Experiments

Experiments

Collection types	Distinctions	Examples
Physical places, online	• Groups of subjects clearly separated • Groups depends on conditions • Isolating from all unneeded causes to influence results • Quantitative in nature but can also produce quantitative results	• Experiment on which styling attracts users attention to commercials • Testing medicine

Experiment is a way to see what impact one feature (condition/change) does

1. Change can come from subjects (within human nature – working environment, demographics).
   - Subjects assign to groups according to the feature and given same task, object to take
   - Prior to experiment information is collected to ensure that each group is unique
2. Change can come from experiment designers (condition of the task to perform, medicine to test)
   - Subjects assign to groups by random. And each group gets different feature to perform
   - Prior to experiment information is collected to ensure that all groups are similar

Groups

- must be isolated
- Conditions/environement should be the same, except related to the task for both groups

Results can be collected:

1. by survey (Likert scale usually, can be open questions)
2. measuring performance (time spent, questions answered, imaginary money won)

Tested:

- Usually by T test to ensure, whether results has significant difference per groups

Experiment Design

1x2

Condition	Test Group	Control Group
Burana affect on sleep	10 subjects (gets Burana)	10 subjects (placebo)

2x2

Condition	Test Group	Control Group
Commercial in blue background	10 subjects (sees blue color)	10 subjects (sees white color)
Commercial in red background	10 subjects (sees red color)	10 subjects (sees white color)

1x3

Condition	Test1 Group	Test2 Group	Control group
Coding interviews with software	10 subjects (Atlas.ti)	10 subjects (MS Word)	10 subjects (uses pen and paper)

What have we done in our experiment?

We tried to answer research question, how to influence better problem solving:

Main hypotheses: if users will get proper education how to tackle problems they will do better in problem solving

Created three groups

Test 1: To give some activity

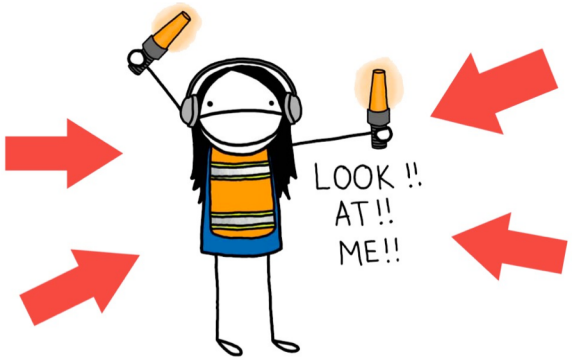
Test 2: To give problem solving information

Control: To give nothing

Our main concentration on “Test2” as its closest to our research question.

But results needs to be isolated and compared:

Maybe its not knowledge but brain activation that allowed to solve problems in experiment (**Test1 group**)

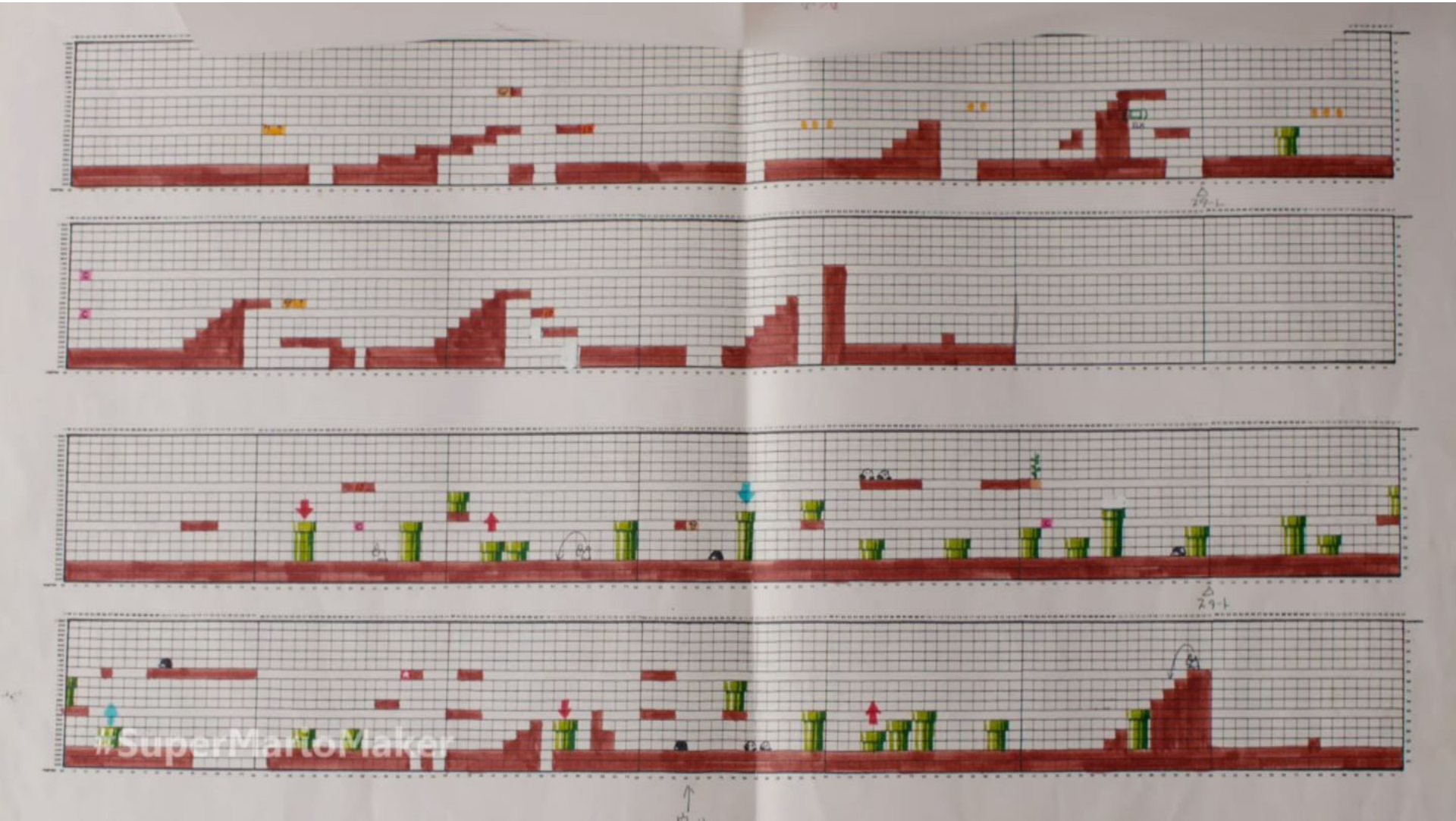


Maybe its attention that users receive, and that’s way they put more efforts (**Control group**)

The point is to isolate environment for testing one single feature

Exploratory statistics

How to get a holistic view in one picture?



Through statistics:

We want to understand our data without looking to each particular unit (or none of them at all) – distant reading.

Who are these people?

What do they do?

What do they talk about?

What do they have in common?

Data Types

Categorical:

- Nominal (Male,Female; Green, Black)
- Ordinal (A,B; First, Second;)

Continuous

Categorical

Nominal – order is not important

- Male = Female
- Green = Black

Ordinal– order is important

- First ≠ Second
- Elements can be grouped under categories

  Male's height in cm: small[0-170], average[170-190], high[190-230]

Continuous

Interval – a number without boundaries

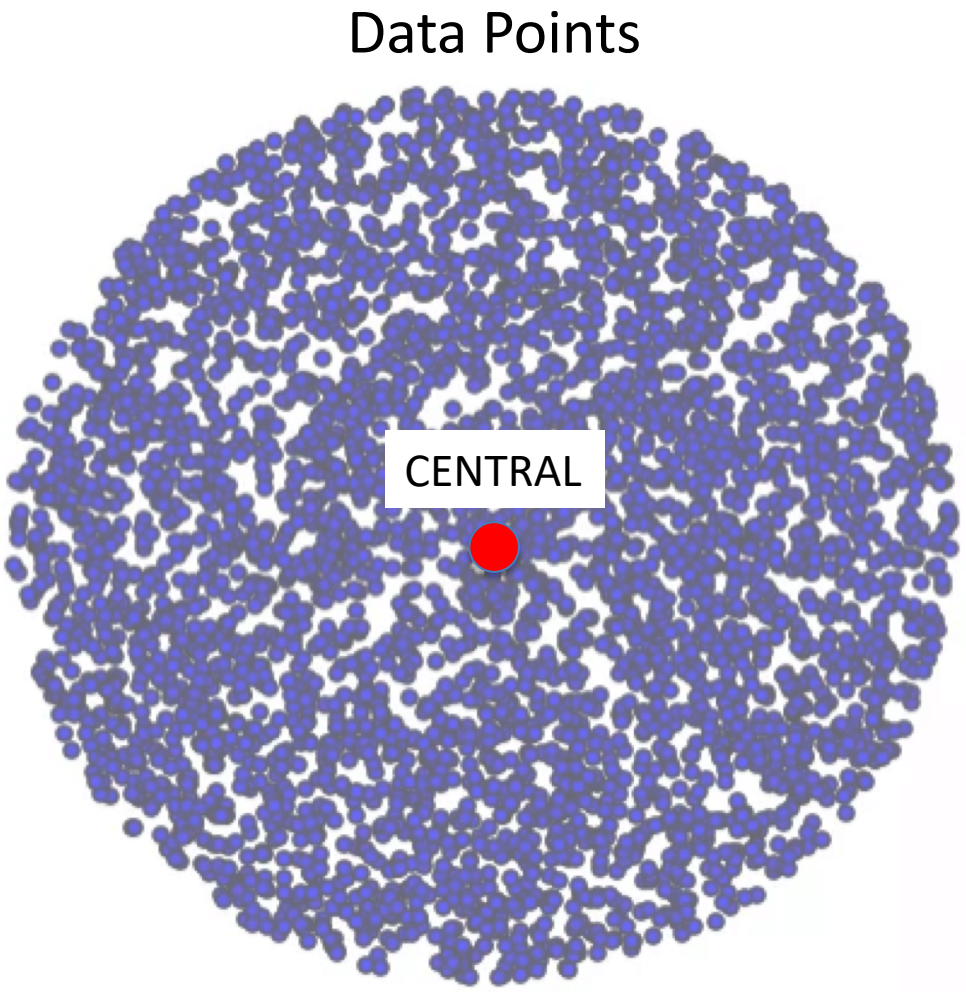
Person’s height, Competition score

	A	B	C	D	E	F	G
1	ID	Last Name	First Name	City	State	Gender	Student Status
2	1	DOE01	JANE01	Los Angeles	California	Female	Graduate
3	2	DOE02	JANE02	Sedona	Arizona	Female	Undergraduate
4	3	DOE01	JOE01	Elmira	New York	Male	Graduate
5	4	DOE02	JOE02	Lackawana	New York	Male	Graduate
6	5	DOE03	JOE03	Defiance	Ohio	Male	Graduate
7	6	DOE04	JOE04	Tel Aviv	Israel	Male	Graduate
8	7	DOE05	JOE05	Cimax	North Carolina	Male	Graduate
9	8	DOE03	JANE03	Liberal	Kansas	Female	Undergraduate
10	9	DOE04	JANE04	Montreal	Canada	Female	Undergraduate
11	10	DOE05	JANE05	New York	New York	Female	Graduate
12	11	DOE06	JOE06	Hot Coffe	Mississippi	Male	Undergraduate
13	12	DOE06	JANE06	Java	Virginia	Female	Graduate
14	13	DOE07	JOE07	Varna	Bulgaria	Male	Graduate
15	14	DOE08	JOE08	Moscow	Russia	Male	Graduate
16	15	DOE07	JANE07	Drunkard Creek	New York	Female	Undergraduate

H	I	J	K	L	M	N
Major	Country	Age	SAT	Average score (grade)	Height (in)	Newspaper readership (times/wk)
Politics	US	30	2263	67	61	5
Math	US	19	2006	63	64	7
Math	US	26	2221	78	73	6
Econ	US	33	1716	78	68	3
Econ	US	37	1701	65	71	6
Econ	Israel	25	1786	69	67	5
Politics	US	39	1577	96	70	5
Politics	US	21	1842	87	62	5
Math	Canada	18	1813	91	62	6
Math	US	33	2041	71	66	5
Econ	US	18	1787	82	67	3
Math	US	38	1513	79	59	5
Politics	Bulgaria	30	1637	79	63	4
Politics	Russia	30	1512	70	75	6
Math	US	21	1338	82	64	5

Central tendency

- Mean
- Median
- Mode



Mean

- Add and divide by number of data

$$\overline{X} = \frac{X_1 + X_2 + X_3 \ldots X_N}{N}$$

Where
$\overline{X}$ = the mean
X_1 = the first value
X_2 = the second value
X_3 = the third value
X_N = the last value
N = the number of valuse

$$\text{Average} = \frac{15 + 45}{2}$$

$$= \frac{60}{2}$$

$$= 30$$

Median

- Order the set of numbers, the median is the middle number

Even Number

9, 3, 1, 8, 4, 3, 6

Order

1, 3, 3, 4, 6, 8, 9

Choose middle

1, 3, 3, 4, 6, 8, 9

Uneven Number

9, 3, 1, 8, 3, 7

Order

1, 3, 3, 7, 8, 9

Choose middle

1, 3, 3, 7, 8, 9

/2

5

Mode

- The most common number

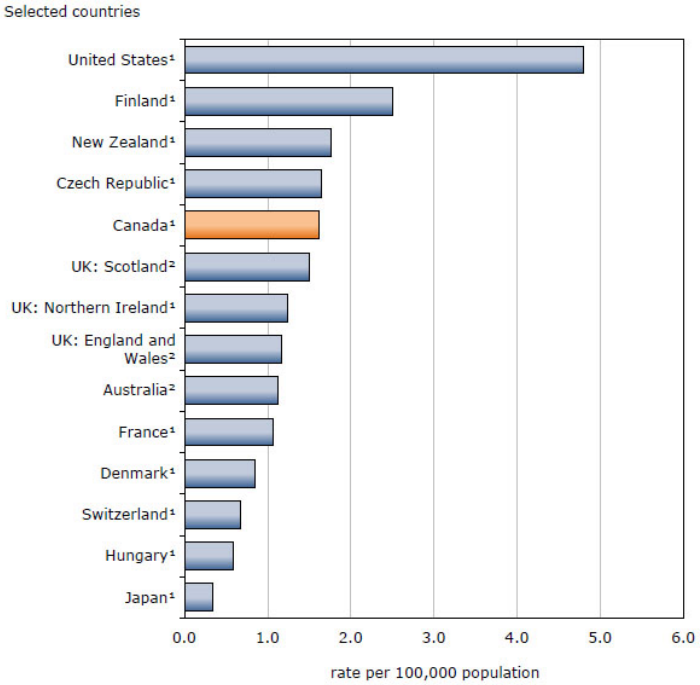
9, 3, 1, 8, 3, 7

1, 3, 3, 7, 8, 9

Describing nominal data

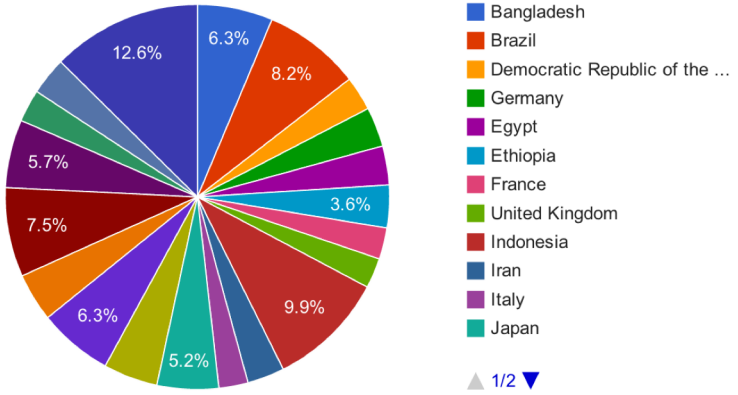
Calculations:

- Counting number of events,
- proportions,
- percentages



Visualizing

- Bar Charts
- Pie Charts



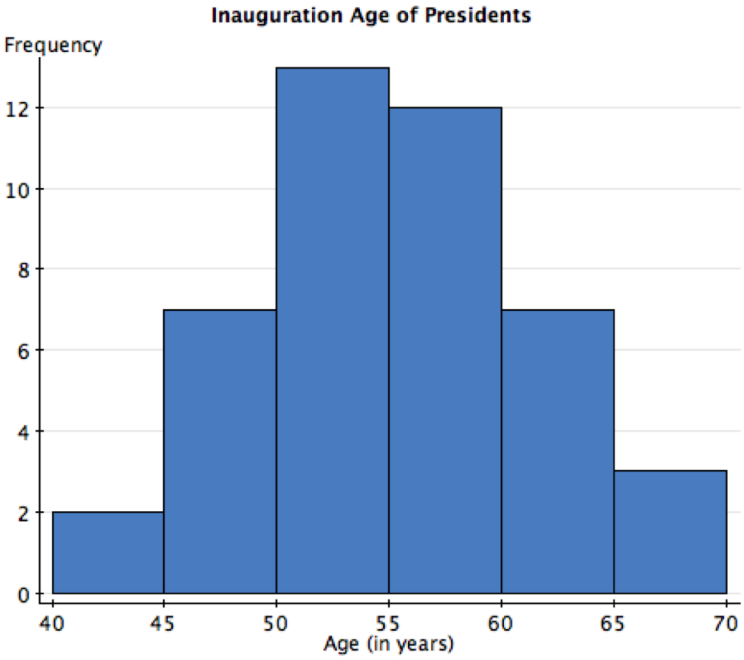
Describing continuous data

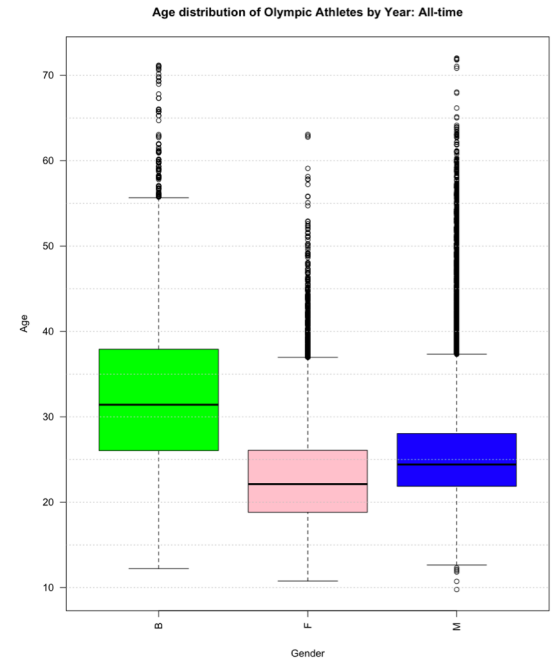
Calculations:

- Counting number of events,
- proportions,
- Percentages,
- Mode
- Median
- Interquartile range
- Mean
- Standard deviation
- Range

Visualizing

- Histogram
- Boxplot





Distance from the center

How much variability is in the group?
How spread is our data?
To what extent the group is similar

???Variability???

- Range
- Standard deviation
- Interquartile range

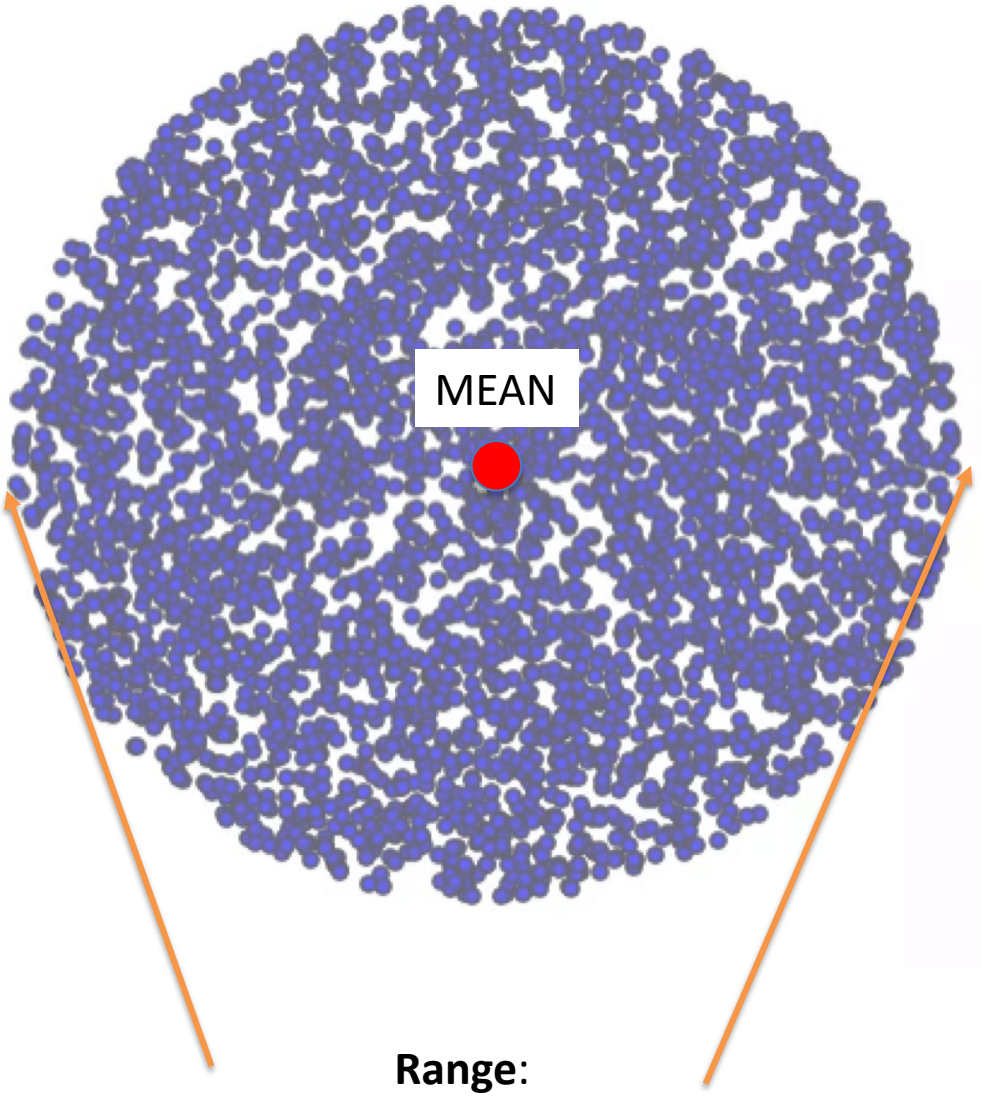
Range

- The difference between the highest and the lowest number

9, 3, 1, 8, 3, 7

9-1 = 8

Data Points



Range:
difference between min and max

(Range) People ages in a party

Party1

[15, 17, 18, 14, 13, 18 ,19]

order

[13, 14, 15, 17, 18, 18, 19]

13-19 = **6**

Party2

[14, 19, 7, 15, 13, 68 ,19]

order

[7, 13, 14, 15, 19, 19, 68]

7-68 = **62**

Does result captures the real difference?

Standard deviation (Dispersion):

Sum of differences between mean and each data point

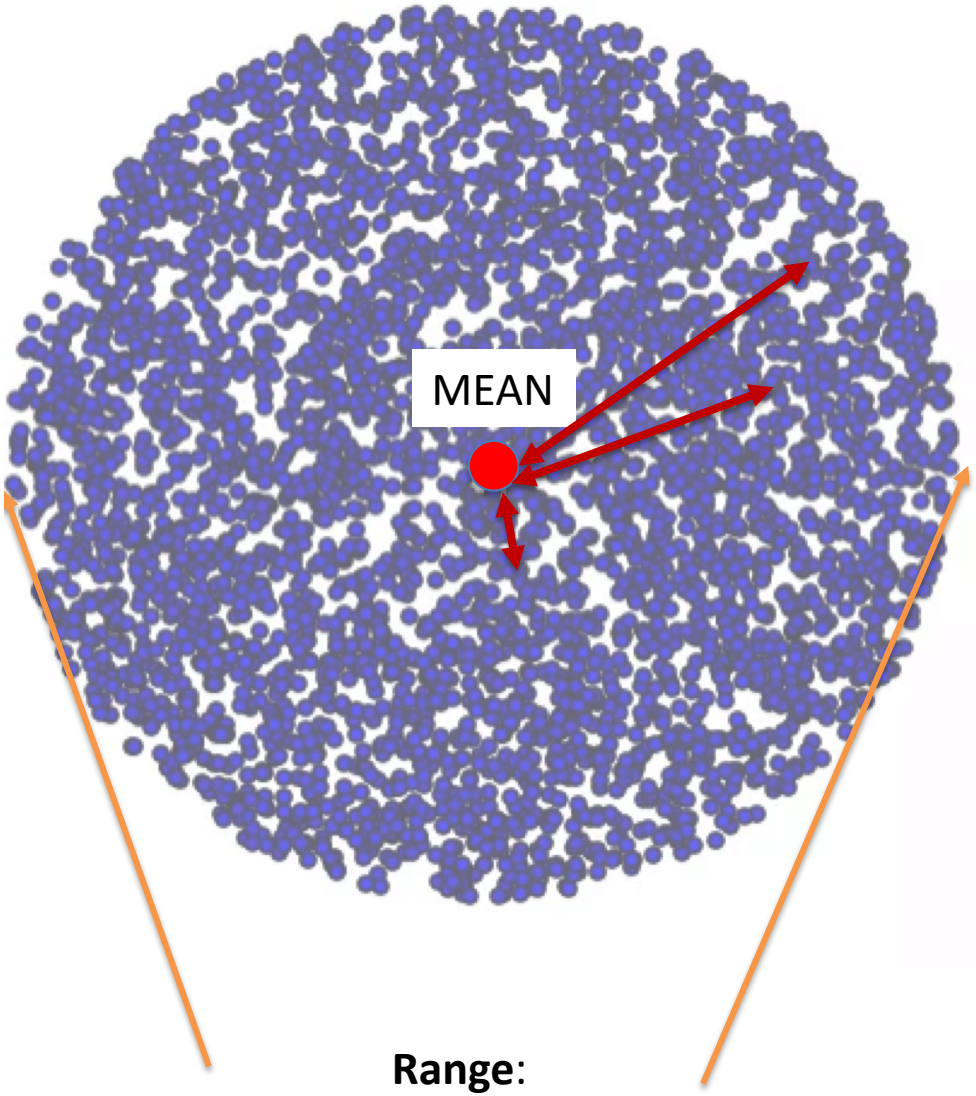
$$\sigma^2 = \sum \frac{(x - x_{mean})^2}{N}$$

σ^2 – *population variance*
x – *each data point*
x_{mean} – *average of all data points*
N – *number of data points*

$$S^2 = \sum \frac{(x - x_{mean})^2}{n - 1}$$

S^2 – *sample variance*
n – *sample number*

Data Points



Range:
difference between min and max

Interpreting Standard deviation

Small Std. Dev. – goal in e.g. manufacturing

Big Std. Dev. – lot of variability in dataset, potential for various explanations e.g. research

http://www.dummies.com/education/math/statistics/how-to-interpret-standard-deviation-in-a-statistical-data-set/

Interpreting Standard deviation: How to interpret the number

If ***Mean*** of people's age in the room is 35 years and
Sample Standard Deviation equals to 5,
Then most of the people in the room fall within 30-40 year range.
OR
The spread is 1/7 of the mean,
OR 15%

- Mean allows to see the center of the data
- Standard Deviation shows how spread with relation to the center data points are
- Usually around 68% of the data lie within 1 **Standard Deviation** from the **Mean**
- Usually around 95% of the data lie within 2 **Standard Deviations** from the **Mean**

T-Test

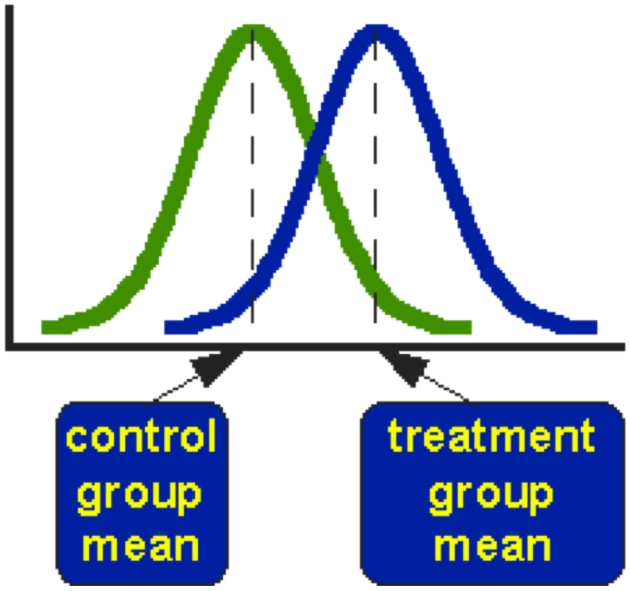
Checks if two means (averages) are different from each other

Is the mean difference significant?

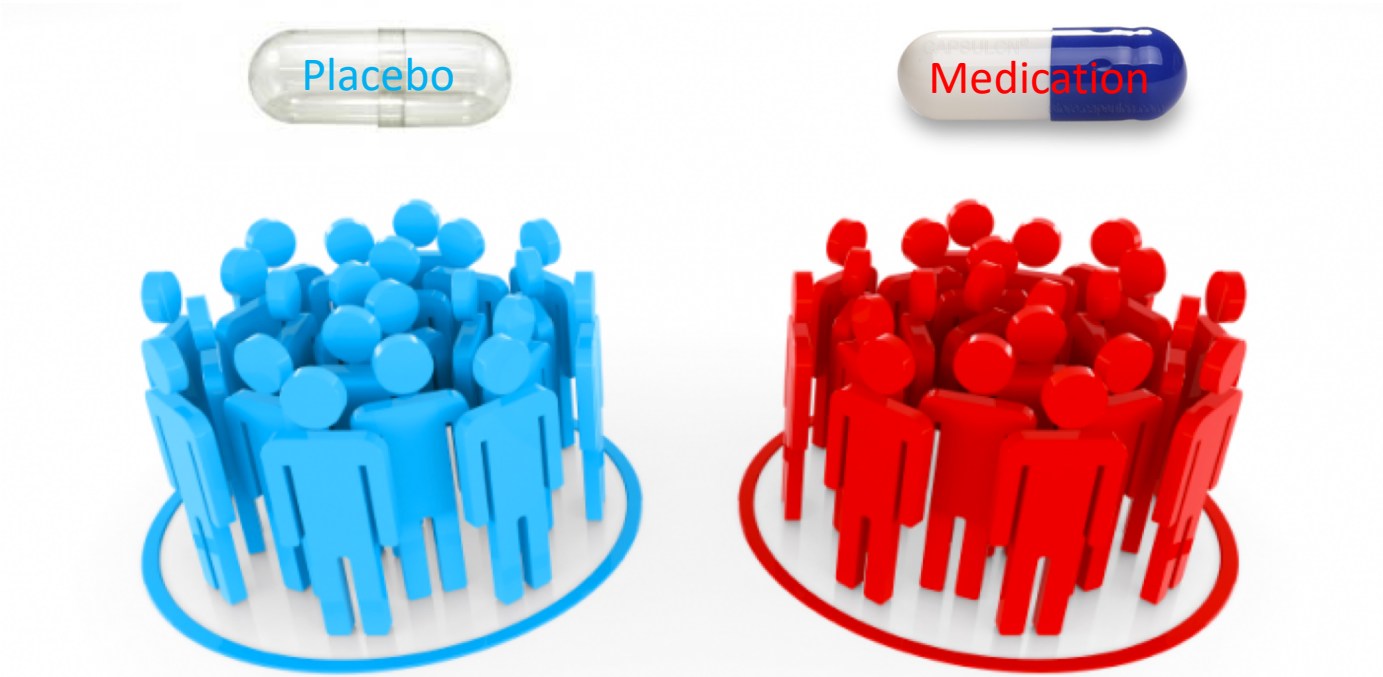
Average height in Group 1: **168** cm

Average height in Group 2: **166** cm

T-test is used to evaluate experiment results



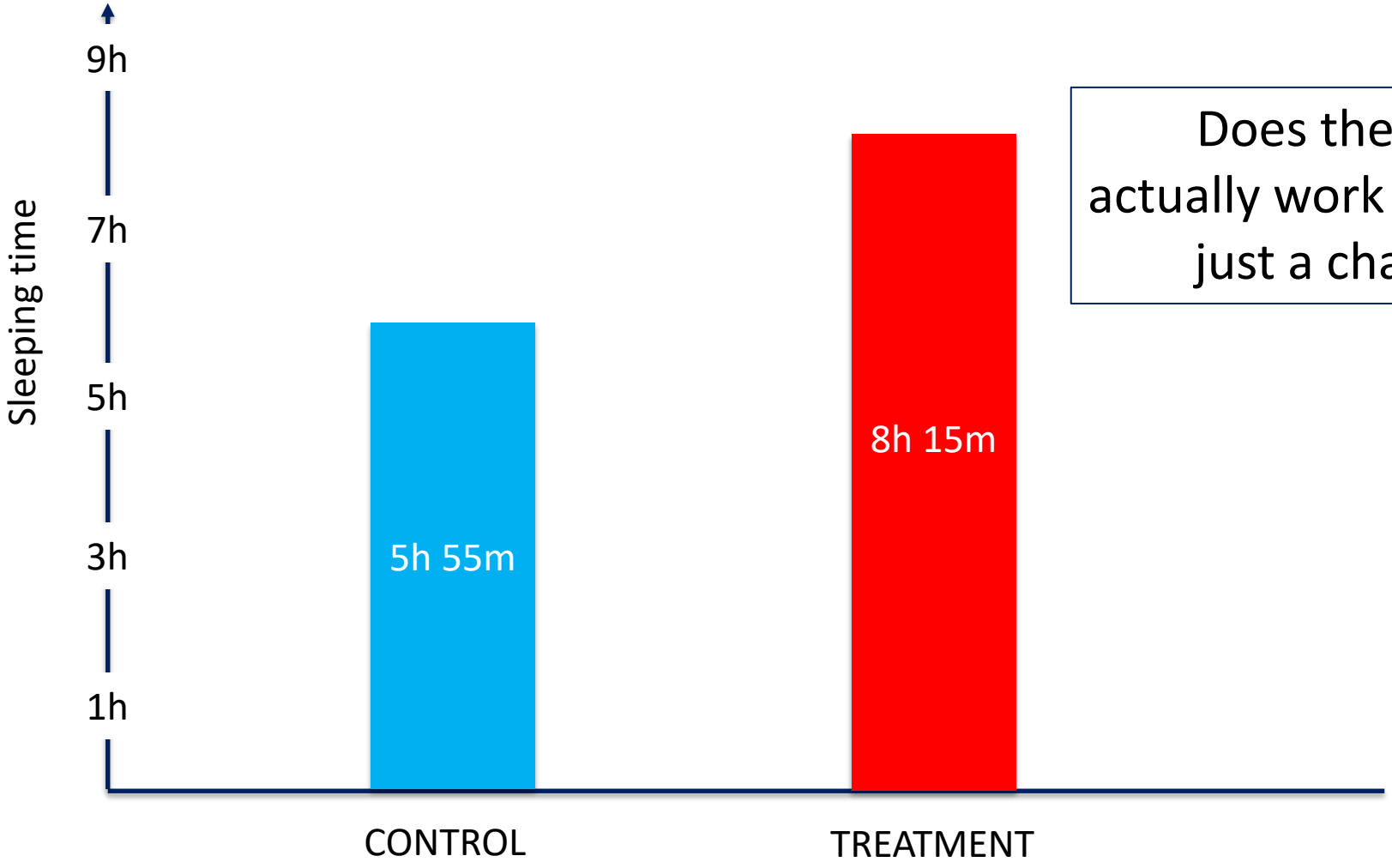
Increasing sleep experiment



CONTROL group

TREATMENT group

On average each group slept….



Does the drug actually work or its just a chance?

T-test

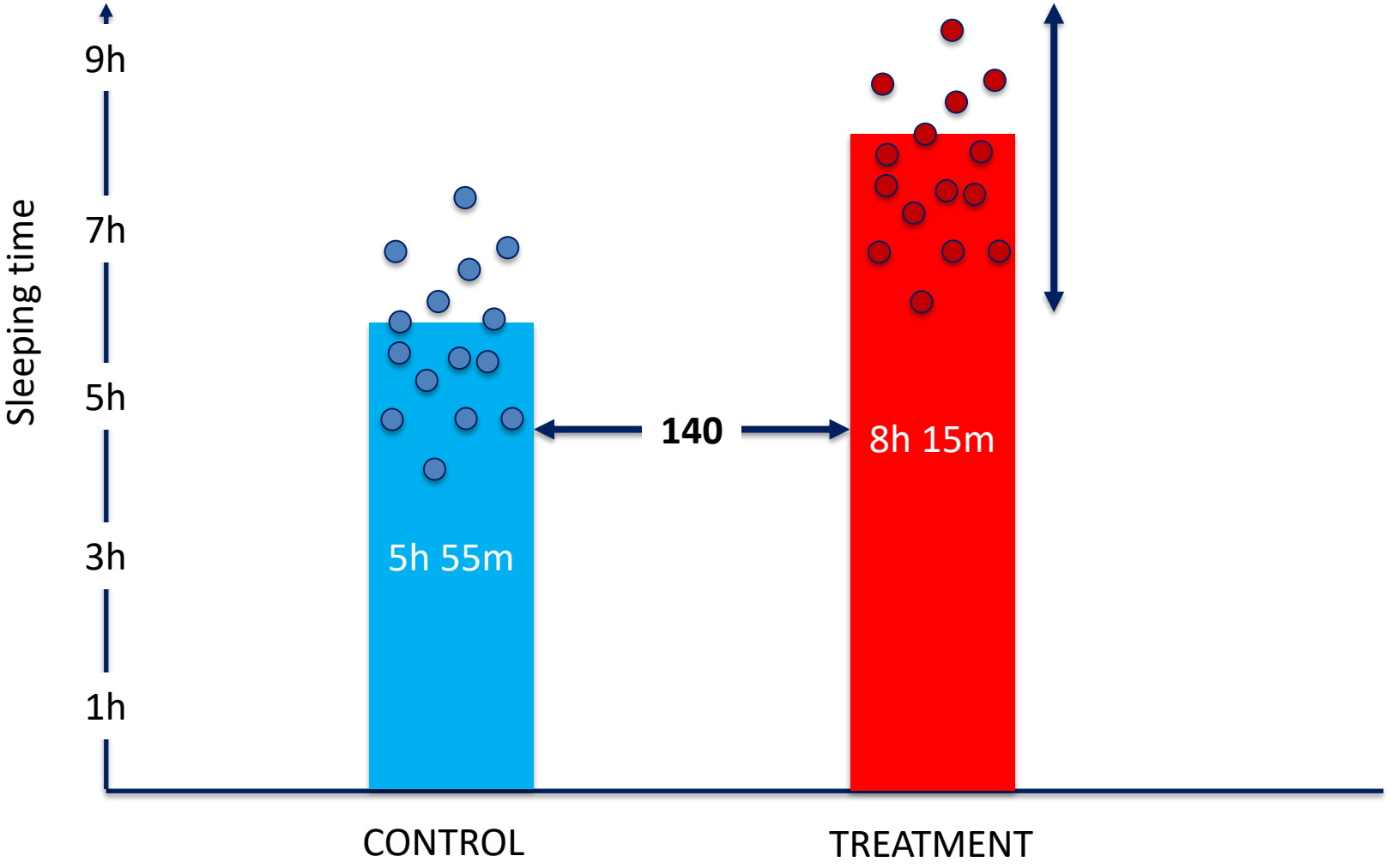
$$t = \frac{variance\ between\ groups}{variance\ within\ groups}$$

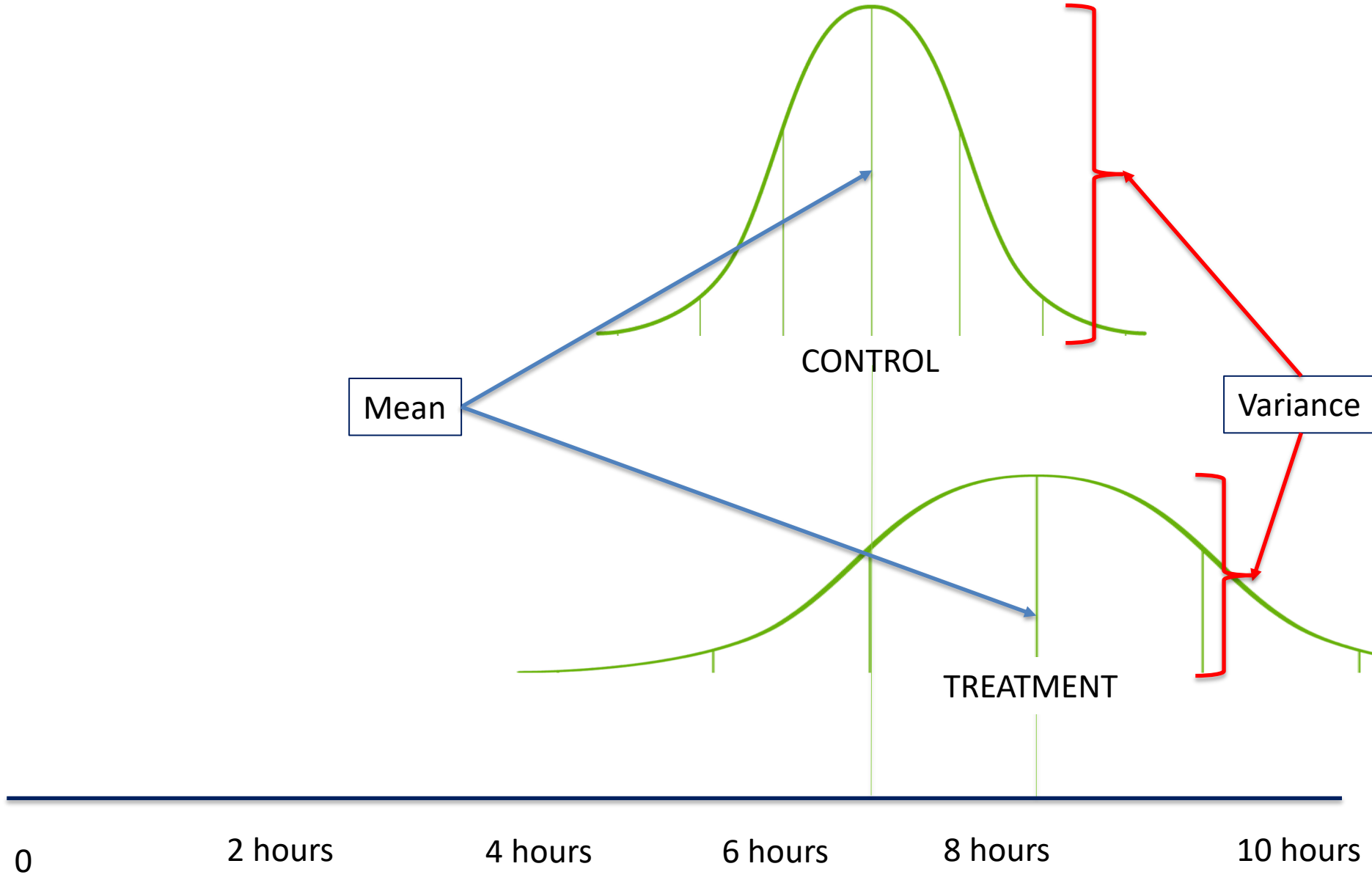
Large ***t*** score – groups are different
Small ***t*** score – groups are identical

Control group slept – **355** minutes on average

Treatment group slept – **495** minutes on average

Variance between groups = 495 – 355 = **140**


Sleeping time
9h
7h
5h
3h
1h
5h 55m
140
8h 15m
CONTROL
TREATMENT


CONTROL
Mean
Variance
TREATMENT
0
2 hours
4 hours
6 hours
8 hours
10 hours

Each **t-value** has corresponding **p-value**

p-value says that the pattern produced by our (experiment) data can be produced by random data

The lower the **p-value** the less likely the difference is caused by chance

If p=**0.05**: there is a **5%** chance the results are caused by chance

If p=**0.01**: there is a **1%** chance the results are caused by chance

p<=0.05

Sample size affects results

If **t-value** = **2.0**

For two groups of **5** subjects **p-value** = 0.04

For two groups of **10** subjects **p-value** = 0.03

20-30 data points for each group

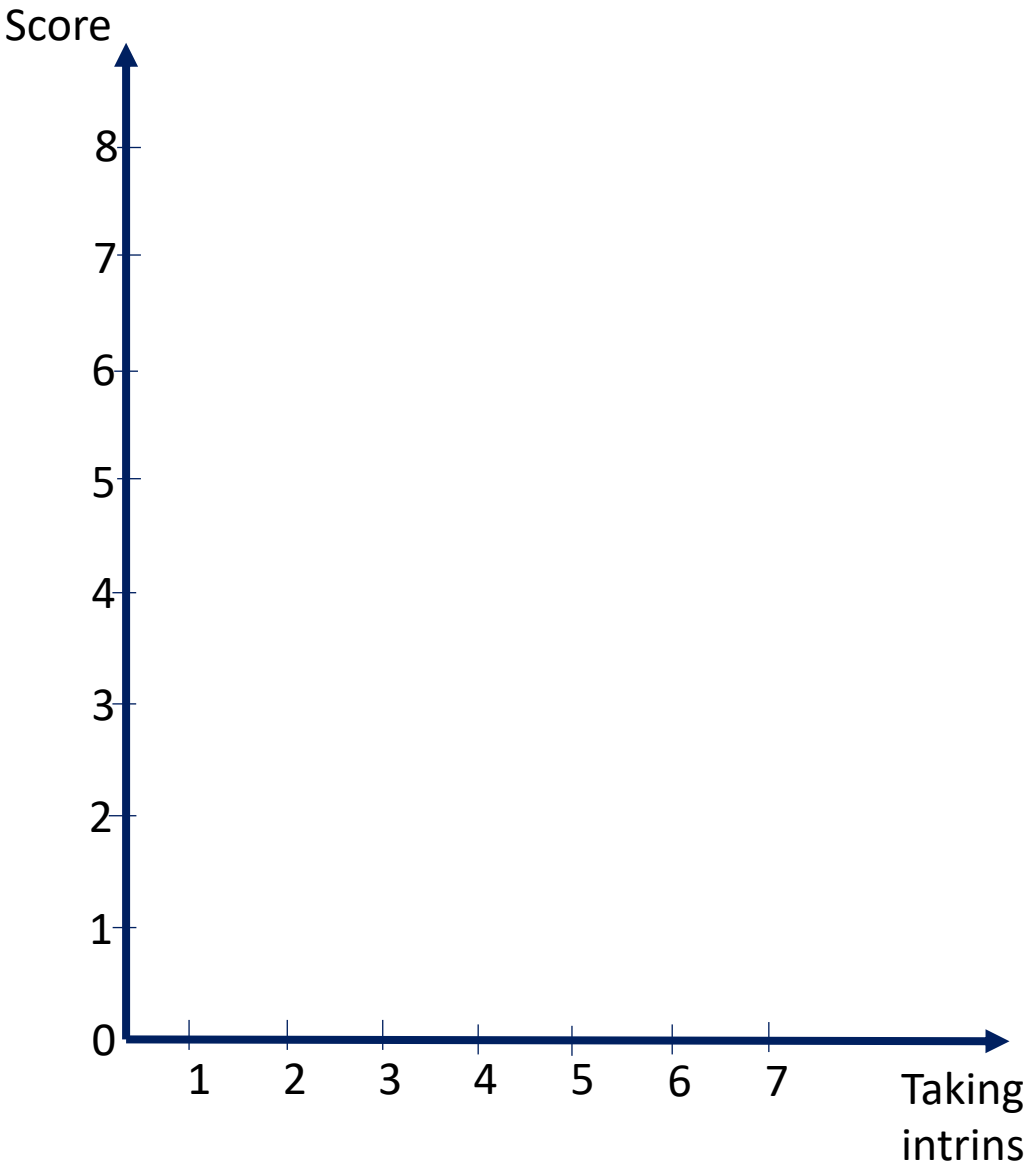
If more than 2 groups - **ANOVA**

Measuring relationships among variables

Represent on a 2d plot:

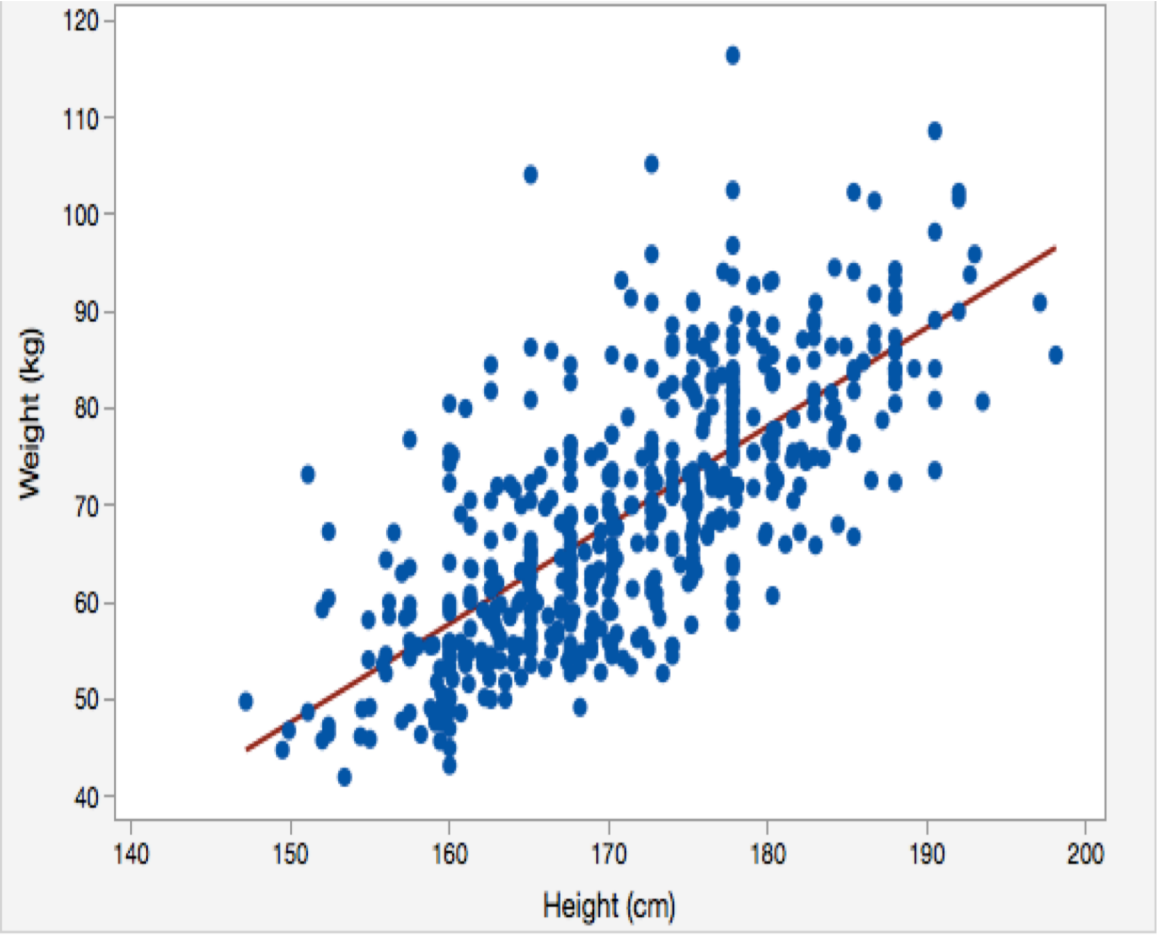
"Score" and "Taking the quiz it was intrinsically rewarding"

Represent on a 2d plot:



SCORE	Taking the quiz it was intrinsically rewarding
7	7
2	6
5	4
6	6
4	5
7	2
3	3
5	7
7	6
8	5
3	5
7	2

Height and Weight correlation



Height (cm)	Weight (kg)
180	75
168	48
195	70
151	72
177	115
190	82
160	42
170	60
…	…

On general correlation is high, but due to few outlier (too skinny, too heavy regardless of height) it will never be equal to 1.0, i.e. all points on the line

Correlations – a pattern between two variables

- How differences in one variable affects changes in another variable

  (to what extent two variables/features are connected)

Connection ranges from "-1" to "1".

If its close to **"1"** variables are **positively** connected – if one is **increasing**, another one is **increasing** too.

If its close to **"-1**" variables are **oppositely** connected – if one is **increasing**, another one is **decreasing**.

If its close to "**0**" variables have **no relationship**

Preparation Homework:
https://www.youtube.com/watch?v=N4mEzFDjqtA

Until 32 minutes

If working on your one machine:
- Install Anaconda https://www.anaconda.com/distribution/
- install requests (in terminal ‘pip install requests’)