DIFFERENCES IN COLOR INTENSITY, but smaller ones don't stand out

Can work well when beginning to explore data and deciding where to dig further

SCATTER PLOT

Good for encoding data simultaneously on two axes to identify what relationships exist

LINE

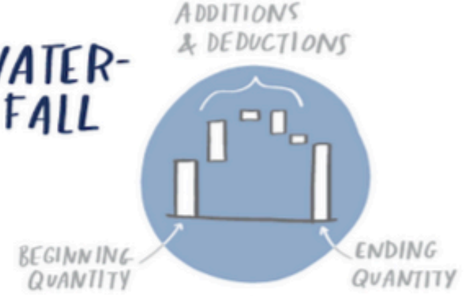
Rule: The lines that connect the dots have to make sense! Most effective with continuous data, often time

SLOPE GRAPH

↑ A FANCY WORD FOR A LINE GRAPH WITH ONLY 2 POINTS

Useful to focus on change between two points in time or difference between groups

WATER-FALL



Often used in finance to show variance to budget

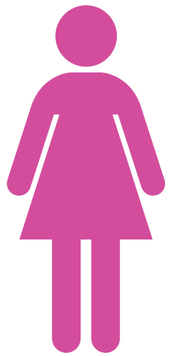
SQUARE AREA

(AKA WAFFLE CHART)

THE GRID IS IMPORTANT BECAUSE WE TEND TO OVERESTIMATE AREAS

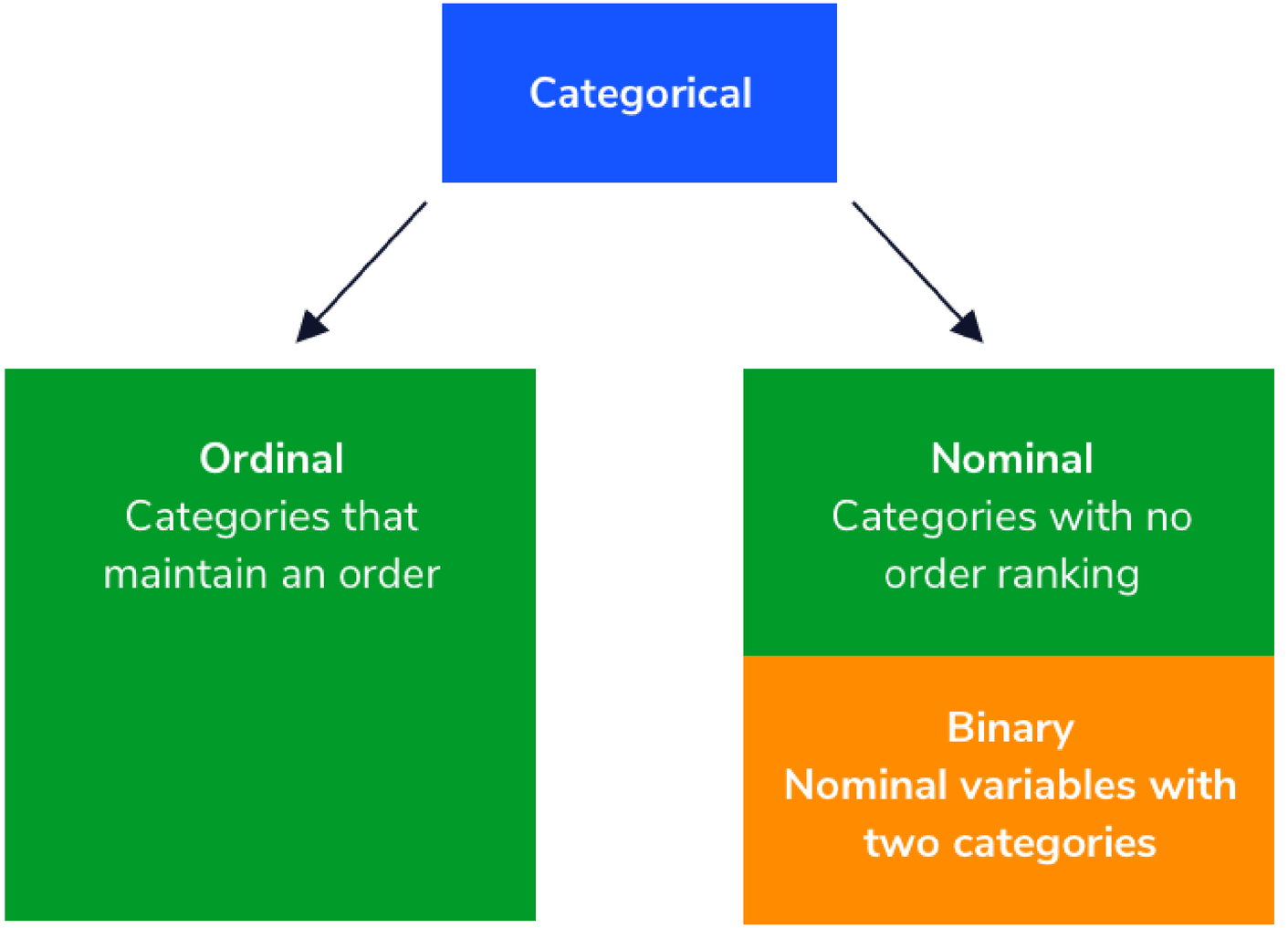
Good for showing numbers of very different magnitudes, or as an alternative to a pie chart

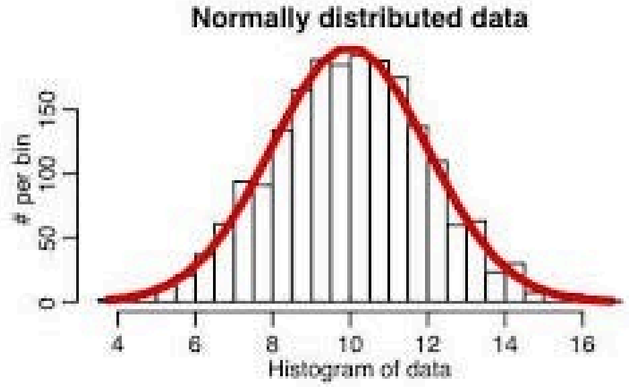
International League Against Epilepsy

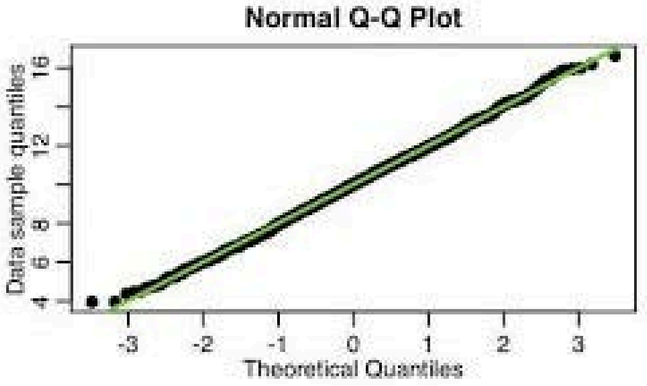


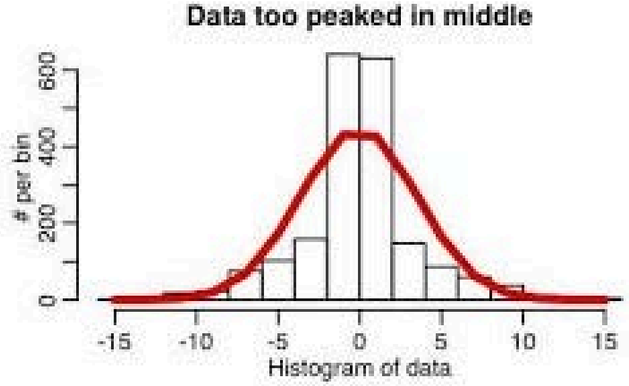
Variable	units	measures
Gender	M/F	freq and %
Age	years	mean±SD
Height Weight BMI	Cm, Kg	mean±SD
Diagnosis	Categ non bin TLE,JME, CAE,GTCA.....	freq and %
Age at diagnosis	years	mean±SD
SF	y/n	freq and %
DR	y/n	freq and %

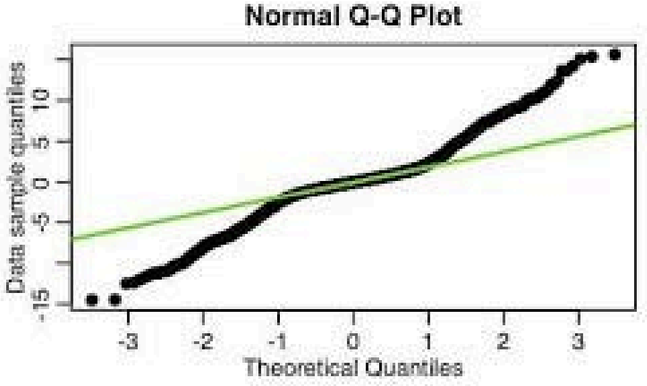
Variable	units	measures
years of education	years	mean±SD
Age	years	mean±SD
Height Weight BMI	Cm, Kg	mean±SD
Seizures/year	number	mean±SD
number of ASM	number	mean±SD
ASM type	Categ non bin LEV,CBZ,LTG,VPA...	freq and %
Hospetalization/y	Number	mean±SD

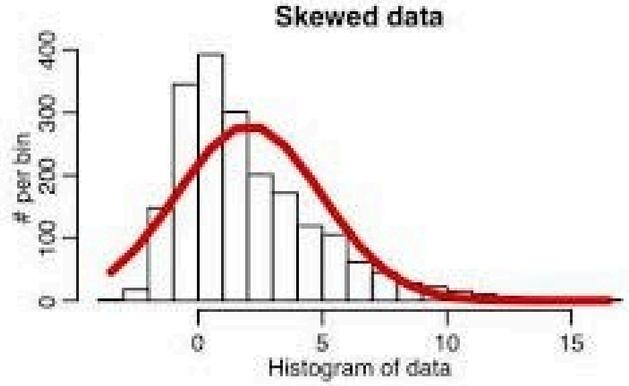

Categorical
Ordinal
Categories that maintain an order
Nominal
Categories with no order ranking
Binary
Nominal variables with two categories

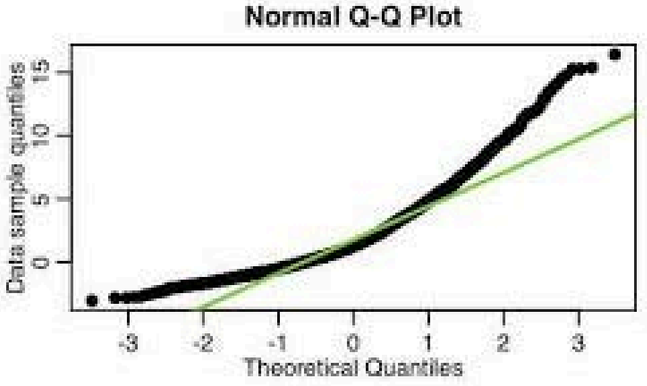

Normally distributed data
per bin
0
50
100
150
4
6
8
10
12
14
16
Histogram of data


Normal Q-Q Plot
Data sample quantiles
4
6
8
12
16
-3
-2
-1
0
1
2
3
Theoretical Quantiles

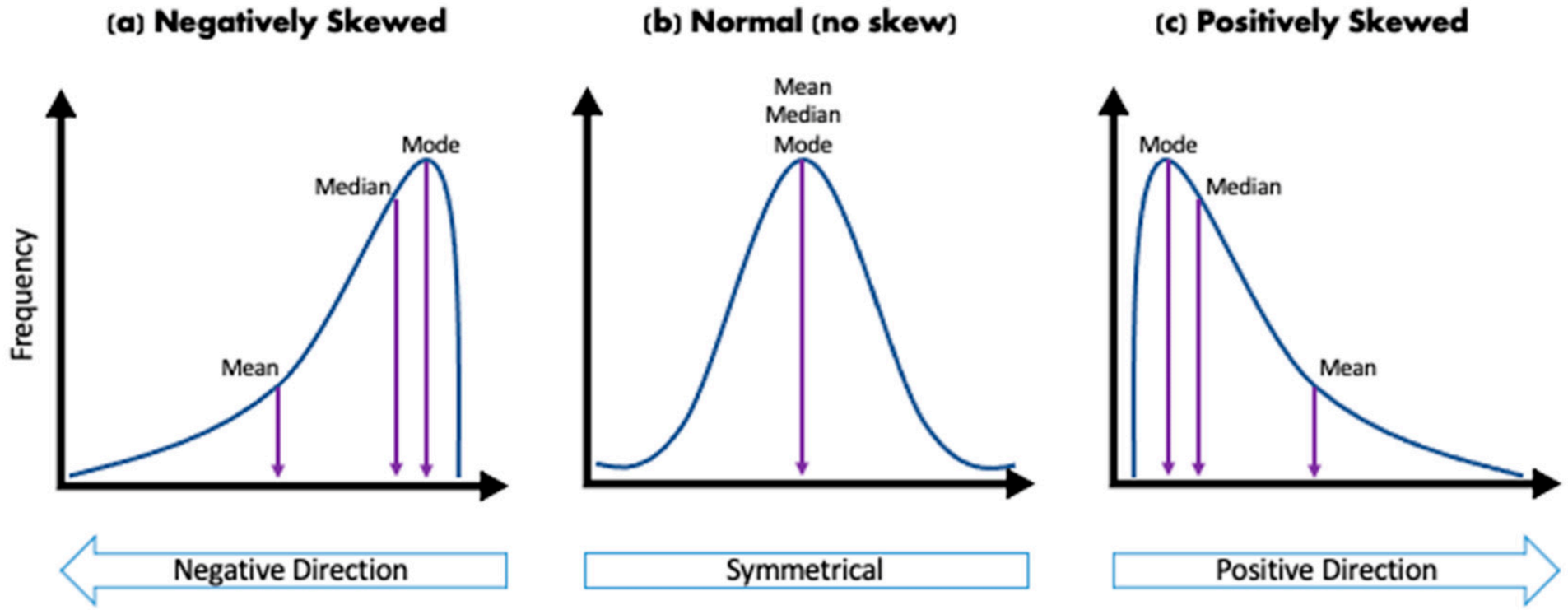

Data too peaked in middle
per bin
0
200
400
600
-15
-10
-5
0
5
10
15
Histogram of data


Normal Q-Q Plot
Data sample quantiles
-15
-5
0
5
10
-3
-2
-1
0
1
2
3
Theoretical Quantiles

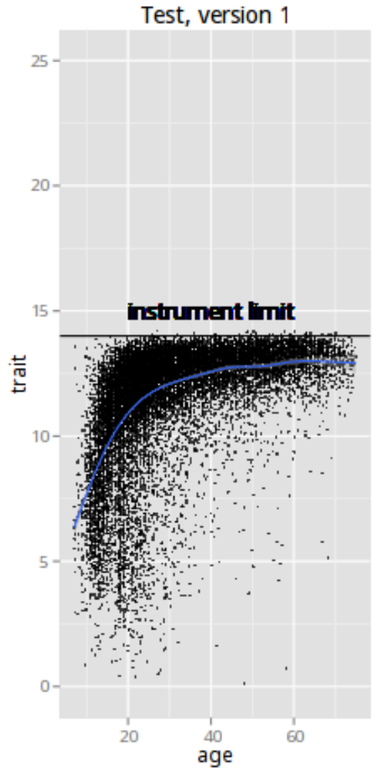

Skewed data
per bin
0
100
200
300
400
0
5
10
15
Histogram of data

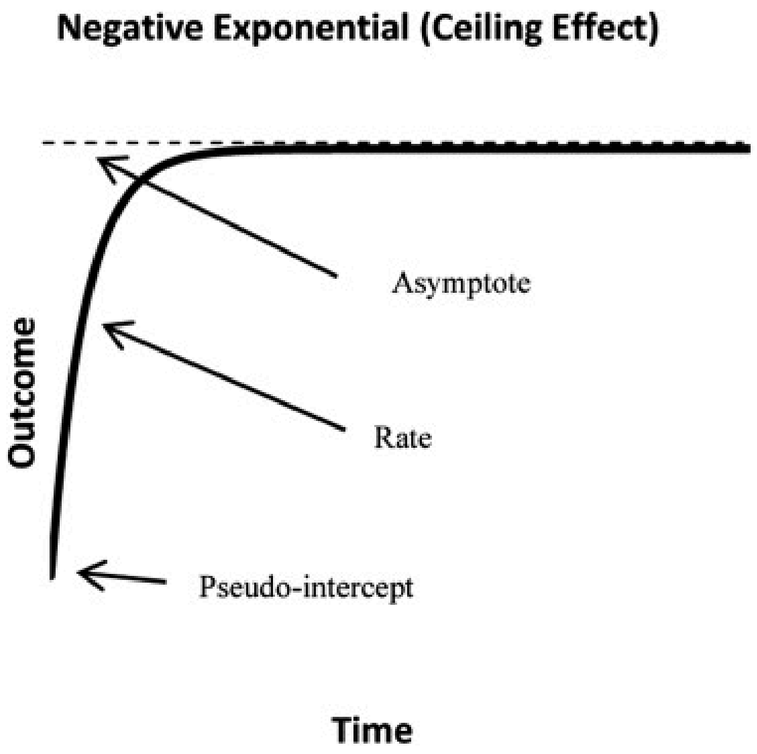

Normal Q-Q Plot
Data sample quantiles
0
5
10
15
-3
-2
-1
0
1
2
3
Theoretical Quantiles

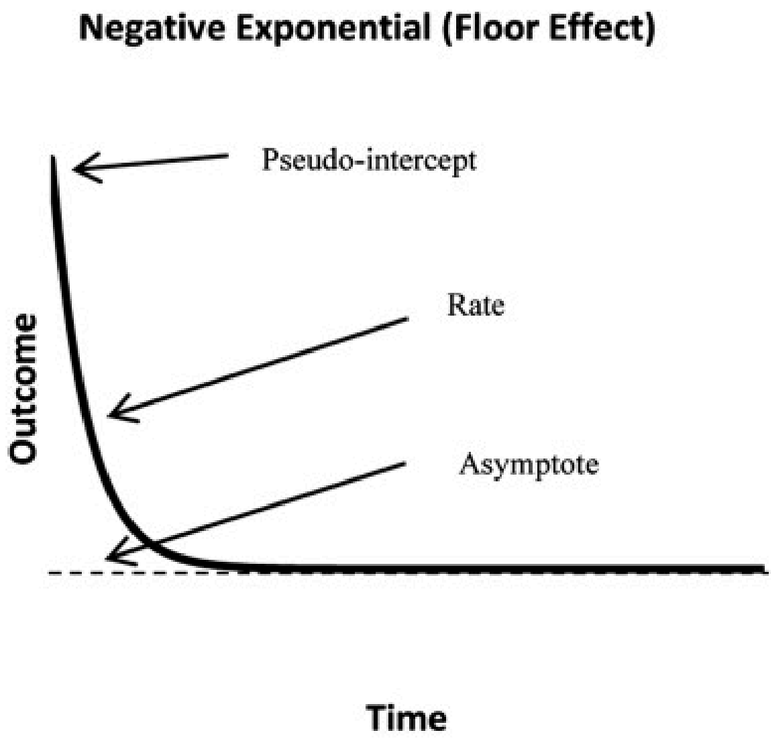
Number of seizures is mostly non normally distributed



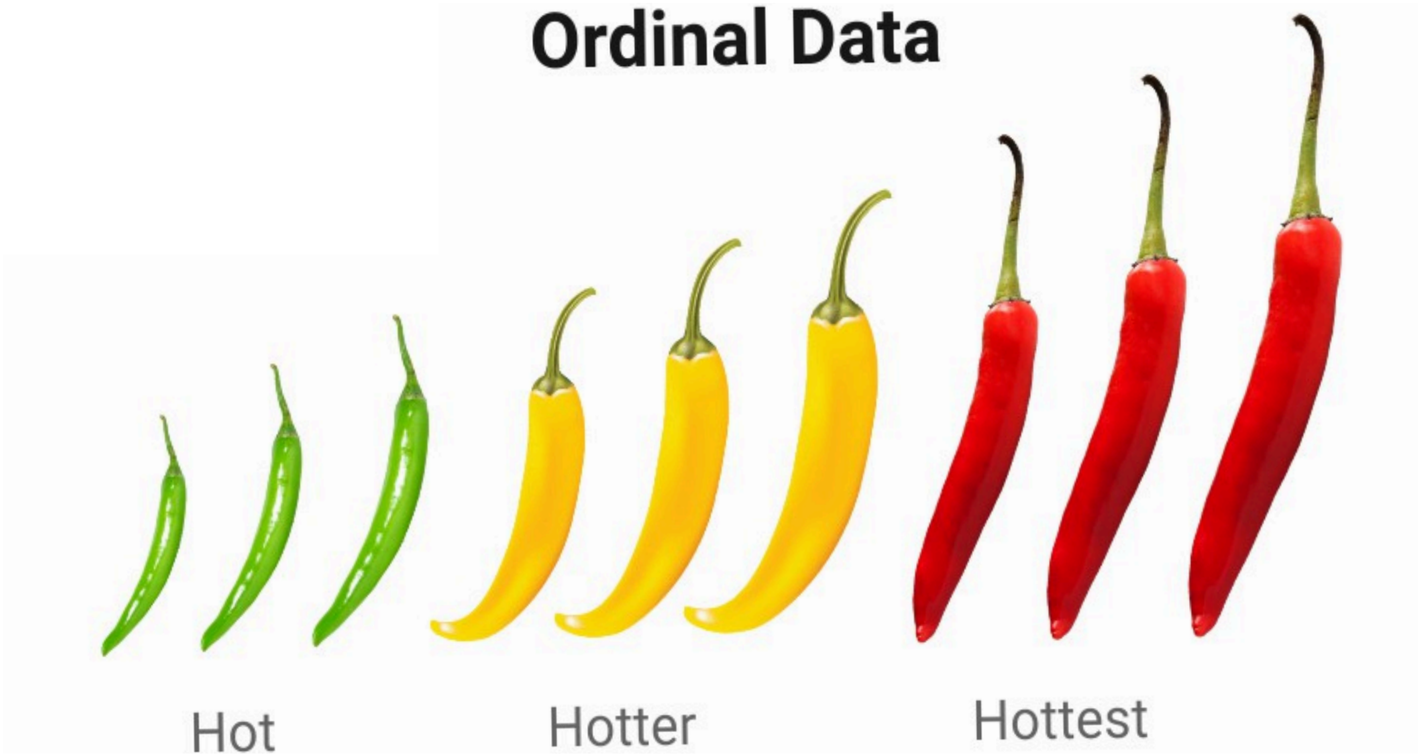
Some measures are limited due to ceiling and floor effects



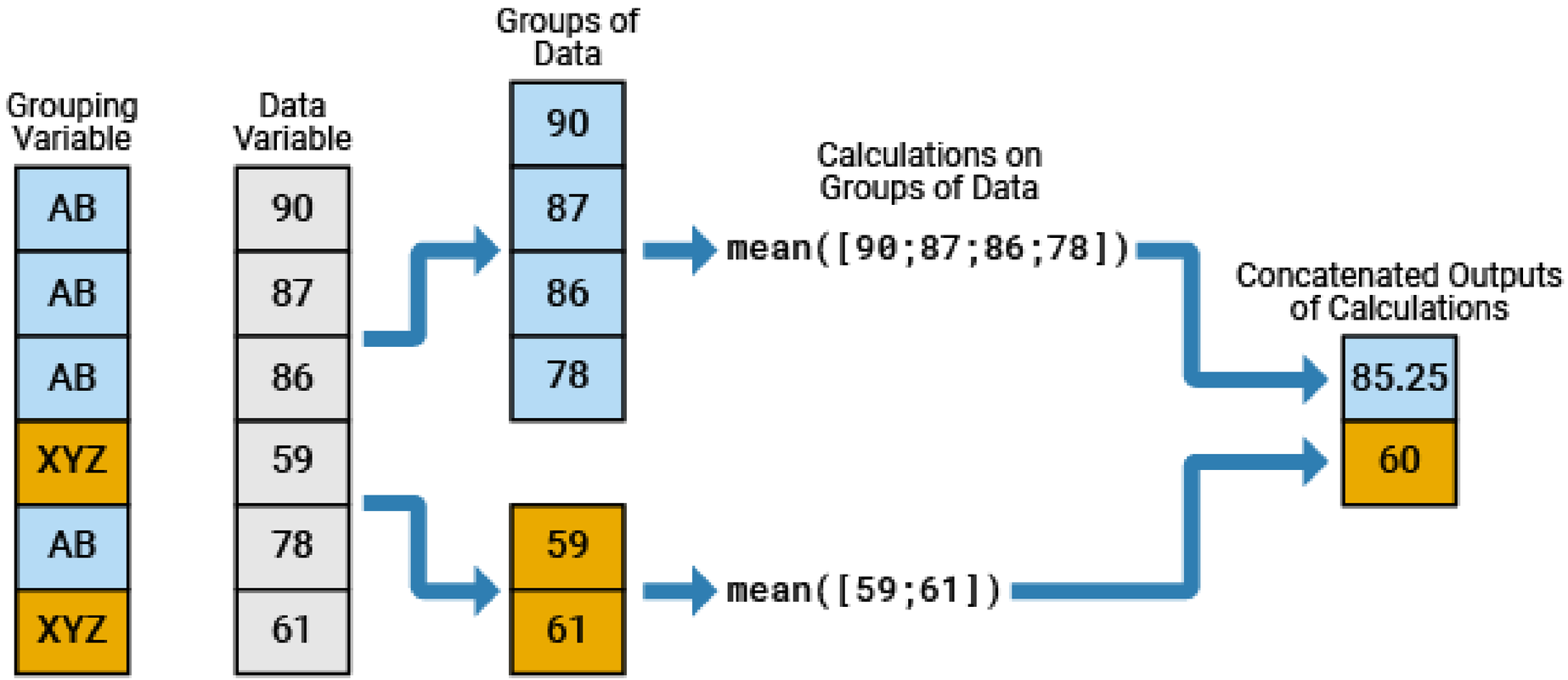




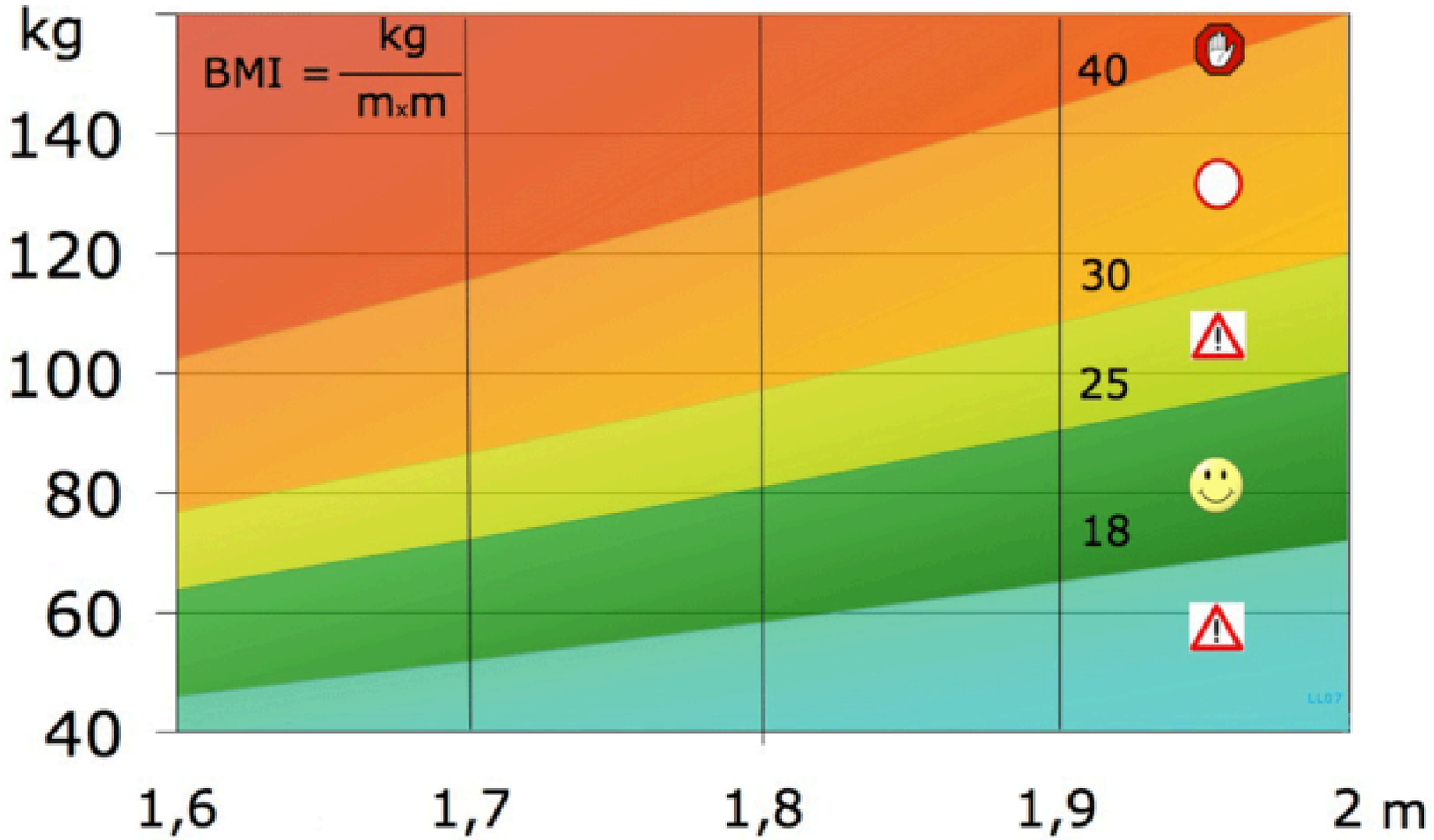
Number of ASM is more an ordinal than continuous variable and is better suited to draw cathegories, simillarly also other variables such as years of education.



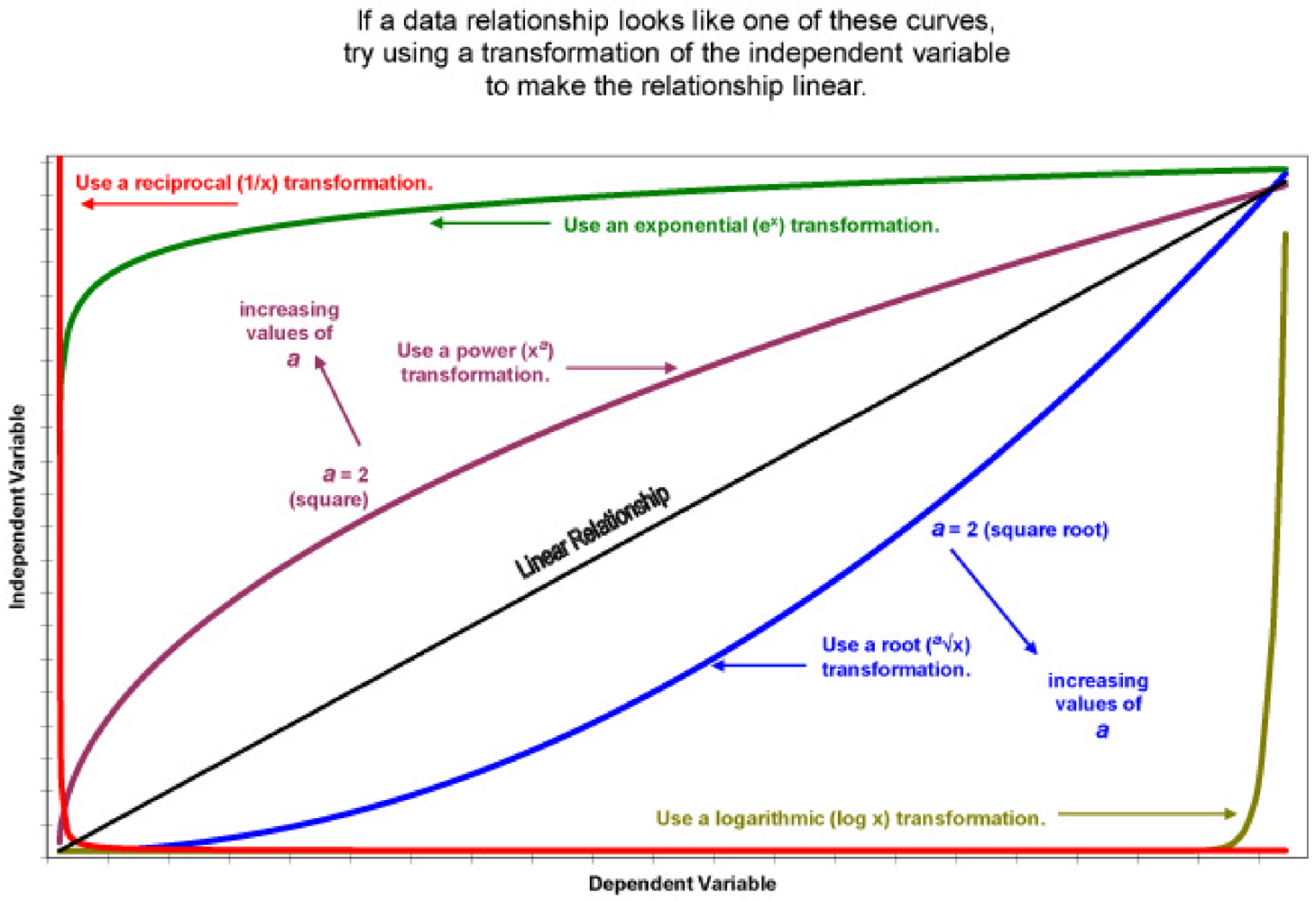
Some variables are most usefull as filtering variables i.e. DR or SF



Variable combination



Variable transformation-normalization



Trade off-normal distribution vs easy to comprehend

Basic formula for sample size - Continuous data

$$\text{Number of samples per group (n)} = \frac{2\,x\,(Z_{(1-\alpha/2)}+Z_{\beta})^2 x\,\sigma^2}{\Delta^2}$$

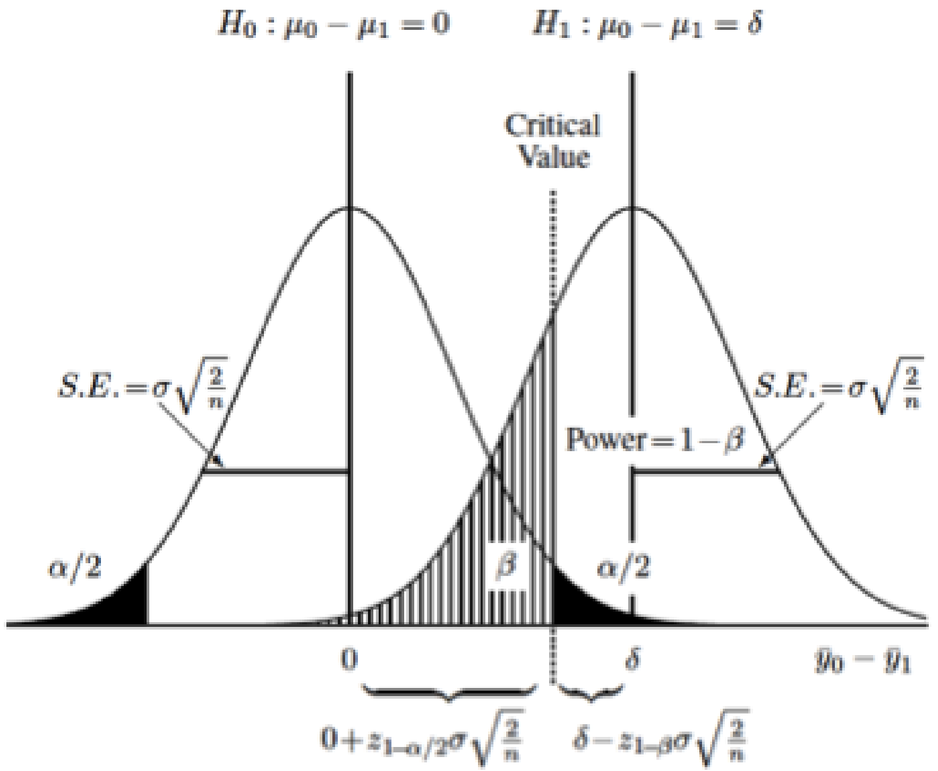
Where Δ = size of difference, minimal effect of interest

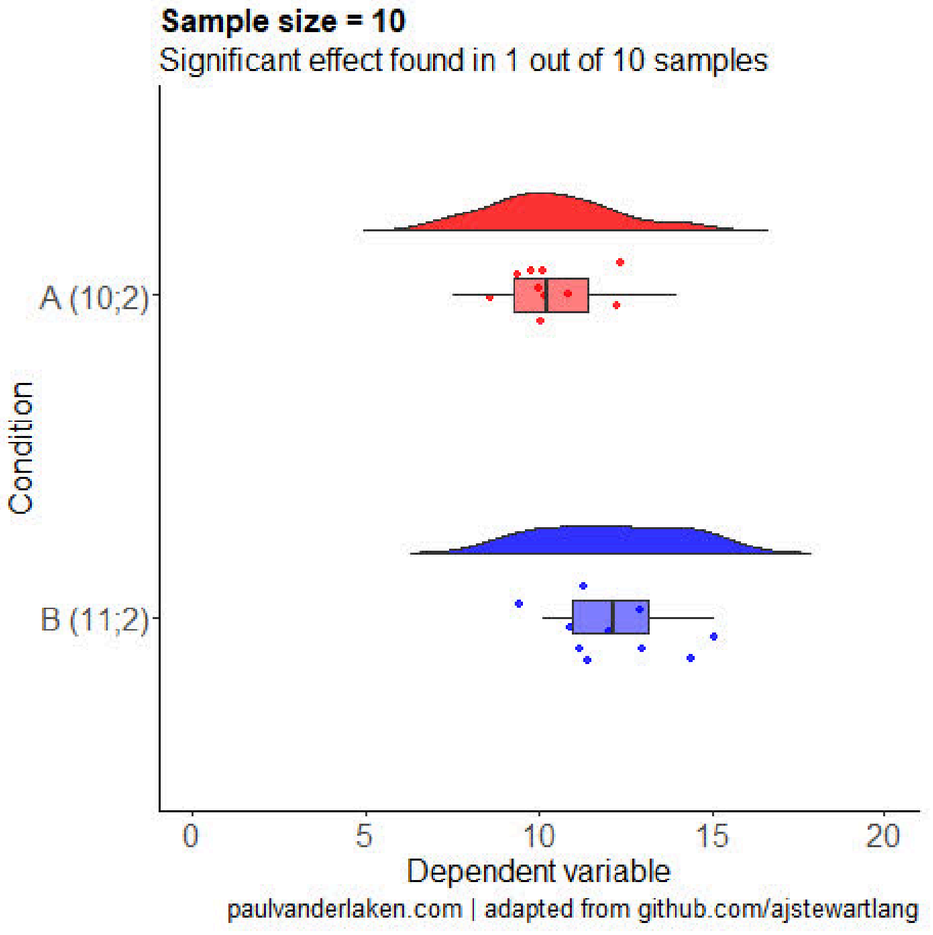
α = significance level (eg 0.05)

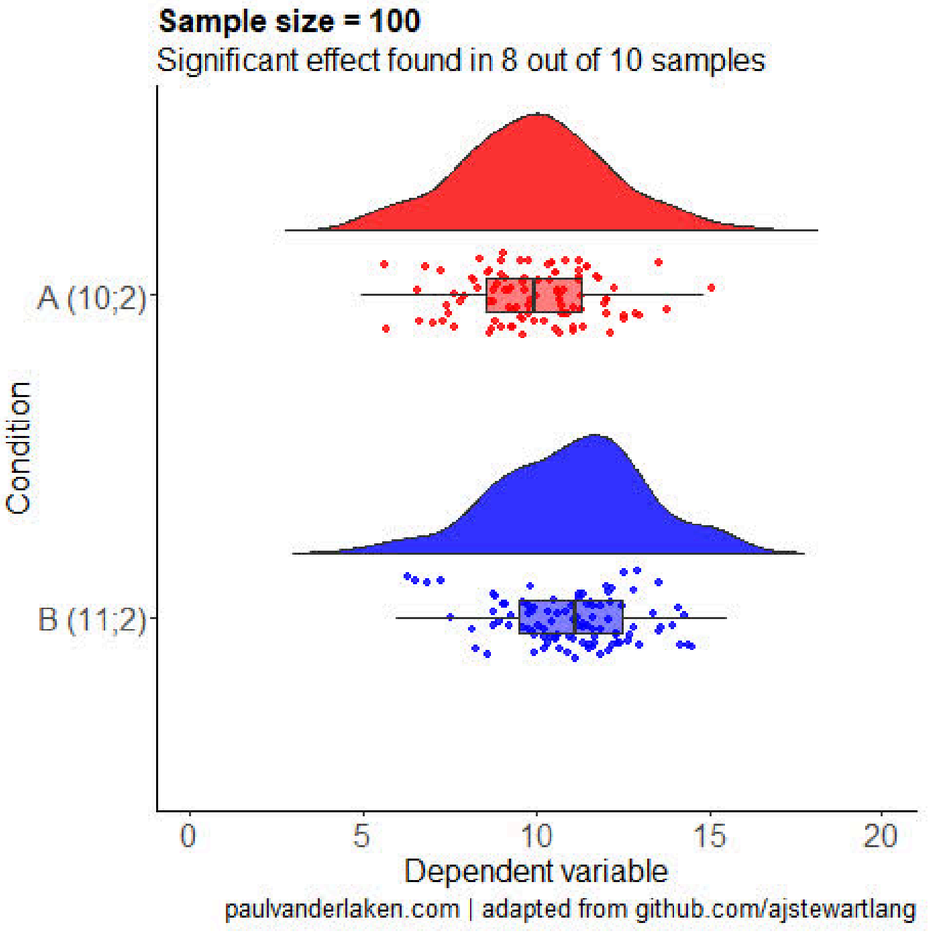
β = power, probability of detecting a significant result (typically 80%, 90%)

σ = SD of data

Z_p = points on normal distribution to give required power and significance




Sample size = 10
Significant effect found in 1 out of 10 samples
A (10;2)
B (11;2)
Condition
0
5
10
15
20
Dependent variable
paulvanderlaken.com | adapted from github.com/ajstewartlang


Sample size = 100
Significant effect found in 8 out of 10 samples
A (10;2)
B (11;2)
Condition
0
5
10
15
20
Dependent variable
paulvanderlaken.com | adapted from github.com/ajstewartlang

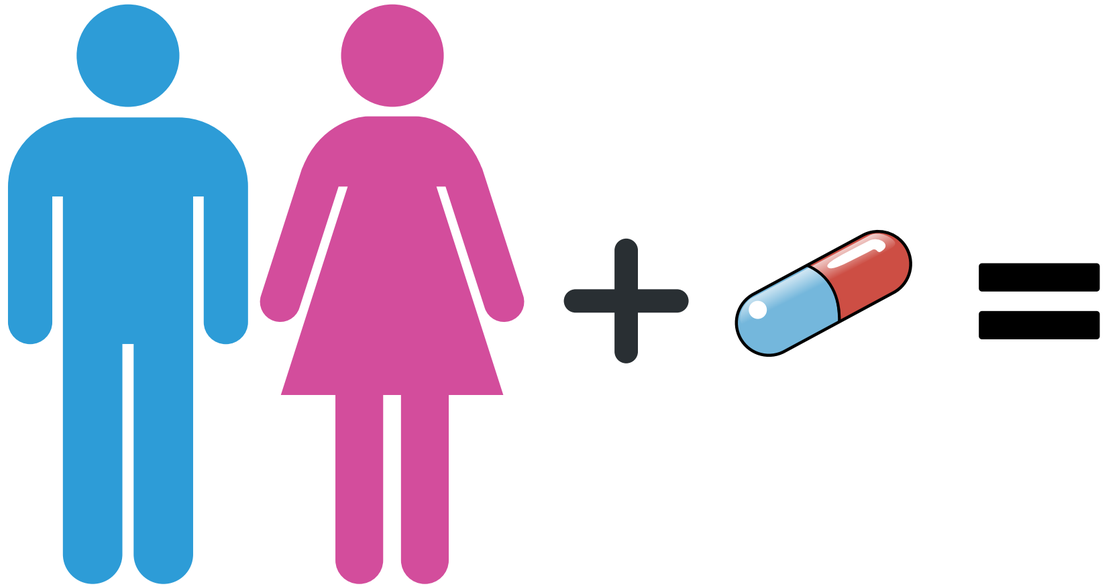
1 What do we want to know

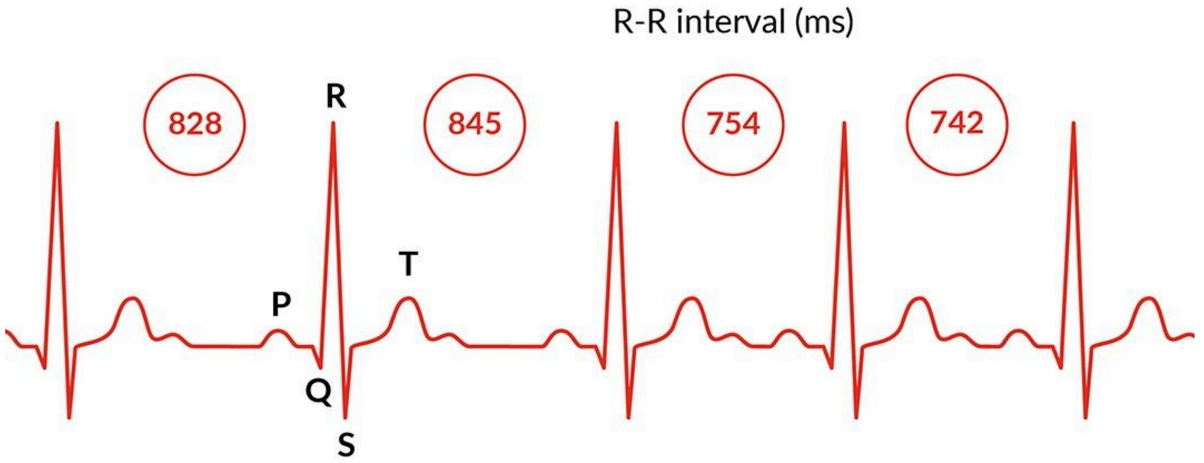
2 How will we assess it

3 What is our statistic plan

1

What are the cardiac effects of a new ASM (Na ch blocker)?

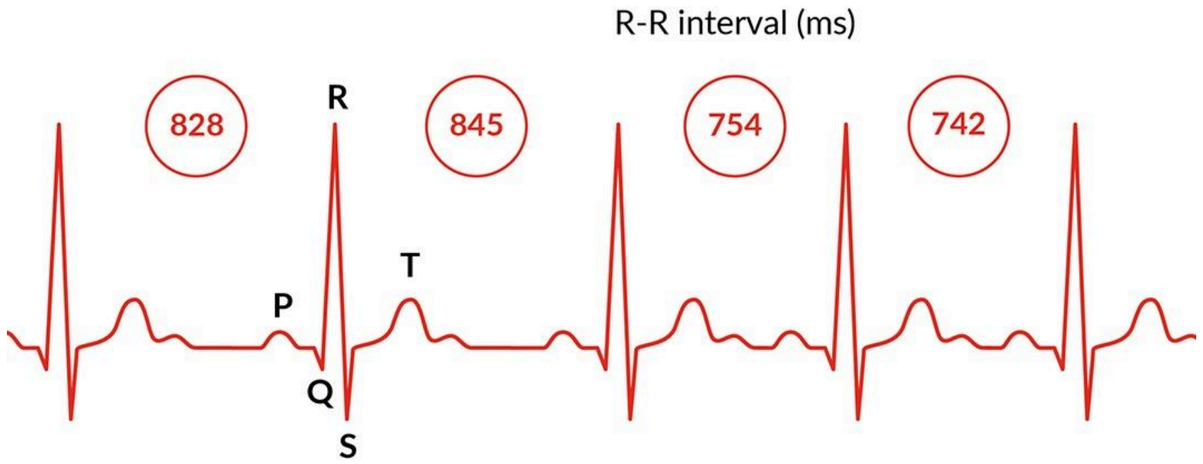




Does the new ASM (Na ch blocker) effect HRV ?

2

EKG HRV



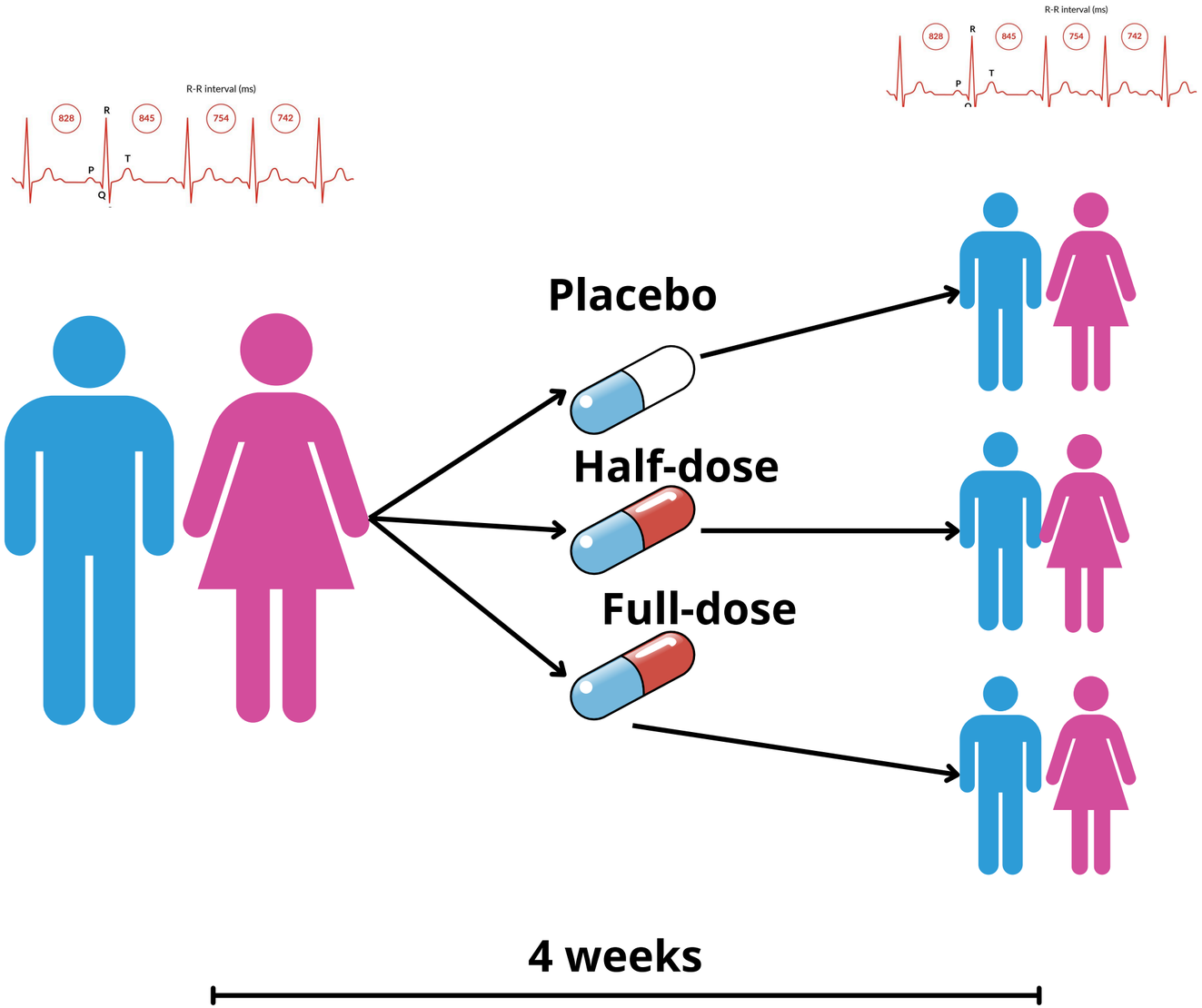
Demographic and Epilepsy related variables

....

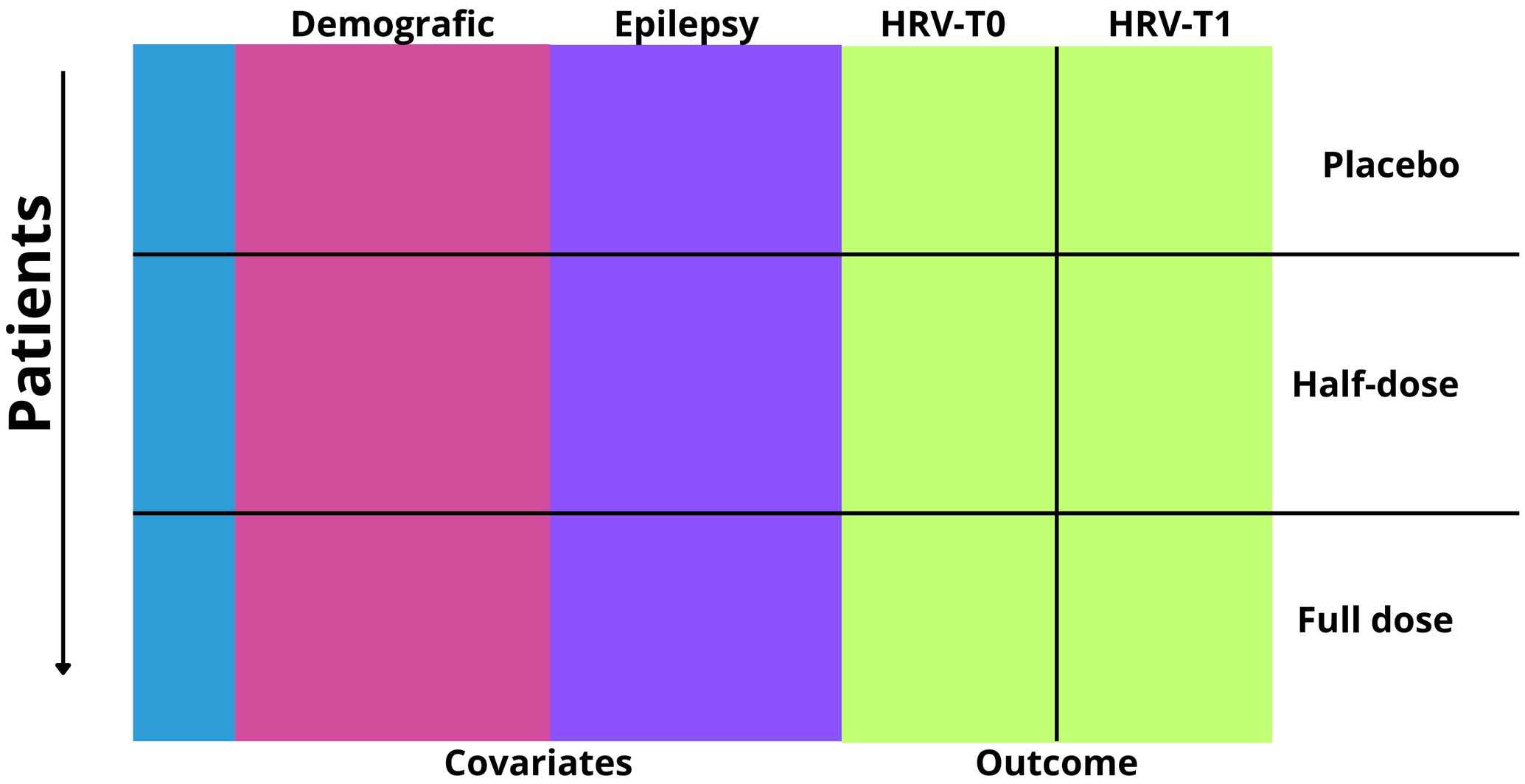
..

.

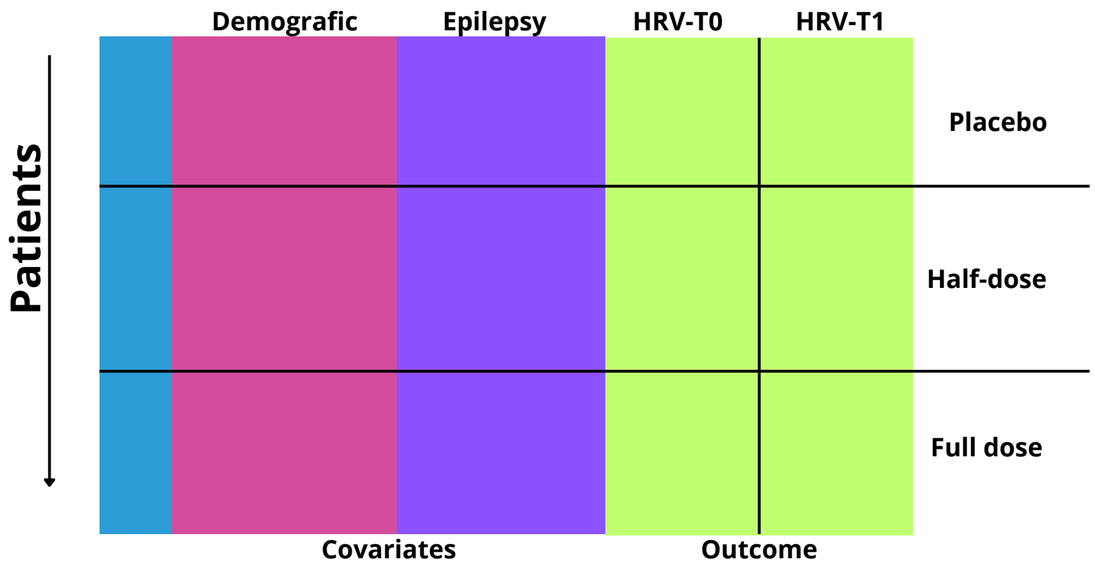
3

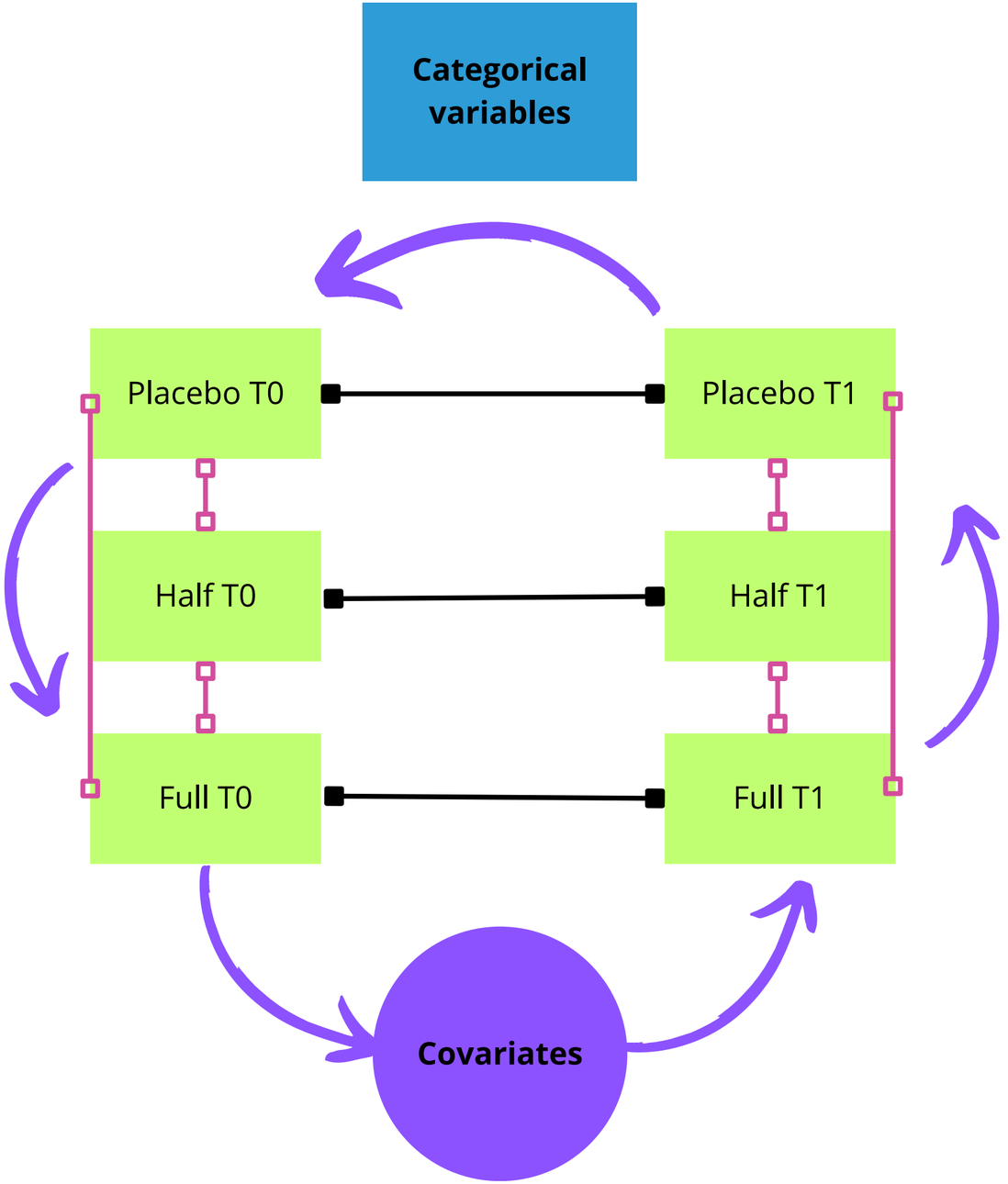

R-R interval (ms)
828
845
754
742
R
P
T
Q
Placebo
Half-dose
Full-dose
4 weeks

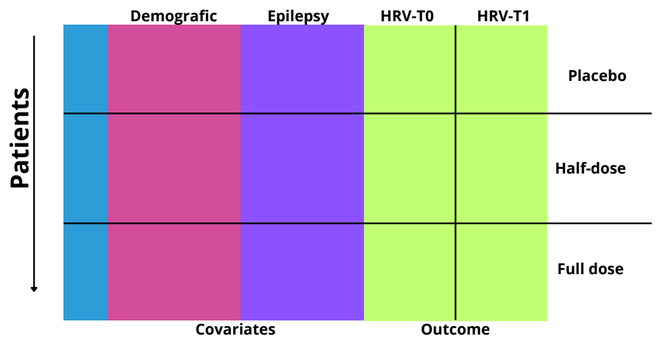
Dimensions of our data



Dimensions of our data





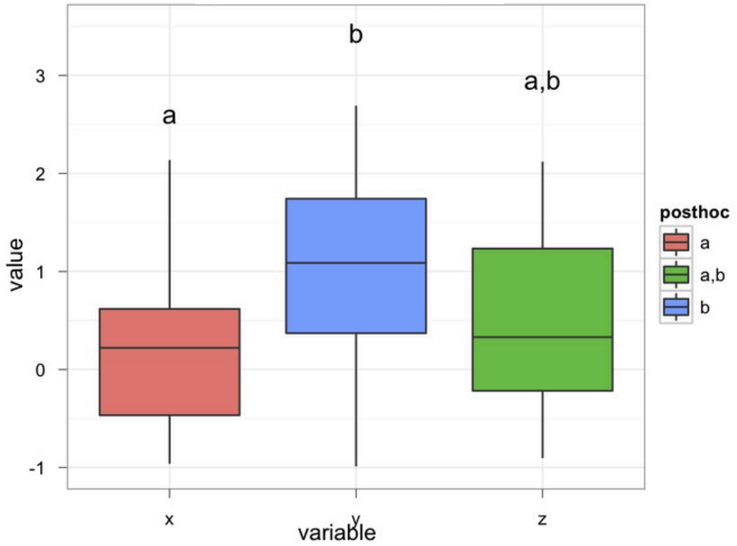


T test takes you all the way?

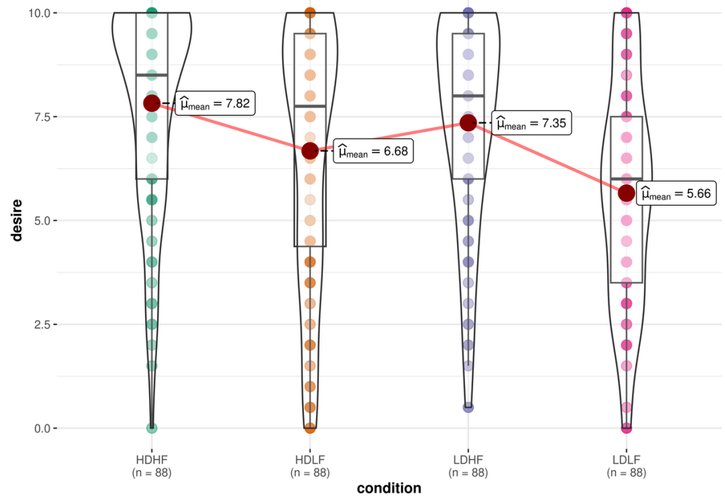
Placebo	**HRV-T0**	**HRV-T1**	**Alpha error 0.05**
			+
Half-dose	**HRV-T0**	**HRV-T1**	**Alpha error 0.05**
			+
Full dose	**HRV-T0**	**HRV-T1**	**Alpha error 0.05**
			=

15% chance of rejecting null hypothesis incorrectly

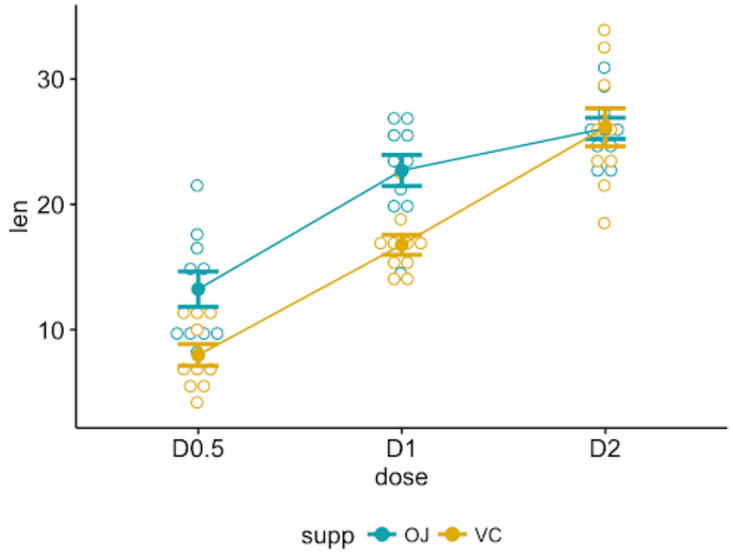
ANOVA



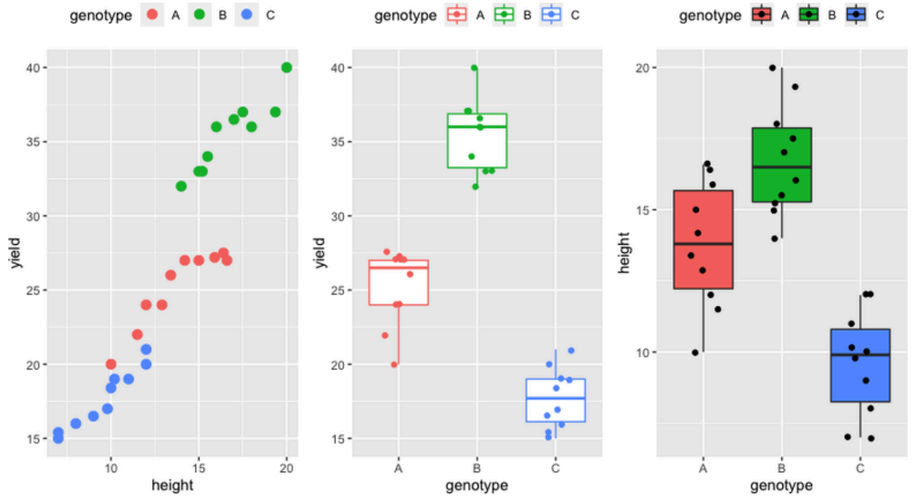
rmANOVA



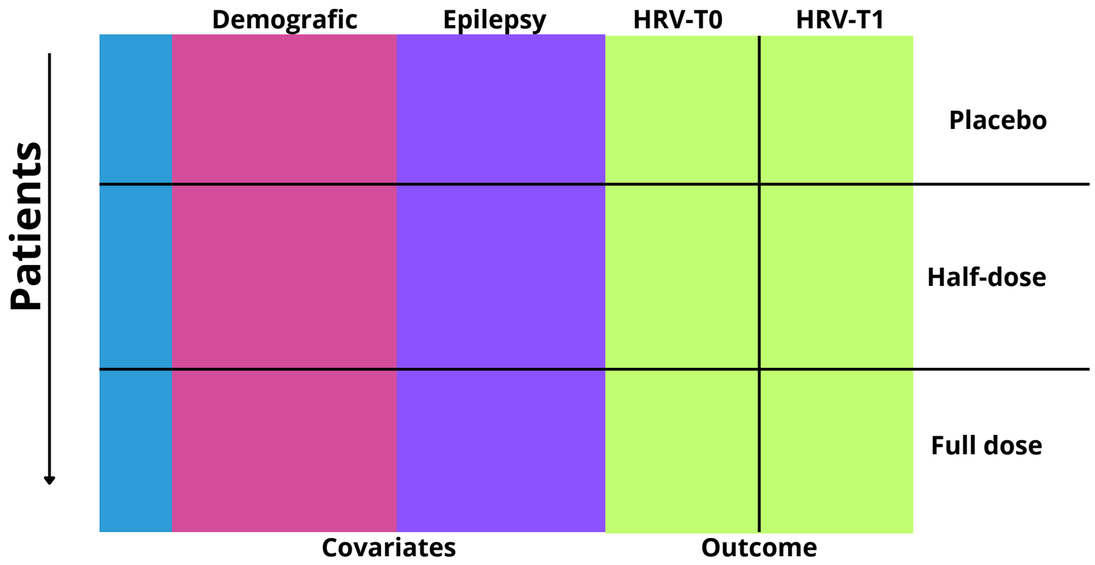
rmANOVA-two way

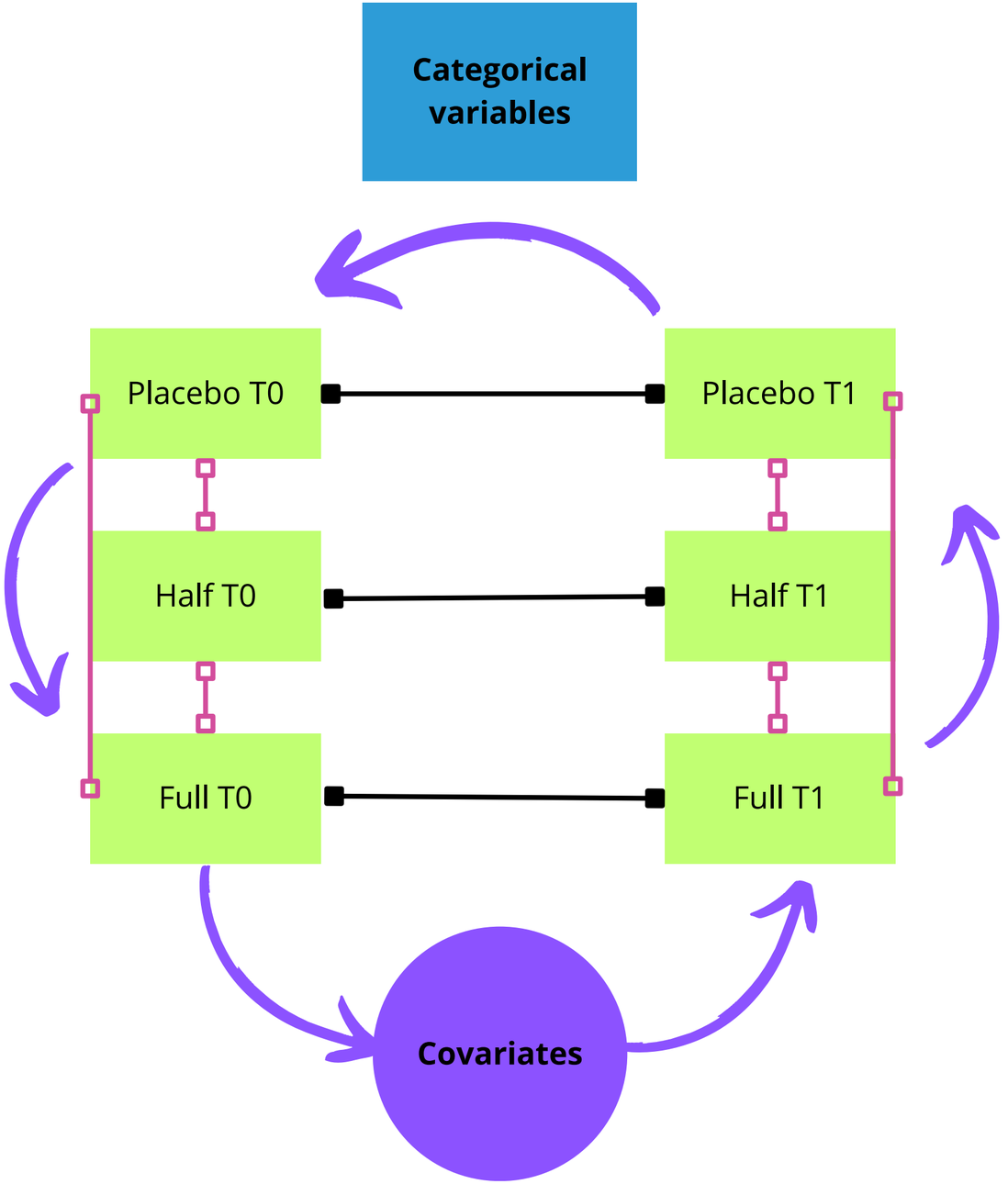


ANCOVA



Dimensions of our data





$$\mathbf{y} = \beta_0 + \sum \beta_i \mathbf{x}_i + \gamma + \varepsilon$$

Response variable

Global intercept

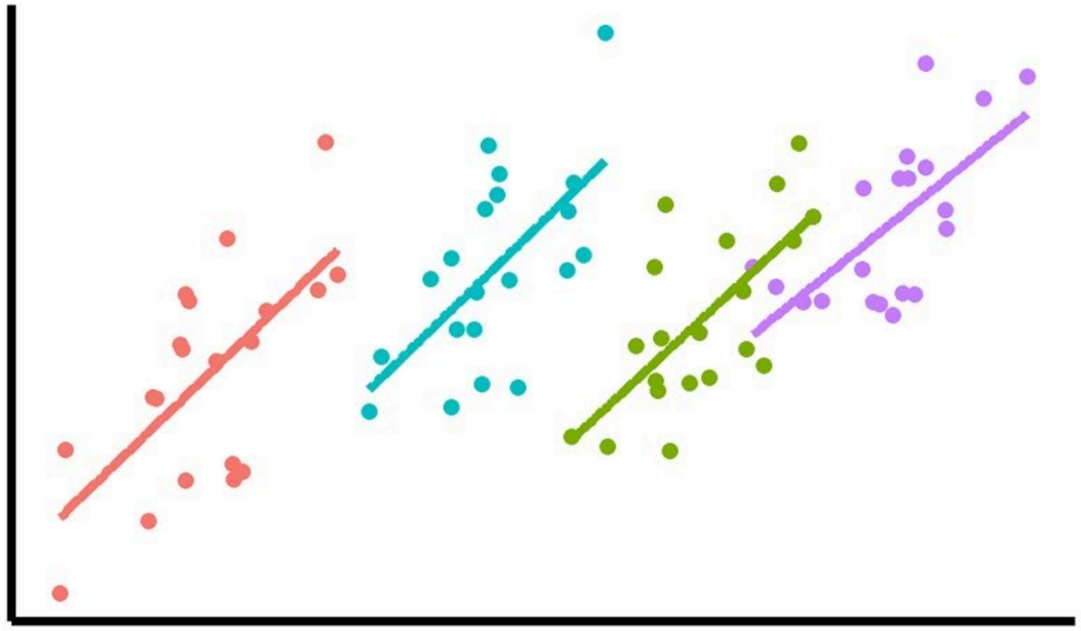
Fixed effect parameters

Fixed effect variables

Random effect variance

Residual variance

Linear Mixed-Effects Models

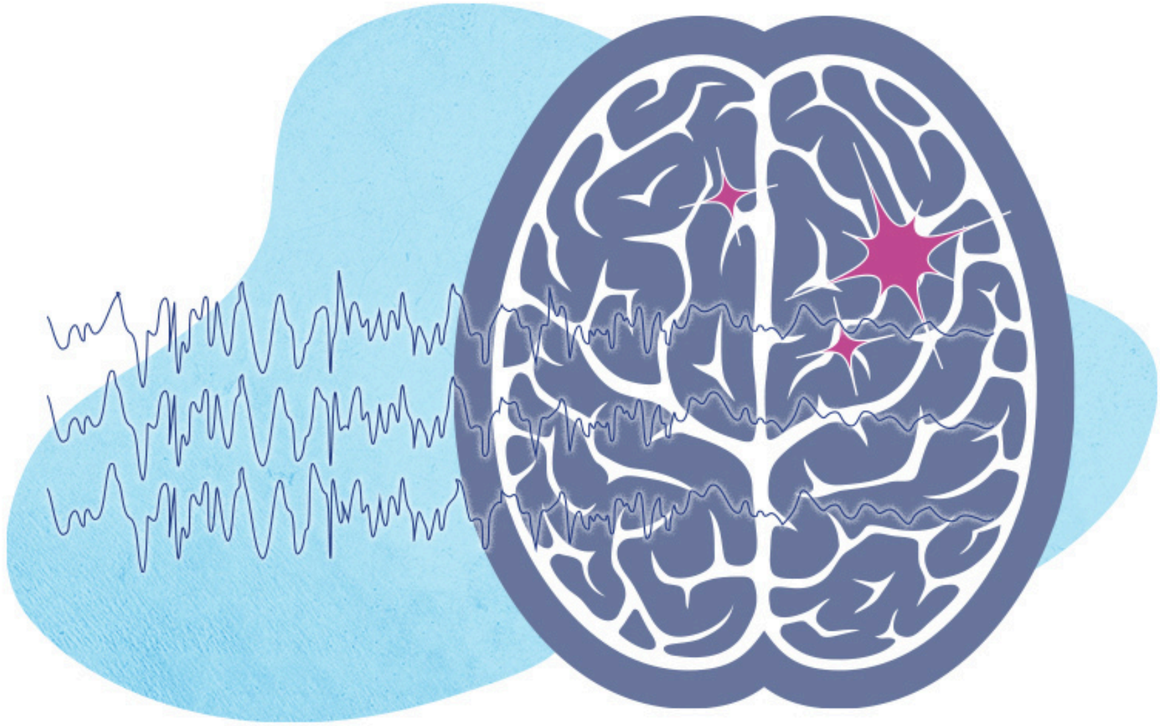


1 What do we want to know

2 How will we assess it

3 What is our statistic plan

1



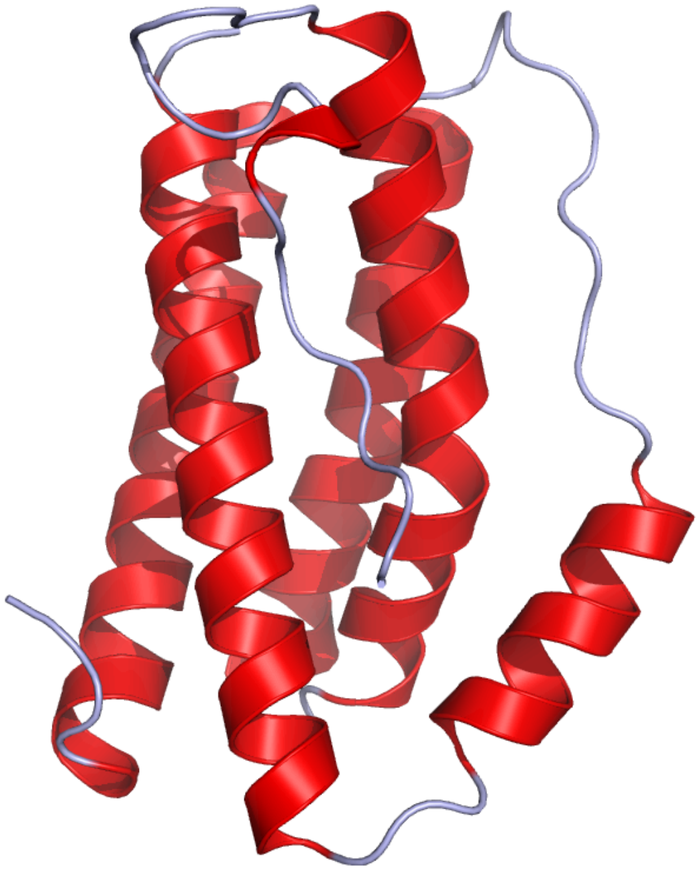
We want to investigate what influences survival rate in SE

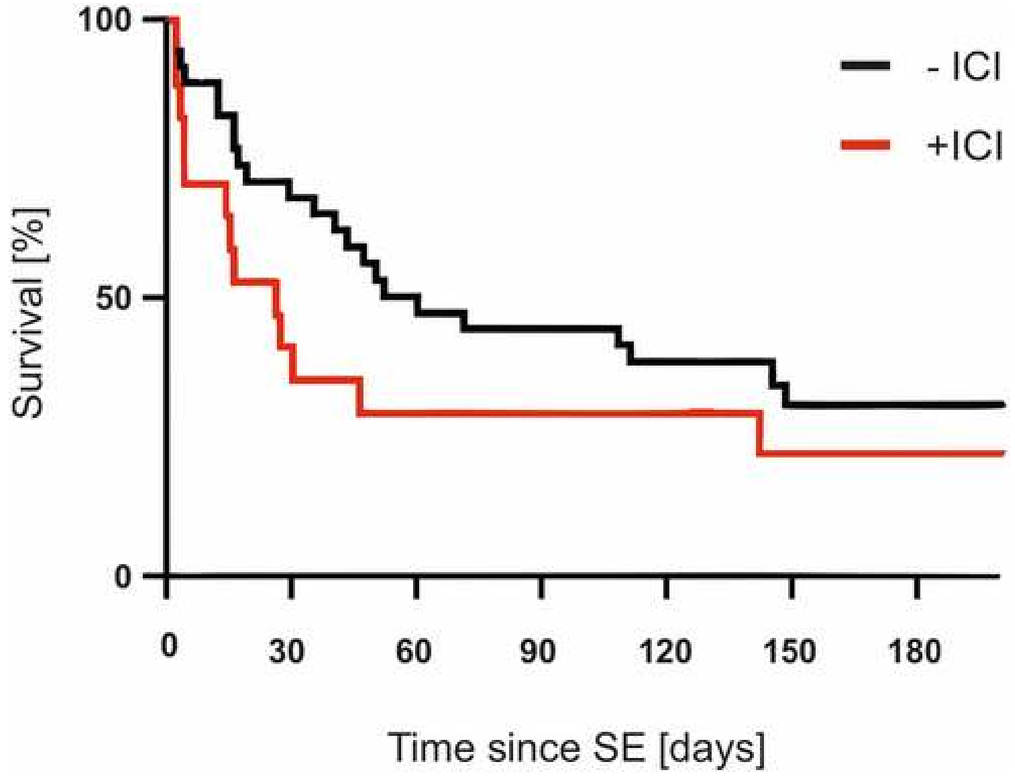
Is Survival different in patients with high vs low IL-6?

What are the factors that relate with survival?

2

IL-6

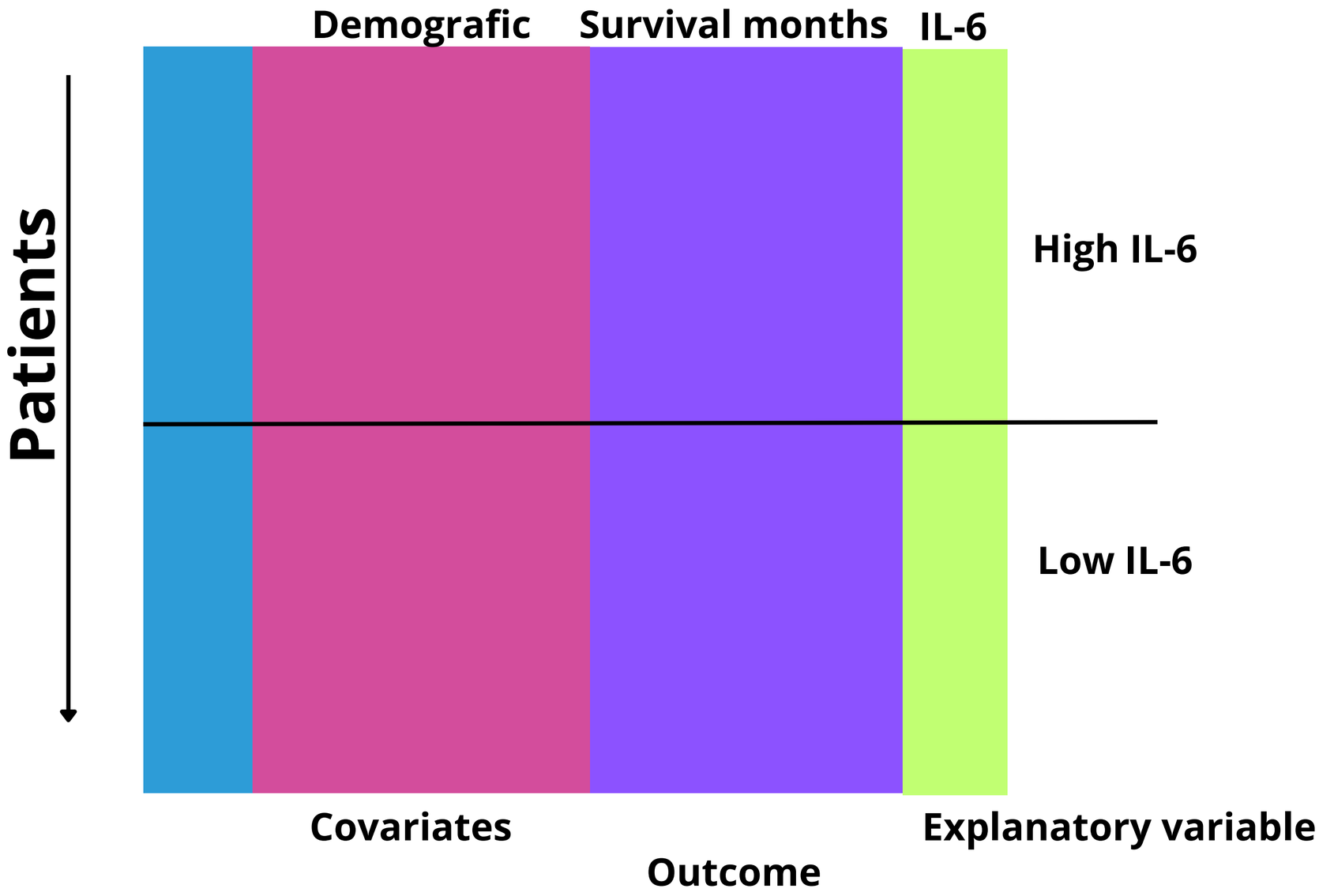


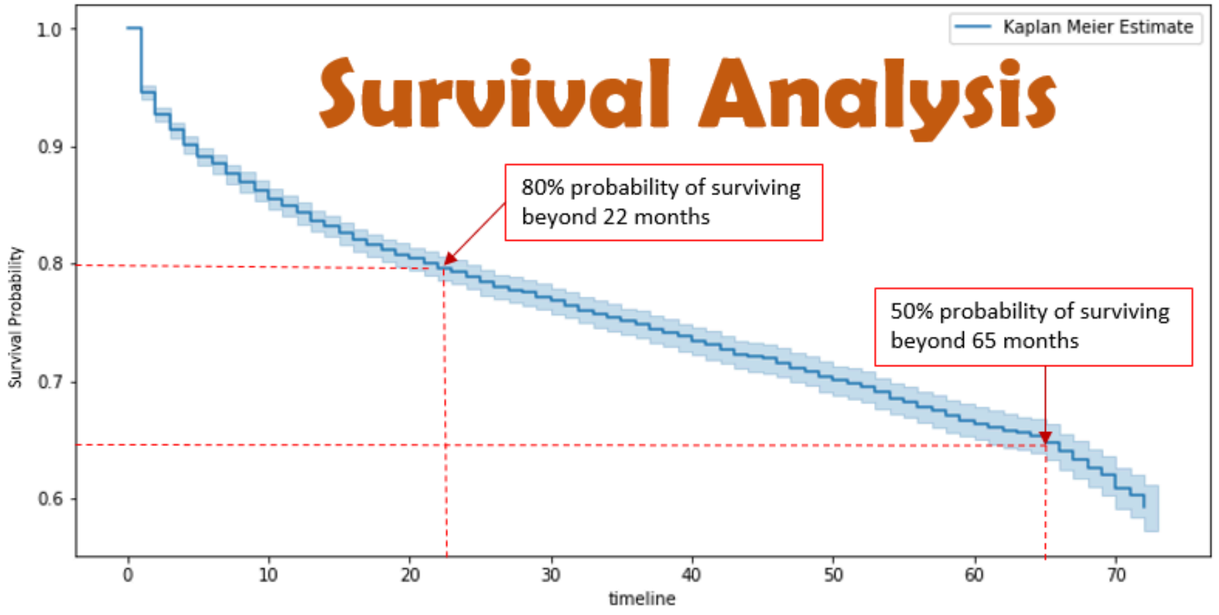


Survival (months)

How do we study survival statistics?

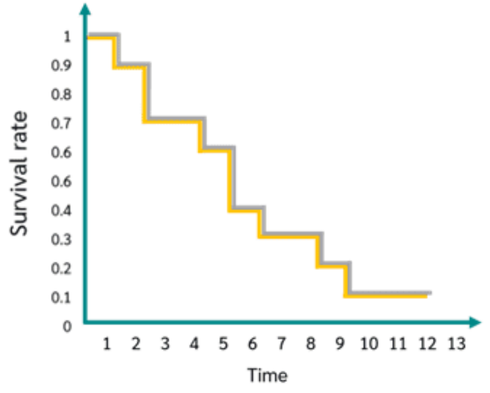
Dimensions of our data





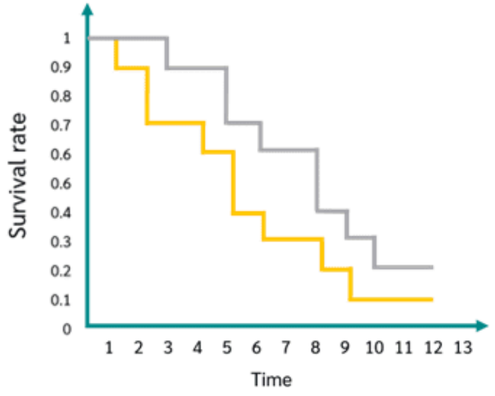
Null hypothesis:

The groups have identical distribution curves.



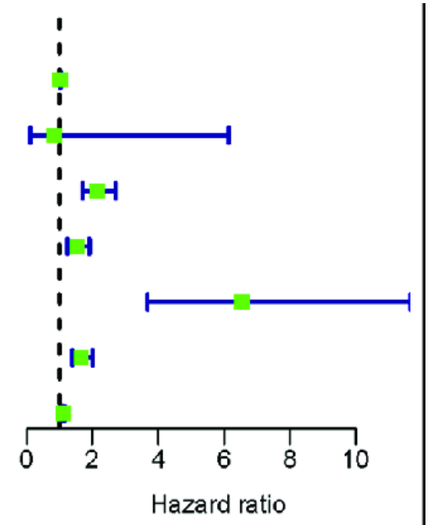
Alternative hypothesis:

The groups have different distribution curves.



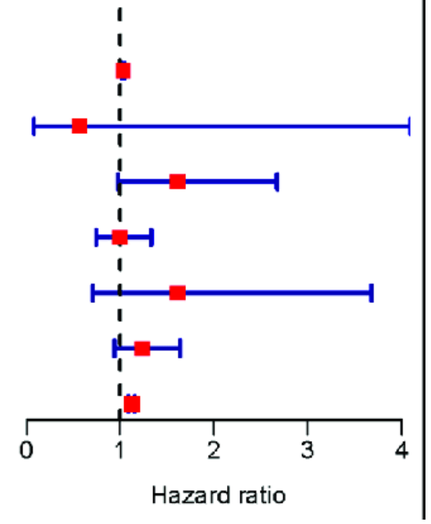
A

	pvalue	Hazard ratio
Age	<0.001	1.033(1.019−1.048)
Gender	0.876	0.855(0.119−6.130)
Stage	<0.001	2.151(1.703−2.717)
T	<0.001	1.553(1.254−1.924)
M	<0.001	6.544(3.675−11.654)
N	<0.001	1.667(1.391−2.000)
riskScore	<0.001	1.121(1.091−1.153)



B

	pvalue	Hazard ratio
Age	<0.001	1.034(1.019−1.049)
Gender	0.571	0.564(0.078−4.093)
Stage	0.063	1.614(0.975−2.672)
T	0.986	0.997(0.746−1.333)
M	0.256	1.612(0.707−3.678)
N	0.136	1.238(0.935−1.640)
riskScore	<0.001	1.124(1.091−1.158)

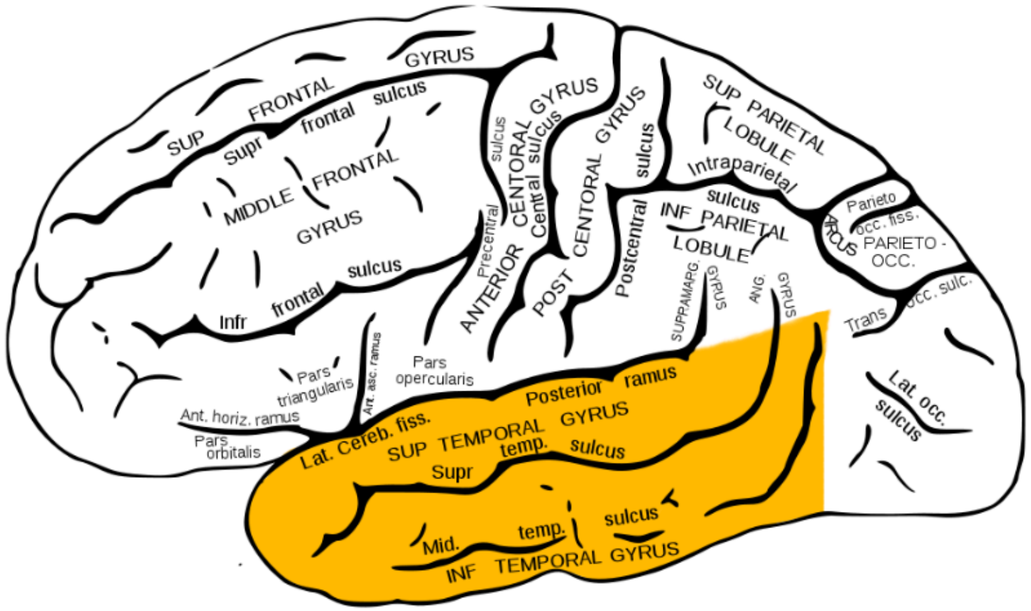


1 What do we want to know

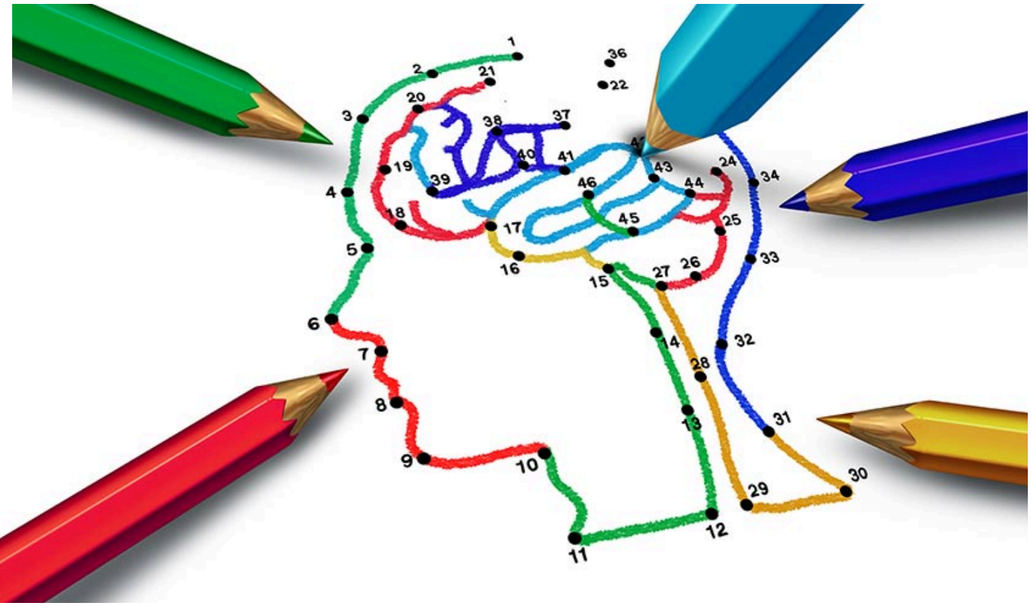
2 How will we assess it

3 What is our statistic plan

1



We want to investigate TLE and NPS functions

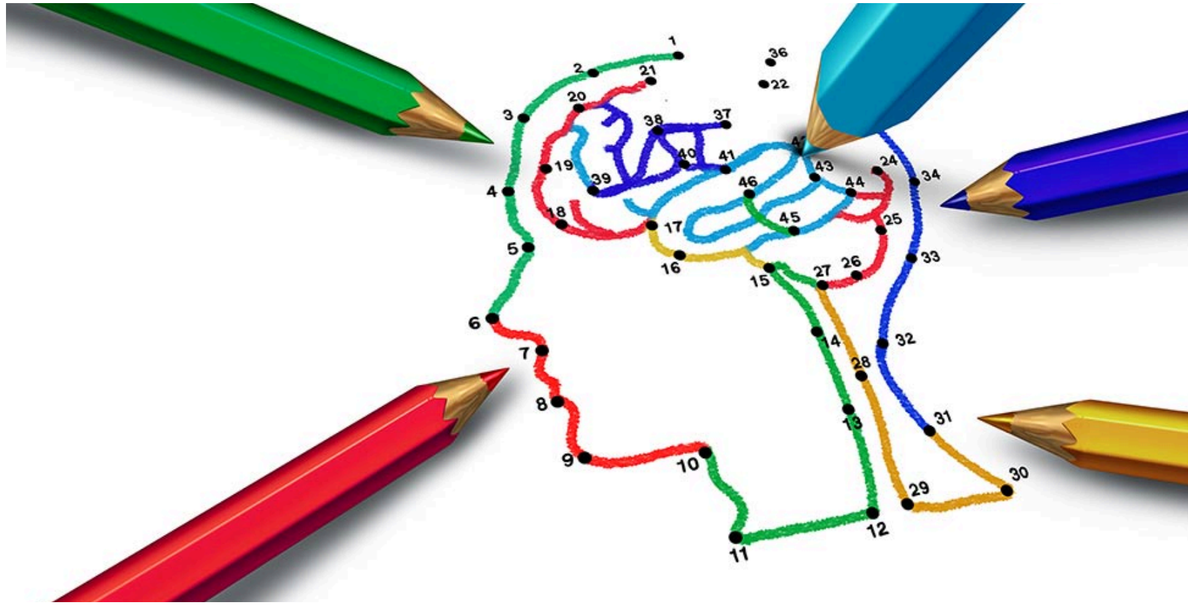


Does TLE affect NPS assessment in PwE

Are EEG alterations in the temporal lobe linked to NPS assessment in TLE patients.

2

HD-EEG
qEEG analisys
Connectivity



NPS test
EPI-Track
ACE-III

....

..

.

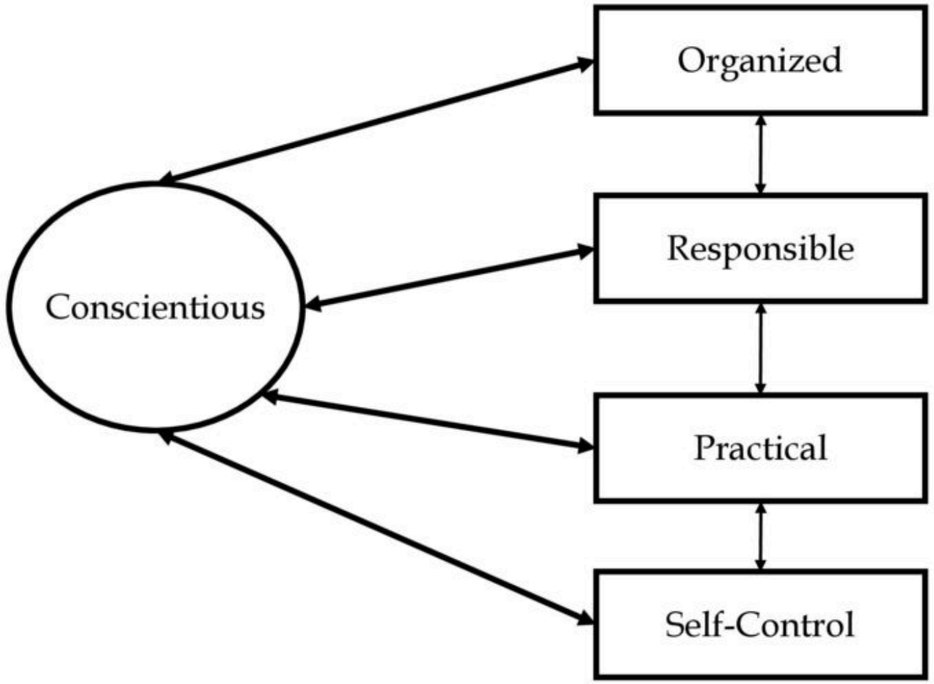
No plan is not an answer
Machine learning is generic and not an answer

Let’s talk about it and try to anticipate possible problems

Questionaires, tests and the concept of latent constructs

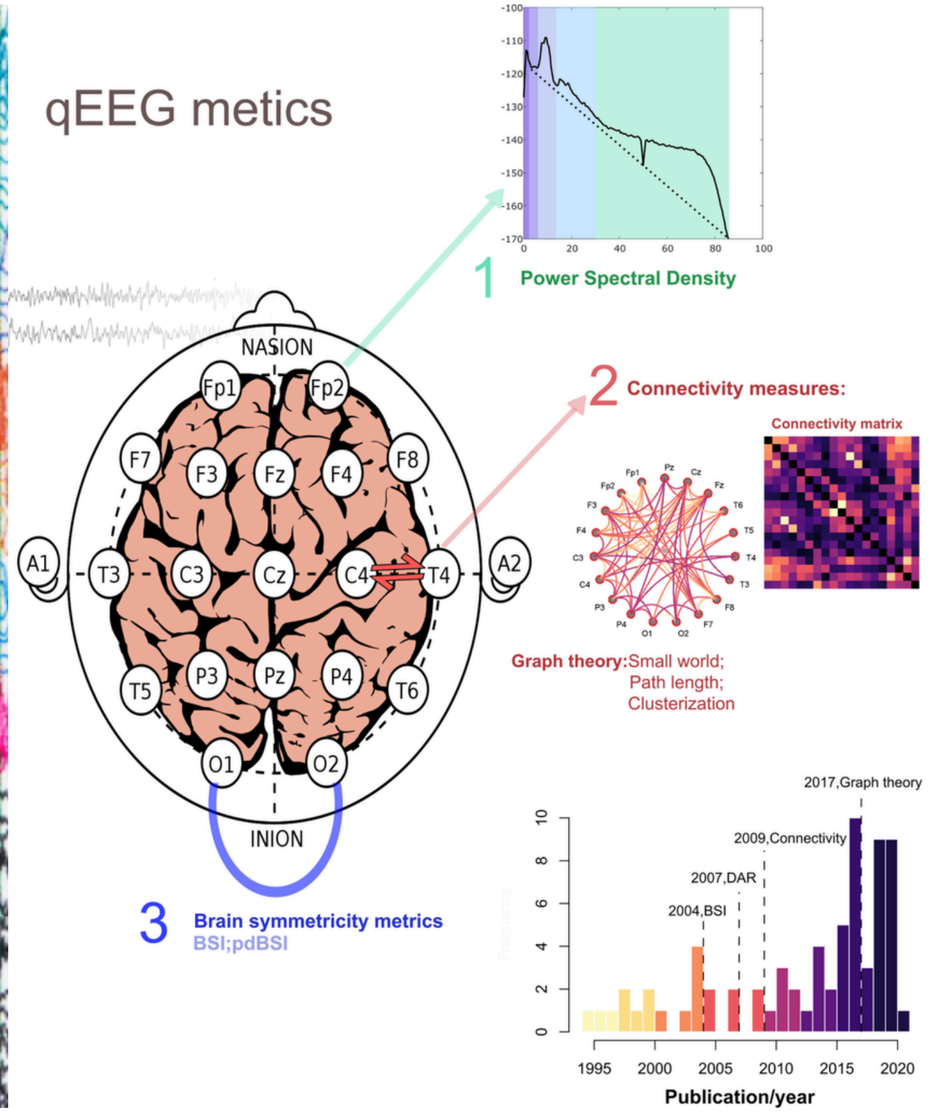
Latent Construct

- A construct defined by other constructs

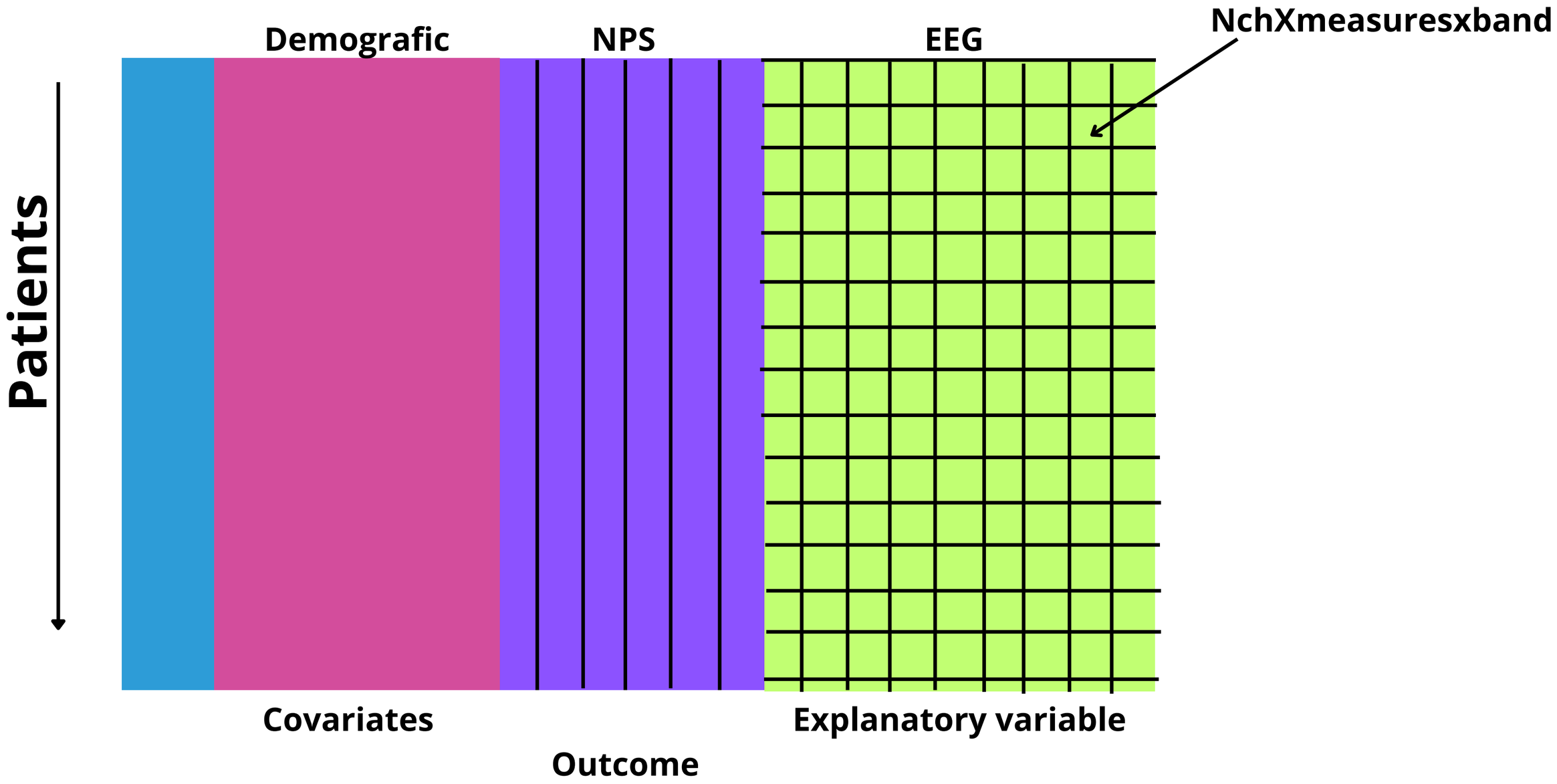


EEG Data Analysis

Benefits and Challenges



Dimensions of our data



Problems?

Multiple outcome measures
Large EEG dataset many entries
No clear classes to perform inference

Solutions?

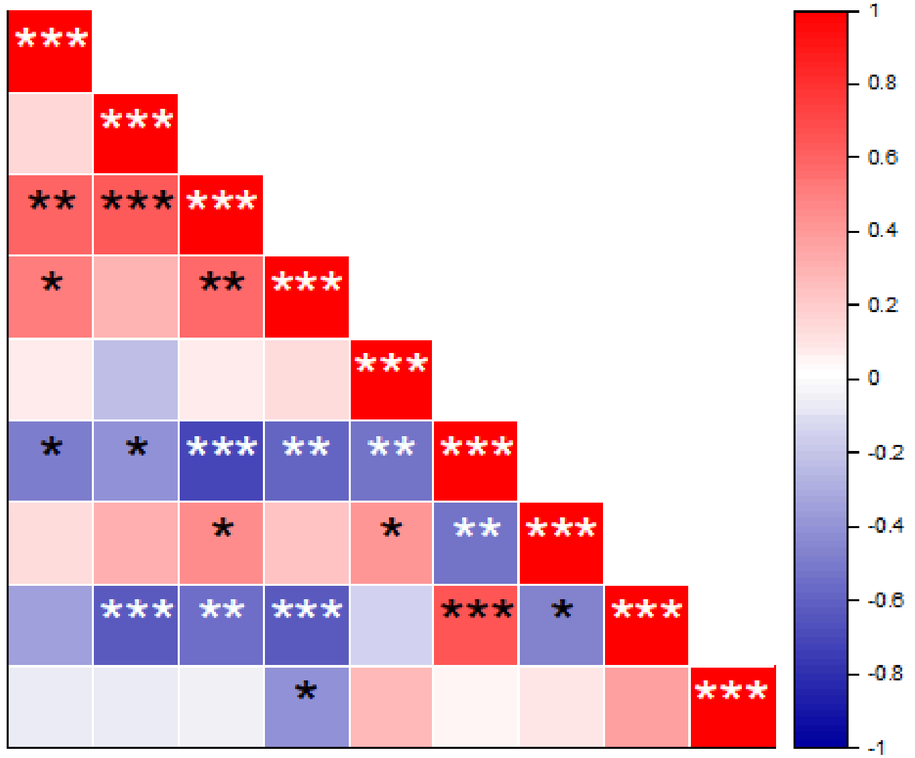
Collapse and reduce outcome measure on one recoded variable.

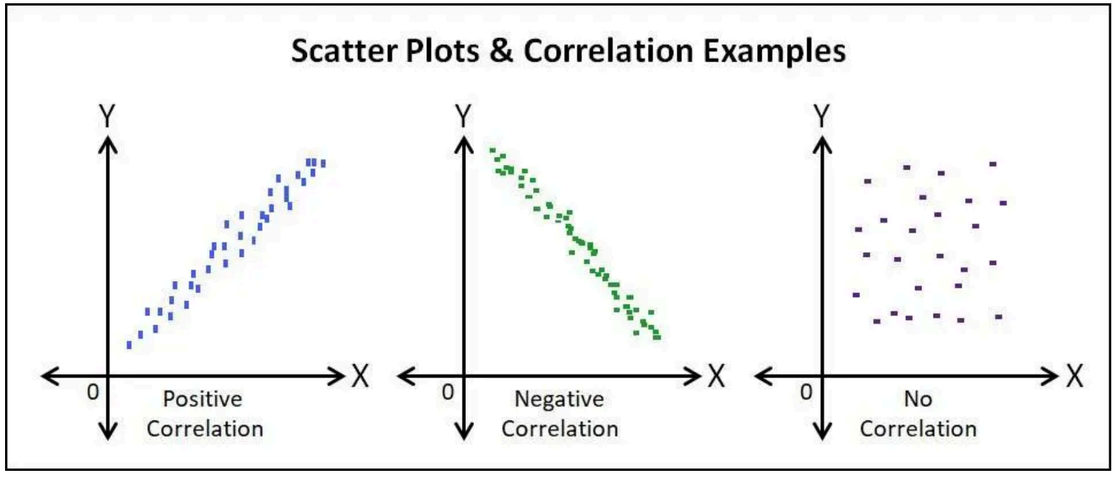
Use ML or formulate clear cut hypothesys and test it.

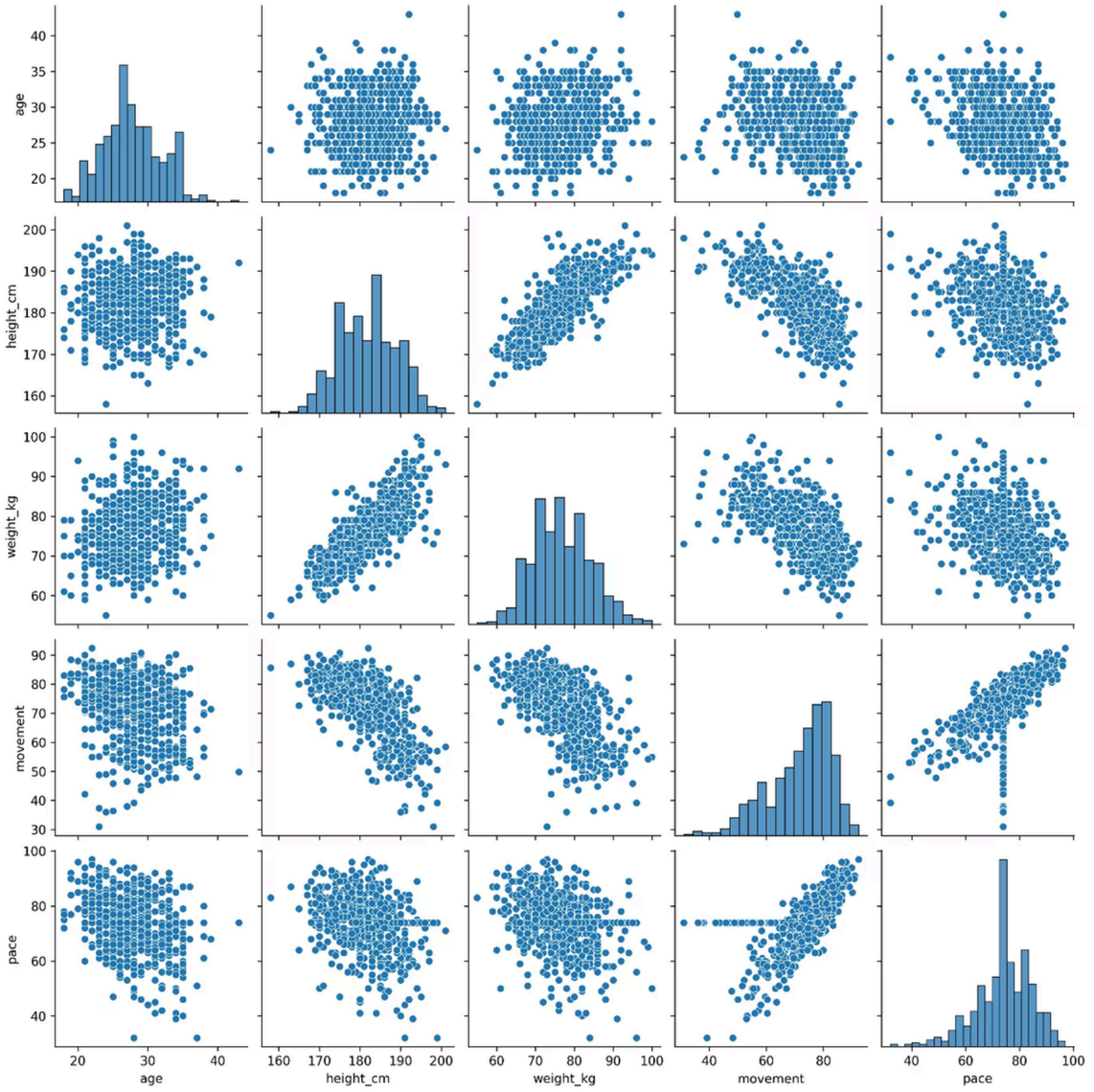
Use normal values of NPS test to define pathologic and non patologic and create multiple binary variables

What if i have to many desired outcomes?

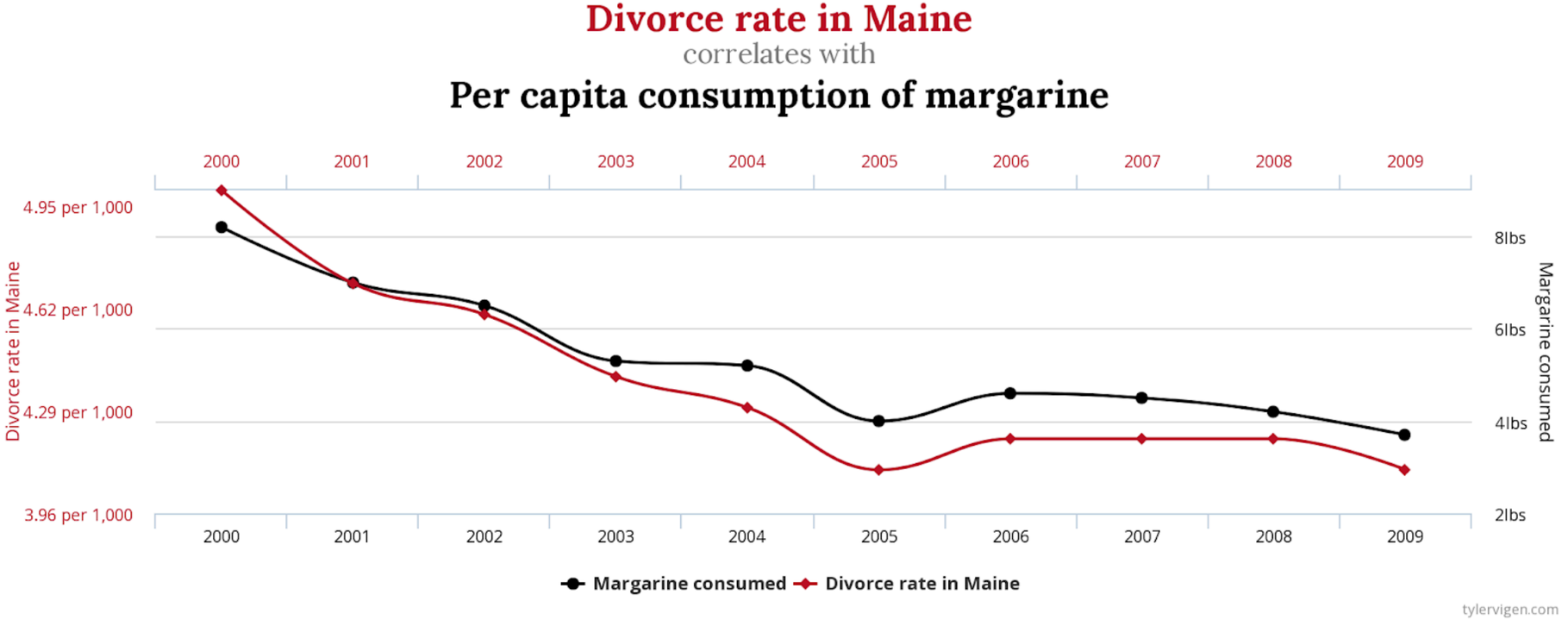
i.e. Quolie31, BDS-II, GAD, MMSE, ACE-R, FAB score


***
***
**
***
***
*
**
***
***
*
*
***
**
**
***
*
***
**
***
**
***
*
***
*
***
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

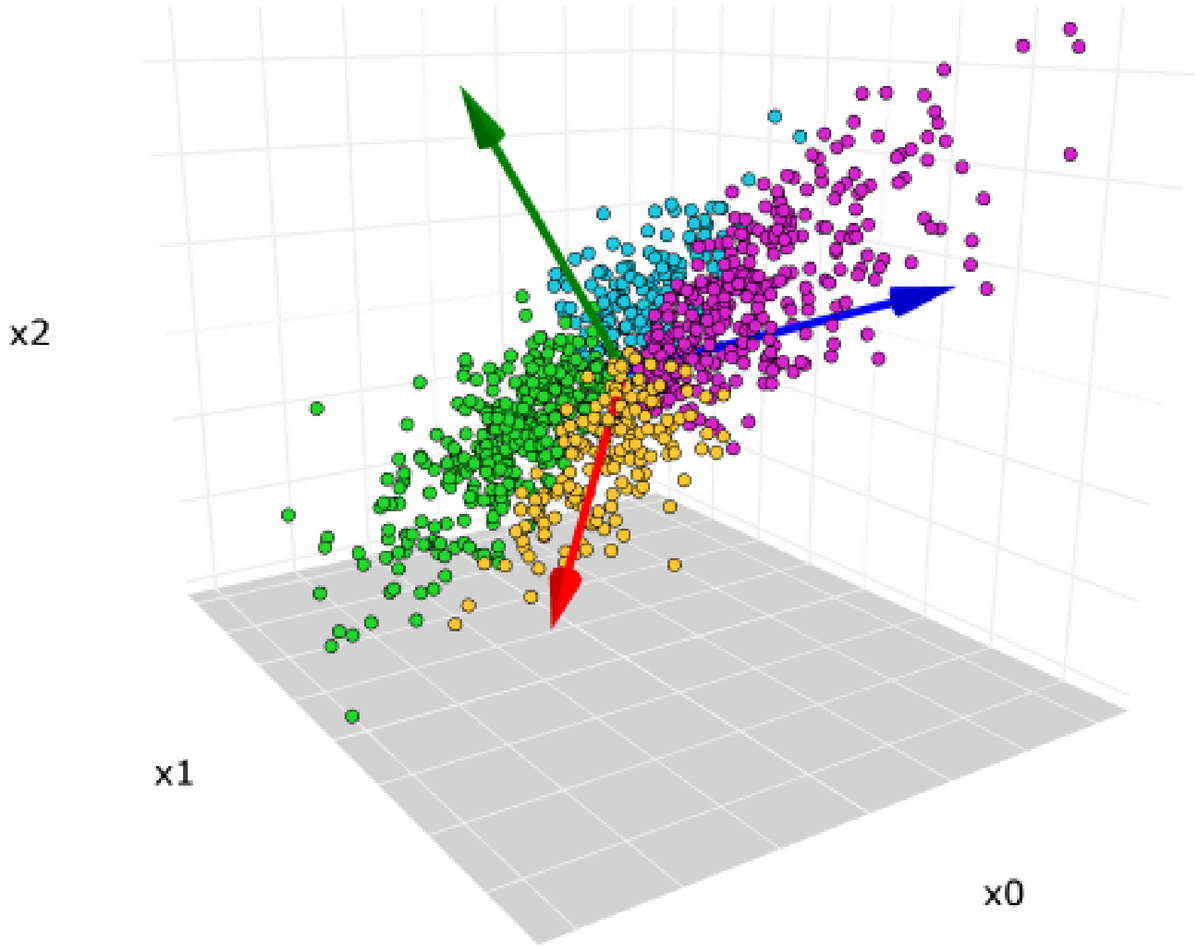

Scatter Plots & Correlation Examples
Y
X
0
Positive
Correlation
Y
X
0
Negative
Correlation
Y
X
0
No
Correlation

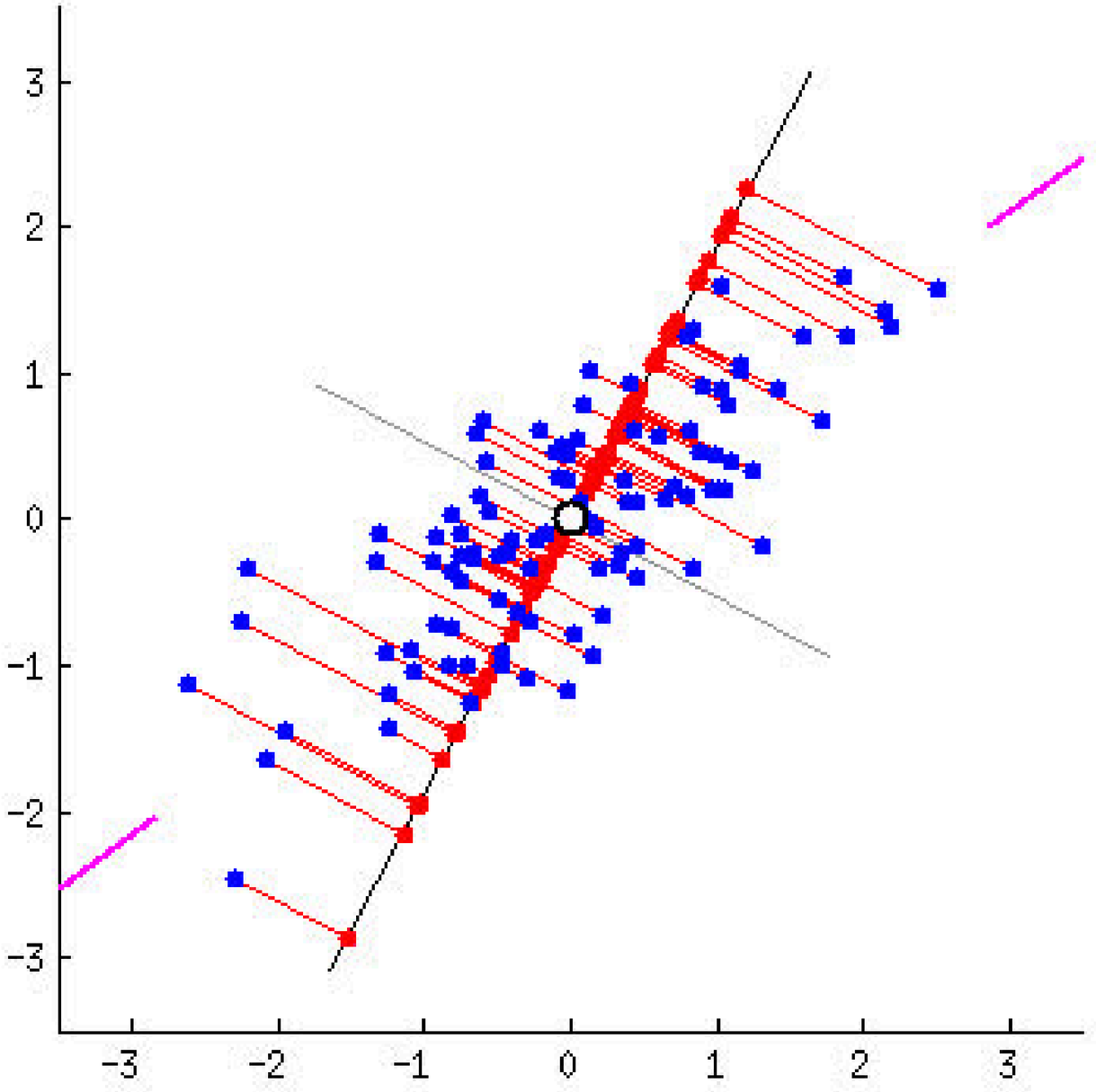

age
height_cm
weight_kg
movement
pace

Correlation does not imply causation

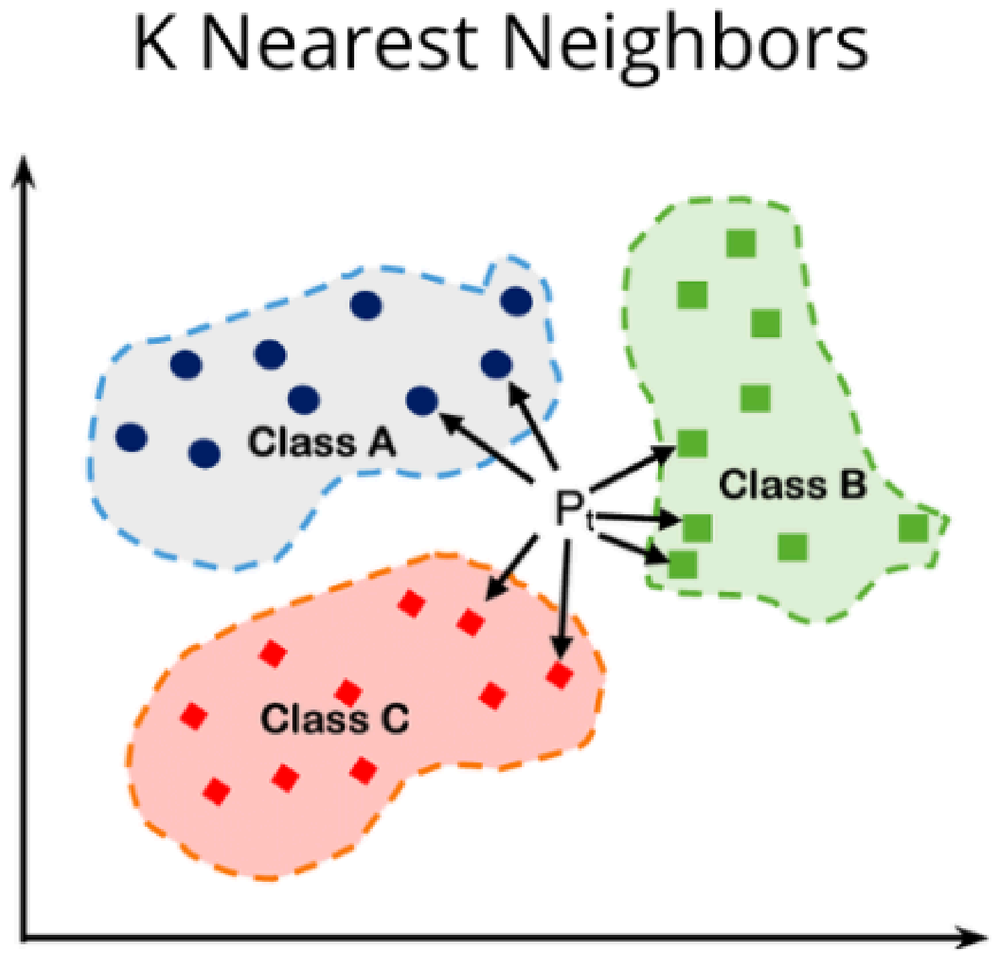


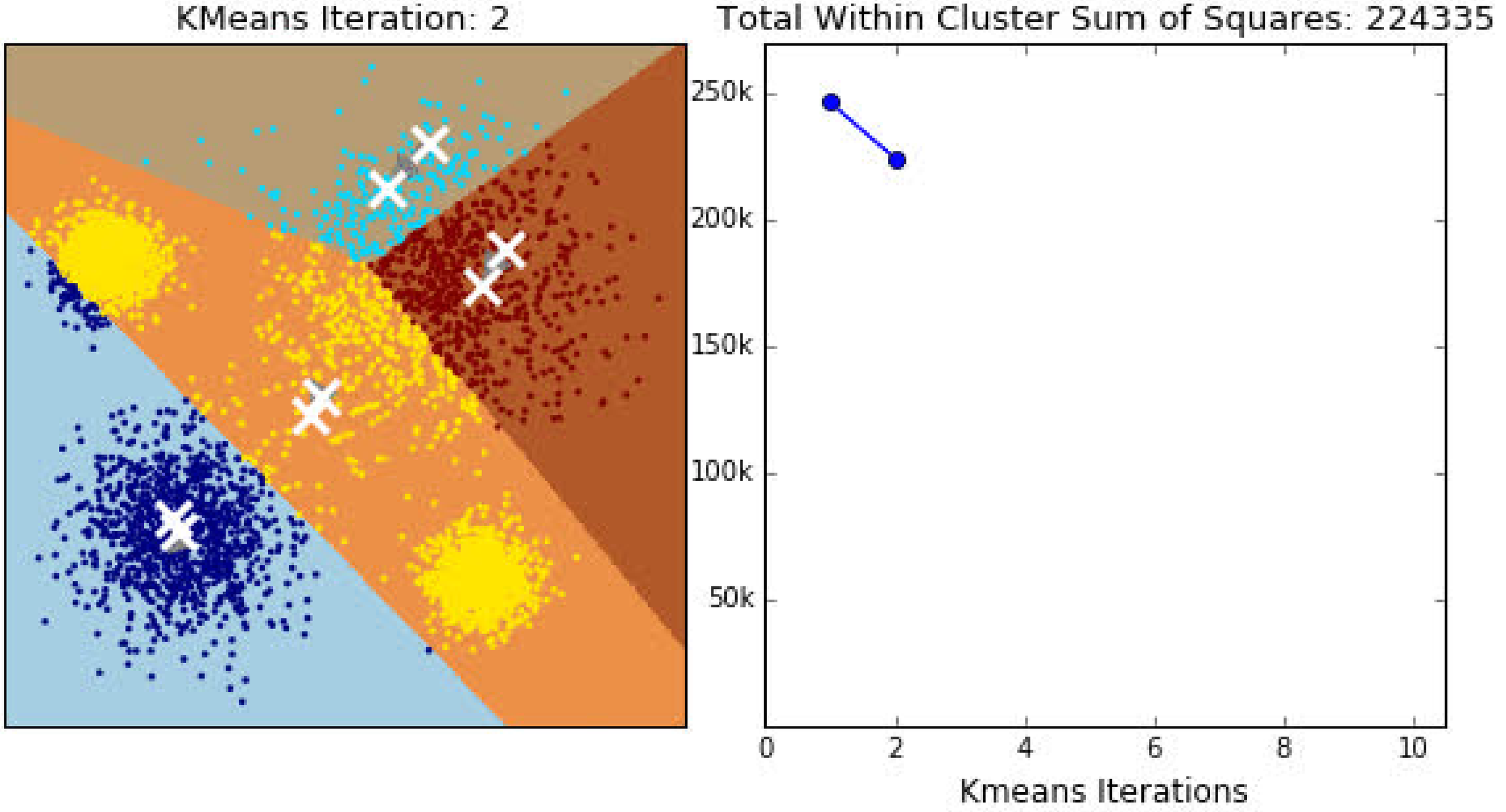
Use PCA to obtain 1 recoded variable that loads most of the explained variance.
We can have a single variable continuouse that summarizes NPS tests useful as an outcome for ML




3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2
3

Use KNN to obtain recoded classes .
We can obtain classes of test performance




KMeans Iteration: 2
Total Within Cluster Sum of Squares: 224335
250k
200k
150k
100k
50k
0
2
4
6
8
10
Kmeans Iterations

What if i have to many explanatory variables?

i.e. clinical variables, EEG (ChannelxMeasurexBand)

Go narrow

Perform a clear Hypothesis based on data and test it.

i.e People with pathological Epitrack will manifest higher Theta and Delta Power

Scientific and clean approach

Multiple hypothesis and cherry-picking is not very scientific

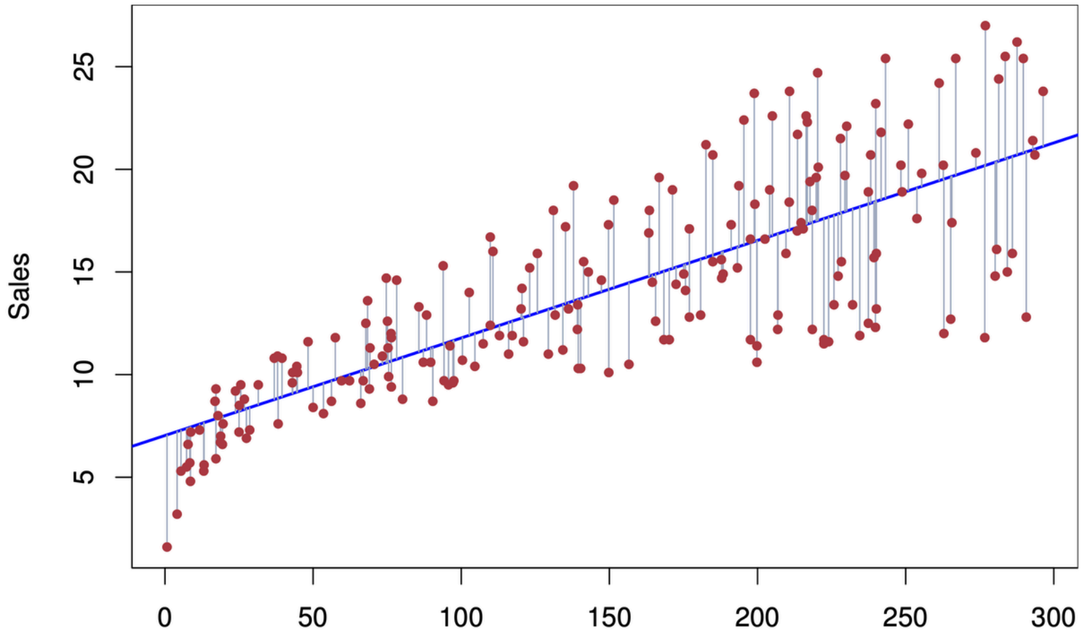
Go Wide

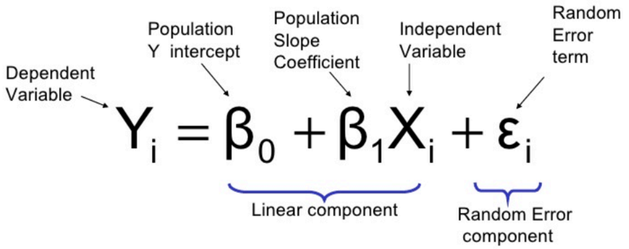
Develop a comprehensive model.

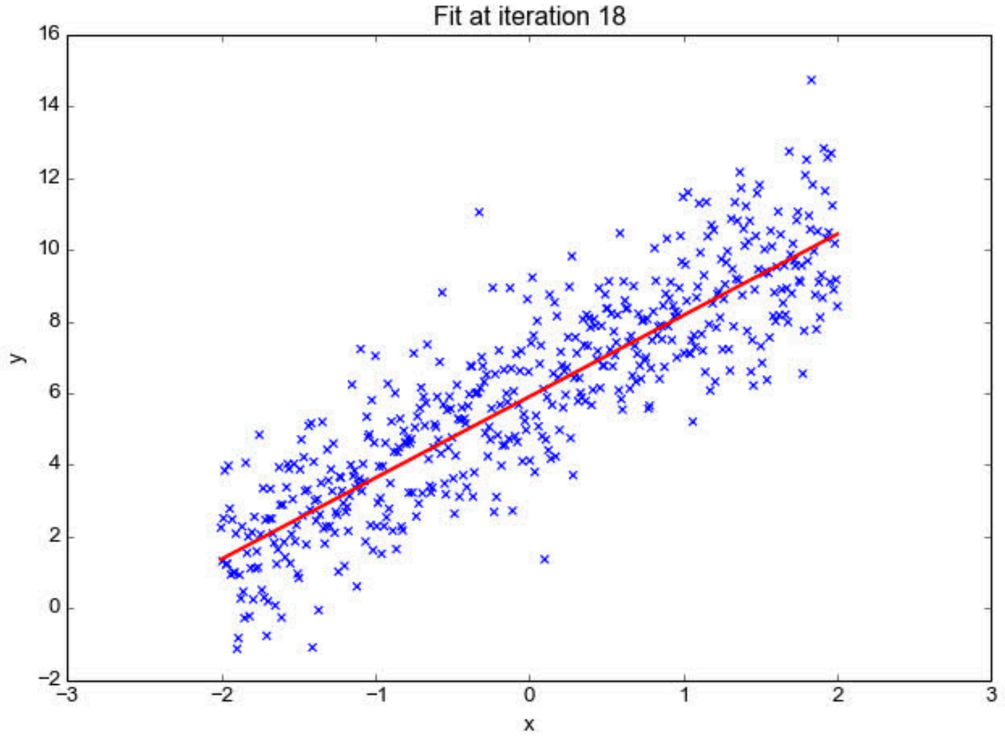
i.e what are the main EEG predictors of my ddesired outcome using machine learning.

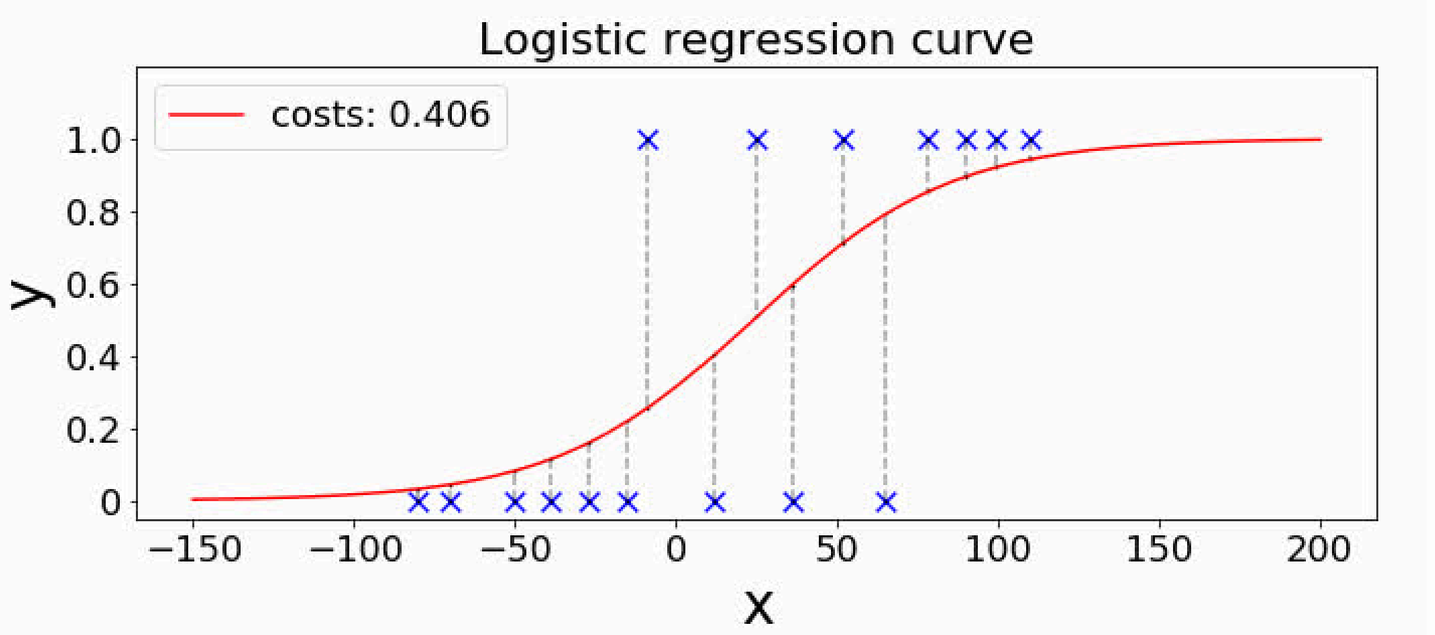
Comprehensive and avoids the problem of multiple comparisons

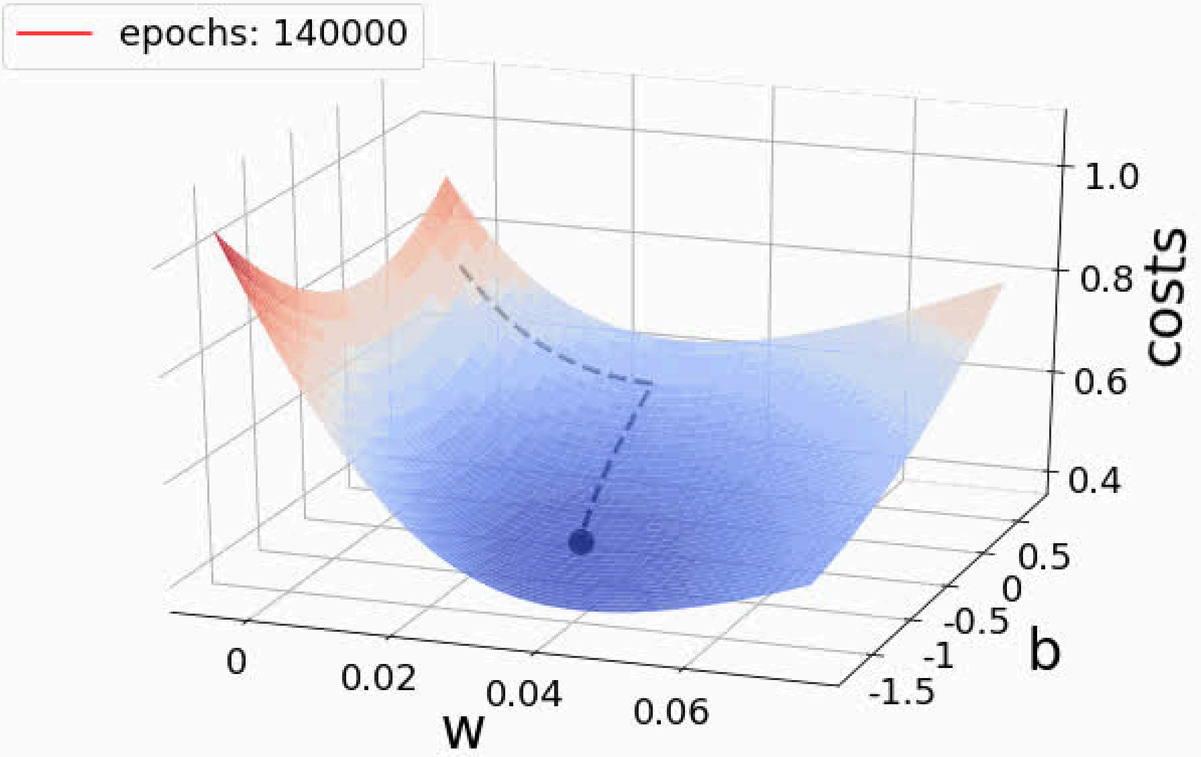
Overfitting can be a serious problem

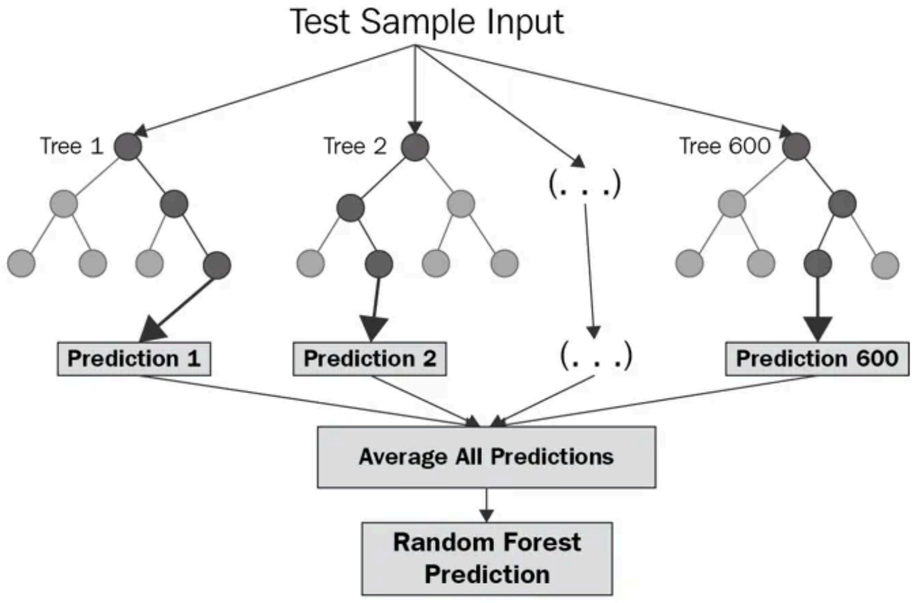

Sales
5
10
15
20
25
0
50
100
150
200
250
300

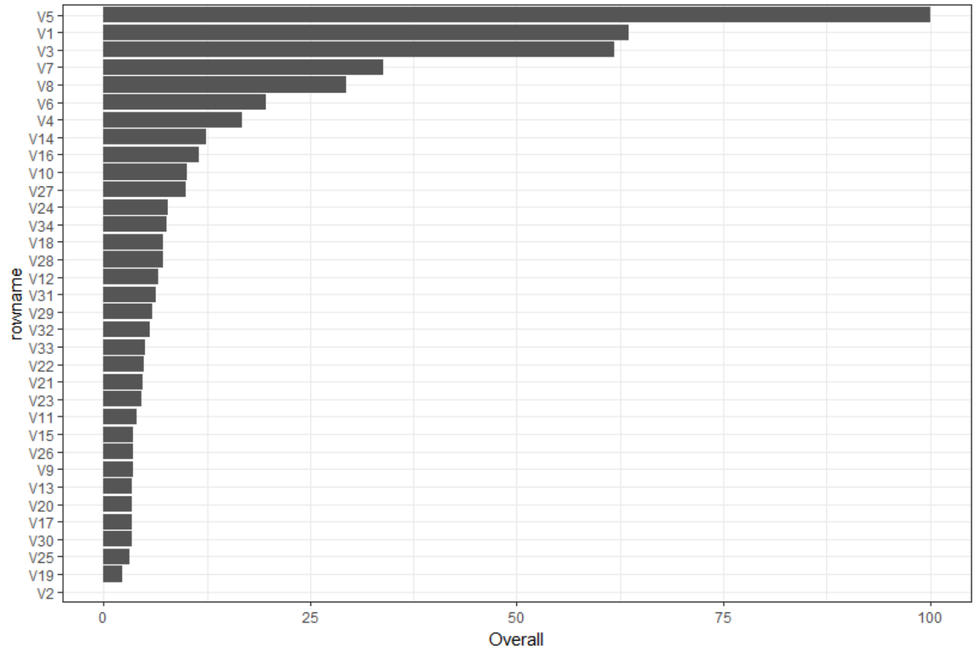

Dependent Variable
Population Y intercept
Population Slope Coefficient
Independent Variable
Random Error term
$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$
Linear component
Random Error component

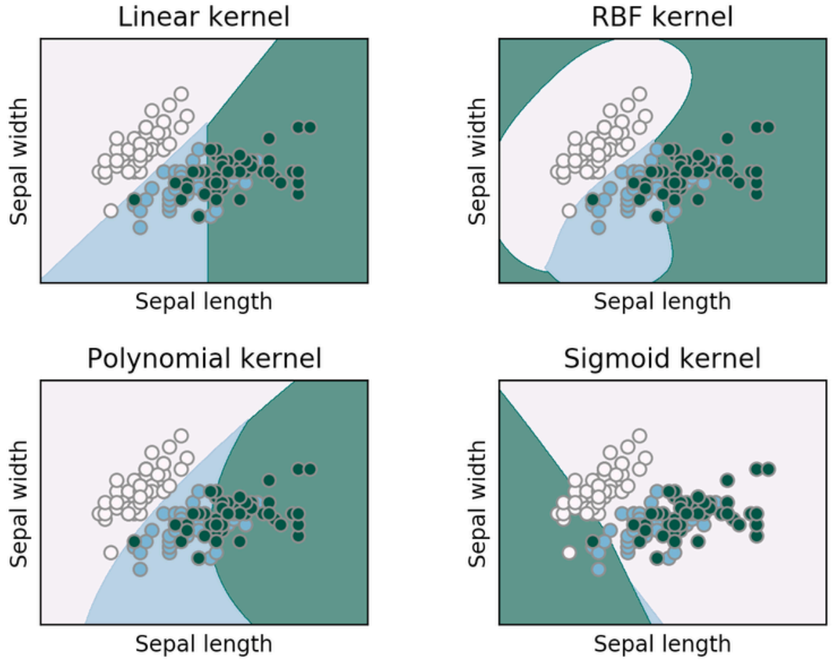

Fit at iteration 18
y
x
16
14
12
10
8
6
4
2
0
−2
−3
−2
−1
0
1
2
3

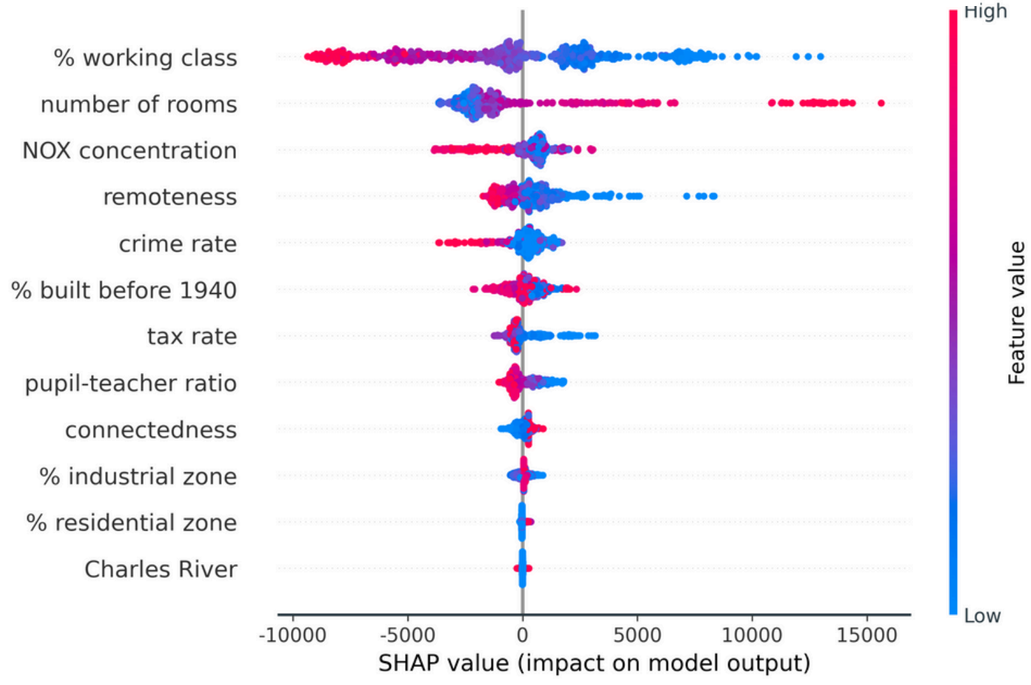

Logistic regression curve
costs: 0.406
1.0
0.8
0.6
0.4
0.2
0
y
−150
−100
−50
0
50
100
150
200
x

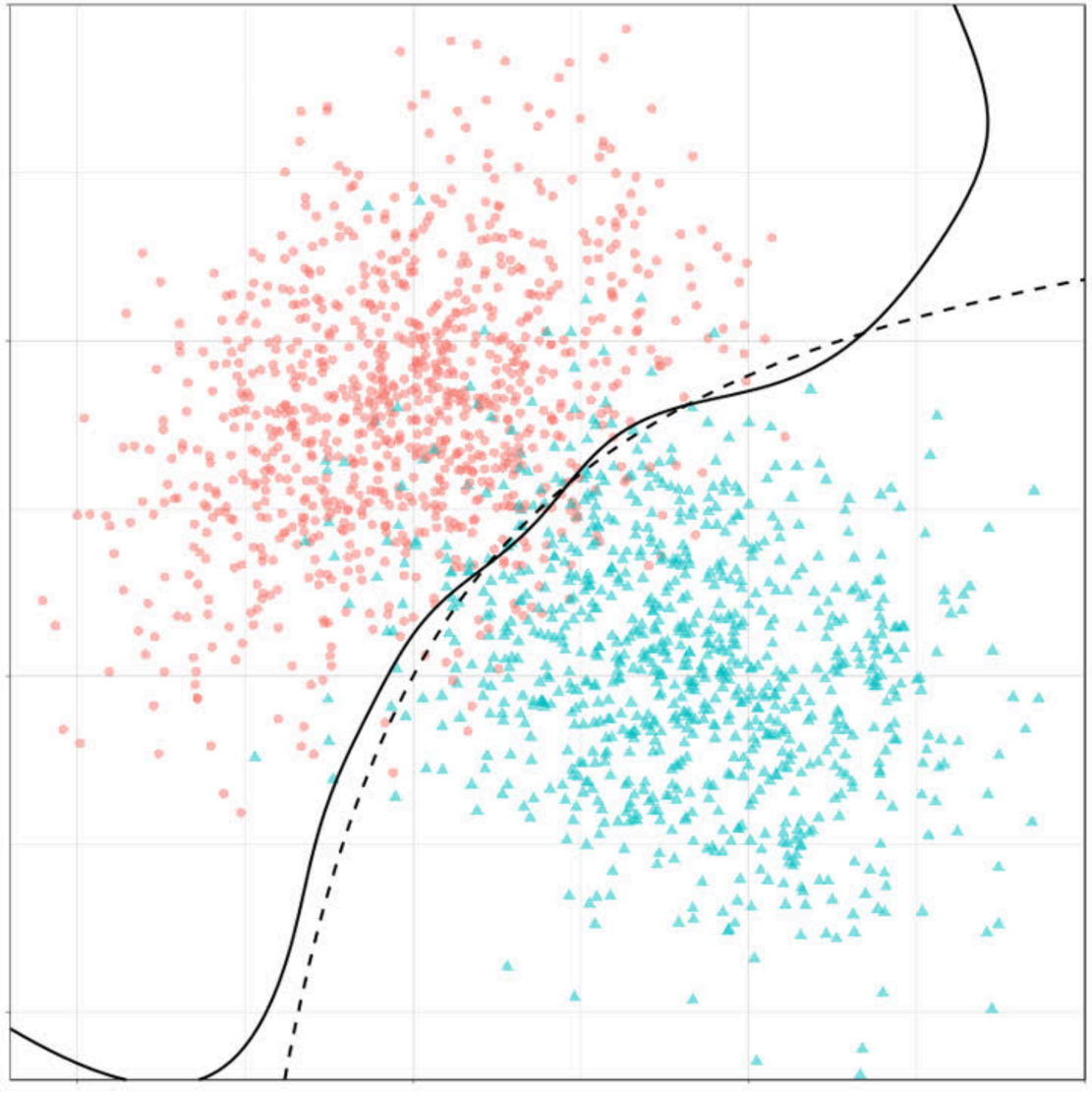

epochs: 140000
1.0
0.8
0.6
0.4
costs
0
0.02
0.04
0.06
w
0.5
0
-0.5
-1
-1.5
b

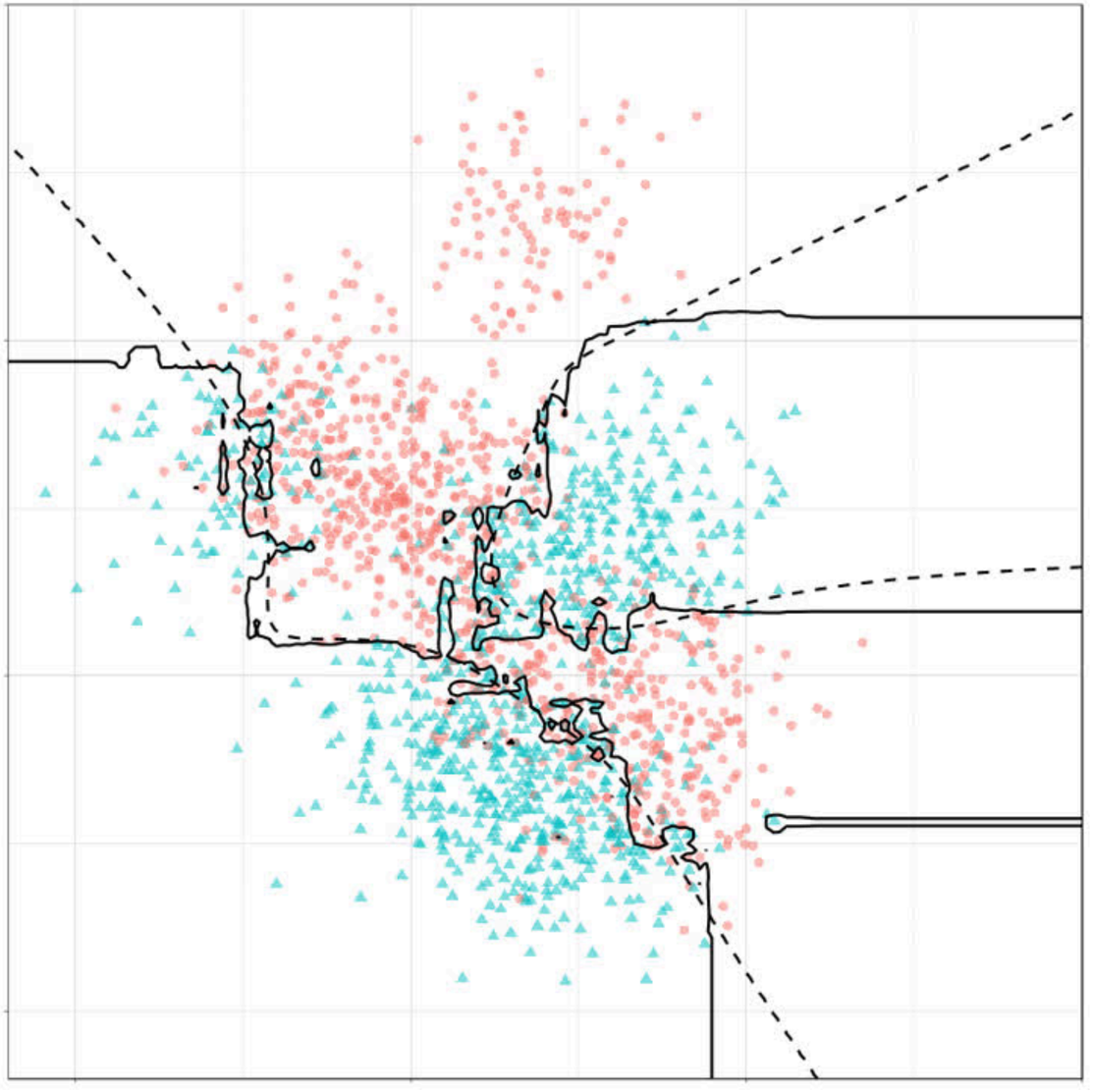

Test Sample Input
Tree 1
Tree 2
(. . .)
Tree 600
Prediction 1
Prediction 2
(. . .)
Prediction 600
Average All Predictions
Random Forest Prediction


rowname
V5
V1
V3
V7
V8
V6
V4
V14
V16
V10
V27
V24
V34
V18
V28
V12
V31
V29
V32
V33
V22
V21
V23
V11
V15
V26
V9
V13
V20
V17
V30
V25
V19
V2
0
25
50
75
100
Overall


Linear kernel
RBF kernel
Polynomial kernel
Sigmoid kernel
Sepal width
Sepal length


% working class
number of rooms
NOX concentration
remoteness
crime rate
% built before 1940
tax rate
pupil-teacher ratio
connectedness
% industrial zone
% residential zone
Charles River
-10000
-5000
0
5000
10000
15000
SHAP value (impact on model output)
High
Low
Feature value





The perfect database

Avoid redundancy: grow in complexity not just in size

Use unique identifiers (i.e Statepi_JL_exp1_T1)

Atomicity: scorporate complex data (i.e LEV1000mg= ASM->LEV, dose->1000mg/day)

Avoid NAs (in large datasets consider imputation)

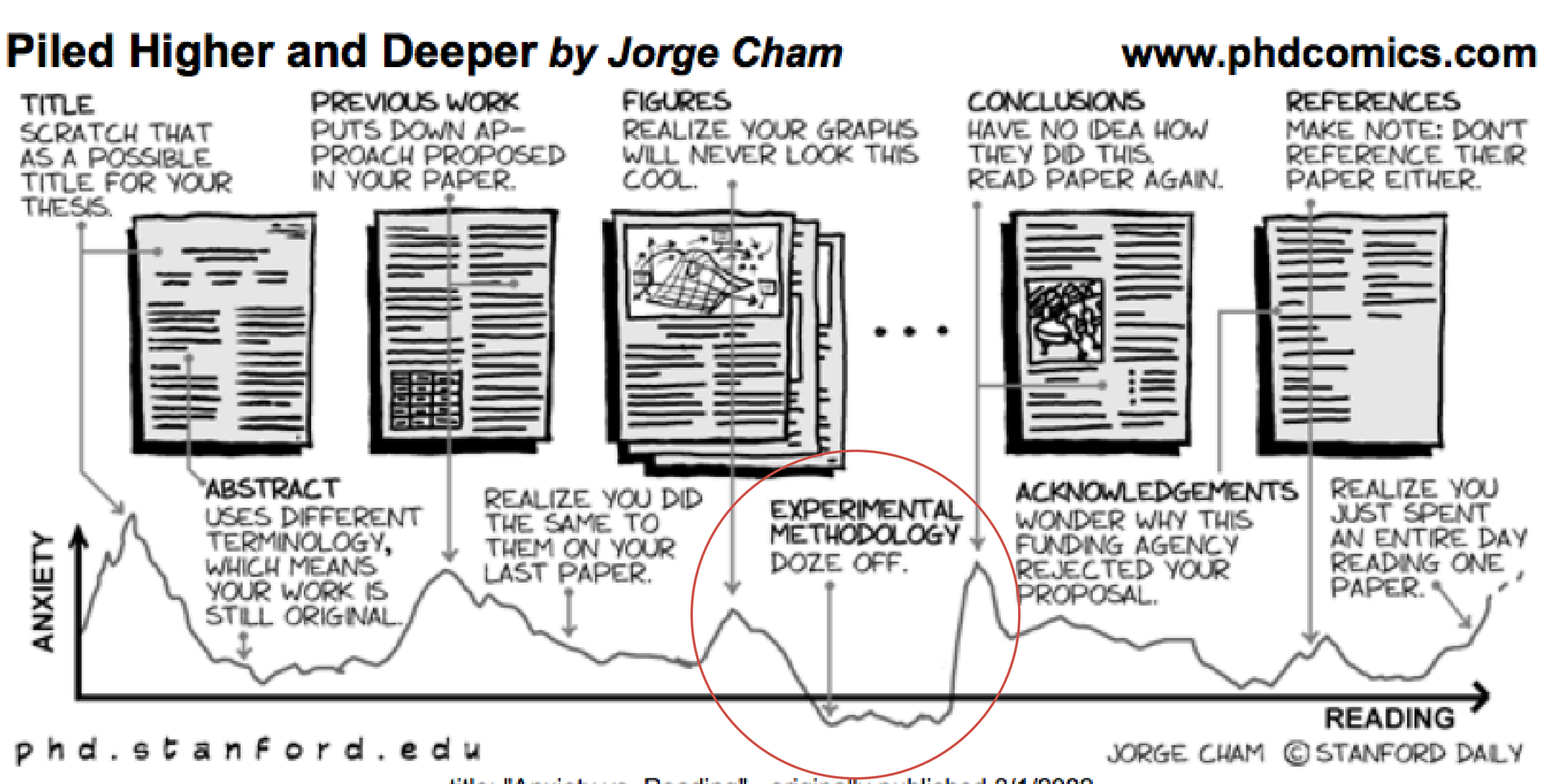
Mitigate typing errors (multiple choices, ore use redcap)

Use easy index variables(index variables are at the heart of statistics, create groups to test, here is where you can think and come up with interesting things)

Do not be afraid to simplify (“Seizure with LOC and subtle clonic jerks of the pinky of the right foot” is not a variable LOC(Y/N) is a varianle)

One subject/One condition/One time = one ROW

Back up the Back ups of your Back up



title: "Anxiety vs. Reading" - originally published 3/1/2002

Thank you