

## Data Manipulation



$$\text{Newton's formula for cooling: } T(t) = T_{env} + (T_0 - T_{env}) e^{-kt}$$

Where:

- $T(t)$  is the temperature of the body at time  $t$ ,
- $T_{env}$  is the temperature of the surrounding environment,
- $T_0$  is the initial temperature of the body (look up average body temperature for this),
- $t$  is the time after death in hours, and
- $k$  is a constant that can be calculated using the formula  $e^{-2k} = \frac{T(t+2) - T_{env}}{T(t) - T_{env}}$ .

## Data Manipulation



Newton's formula for cooling:  $T(t) = T_{env} + (T_0 - T_{env}) e^{-kt}$

Where:

- $T(t)$  is the temperature of the body at time  $t$ ,
- $T_{env}$  is the temperature of the surrounding environment,
- $T_0$  is the initial temperature of the body (look up average body temperature for this),
- $t$  is the time after death in hours, and
- $k$  is a constant that can be calculated using the formula  $e^{-2k} = \frac{T(t+2) - T_{env}}{T(t) - T_{env}}$ .

Let  $k$  represent a constant that is necessary to resort to Newton's formula of cooling. Exponential negative  $2k$  (or  $e$  to the negative  $2k$ ) equals the temperature of the surroundings subtracted from the temperature of the body two hours after  $t$  when the body temperature was taken, all over the temperature of the body at time  $t$  minus the room temperature.

# Data Manipulation

## First Murder Mystery: A Warm-Up

Mrs Campbell was found dead in her physics classroom; a bullet to the brain seemed to be the cause of her death. It was not suicide for the bullet wound was on the left side and Mrs Campbell was right handed. Also, Mrs Campbell was too arrogant to end her own life.

After much poking around, Sherlock was able to prove that Mrs Campbell was the most hated teacher at St. Duke's school of scientific arts. Which made it a bit hard to narrow down the suspects. The only clue Sherlock found on the blackboard, in the room where she was murdered, was:

E=MC<sup>2</sup>

The suspects are:

**Simon Thompson:** The student body president, who Mrs Campbell had been blackmailing.

**Mary Campbell:** Mrs. Campbell's twin, who had always been jealous of her twin sister.

**Ruby Griffith:** The principal, who had been looking for an excuse to fire her since day one.

**Meagan Flakstodon:** The left-handed best friend, who had always been stuck in the shadow of Mrs Campbell.

## Who was the culprit?

Answer:

Mrs>Mary Campbell

2>Twin

Mary Campbell is the culprit.

## Second Murder Mystery

You and your class have rented out a mansion for an end-of-year celebration. Everyone is having a lot of fun hanging out, playing games, and eating food. Suddenly you all hear a loud scream coming from the upstairs of the mansion. Everyone rushes upstairs and you all see that on their way to the bathroom one of your classmates had uncovered a dead body in the hallway outside of the library. The person who had died is the owner of the mansion and no one seems to know who killed him. The only thing everyone agrees on is that it must be someone in the class because there was no one else in the mansion. At the time the body was found the entire class is present.

One of your friends comes up with the suggestion that maybe if they know the time of death you can determine who must be the murderer based on who was most likely to be with the mansion owner at that time.

## How would you proceed?

Another one of your friends has the idea but knowing the temperature of the body would help to figure out the time of his death, so immediately a thermometer is found and the corpse's body temperature is taken. According to these measurements the body was 33.89 degrees Celsius at 6 pm when the body was found. Two hours later the body's temperature is taken again but it is now found to be 30.33 degrees Celsius.

## Newton's formula for cooling: $T(t) = T_{\text{env}} + (T_0 - T_{\text{env}}) e^{-kt}$

Where:

- $T(t)$  is the temperature of the body at time  $t$ ,
- $T_{\text{env}}$  is the temperature of the surrounding environment,
- $T_0$  is the initial temperature of the body (look up average body temperature for this),
- $t$  is the time after death in hours, and
- $k$  is a constant that can be calculated using the formula  $e^{-2k} = \frac{T(t+2) - T_{\text{env}}}{T(t) - T_{\text{env}}}$ .

Answer:

According to the wording of the problem treated, after substituting the values and processing the date, it can be said  $k$  approximately equals 0.07. So the time  $t$  elapsed before the body was found at 6pm is about 1.38 hours: the owner was killed around 4.62 pm or 4:37pm (4:40pm).

Example of description: To find exponential minus 2k, we subtract the temperature of the surrounding environment from the temperature of the body at time  $t$  plus 2, and we divide the result by the temperature of the body at time  $t$  minus the temperature of the surrounding environment.

<b>(a)</b>	negative b
<b>a + b</b>	a and b add up to add a and b
<b>*</b>	plus
<b>*</b>	minus
<b>a - b = c</b>	to subtract b from a, to deduct b; c is the difference between a and b
<b>a x b = c</b>	a times b; to multiply a by b; c is the product of a by b
<b>abc + abd = ab(c+d)</b>	to factorize ab to factor out ab; ab is the common factor
<b>a(b+c)</b>	a factor b plus c
<b>a/b</b>	a divided by b; a over b, numerator, denominator
<b>(b+c)/a</b>	b/c, all over a
<b>a<sup>n</sup></b>	a raised to the power n; a raised to the nth power; a to the n; a to the nth; a to the nth power
<b>a<sup>-n</sup></b>	a to the minus/negative n
<b>l'exposant n</b>	exponent n
<b>a<sup>2</sup>, a<sup>3</sup></b>	to square a; to cube a; a squared, a cubed
<b>sin<sup>2</sup>x, cos<sup>3</sup>y</b>	sine squared x, cosine cubed y (pron. /sain'/)
<b>a<sup>1/3</sup></b>	a to the one third
<b>racine carrée de a</b>	the square root of a
<b>racine cubique de a</b>	the cube root of a
<b>a</b>	a to the x; exponential x
<b>e<sup>x</sup></b>	natural/Napierian logarithm of y (dans la lecture d'une équation, ln y et log P se lisent sous forme abégée comme en français)
<b>x=ln y</b>	natural/exponentiel logarithme de y (dans la lecture d'une équation, ln y et log P se lisent sous forme abégée comme en français)
<b>n=log P</b>	decimal/common logarithm of P dans la lecture d'une équation, ln y et log P se lisent sous forme abégée comme en français
<b>log2=0.30103</b>	logarithme à la base 2 à 3 chiffres – à 3 places logarithme, 0 caractéristique – caractéristique, 0(0103)-mantisse, mantissa, table de log – log table, table of logs, règle à calcul – slide rule, échelle, grille, papier – scale, graph, paper
<b>Produit en croix</b>	Cross-multiplication

## Data Manipulation

How Math Can Save  
Democracy



Can Math  
Solve the



Gerrymandering  
Problem?

What is Gerrymandering?



## Data Manipulation

How Math Can Save  
Democracy



Can Math  
Solve the



Gerrymandering  
Problem?

What is Gerrymandering?



Gerrymandering is the process of drawing voting districts unfairly in favor of one's political party.

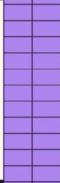
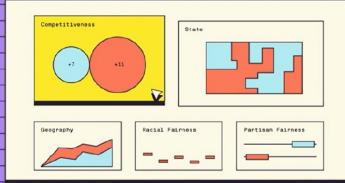
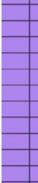
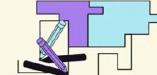
## Data Manipulation

### How Math Can Save Democracy



Which parameters and data can a mathematician look at to check it?

How it works



Explain in one sentence the use of each equation.

$$EG = \frac{|(\text{one party's wasted votes}) - (\text{other party's wasted votes})|}{\text{total number of votes}}$$

$$PP(S) = 4\pi \frac{\text{area enclosed by the district}}{(\text{perimeter of the district})^2}$$

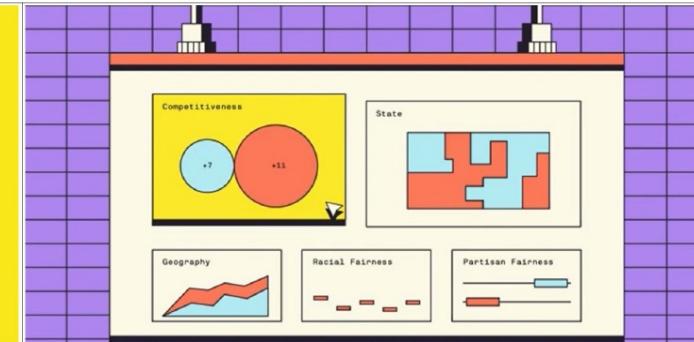
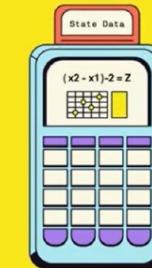
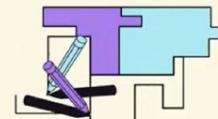
## Data Manipulation

How Math Can Save  
Democracy



Which parameters and data can a mathematician look at to check it?

How it works



Population equity; Compactness and contiguity; Votes;  
Traditional districting principles (respecting communities)

## Data Manipulation

### How Math Can Save Democracy

Explain in one sentence the use of each equation.

#### THE MATHEMATICS OF DEMOCRACY

##### WHAT IS *Polsby-Popper Compactness Score*?

We can try to quantify gerrymandering by measuring how "reasonably shaped" a voting district is. The **Polsby-Popper Compactness Score** of district  $S$ ,  $PP(S)$ , measures how much "unnecessary" perimeter  $S$  has.

$$PP(S) = 4\pi \frac{\text{area enclosed by the district}}{(\text{perimeter of the district})^2}$$

If there is too much perimeter for the area,  $PP(S)$  is closer to 0.

If  $PP(S)$  is close to 0, this indicates that the district might have been gerrymandered.

#### THE MATHEMATICS OF DEMOCRACY

##### WHAT IS *the Efficiency Gap*?

**Wasted votes** are those that do not affect an election: either votes above the 50% majority threshold for a winning candidate or any for a losing candidate. The **efficiency gap (EG)** is one method for measuring gerrymandering by adding up the wasted votes of each party's candidates over all electoral districts.

$$EG = \frac{|(\text{one party's wasted votes}) - (\text{other party's wasted votes})|}{\text{total number of votes}}$$

## Data Manipulation

### How Math Can Save Democracy



#### Comment traduire le verbe “représenter”?

*Traduisez les phrases suivantes en faisant particulièrement attention au sens du verbe “représenter”. Évitez le verbe “to represent” chaque fois que possible.*

1. Le terrorisme représente une grave menace pour la démocratie.
2. Écrire un livre comme celui-ci représente beaucoup de travail.
3. Les matières premières représentent 50% de nos coûts de production.
4. Que représentent les lettres OMS?
5. La photo représente deux amoureux embrassés, et non deux amoureux en train de s'embrasser.

## Data Manipulation

### How Math Can Save Democracy



#### Comment traduire le verbe “représenter”?

*Traduisez les phrases suivantes en faisant particulièrement attention au sens du verbe “représenter”. Évitez le verbe “to represent” chaque fois que possible.*

1. Le terrorisme représente une grave menace pour la démocratie.

Terrorism **is/poses** a serious threat to liberal democracy.

2. Écrire un livre comme celui-ci représente beaucoup de travail.

Writing a book like this **involves/means** a lot of work.

3. Les matières premières représentent 50% de nos coûts de production.

Raw materials **account for** 50% of our production costs.

4. Que représentent les lettres OMS?

What do the letters WHO **stand for**?

5. La photo représente deux amoureux embrassés, et non deux amoureux en train de s'embrasser.

The picture **shows** two lovers **kissed**, not two lovers **kissing**.

## Data Manipulation

### Les prépositions qui accompagnent des adjectifs

**Complétez les phrases suivantes avec l'adjectif et la préposition qui conviennent.**

1. The storm was ..... a lot of damage. (*responsable de*)
2. Don't get ..... me, I'm not to blame. (*être en colère, fâché*)
3. Everyone was ..... your language skills. (*impressionné par*)
4. Are you ..... economics? (*s'intéresser à*)
5. The shareholders will be ..... the end-of-year results. (*déçu par*)
6. The climate here is very ..... California's. (*similaire à*)
7. Denmark is very ..... its Nordic neighbours. (*différent de*)
8. The police station ..... the metro. (*près de*)
9. My flat isn't very ..... the station. (*loin de*)
10. Self-governance is ..... autonomy. (*synonyme de*)
11. The voters were ..... the government's decision to cancel its plans. (*content de*)
12. Many aren't ..... the seriousness of the problem. (*conscient de, au courant de*)
13. People ..... their weight may become anorexic. (*obsédé par*)
14. The maths teacher was ..... my homework. (*satisfait de*)
15. Most people feel ..... the rise in temperature. (*inquiet de, préoccupé par*)
16. They are ..... outside financial support. (*dépendant de*)

## Data Manipulation

### Les prépositions qui accompagnent des adjectifs

*Complétez les phrases suivantes avec l'adjectif et la préposition qui conviennent.*

1. The storm was **responsible for** a lot of damage. (*responsable de*)
2. Don't get **angry or cross with** me, I'm not to blame. (*être en colère, fâché*)
3. Everyone was **impressed with** your language skills. (*impressionné par*)
4. Are you **interested in** economics? (*s'intéresser à*)
5. The shareholders will be **disappointed with** the end-of-year results. (*déçu par*)
6. The climate here is very **similar to** California's. (*similaire à*)
7. Denmark is very **different from** its Nordic neighbours. (*différent de*)
8. The police station **is close to** the metro. (*près de*)
9. My flat isn't very **far from** the station. (*loin de*)
10. Self-governance is **synonymous with** autonomy. (*synonyme de*)
11. The voters were **pleased / happy with** the government's decision to cancel its plans. (*content de*)
12. Many aren't **aware of** the seriousness of the problem. (*conscient de, au courant de*)
13. People **obsessed with** their weight may become anorexic. (*obsédé par*)
14. The maths teacher was **satisfied with** my homework. (*satisfait de*)
15. Most people feel **concerned about** the rise in temperature. (*inquiet de, préoccupé par*)
16. They are **dependent on** outside financial support. (*dépendant de*)

## Data Manipulation

### Les prépositions qui accompagnent des adjectifs

**Complétez les phrases suivantes avec l'adjectif et la préposition qui conviennent.**

1. The storm was **responsible for** a lot of damage. (*responsable de*)
2. Don't get **angry or cross with** me, I'm not to blame. (*être en colère, fâché*)
3. Everyone was **impressed with** your language skills. (*impressionné par*)
4. Are you **interested in** economics? (*s'intéresser à*)
5. The shareholders will be **disappointed with** the end-of-year results. (*déçu par*)
6. The climate here is very **similar to** California's. (*similaire à*)
7. Denmark is very **different from** its Nordic neighbours. (*différent de*)
8. The police station **is close to** the metro. (*près de*)
9. My flat isn't very **far from** the station. (*loin de*)
10. Self-governance is **synonymous with** autonomy. (*synonyme de*)
11. The voters were **pleased / happy with** the government's decision to cancel its plans. (*content de*)
12. Many aren't **aware of** the seriousness of the problem. (*conscient de, au courant de*)
13. People **obsessed with** their weight may become anorexic. (*obsédé par*)
14. The maths teacher was **satisfied with** my homework. (*satisfait de*)
15. Most people feel **concerned about** the rise in temperature. (*inquiet de, préoccupé par*)
16. They are **dependent on** outside financial support. (*dépendant de*)

## Misinformation

**Misleading Statistics  
in the Media:  
Examples  
and How to  
Spot Them**



### PART 1: How false news can spread

1. Why should we be more careful today about information and data displayed in the news?

2. Oral comprehension.

*Watch TED-Ed's video about how false news can spread (00:00-03:02) and answer the following questions.*



Read the questions before watching the video so you know what to expect.



## Misinformation



*Right or wrong? You may have to justify your answer.*

1. Decades ago there were fewer newspapers so the information was less trustworthy: .....
2. Aggregated sources of news like the Associated Press (AP) are a safe way to avoid the circulation of misinformation: .....
3. Since information travels fast and everybody can participate, false data are easily corrected: .....
4. Children that should have been vaccinated were contaminated: .....
5. The video suggests serious journals should not make jokes: .....
6. In terms of information, new technologies have had positive effects: .....

*Answer the following question through complete sentences.*

7. Why can it be said that Mark Twain's quotation proves the precise point it is making?

## Misinformation

*Right or wrong? You may have to justify your answer.*

1. Decades ago there were fewer newspapers so the information was less trustworthy: **Wrong**. In previous decades most media with global reach consisted of several major newspapers and networks which had the resources to gather information directly.
2. Aggregated sources of news like the Associated Press (AP) are a safe way to avoid the circulation of misinformation: **Wrong**. The speed with which information spreads now has created the ideal conditions for “circular reporting,” when publication A publishes misinformation, publication B reprints it, and publication A then cites B as the source for the information.
3. Since information travels fast and everybody can participate, false data are easily corrected: **Wrong**. The pseudo-scientific article linking vaccination and autism has repeatedly been discredited by the scientific community but the belief persists. It is also a form of problematic circular reporting when multiple publications report on the same initial piece of false information which then appears to another author as having been verified by multiple sources. Moreover an unverified fact in a wiki page can make its way into a published article that may later be added as a citation for the very same wiki information.
4. Children that should have been vaccinated were contaminated: **Right**. Deliberately unvaccinated children are now contracting contagious diseases that had been virtually eradicated in the US.
5. The video suggests serious journals should not make jokes: **Wrong**. A joke article in the British Medical Journal has been referenced in serious science publications over 400 times. But the problem is not the joke, it is the fact outlets pick it up not in on the joke.
6. In terms of information, new technologies have had positive effects: **Right**. User-generated content may entail circular reporting, but recent advances in communication technologies have had immeasurable benefits, such as breaking down the barriers between information and people.

*Answer the following question through complete sentences.*

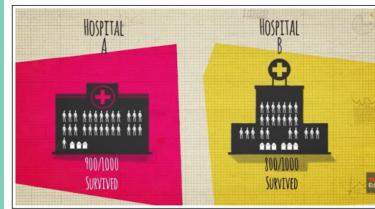
7. Why can it be said that Mark Twain's quotation proves the precise point it is making? The quote usually attributed to Mark Twain says that “A lie can travel halfway around the world while the truth is putting on its shoes.” Since there is reason to doubt that Mark Twain ever said this at all, the attribution of the quote is an example of misinformation that “travelled halfway around the world.”

## Misinformation

**Misleading Statistics  
in the Media:  
Examples  
and How to  
Spot Them**



### PART 2: How statistics can be misleading



1. Consider the following problem and make a decision according to the data provided. Justify.

*Imagine you need to choose between two hospitals for an elderly relative's surgery. Out of each hospital's last 1000 patients, 900 survived at Hospital A, while only 800 survived at Hospital B. Which is the better choice?*

2. What could be wrong about one's spontaneous decision in that particular case?

## Misinformation

**Misleading Statistics  
in the Media:  
Examples  
and How to  
Spot Them**



3. Actually, not all patients arrive at the hospital with the same level of health.

*Let's divide each hospital's last 1000 patients into those who arrived in good health and those who arrived in poor health. Hospital A had only 100 patients who arrived in poor health, of which 30 survived. But Hospital B had 400, and they were able to save 210. Which hospital should one choose? Describe your calculations.*

4. What contradiction do the results raise? Do you know what to call such a phenomenon?

5. What factor should you take into account in the following cases before drawing definitive conclusions? Which is the lurking variable?



a. One study in the UK appeared to show that smokers had a higher survival rate than non-smokers over a twenty-year time period.



b. An analysis of Florida's death penalty cases seemed to reveal no racial disparity in sentencing between black and white defendants convicted of murder.

## Misinformation

**Misleading Statistics  
in the Media:  
Examples  
and How to  
Spot Them**



3. Actually, not all patients arrive at the hospital with the same level of health.

*Let's divide each hospital's last 1000 patients into those who arrived in good health and those who arrived in poor health. Hospital A had only 100 patients who arrived in poor health, of which 30 survived. But Hospital B had 400, and they were able to save 210. Which hospital should one choose? Describe your calculations.*

Hospital B is the better choice for patients who arrive at hospital in poor health, with a survival rate of 52.5% (210/400). What if your relative's health is good when she arrives at the hospital? Strangely enough, Hospital B is still the better choice, with a survival rate of over 98% ((800 total survivors – 210 survivors originally in poor health)/600 who arrived in good health).

6. What contradiction do the results raise? Do you know what to call such a phenomenon?

How can Hospital A have a better overall survival rate if Hospital B has better survival rates for patients in each of the two groups? Here is a case of Simpson's paradox, where the same set of data can appear to show opposite trends depending on how it is grouped. Or in other words, it is possible to draw two opposite conclusions from the same data depending on how you divide things up. It often occurs when aggregated data hides a conditional variable, sometimes known as a lurking variable, which is a hidden additional factor that significantly influences results. Here the hidden factor is the relative proportion of patients who arrive in good or poor health.

## Misinformation

**Misleading Statistics  
in the Media:  
Examples  
and How to  
Spot Them**



3. Actually, not all patients arrive at the hospital with the same level of health.

*Let's divide each hospital's last 1000 patients into those who arrived in good health and those who arrived in poor health. Hospital A had only 100 patients who arrived in poor health, of which 30 survived. But Hospital B had 400, and they were able to save 210. Which hospital should one choose? Describe your calculations.*

4. What contradiction do the results raise? Do you know what to call such a phenomenon?

5. What factor should you take into account in the following cases before drawing definitive conclusions? Which is the lurking variable?



a. One study in the UK appeared to show that smokers had a higher survival rate than non-smokers over a twenty-year time period.



b. An analysis of Florida's death penalty cases seemed to reveal no racial disparity in sentencing between black and white defendants convicted of murder.

## Misinformation

**Misleading Statistics  
in the Media:  
Examples  
and How to  
Spot Them**



7. What factor should you take into account in the following cases before drawing definitive conclusions? Which is the lurking variable?



a. One study in the UK appeared to show that smokers had a higher survival rate than non-smokers over a twenty-year time period.

That is until dividing the participants by age group showed that the non-smokers were significantly older on average, and thus more likely to die during the trial period, precisely because they were living longer in general. Here the age groups are the lurking variable and are vital to correctly interpret the data.

b. An analysis of Florida's death penalty cases seemed to reveal no racial disparity in sentencing between black and white defendants convicted of murder.

Dividing the cases by the race of the victim told a different story. In either situation, black defendants were more likely to be sentenced to death. The slightly higher overall sentencing rate for white defendants was due to the fact that cases with white victims were more likely to elicit a death sentence than cases where the victim was black, and most murders occurred between people of the same race.

