



# Techniki Wizualizacji Danych

**Politechnika Warszawska**

Anna Kozak | Mateusz Krzyżiński  
Mikołaj Spytek | Katarzyna Woźnica

# Cel zajęć projektowych

- wykorzystanie i utrwalenie zdobytej wiedzy z wykładu oraz laboratoriów
- praktyczna praca z danymi
- ćwiczenie sposobu prezentacji wyników

# Zasady

- 2 projekty w ciągu semestru
- zespoły 3 osobowe, różne podczas 1 i 2 projektu
- projekt trwa 7-8 tygodni
- 25 lub 29 punktów (w tym 5 punktów za pracę na zajęciach projektowych)

# Projekt 1

Zadanie: Przygotowanie plakatu na zadany temat.

Rezultat: Plakat w formacie A2

Stacjonarnie: wydrukowany + sesja plakatowa podczas wykładu

Zajęcia:

- wspólne dyskusje
- prezentacje kolejnych etapów

# Ocena

Za projekt można otrzymać od 0 do 25 punktów, z czego:

- do 2 punktów uzyskuje się za przedstawienie postępu prac w danym tygodniu (3x),
- do 5 pkt. uzyskuje się za przygotowanie poprawnych wykresów (trzy lub więcej),
- do 5 pkt. uzyskuje się, jeżeli przygotowane wykresy mają wszystkie niezbędne elementy do trafnego odczytania danych (tytuł, podtytuł, adnotacje na osiach, legenda, jednostki, opis jak czytać wykres),
- do 5 pkt. uzyskuje się za znalezienie ciekawych danych, tematów mniej popularnych w mediach, ale interesujących,
- do 5 pkt. uzyskuje się za estetykę, spójność i pomysłowość aranżacji wykresów oraz ich opisów w jedną całość, umieszczenie informacji (tytuł, autorzy, źródło danych).

# Za tydzień

- podział na zespoły 3 osobowe
- "burza mózgów"

# Kamienie milowe

Tydzien	Data	Zadanie	Punkty
1	05-10	Wprowadzenie do projektu, podział na zespoły.	
2	12-10	Praca w zespołach, burza mózgów, określenie tematyki plakatu.	1
4	02-11	Na zajęcia należy przygotować pomysły oraz pierwsze wizualizacje bazujące na znalezionych danych (prezentacja/raport w postaci pliku .pdf)	2
5	09-11	Na zajęciach zespoły prezentują swoje zaawansowane wizualizacje oraz prezentują prototyp plakatu (prezentacja/raport w postaci pliku .pdf)	2

**Temat projektu to...**

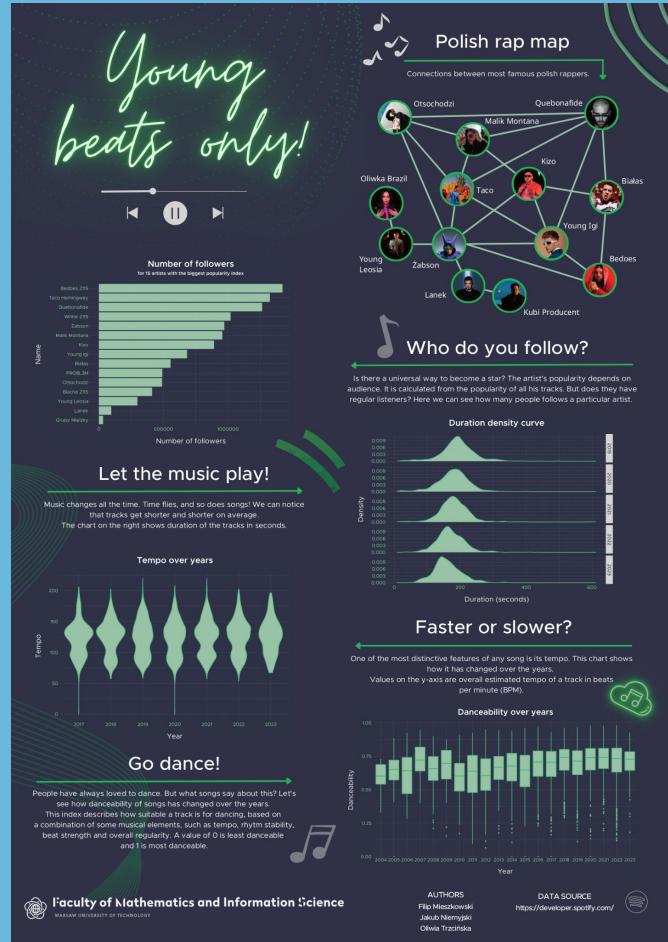
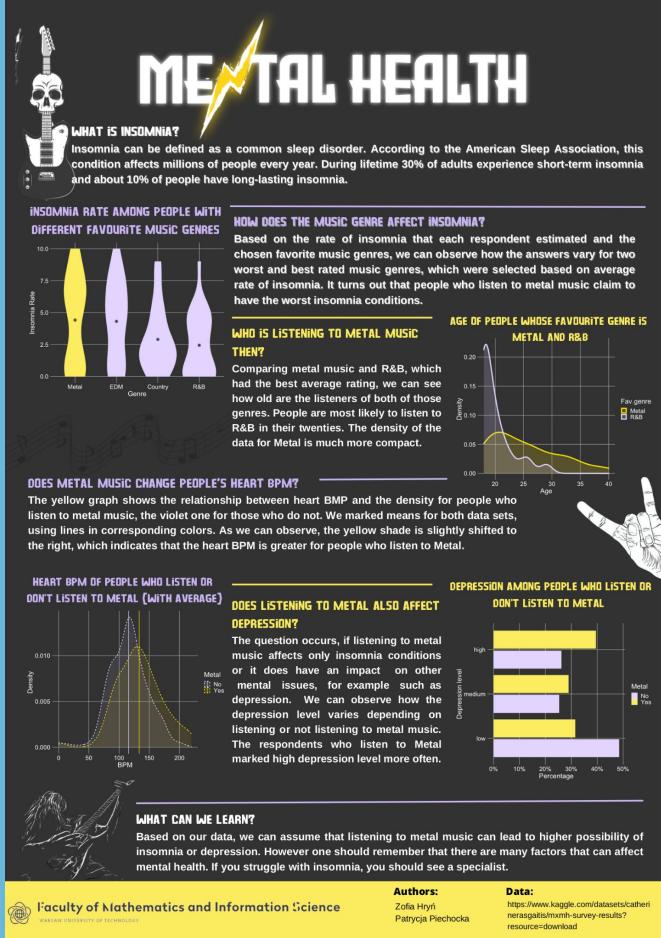
# ŻYWNOŚĆ

# **Jakie prace zostały wykonane w poprzednich latach?**

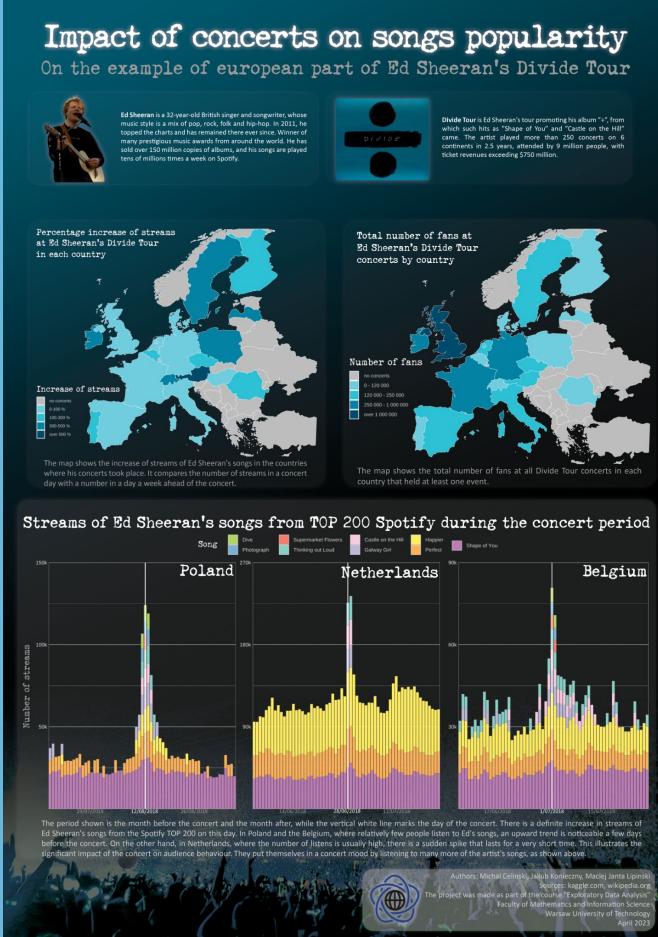
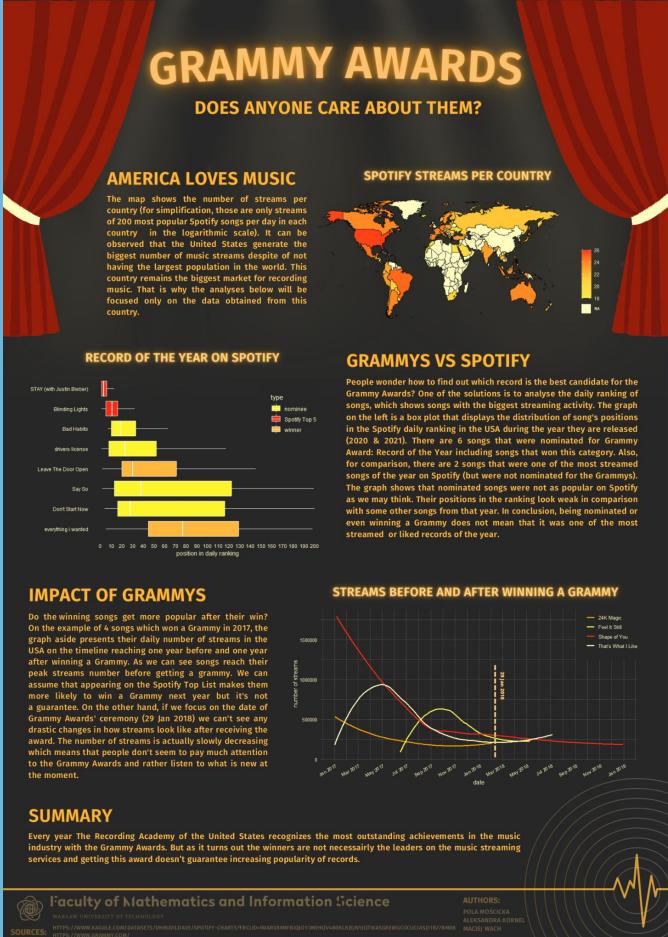
2022/2023 Muzyka, Sport

2021/2022 Plakaty, które zmieniają spojrzenie na klimat i środowisko

2020/2021, 2018/2019 Filmy, seriale, książki, audiobooki, gry



<https://medium.com/@kozaka/data-visualization-posters-let-the-music-speak-a52fbcd5687>

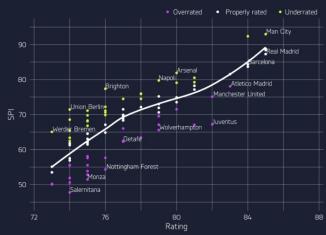


<https://medium.com/@kozaka/data-visualization-posters-let-the-music-speak-a52fbcd45687>

# FIFA VS REALITY

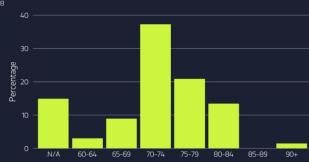
With the release of every single FIFA people argue about the ratings of players and call the creators biased. But is FIFA really that bad in terms of ratings and statistics?

## FIFA TEAM RATINGS VS SPI



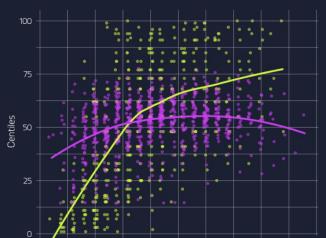
The first graph shows the dependence of the teams' overall rating in FIFA 23 on the soccer power index as of the start of the 22/23 season. SPI determines how good a team is based on match statistics from the last few months. The trend is that the higher the SPI, the higher the rating. However, at the extreme values of the rating, the trend line is a bit more vertical, which indicates that FIFA tries not to create extremely strong and extremely weak teams. It can be also observed that the greatest discrepancies between the ratings and the SPI are in case of the weaker teams.

## CURRENT RATINGS OF PERSPECTIVE PLAYERS



Can FIFA accurately predict how good a player will become? The chart on the right presents those players whose potential in years 2015-2017 was rated above 83 and had an overall rating lower than 72. As you can see, the creators' predictions were not very precise. In fact, only about 15% of those players have become world-class footballers.

## DEPENDENCE OF FORWARDS STATISTICS ON AGE



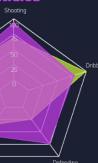
Players in FIFA are rated also on factors unrelated to their skills such as age of a player. The graph on the left compares the strikers' statistics in reality and in FIFA. According to the trend lines of these statistics, older players, despite a significant decrease in technical skills and physical ability, do not lose their statistics in FIFA. Similarly, young players whose form has exploded receive inadequate and underestimated statistics.

Great examples are Vinicius Jr (born in 2000) and Cristiano Ronaldo (born in 1985). Below you will find a comparison of their statistics in percentiles against other footballers playing in the same position.

### RONALDO



### VINICIUS



Authors:

Tomasz Urban  
Sławomir Urban  
Sławomir Trojan  
Mikołaj Wójcik

Source:

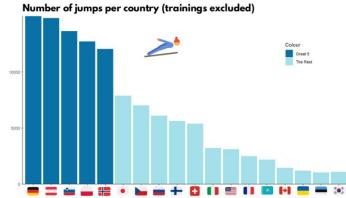
Faculty of Mathematics and Information Science  
Wojciech M. Zajączkowski  
Faculty of Mathematics and Information Science  
Wojciech M. Zajączkowski  
<https://www.wmi.krajobraz.pl/~zajaczkowski/fifa23/fifa23.html>

# See how fast they fly

## The Great 5

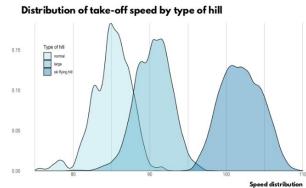
Ski jumping is still rather a niche sport - there is a huge difference between ski jumping popularity between Germany, Austria, Slovenia, Poland and Norway and the rest of the world. Because of that popularity, those countries have won 20 consecutive Nations Cups, and are therefore called "The Great 5".

To get acquainted with this weird sport, let's have a look at one of the key aspects of a great, long ski jump: the take-off speed.



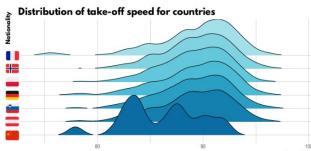
## The bigger the hill, the faster you go!

First thing we need to refer to when measuring take-off speeds, is of course the size of the hill. As it turns out, ski flying hills give competitors as much as 10-20 km/h faster take-off than most commonly used, large hill. Needless to say, this extra ordinary velocity results in extra ordinary distance - and possibly even world records!



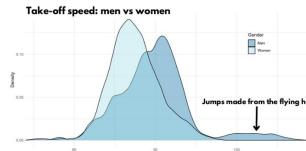
## Speed by country

Due to different techniques and proficiency of their technical skills, jumpers from different countries achieve different take-off speeds. On the chart we represented The Great 5 nations, China and France as they have interesting distributions.



## Who goes faster: men or women?

As we see on the chart below: men athletes more often have high take-off speeds in comparison to women. This effect is mostly caused by the type of hills they compete in - as now, there are no women's competitions on flying hills.



**Faculty of Mathematics and Information Science**  
WARSAW UNIVERSITY OF TECHNOLOGY

Sources:  
[https://github.com/wrotki8778/Ski\\_jumping\\_data\\_center](https://github.com/wrotki8778/Ski_jumping_data_center)  
<https://www.wi-ski.com/DB/general/biographies.html>

Authors:  
Michał Gęska  
Rafał Pyzowski  
Jakub Seliga

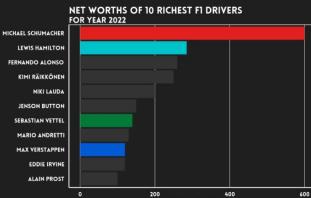
# GREATEST F1 DRIVERS



Lewis Hamilton  
37 y.o.  
Max Verstappen  
25 y.o.  
Michael Schumacher  
53 y.o.  
Sebastian Vettel  
35 y.o.



Undoubtedly, one of the most important and prestigious indicators of a Formula One driver's success must be the total number of Grand Prix won in the career. By looking at the rate at which this number grew across the seasons, one can achieve valuable insight into the driver's career. For each driver, the graphs follow a similar shape, but in the case of Vettel and Verstappen, a sharp increase in wins begins earlier. Will Verstappen be able to keep such pace?

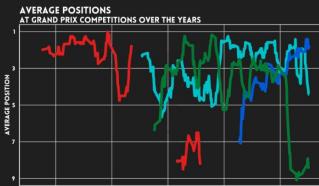


In each Formula One Grand Prix, only 20 drivers compete, making it one of the most competitive sports to enter. In addition to contracts with teams, getting to the top guarantees numerous sponsorships. For this reason, drivers accumulate a substantial fortune throughout their careers.

Formula One is an exhilarating sport with plenty of unique careers in its history. The discussion regarding which drivers deserve to be called the best could last for hours. The following list, however, focuses on the subjectively selected four drivers, which, in our opinion, are the most popular and entitled in modern history of Formula One.



Many circuits in Formula One are known for their character. Nuances in construction, cheering of the audience, or drivers' individual preferences determining which podium composition is predicted on a given track. But is it justified? The ratio between the number of races won and completed shows that certain tracks have their favourites, which is not without significance when planning the strategy for the season.



Each driver celebrated numerous victories in a certain golden period of their career. We tracked drivers' performances over the years by looking at their race positions using a 12-months rolling average. Such a comparison shows the long-term interaction between drivers and allows one to compare the stability of their performance on the track.



Authors:

- Wiktor Wierzbicki
- Tymoteusz Kowalewski
- Marcin Nowakowski

Sources:

- <https://www.kaggle.com/datasets/kharlamafm/formula-one-championship-1950-2020>
- <https://www.kaggle.com/datasets/kharlamafm/best-formula-one-drivers-and-their-careers-max-verstappen-is-177-times-more-earning/>
- <https://www.scmp.com/magazines/style/celebrity/article/3182964/10richest-f1-drivers-all-time-net-worths-ranked-lewis>

# Wyścigi konne



## Popularność wyścigów konnych w latach 2010-2020

mierzona przez liczbę goniów na terenie krajów



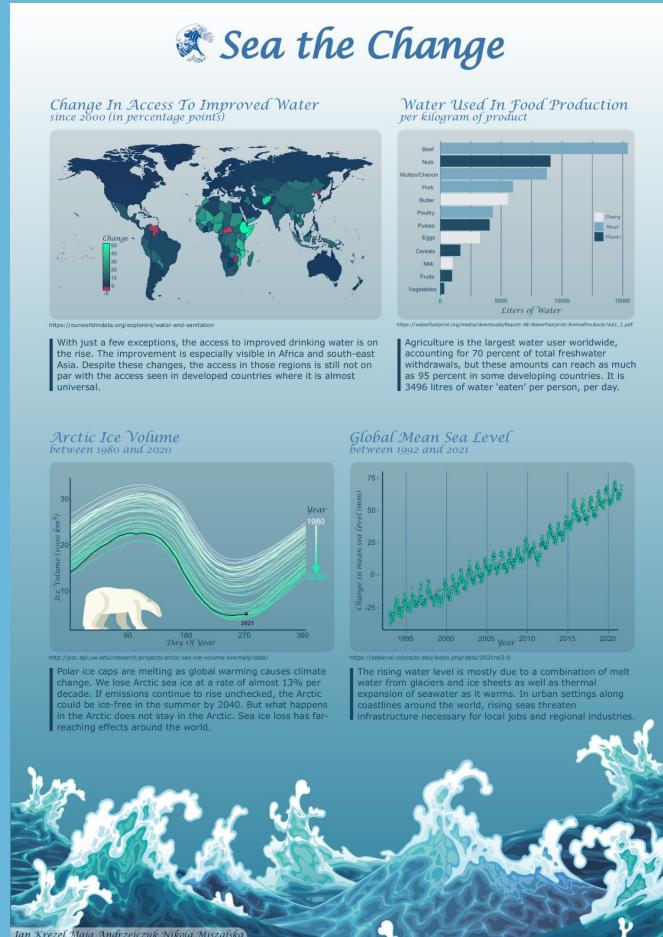
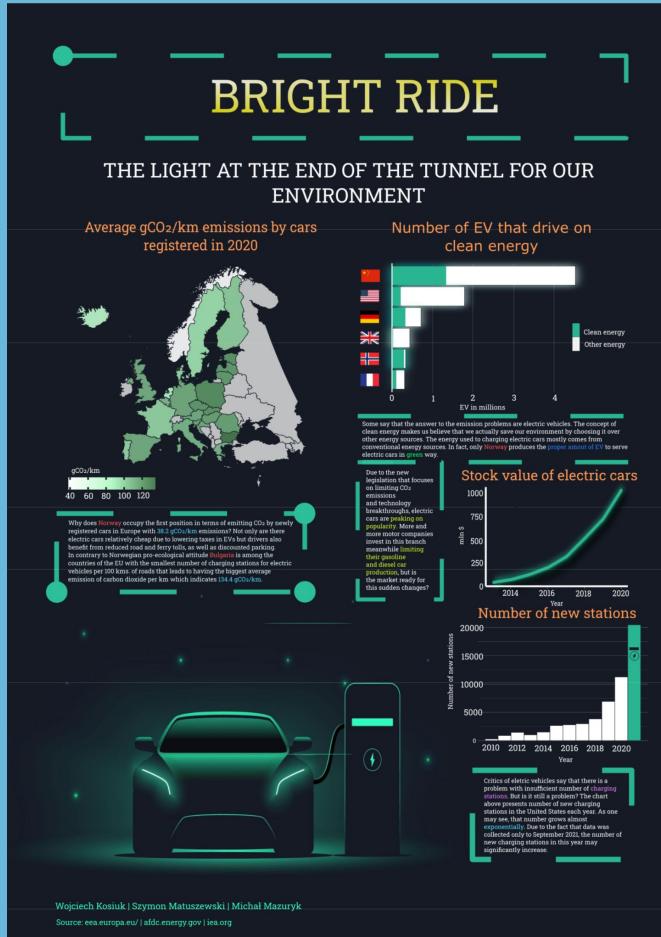
Wyścigi konne cieszą się popularnością na całym świecie, przy co najmniej 1,5 milionami koni. Wielka Brytania, gdzie organizowane jest około 10 tysięcy wyścigów rocznie, jest to 4 najbardziej popularny sport w UK bazując na liczbach goniów i wyników. Brytyjski rynek wyścigów konnych jedz szacowany rocznie na wartości 4,1 miliarda funtów. Nalejkowy wyścig the Grand National ogląda około 600 milionów widzów na całym świecie.



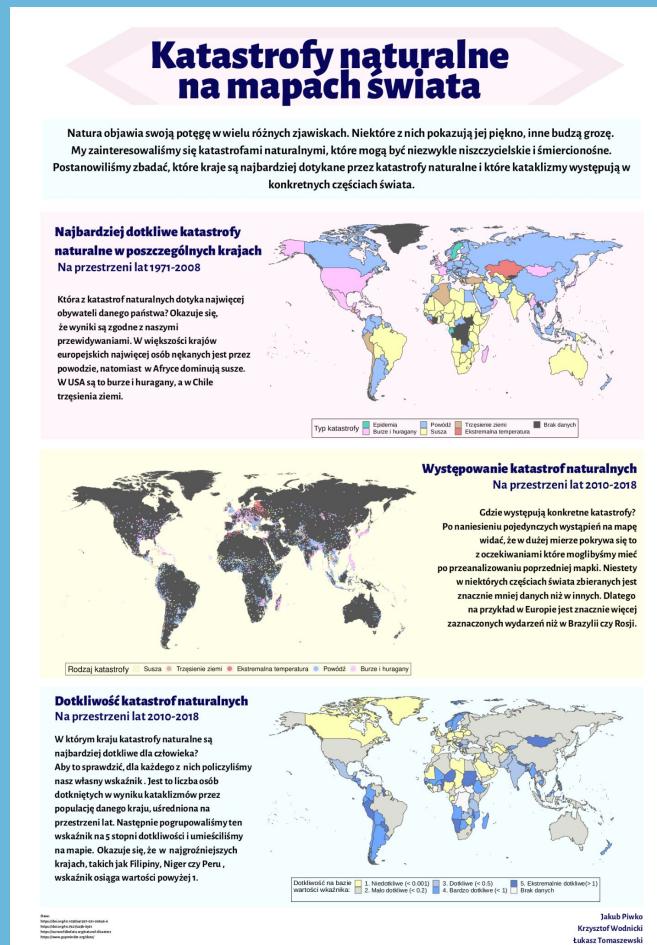
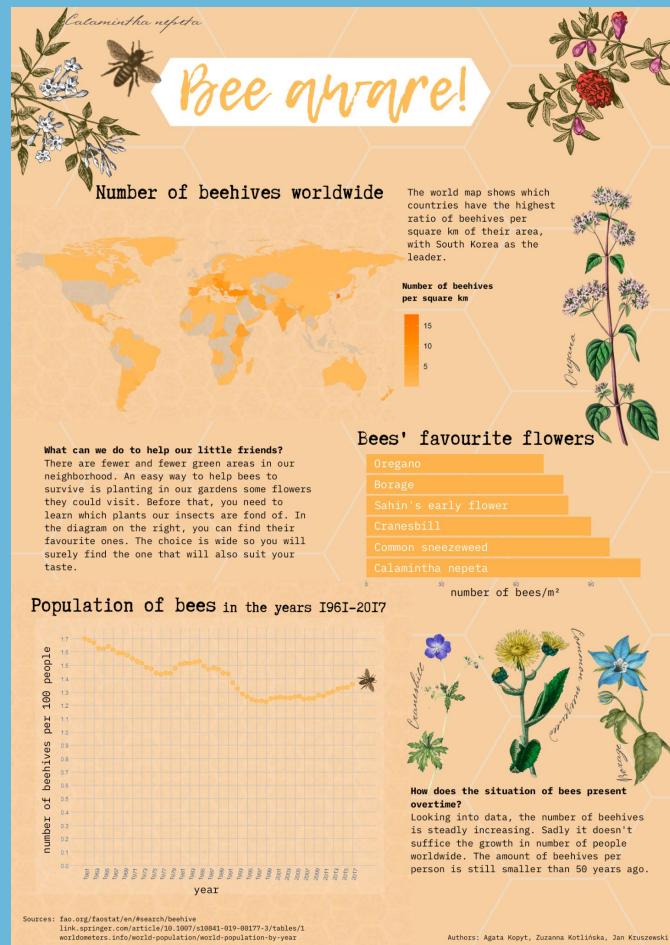
Ostatnio wyścigi konne jest bardzo popularne na całym świecie. Badania sugerują, że szansa na wygranie favorita wynosi około 32%, jednak konie, który był najbardziej wskazywane do wygrania wyścigu osiąga jeszcze lepsze wyniki. Rekordowy - klucz Winx - na 36 wyścigach, w których była favoritem, wygrała 34 z nich.

Najczęstszym typem obstawiania zakładów bukmacherskich jest zakład zwyczajny polegający na wytypowaniu konia, który zwycięży goniówkę. Jest też mniej ryzykowny, zakład porządkowy, podczas którego typuje się dwa konie, które jako pierwsze prześlą metaę. Nie ma wtedy znaczenia, który z nich zwycięże. Są też zakłady typu dwójka, trójka i czwórka, w których kolejność ma już znaczenie.





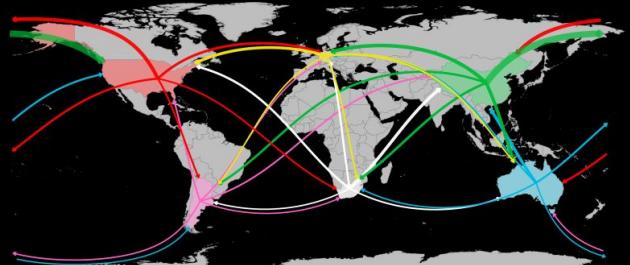
<https://medium.com/responsibleml/posters-that-change-the-perspective-on-climate-and-the-environment-c3682c0f6c39>



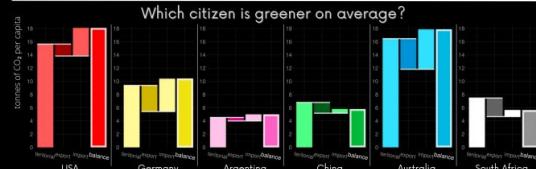
<https://medium.com/responsibleml/posters-that-change-the-perspective-on-climate-and-the-environment-c3682c0f6c39>

# The hidden emissions

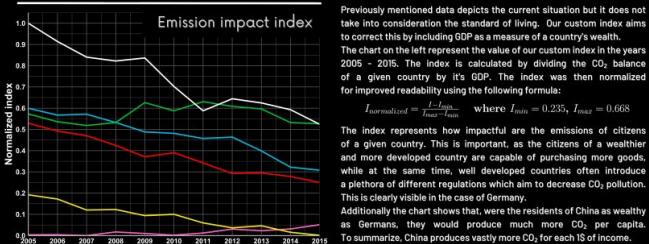
CO<sub>2</sub> in production and trade



Arrows coming out of a country have the same color as the country and they represent export from it to all others. The width of the arrow represents the amount of CO<sub>2</sub> exported in 2015. From this chart we can clearly see that some countries, like China, export much more than they import; while other countries, like Australia, do the opposite. Those exports and imports are usually omitted when calculating countries' emissions. This leads to significant changes in total emissions.



The other factor that usually is not taken into account is the number of people living in the country. Although the map would imply that China is the biggest polluter in the world, it does not necessarily reflect the whole truth. The chart above presents carbon dioxide values per capita in 2015. From the per capita perspective there seems to be a discrepancy between what general public reckons and what the numbers say. As one can observe, it would appear that USA CO<sub>2</sub> demand surpasses the Chinese by a whooping 10 tonnes disparity.



Faculty of Mathematics and Information Science  
University of Warsaw

Piotr Rakus Kacper Trębacz Małwina Wojewoda

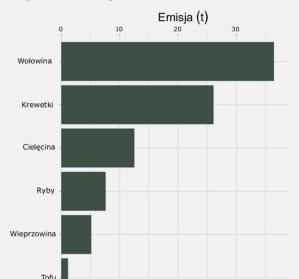
ourworld-in-data.org Data source: stats.oecd.org

# Produkcja żywności a ekologia

Emisja CO<sub>2</sub> per capita w roku 2013  
w wyniku produkcji żywności



Emisja CO<sub>2</sub> w wyniku produkcji żywności w przeliczeniu na 1000 kcal



Spożycie wybranych produktów w 2013 roku  
W kilogramach per capita

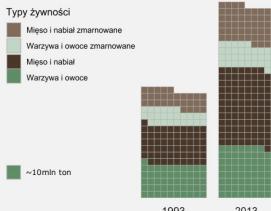


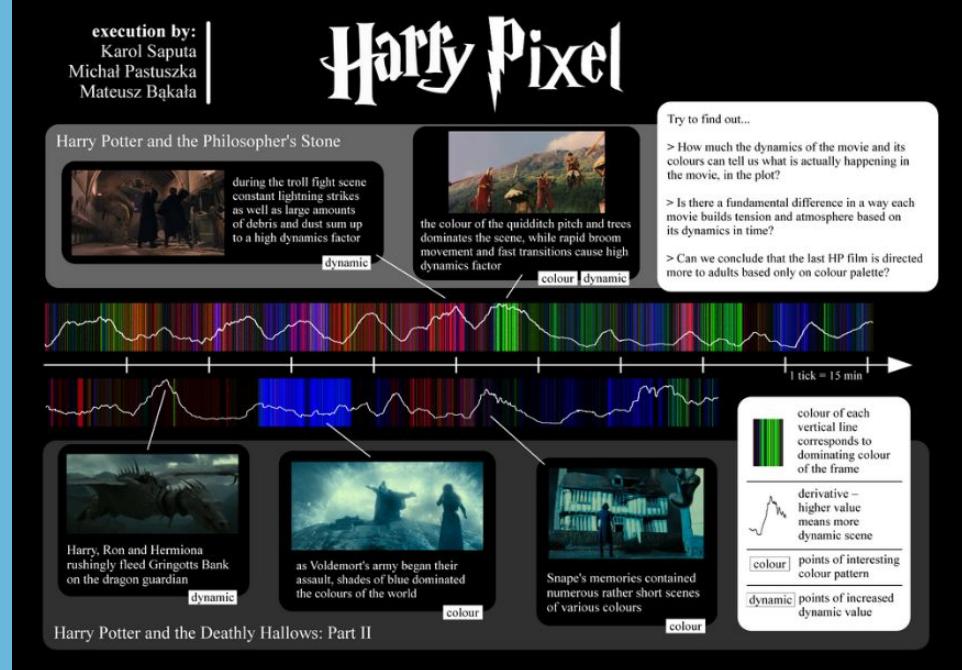
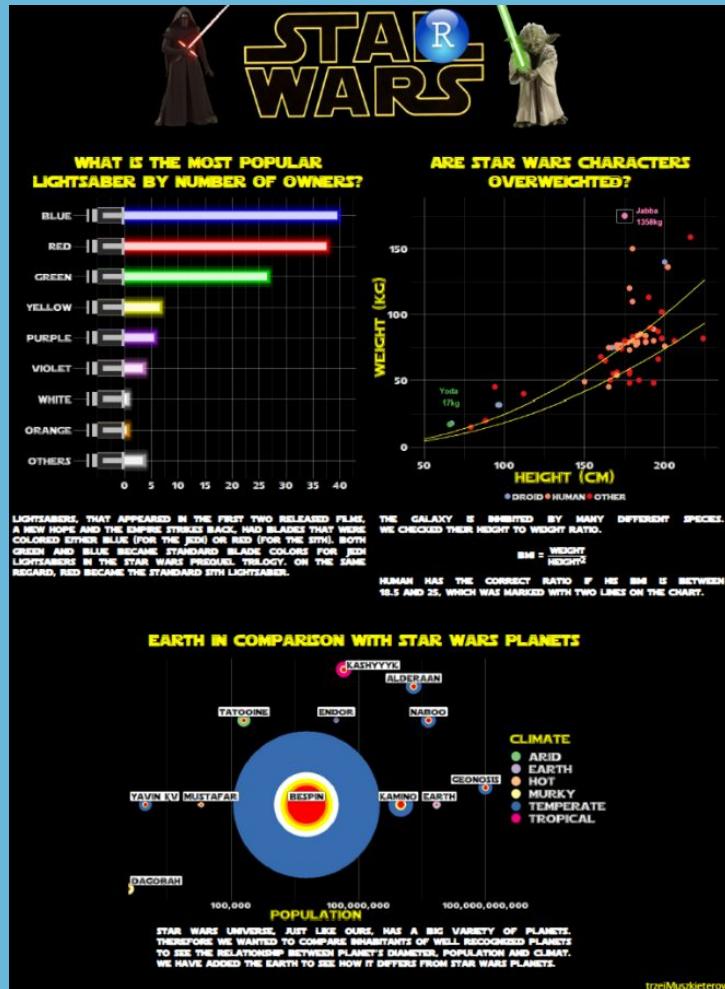
Dawid Płudowski  
Antoni Zajko  
Grzegorz Kiersnowski  
Źródła: FAOSTAT, Kaggle, OWID

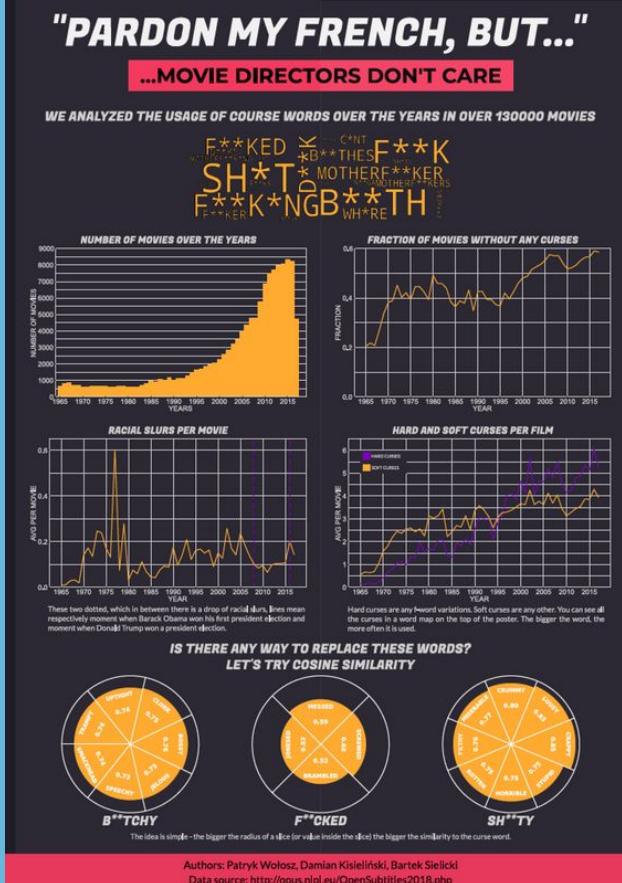
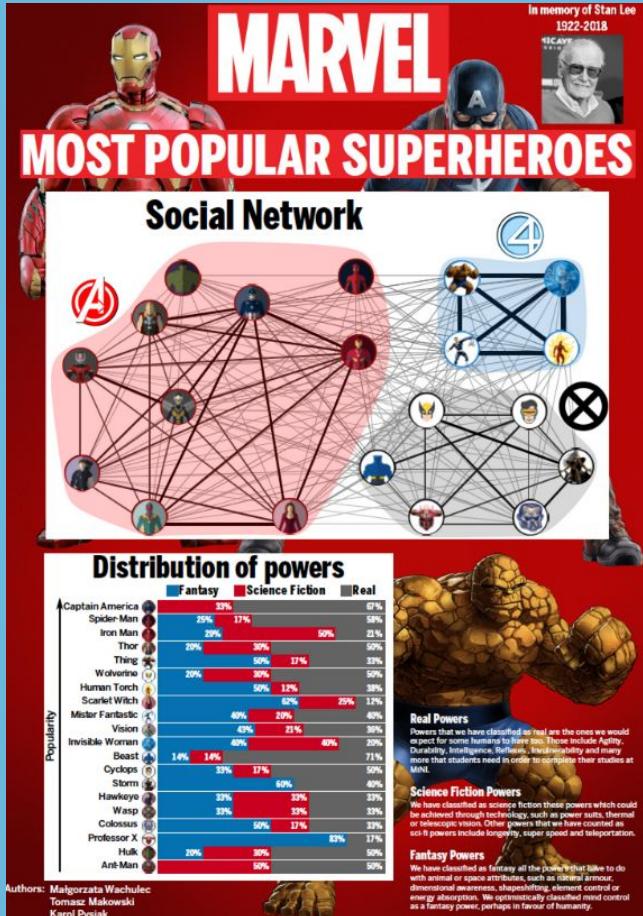
CO<sub>2</sub>[t]

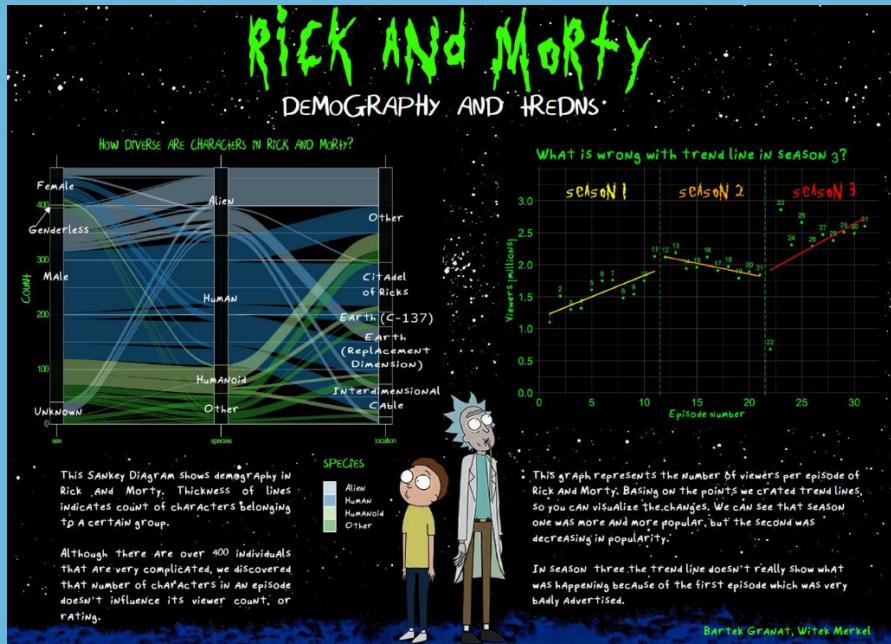
Bardzo sugerują, że 20% dwutlenku węgla, który wytwarzamy, pochodzi z produkcji żywności. Marnowanie jedzenia również przyczynia się do transmisji nadmiarowego dwutlenku węgla do atmosfery. Najbardziej odpowiedzialne za ten stan rzeczy są kraje rozwinięte.

Produkcja i marnowanie żywności  
W latach 1993 i 2013





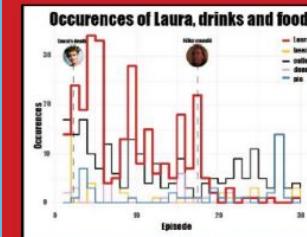




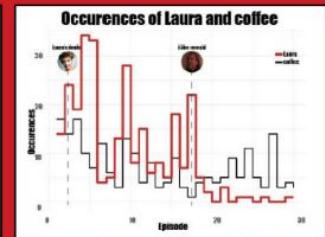
# Did coffee help to solve the murder case?

The main theme of series called Twin Peaks is murder of Laura Palmer. Investigation is lead by FBI agent Dale Cooper, who is a huge fan of coffee and cherry pie. Actually everyone who lives in Twin Peaks is a huge fan of coffee and cherry pie. Let's find out if their favourite food and drinks helped them with finding the murderer.

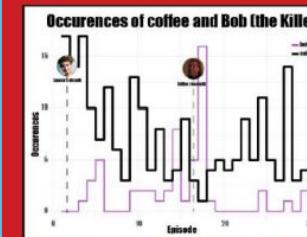
The most important thing on charts below are 2 marked moments: Laura's death and killer revelation. Apart from that, we can observe occurrences of specific words in subtitles per episode. There were 30 episodes in total in Twin Peaks series.



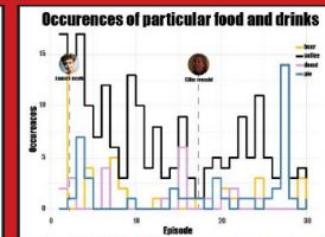
On this plot we can see all most popular food and drinks that were consumed during the story. We can observe the fact that coffee is the most consumed during the whole series. Which is pretty obvious - it's Twin Peaks favorite!



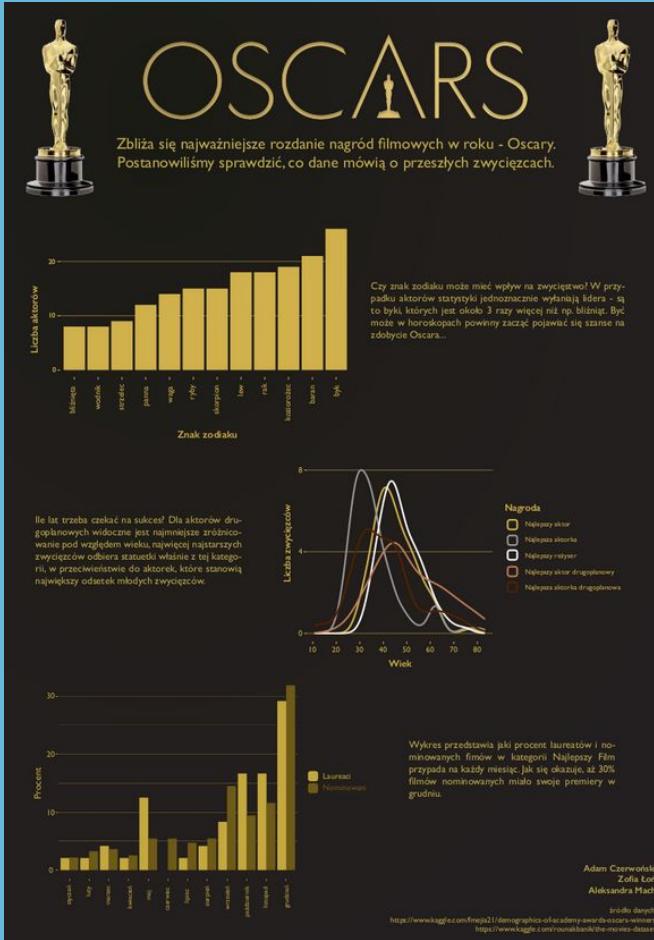
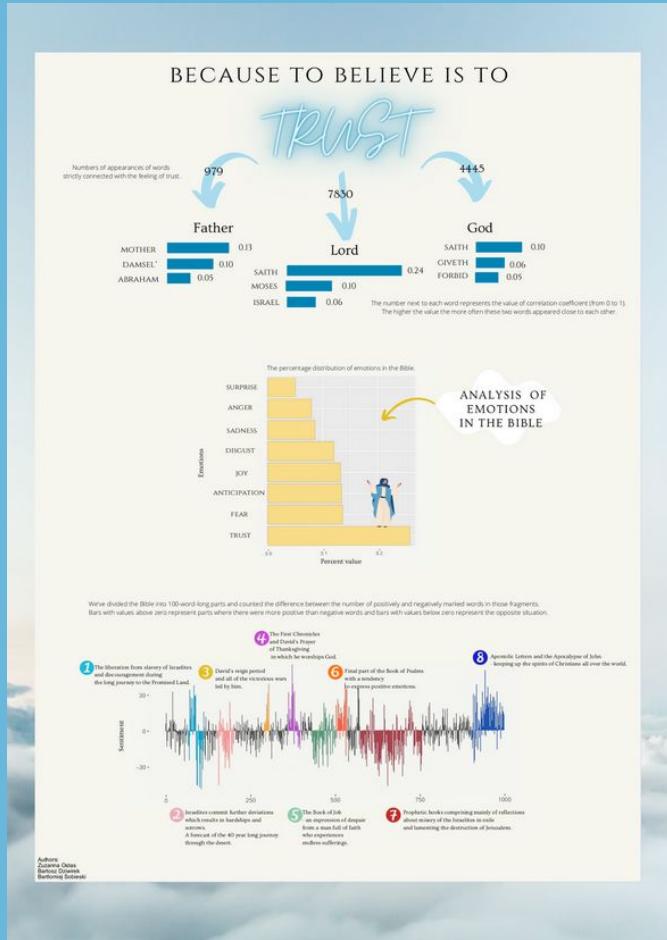
On this plot we can observe mentions of Laura and coffee. We can see that at the beginning of investigation, when everybody is excited about it, there are much more occurrences of Laura name and more coffee consumption.



On this plot we can see how coffee motivated people to find Bob, who was Laura's killer. There is obvious connection between their occurrences.



On this plot we have just food and drinks. It's supposed to show us the influence of particular food during 30 episodes.



<https://medium.com/responsibleml/poster-make-movie-book-series-3ac2c8a01180>

