From Reflame your English. English for Science by prof. dr Milica Vuković Stamatović

**DATA VISUALISATION**

**In this unit:**

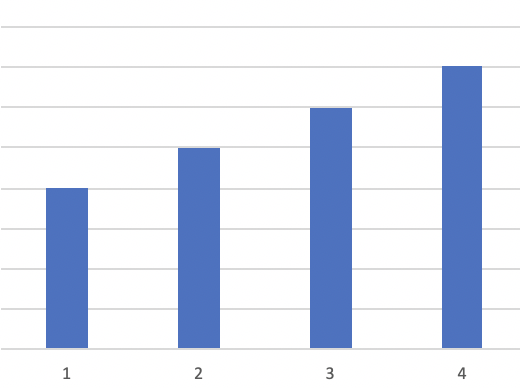
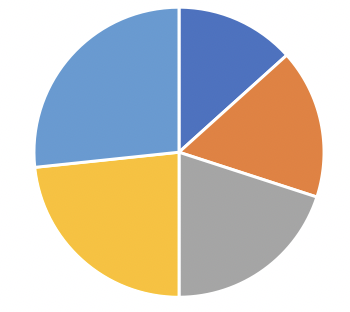
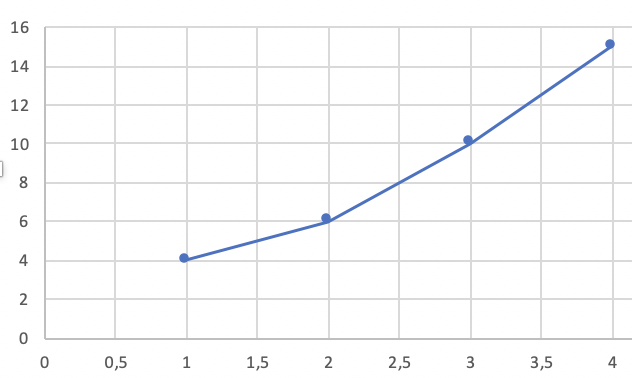
* *describing various types of graphs*
* *describing upward and downward trends*
* *vocabulary used for describing graphs and trends*
* *transitive and intransitive verbs describing trends*
* *listening to lectures for information*

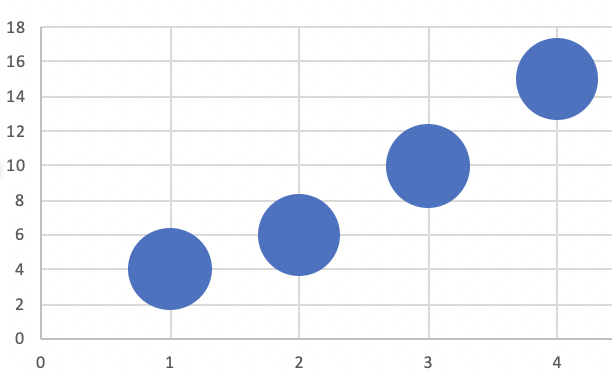
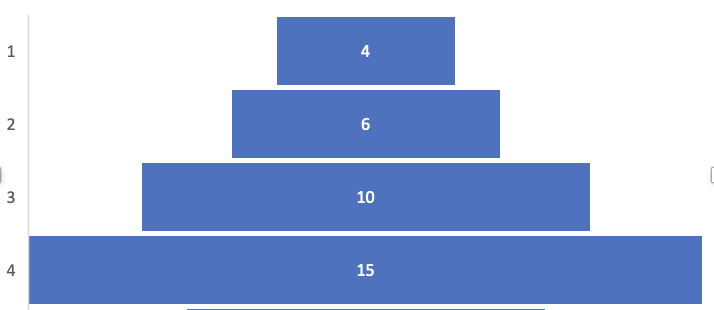
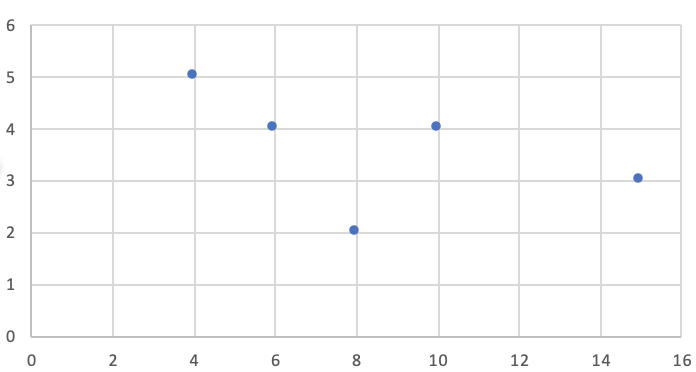
**I WARM-UP** Warm Up Icons - Download Free Vector Icons | Noun Project

**I. a. Answer the questions:**

1. Some people learn best through hands-on activities, some like to hear the spoken word, still others like to see written texts to grasp concepts. On the other hand, some learners need visual information to understand concepts. Which type of a learner are you? How important are visuals for you?
2. Have you ever produced a chart? What kind of a chart was it?
3. What types of charts are there? How do you choose which chart to use?
4. Can you make any advanced charts? What programmes can you use for that?

**I. b. Name these basic types of charts:**

**  **

**  **

**II READING** Reading Icon - Free Download, PNG and Vector

**Read the introductory part of the article written by Betsy Mason and published in the *Knowable Magazine*, and then do the exercises below.**

***Why scientists need to be better at data visualization*[[1]](#footnote-1)**

*The scientific literature is riddled with bad charts and graphs, leading to misunderstanding and worse.*

*Avoiding design missteps can improve understanding of research.*

Imagine a science textbook without images. No charts, no graphs, no illustrations or diagrams with arrows and labels. The science would be a lot harder to understand. That’s because humans are visual creatures by nature. People absorb information in graphic form that would elude them in words. Images are effective for all kinds of storytelling, especially when the story is complicated, as it so often is with science. Scientific visuals can be essential for analyzing data, communicating experimental results and even for making surprising discoveries.

Visualizations can reveal patterns, trends and connections in data that are difficult or impossible to find any other way, says Bang Wong, creative director of MIT’s Broad Institute. “Plotting the data allows us to see the underlying structure of the data that you wouldn’t otherwise see if you’re looking at a table.”

And yet few scientists take the same amount of care with visuals as they do with generating data or writing about it. The graphs and diagrams that accompany most scientific publications tend to be the last things researchers do, says data visualization scientist Seán O’Donoghue. “Visualization is seen as really just kind of an icing on the cake.”

As a result, science is littered with poor data visualizations that confound readers and can even mislead the scientists who make them. Deficient data visuals can reduce the quality and impede the progress of scientific research. And with more and more scientific images making their way into the news and onto social media — illustrating everything from climate change to disease outbreaks — the potential is high for bad visuals to impair public understanding of science.

The problem has become more acute with the ever-increasing amount and complexity of scientific data. Visualization of those data — to understand as well as to share them — is more important than ever. Yet scientists receive very little visualization training. “The community hasn’t by and large recognized that this is something that really is needed,” says O’Donoghue, of the University of New South Wales and lead author of a paper about biomedical data visualization in the 2018 *Annual Review of Biomedical Data Science*.

There are signs of progress, however. At least two annual conferences dedicated to scientific data visualization have sprung up in the last decade. And the journal *Nature Methods* ran a regular column from 2010 to 2016 about creating better figures and graphs, which was then adapted into guidelines for scientists submitting papers to that journal. But so far, few scientists are focusing on the problem.

Improving scientific visualization will require better understanding of the strengths, weaknesses and biases of how the human brain perceives the world. Fortunately, research has begun to reveal how people read, and misread, different kinds of visualizations and which types of charts are most effective and easiest to decipher. Applying that knowledge should lead to better visual communication of science.

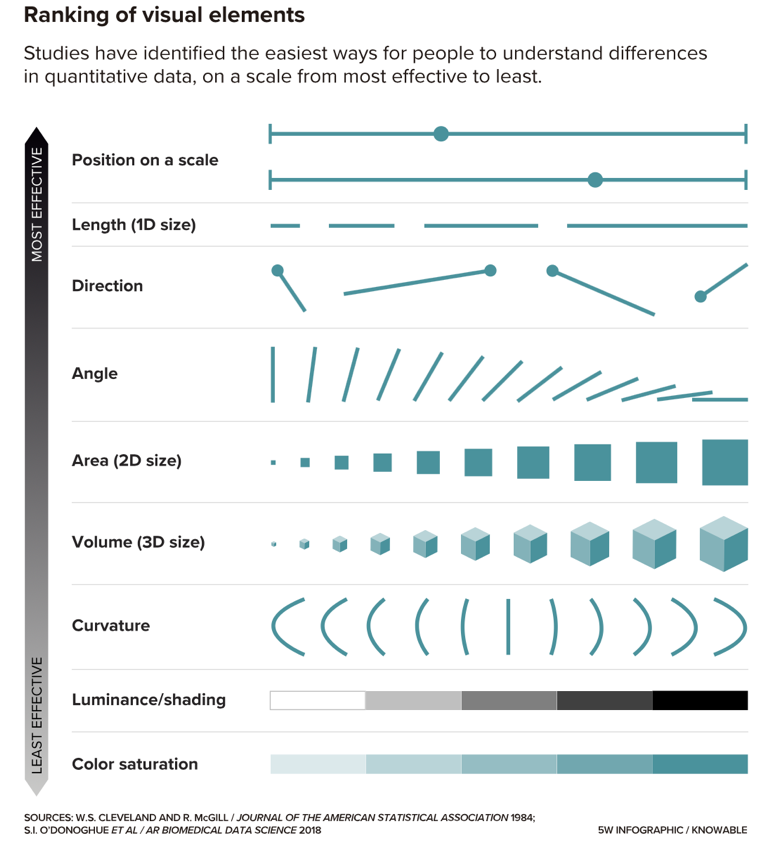
“We have a lot of practical knowledge about what works and what doesn’t,” says computer scientist Miriah Meyer of the University of Utah. “There are a lot of principles that have been through the test of time and have been shown over and over again to be really effective.”

**Chart choice**

The human visual system evolved to help us survive and thrive in the natural world, not to read graphs. Our brains interpret what our eyes see in ways that can help us find edible plants among the toxic varieties, spot prey animals and see reasonably well in both broad daylight and at night. By analyzing the information we receive through our eyes to serve these purposes, our brains give us a tailored perception of the world.

In the early 1980s, Bell Labs statisticians William Cleveland and Robert McGill began researching how the particulars of human perception affect our ability to decipher graphic displays of data — to discover which kinds of charts play to our strengths and which ones we struggle with. In a seminal paper published in 1984 in the *Journal of the American Statistical Association*, Cleveland and McGill presented a ranking of visual elements according to how easily and accurately people read them.

People are better at discerning subtleties in some types of visuals than others — the length of two lines, for example, or the direction of a line are easier to tell apart than shades of gray or the intensity of a color. Studies show that graphs using visual elements high on this list are easier to read and more effective than those near the bottom.



Their experiments showed that people are best at reading charts based on the lengths of bars or lines, such as in a standard bar chart. These visualizations are the best choice when it’s important to accurately discern small differences between values.

Study participants found it somewhat harder to judge differences in direction, angle and area. Figures using volume, curvature or shading to represent data were even tougher. And the least effective method of all was color saturation.

“The ability of the audience to perceive minute differences is going to get worse and worse” as you move down the list, says computer scientist Jeffrey Heer of the University of Washington in Seattle. In general, it’s best practice to use the highest graphical element on the list that meets the needs of each type of data.

For example, if it’s important to show that one particular disease is far more lethal than others, a graphic using the size of circles to represent the numbers of deaths will do fine. But to emphasize much smaller differences in the numbers of deaths among the less-lethal diseases, a bar chart will be far more effective.

In 2010, Heer used Amazon’s Mechanical Turk crowdsourcing service to confirm that Cleveland and McGill’s ranking holds true in the modern digital environment. Since then, Heer, O’Donoghue and others have used crowdsourcing to test many other aspects of visualization to find out what works best. “That has huge power going forward to take this whole field and really give it a solid engineering basis,” O’Donoghue says.

|  |  |
| --- | --- |
| A slide show | A slide show |
| A slide show | A slide show |
| A slide show |  |

**II. a. Decide whether the following statements are true or false.**

1. It’s easier to understand texts without visualizations compared to those accompanied by graphs.
2. Plotting the data enables us to see patterns.
3. Scientists are typically trained to visualize data.
4. Joint efforts are invested into improving scientific visualization.
5. Humans have an innate ability to read graphs.
6. Humans are best at reading charts based on lines or bars.
7. Charts using colour saturation are typically easy to interpret.
8. Heer, O’Donoghue and their colleagues obtained a project grant in the visualization field.

**II. b. Answer the questions below.**

1. Why is visualization of data important?
2. Why are some scientists not so good at visualizing data?
3. Are all types of data visualizations equally effective?
4. Why do human brains experience difficulties in interpreting graphs?
5. Are all types of bar charts equally applicable in all cases?

**II. c. Find the words in the text which mean the following (the first letter of the word has been provided for you).**

1. to escape e…………………
2. to draw lines/marks… to present facts, numbers etc. p…………………
3. to make it difficult for to move forward/make progress i…………………
4. something that can be eaten is e…………………
5. to confuse or mix c…………………
6. to suddenly appear or start existing s………………… u………
7. to understand something that is hard to understand d…………………
8. to become successful, make a lot of progress t…………………
9. customized t…………………
10. notice something by observing it; to differentiate d…………………
11. the degree to which something is mixed into something else s…………………
12. stress, point out, accentuate e…………………

**II. d. Use the words from I. c. in the sentences below.**

1. Pulse oximetry is used to check oxygen ……………………. and heart rate.
2. The problem with electrical sensors is that they are easily ……………………. by electromagnetic

interference.

1. Instead of clarifying the problem, the maths professor ……………………. the students.
2. The conference focuses on *Physical Biology* and illustrates how physics ideas and concept

……………………. up in the fields such as viral evolution.

1. Scientists are trying to ……………………. the mechanisms which underlie the biology of aging.
2. Treatment is ……………………. to the needs of each patient.
3. You can use italics to ……………………. a word in a piece of academic writing.
4. Mathematical skills involve the ability to ……………………. mathematical relationships and patterns in the world around them.
5. The claim that any integer N can be uniquely presented as a product of prime numbers

……………………. Euclid and it also ……………………. other mathematicians for another 2,000 years.

1. In the fissures found deep in the ocean microbes ……………………. in spite of mineral concentrations and extreme heat.

**II. e. Find the word in the text which goes together with the ones provided below, forming a collocation or an idiom:**

1. ……………………. change
2. ……………………. progress
3. ……………………. outbreak
4. ……………………. on the cake
5. ……………………. the data
6. ……………………. patterns
7. ……………………. daylight
8. ……………………. conference
9. ……………………. a discovery
10. ……………………. plants
11. go through the test of ………………….
12. perceive …………………. differences
13. ……………………. and large
14. a…………………... information

**III SPEAKING AND WRITING** Speaking - Free people icons Write - Free interface icons

**Describing Graphs**

**Vocabulary for describing graphs**

- Verbs: grow, go up, rise, soar, increase, surge, improve, shoot up, jump, surge, soar…

- Verbal phrases: reach a peak, reach the highest point, reach a plateau

- Nouns: an increase, a rise, growth, an increasing trend, a rising trend, an upward trend, an improvement, a surge, a jump

- Verbs: go down, decrease, fall, drop, decline, slump, plummet, plunge

- Verbal phrases: reach a low, hit a low/the lowest point

- Nouns: a decrease, a fall, a decline, a downward trend, a falling trend, a decreasing trend, a slump, a plunge

No change

- Verbs: remain stable, remain constant, stabilize, stay at the same level

- Verbs: fluctuate, zig-zag, flutter

- Nouns: a fluctuation, a zig-zag

Adjectives:

Small changes: moderate, steady, gradual, slow, gentle

Big changes: rapid, sudden, dramatic, considerable, sharp, significant

Adverbs:

Small changes: moderately, steadily, gradually, slowly, gently

Big changes: rapidly, suddenly, dramatically, considerably, sharply, significantly

Prepositions:

a growth **from** 25% **to** 75%

to increase **by** 50%

an increase **of** 50% **in** the drug addiction rate

Small changes: moderate, steady, gradual, slow, gentle

Useful phrases:

The slices of the pie chart compare…

The chart is divided into 3 parts.

The chart highlights...

only a third ...

less than a third…

**III. a. Describing trends. Match the words with the figures.**

a) fluctuate b) plunge c) increase d) decline e) rocket

f) recover g) dip h) peak i) stabilise

1.  2.  3.  4.  5. 

6.  7.  8.  9. 

**III. b. Explain the type of movement/trend suggested by the words given in bold.**

1. Feynman says, "From the hypothesis that the world is a **fluctuation**, all the predictions are that if we look at a part of the world we've never seen before, we will find it mixed up, and not like the piece we've just looked at — high entropy. If our order were due to a **fluctuation**, we would not expect order anywhere but where we have just noticed it. We therefore conclude the universe is not a **fluctuation**."[[2]](#footnote-2)
2. Field surveys carried out at three sites in Panama showed that, after the chytrid fungus arrived from 2004 onwards, the populations of many species **plummeted**. Among the victims were variable harlequin frogs, which are now critically endangered.[[3]](#footnote-3)
3. The Arctic climate is changing rapidly, breaking at least a dozen major records in the past three years. Sea ice is disappearing, air temperatures are **soaring**, permafrost is thawing and glaciers are melting.[[4]](#footnote-4)
4. Research funding had been **declining** and although people often said they loved science, they would then say how acupuncture had cured their back pain, or produce a salad of words like *quantum* and *consciousness* with no regard to physics or neuroscience. Science was well

loved, but much abused and rarely understood.[[5]](#footnote-5)

1. Smokers face more than twice a nonsmoker’s risk of pancreatic cancer, and even though

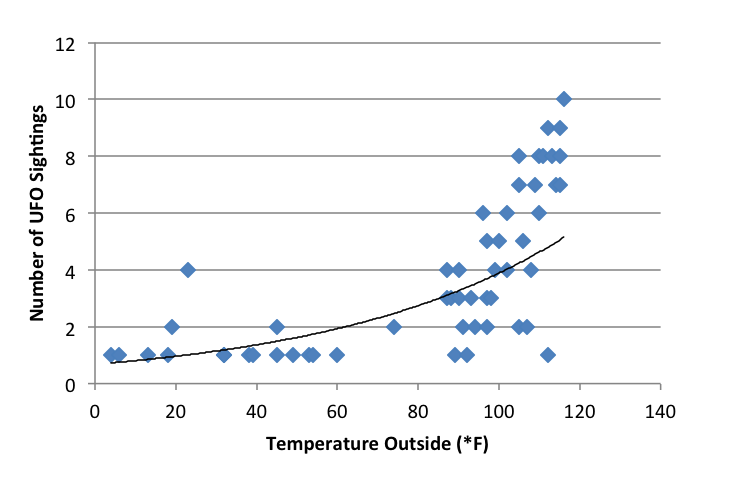
smoking has **slumped** in the U.S., there is a 30- to 40-year lag time before we see a corresponding **drop** in cancer rates.[[6]](#footnote-6)

1. The U.S. hip fracture rate **plateaued**, leading to 11,000 more fractures between 2013 and 2015 than predicted, according to a new study by Siris and her colleagues.[[7]](#footnote-7)
2. A tentacled creature called Mesodinium at 22 microns, a giant next to some of the three-micron sun-gathering plankton—comes **zigzagging** through the waters, drawn by sugars and amino acids leaking from the smaller organisms.[[8]](#footnote-8)

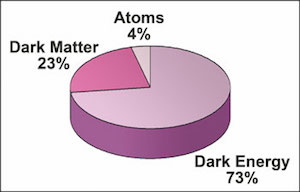
**III. c. Complete the sentences below with the correct preposition.**

1. The instantaneous neutron flux is expected to increase ……….. a factor of x.
2. The distance is increased with an increase …………. the magnetic field.
3. The gel was calcined at various temperatures ranging ………. 500 ……… 800º C for 5 h and cooled naturally to room temperature.
4. The group was divided …………. several distinct sets.
5. Some cases reported less …………… 100% efficiency.
6. The significant rise …………. the pulsed applied voltage during the pre- and propagation phase enhances the charge carrier generation.
7. ………….. can be seen ……….. Fig. 2, the temperature rises.

**III. d. Now look at some graph descriptions. Study the words and phrases given in bold.**

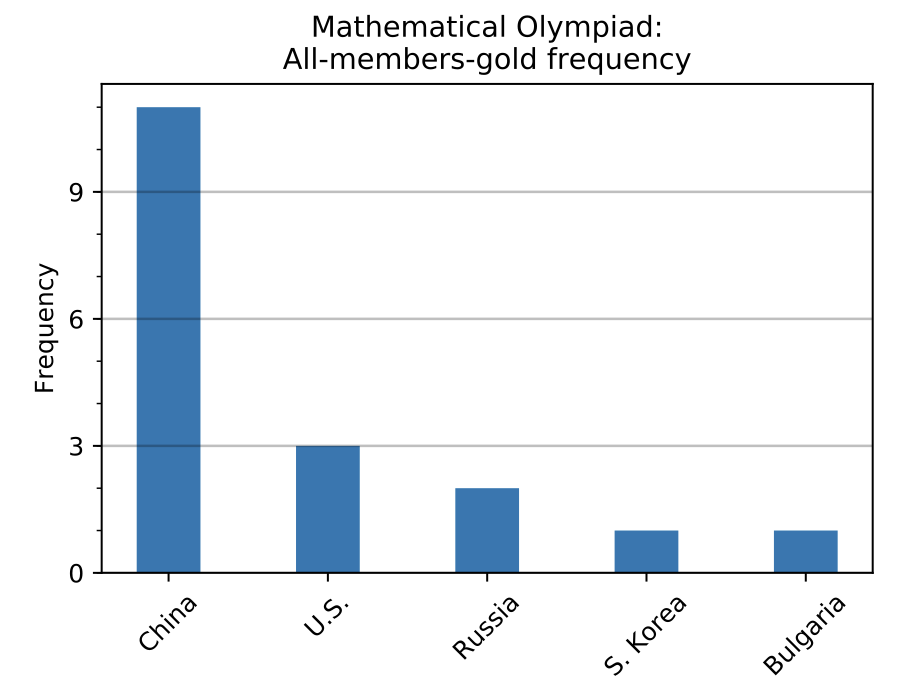


**The scatter plot shows the relationship between** temperature (x-axis, independent variable) **and** the number of UFO sightings (y-axis, dependent variable) for 53 separate data points. The temperature **ranges from** about 0°F **and** 120°F, and the number of UFO sightings **ranges from** 1 **to** 10. **The plot shows** a low number of UFO sightings (ranging from 1 to 4) at temperatures below 80°F and **a much wider range** of the number of sightings (from 1 to 10) at temperatures above 80°F. **It appears that** the number of sightings **tends to increase** as temperature increases, though there are many cases where only a few sightings occur at high temperatures.[[9]](#footnote-9)

****

**This pie chart represents the composition of** the universe. Scientists have found that **just** 4% of the universe **is made of** typical atoms, **whereas just under a quarter** (23%) of the universe is made of dark matter (which does not give off or absorb light). **Almost three quarters** (73%) are made of dark energy (which is makes the universe expand faster).

Image credit: NASA / WMAP Science Teams

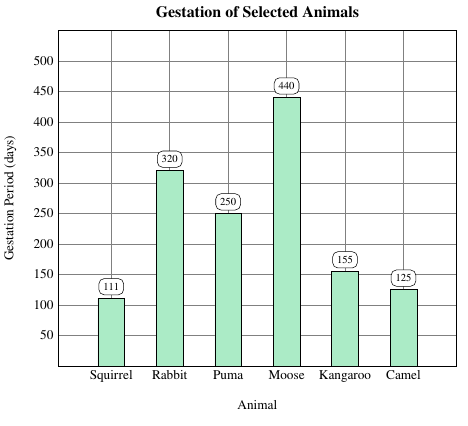
****

**This bar chart shows the distribution of** Mathematical Olympiad gold medal winners **according to** their country of origin. **The vast majority** of winners come from China – in fact, more winners come China than from the USA, Russia, South Korea and Bulgaria combined, which rank the second, third, fourth and fifth amongst the countries with the biggest number of winners. The USA **ranks second, trailing far behind** China. Russia **occupies the third position**, **whereas** South Korea and Bulgaria have **approximately the same number** of gold medal winners.

Image credit: Wikipedia, author: A. Dagur,

licence CC0 1.0

**III. e. Complete the following chart descriptions.**



This ……………………… shows the duration of gestation of different selected animals in days. ……… can be seen, ……… far the longest gestation period is that of ……………… with a total of 440 days, followed …………. rabbit, which ………….. second. Puma ……………… the ………….. place. Whereas the gestation of kangaroos lasts ……… 155 days, the gestation of camels and squirrels is ………………… similar.

Image credit: Wikipedia, author: S. Alshazly,

licence CC BY-SA 4.0

This chart ……….. the average annual temperature in the period ………….. 1878 …….. 2005. While the temperature fluctuated throughout the given period, a ……………… trend over the years is notable. The temperatures reached a ………….. in the 1980s and the 1990s. This was ……………… by a relative ……….. in the temperature in the early 2000s. However, the temperature in the 2000s was significantly ………… than it was back in the 19th century.

Describe the following charts yourself:

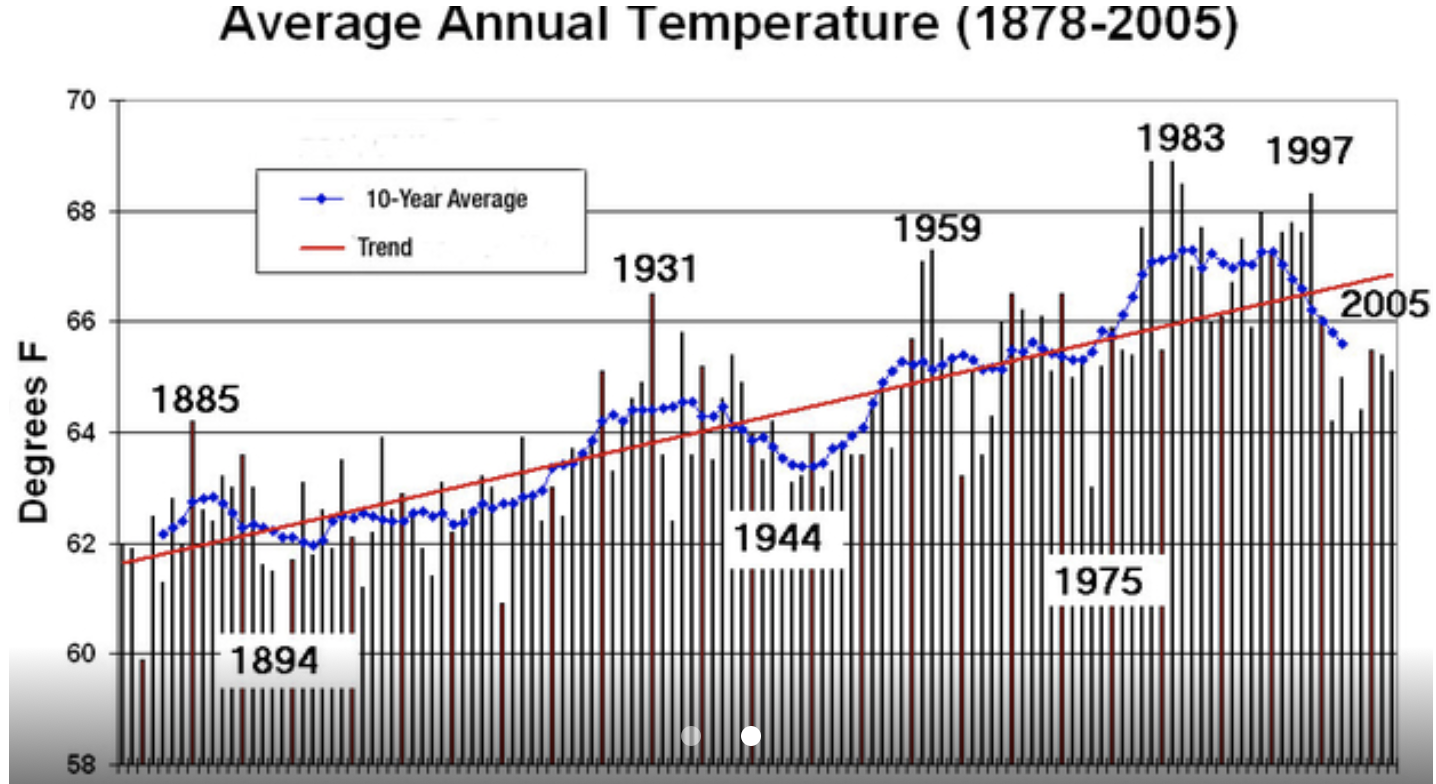
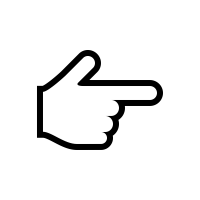
****

Image credit: NASA/JPL/Cal State L.A.

**III. f. Describe the following charts yourself. Use the words and phrases from the previous exercises.**

|  |  |
| --- | --- |
| Image credit: Wikipedia, author: O. Soul, CC BY-SA 3.0 | The types of vertebrate animals used in lab research in Europe    Image credit: Wikipedia, author: L. Lukomski, CC BY-SA 3.0 |
| Line chart - Wikipedia  Image credit: Wikipedia, author: Urocyon, CC | Image credit: *Kevin’s Amazing blog[[10]](#footnote-10)* |

 **GRAMMAR BOX**

**Transitive and intransitive verbs**

*Transitive* verbs require and object, while the *intransitive* ones do not.

Typical transitive verbs used for describing graphs include: *reduce, lower, cut, raise.*

***Reducing*** *poverty in developing countries will also help to* ***reduce*** *environmental destruction.*

*The UK government looking to* ***cut*** *science funding.*

Typical intransitive verbs used for describing graphs include: *go up/down, be up/down, rise, fall, grow, collapse, decline, slump.*

*The researchers noticed that rotation speeds* ***did not******decline*** *with increasing distance.*

*The proportion of people dying from stroke, heart disease and pneumonia/influenza* ***has fallen*** *substantially over the last sixty year.*

Some verbs, however, can be both transitive and intransitive, depending on the context: *drop, increase, decrease, extend, expand.*

*Galileo* ***dropped*** *things from the Leaning Tower of Pisa. He realized that objects* ***do not drop*** *at the same speed.*

**III. g. Complete the sentences using the appropriate transitive or intransitive verb. Use the verbs from the grammar box.**

1. To ………………………… (+) as a scientist, I need some more experience.

2. These discoveries ……………………. (-) the probability that fuel cells will be a feasible and long-term option for transportation.

3. The science budget has stopped ………………………. (-), and even …………………… (+) a little, which allowed for yet another version of the mission to Pluto.

4. This combination worked in approximately one-quarter of patients, ……………………. (-) the levels of a myeloma protein detected in their blood.

5. In fractions which can be ……………………. (-) (fractions not in lowest terms), the numerator and the denominator share at least one common factor.

**IV LISTENING** Listen Icons - Download Free Vector Icons | Noun Project

**Lecture on Data Visualization**

**IV. a. Listen to the introductory part of the lecture delivered at the MIT University[[11]](#footnote-11) and answer the following questions:**

1. What topic will be discussed this week?
2. What will the first lecture be about?
3. What will the second lecture be about?
4. What programme will be used to create visualizations? Have you ever used this programme?

**IV. b. Listen to part 2 of the lecture. Try to complete the sentences below. Which graph is described first, A or B?**

In this lecture, we'll ………………. the idea of using visualization to better understand ………………… and to provide insights on the problem we're ………………... Why visualization?

People often say that a picture is like a thousand words. In the same ………………., John Tukey, a major ………………… at Princeton, wrote that "the picture-examining eye is the best finder we have of the wholly unanticipated." Visualizing data allows us to ………………. relationships, structures, distributions, outliers, ………………….., behaviours, dependencies, and ………………….. .

Visualization is further useful for initial data exploration, for ………………….. models, and for communicating results effectively. Let us give some examples of different modes of visualization that illustrate these points.

|  |  |
| --- | --- |
| A. | B. |

**Follow-up task:**

*Choose a graph from the textbook which you use in you use in your content classes. Write a description of it using the vocabulary suggested in this unit, minding how transitive and intransitive verbs of trends are used.*

**Closing thought**

“There is no such thing as information overload. There is only bad design.”

Edward Tufte

1. This is the beginning of an article published by Knowable Magazine (knowablemagazine.org), available under CC BY-ND licence (*10.1146/knowable-110919-1*) at: https://www.knowablemagazine.org/article/mind/2019/science-data-visualization [↑](#footnote-ref-1)
2. Sean Carrol, "Distant time and the hint of a multiverse", TEDxCaltech 2011, available at: https://www.ted.com/talks/sean\_carroll\_distant\_time\_and\_the\_hint\_of\_a\_multiverse [↑](#footnote-ref-2)
3. Michael Le Page, "The frogs bouncing back after almost being wiped out by disease", *New Scientist*, March 2018, available at:

   Read more: https://www.newscientist.com/article/2165004-the-frogs-bouncing-back-after-almost-being-wiped-out-by-disease/#ixzz6R9YZIplT [↑](#footnote-ref-3)
4. Jennifer A. Francis, "The Arctic Is Breaking Climate Records, Altering Weather Worldwide ", *Scientific American*, April 2018, available at: https://www.scientificamerican.com/article/the-arctic-is-breaking-climate-records-altering-weather-worldwide/ [↑](#footnote-ref-4)
5. Johnatan Breman, "Science fans have many reasons to take to the streets again", *New Scientist*, April 2018, available at: https://www.newscientist.com/article/mg23831732-700-science-fans-have-many-reasons-to-take-to-the-streets-again/ [↑](#footnote-ref-5)
6. Claudia Wallis, "Why Pancreatic Cancer Is on the Rise", *Scientific American,* April 2018, available at: https://www.scientificamerican.com/article/why-pancreatic-cancer-is-on-the-rise/ [↑](#footnote-ref-6)
7. Claudia Wallis, "Osteoporosis: An Avoidable Crisis: Bone fractures are rife after age 50, so why aren’t we doing more to prevent them?“, *Scientific American,* March 2018, available at: https://www.scientificamerican.com/article/osteoporosis-an-avoidable-crisis/ [↑](#footnote-ref-7)
8. Adite Mitra, "The Perfect Beast: Mixotrophs, tiny sea creatures that hunt like animals but grow like plants, can change everything from fish populations to rates of global warming", *Scientific American,* Fall 2018, p. 38. [↑](#footnote-ref-8)
9. The Writing Center • University of North Carolina at Chapel Hill, "Figures and Charts“, Licenced under Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License, available at: https://writingcenter.unc.edu/tips-and-tools/figures-and-charts/ [↑](#footnote-ref-9)
10. https://kwtan16.wordpress.com/subjects/mathematics/ [↑](#footnote-ref-10)
11. The full course with audio and transcripts is available as part of MIT OpenCourseWare (https://ocw.mit.edu/courses/sloan-school-of-management/15-071-the-analytics-edge-spring-2017/visualization/welcome-to-unit-7/), under CC BY NC SA licence. [↑](#footnote-ref-11)