

PROJEKTY

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
- 24p (P1) i 20p (P2) za projekt (w tym do 5p za pracę na zajęciach projektowych)

PROJEKT 1

Zadanie: Przygotowanie plakatu na zadany temat

Rezultat: Plakat w formacie A2 wydrukowany + sesja plakatowa podczas wykładu

Zajęcia:

- wspólnie dyskusje
- prezentacje kolejnych etapów

Ocena

Za projekt można otrzymać od 0 do 24 punktów, z czego:

- 5p (1 x 1p, 2 x 2p) uzyskuje się za przedstawienie postępu prac w danym tygodniu
- 5p uzyskuje się za przygotowanie estetycznych wykresów (dwa lub więcej)
- 5p uzyskuje się, jeżeli przygotowane wykresy mają wszystkie niezbędne elementy do poprawnego odczytania danych (tytuł, podtytuł, adnotacje na osiach, legenda, jednostki, opis jak czytać wykres)
- 5p uzyskuje się za estetykę i pomysłowość aranżacji wykresów i opisów w jedną całość
- 4p uzyskuje się za prezentację projektu

Za tydzień

- podział na zespoły 3 osobowe
- "burza mózgów"

Temat projektu to...

**FILM, SERIAL,
KSIĄŻKA, AUDIOBOOK**

Jakie prace zostały wykonane w poprzednich latach?

2023/2024 Żywność

2022/2023 Sport, Muzyka

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

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

A SIP OF LUXURY

10.5 billion kilograms of coffee were consumed in 2022, with Europeans averaging 5 kilograms per person. Specialty coffees make up 20% of the market by weight but over 40% by value, with a projected growth to \$5.1 billion by 2030. Young adults, 18-24 years old, are notable consumers, accounting for 32% of specialty coffee consumption.

Coffee profile based on SCA Scores

A radar chart categorizes coffees into three quality groups, showing that specialty coffees nearly universally score high points for sweetness, cleanliness, and consistency. Specialty coffees demonstrate a higher rating in attributes such as acidity and flavor, distinguishing them significantly from coffees scored around 80 points.

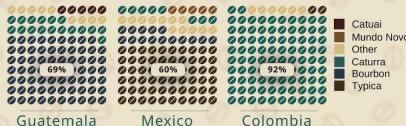


PRICE INSIGHTS

The detailed violin chart lays out the price distribution for different coffee varieties, pinpointing median cost for each. Notably, the Geisha variety stands out with its exquisitely complex and rich flavor profile, which includes fruit, floral, and chocolate notes, making it a high-value bean in the coffee market. The Kona variety, cultivated in Hawaii's unique volcanic soil, offers a distinctive chocolate-caramel taste and fetches a high price due to its limited production and superior quality standards.

Cultivation Patterns

The waffle chart shows coffee varieties by region: Typica leads in Mexico, Bourbon varieties are favored in Guatemala, and Caturra is preferred in Colombia, each selected for their distinct traits and local adaptability.



Trade Map

The map indicates Brazil as the largest coffee exporter with 1.98 billion kilograms, followed by Vietnam. The United States emerges as the largest importer with 1.6 billion kilograms, with the map's color shades—brown for exporters and blue for importers—visually quantifying the scale of the global coffee trade.

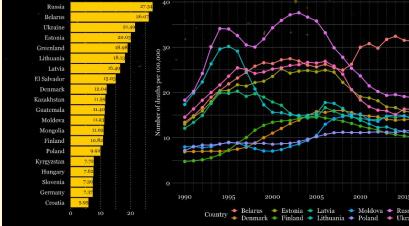


Adam Kaniasty | Hubert Kowalski | Igor Kalodziej

Source: kaggle/datasets
ICO Coffee Dataset
Coffee Quality database from CQI
Coffee Reviews Dataset

ALCOHOL CONSUMPTION

Death rate because of excessive alcohol use depending on Country and Year



Percent of people who began drinking before age 13

Despite the legal drinking age being 18 years (for many countries), the majority start drinking at an earlier age. The map of Europe illustrates the prevalence of individuals who first tried alcohol before age 13. The study was conducted among 15-year-olds by the World Health Organization (WHO). The highest percentage of individuals consumed alcohol before the age of 13 in Estonia, at 49%, Greece and Lithuania follow closely with 43%. Conversely, Iceland reports the lowest percentage, a mere 5%, while Poland has 32%.



Percentage of men vs. women who drank alcohol in the past year

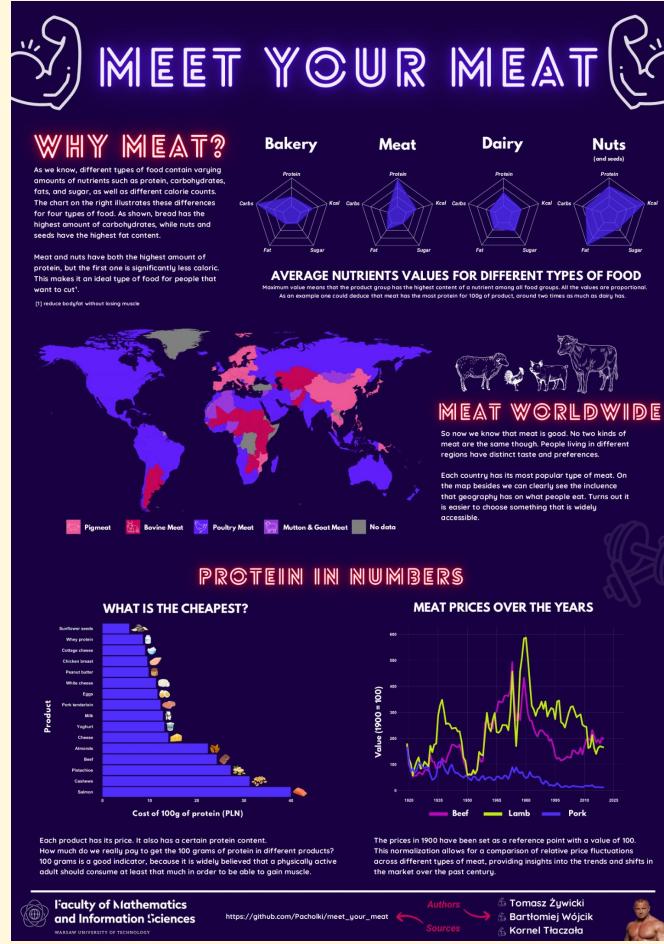


This graph shows for each country the percentage of women and men who drink alcohol at least once in the past year (2016). It is noticeable that in all countries the percentage of men who drank in the past year is higher than the percentage of women. We can see the highest percentages in Luxembourg, Russia, Ukraine, Belarus line up in something similar to a parabola branch. For the 4 outliers, it is observed that women drink almost as much alcohol as men. If we look by continent, we can see that the highest percentage for both genders is in Europe, and the lowest in Asia, Africa and Oceania (except Australia and New Zealand).

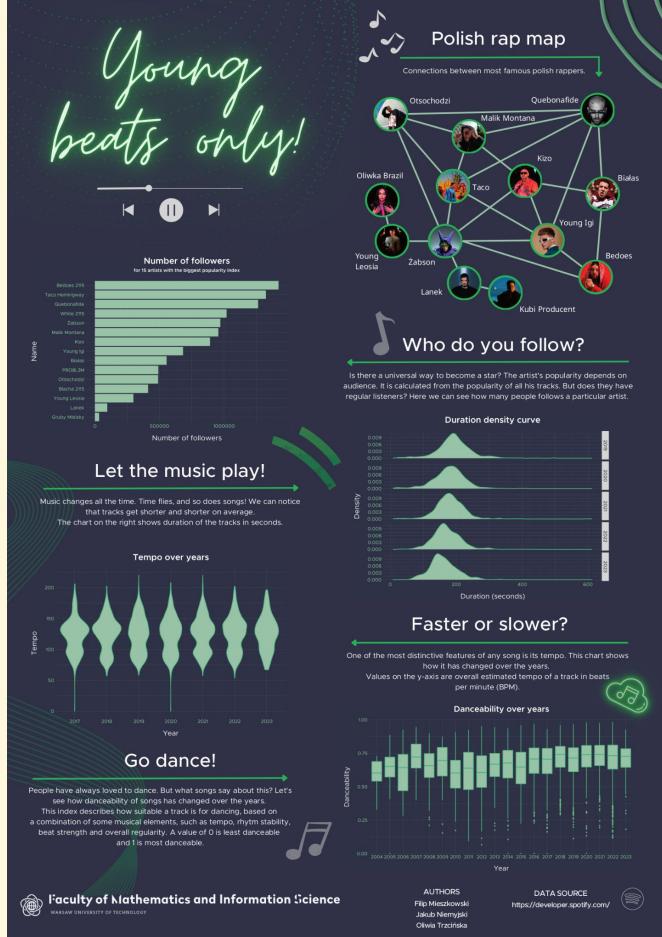
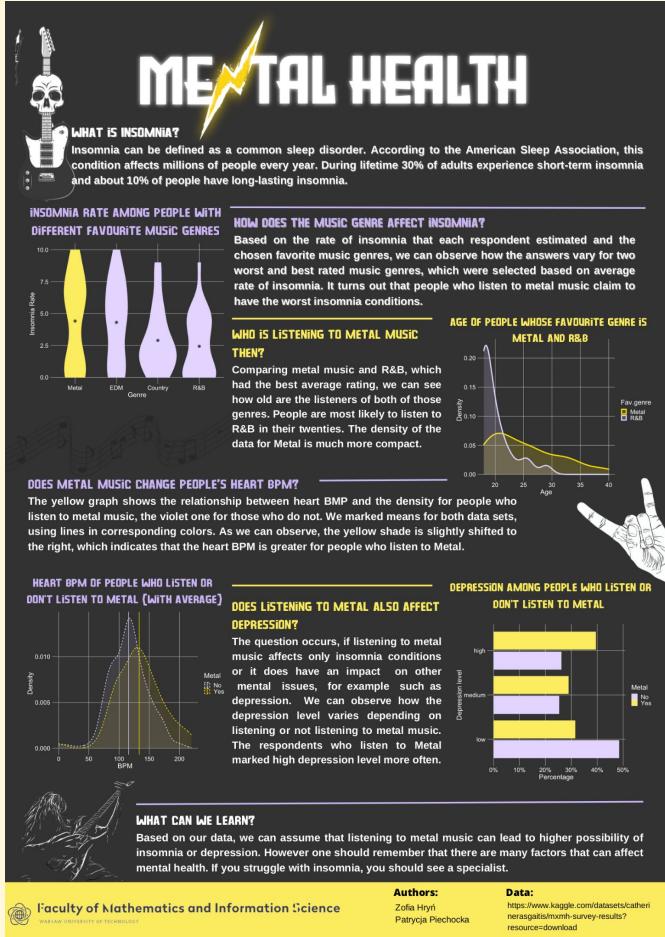
Faculty of Mathematics and Information Science
WARSAW UNIVERSITY OF TECHNOLOGY

DATA SOURCES:

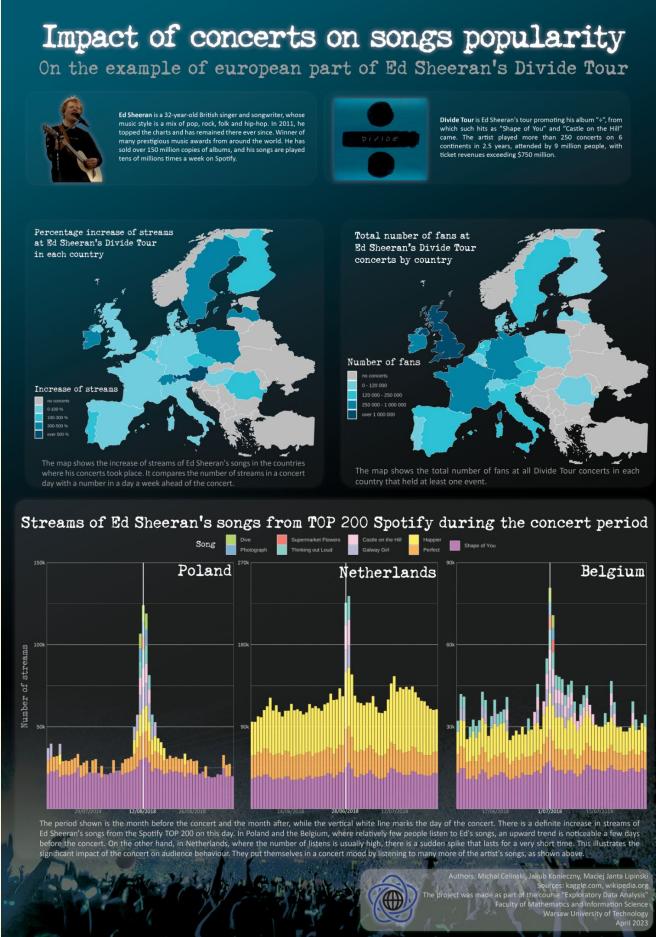
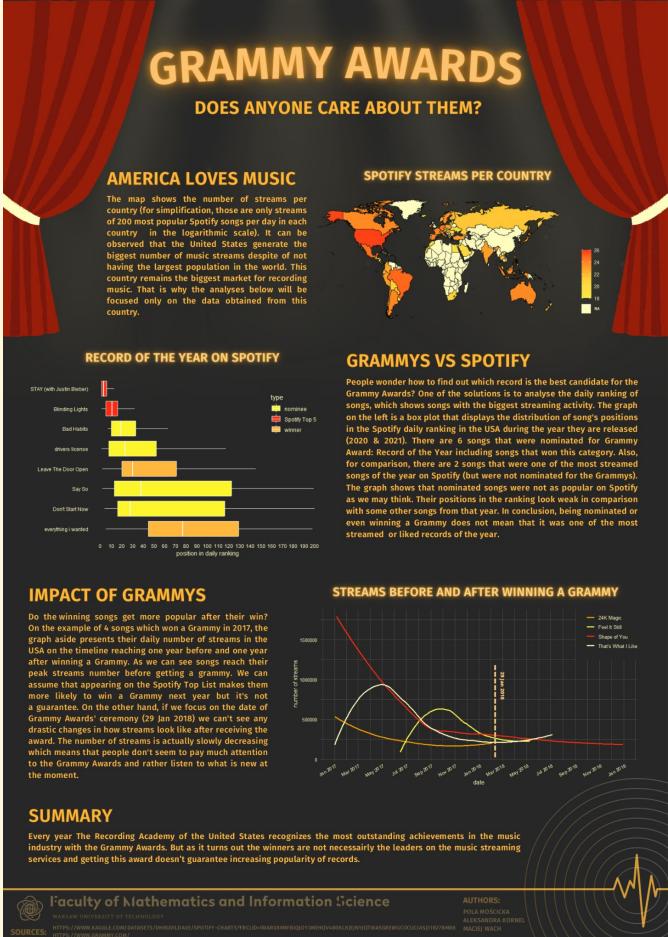
BAZIERA ULEK
MARIAN SAWIŃSKA
PAHIAN MILANA
AUTHORS:



<https://medium.com/@kozaka/a-flavour-of-posters-posters-about-food-2a1786c115dc>



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

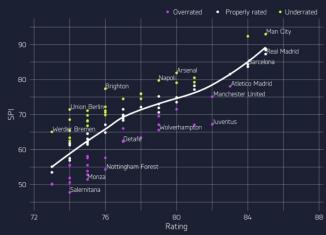


<https://medium.com/@kozaka/data-visualization-posters-let-the-music-speak-a52fbcd5687>

FIFA VS REALITY

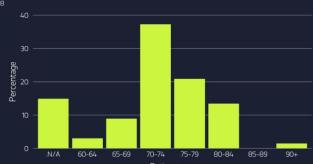
With the release of every single FIFA people argue about the ratings of players and call the creators biased. But is FIFA really that bad in terms of ratings and statistics?

FIFA TEAM RATINGS VS SPI



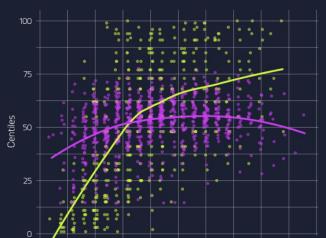
The first graph shows the dependence of the teams' overall rating in FIFA 23 on the soccer power index as of the start of the 22/23 season. SPI determines how good a team is based on match statistics from the last few months. The trend is that the higher the SPI, the higher the rating. However, at the 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.

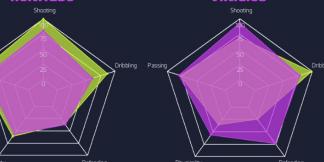
DEPENDENCE OF FORWARDS STATISTICS ON AGE



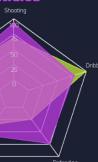
Players in FIFA are rated also on factors unrelated to their skills such as age of a player. The graph on the left compares the strikers' statistics in reality and in FIFA. According to the trend lines of these statistics, older players, despite a significant decrease in technical skills and physical 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:

Florian Urban
Sebastian Trojan
Wojciech Wieliczka

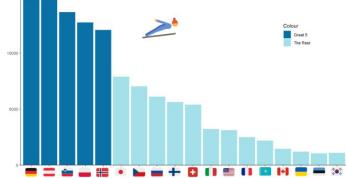
See how fast they fly

The Great 5

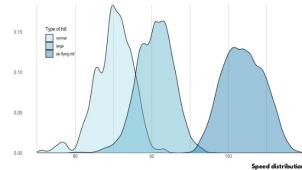
Ski jumping is still rather a niche sport - there is a huge difference between ski jumping popularity between Germany, Austria, Slovenia, Poland and Norway and the rest of the world. Because of that popularity, those countries have won 20 consecutive Nations Cups, and are therefore called "The Great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

Number of jumps per country (trainings excluded)



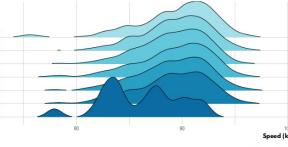
Distribution of take-off speed by type of hill



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.

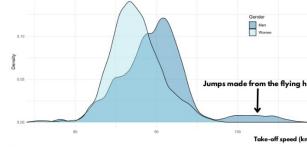
Distribution of take-off speed for countries



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.

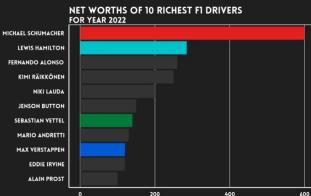
Take-off speed: men vs women



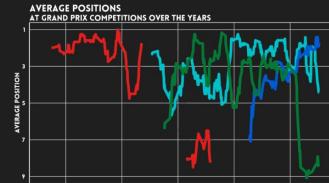
GREATEST F1 DRIVERS



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.



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:

- Wojciech Wierzbicki
- Tymoteusz Kowalewski
- Marcin Nowakowski

Sources:

- <https://www.google.com/search?q=letharose+formula+1+world+championship+1950-2022>
- <https://www.google.com/search?q=fernando+alonso+and+max+verstappen+4.737+kids+earns+1000000000>
- <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



W latach 2010-2020 pojed. 1,5 miliona koni wieku udział w wyścigach konnych na całym świecie. Należącym z nich zwycięzcom w roku 2020 w wykigu konia, jego wiekiem, jednak liczba goniów, w których uczestniczy nie przekracza mety, nie zajmuje przy tym żadnej pozycji, zaznacząc się zwiększa czym starszy koń. Dla grupy wiekowej 12+ jest to ponad 35%.



Szansa na zajęcie pierwszego miejsca dla pięciu najczęstszych favoritów



Ostatnio wyścigi konne jest bardzo popularne na całym świecie. Badania sugerują, że szansa na wygranie favorita wynosi około 32%, jednak konie, które były najczęściej wskazywane do wygrania wyścigu osiągają jeszcze lepsze wyniki. Rekordowy - klucz Winx - na 36 wyścigach, w których była favoritem, wygrał je 34 z nich.

Najczęstszym typem obstawiania zakładów bukmacherskich jest zakład zwyczajny polegający na wytycaniu konia, który zwycięży goniówkę. Jest też mniej ryzykowny, zakład porządek, podczas którego typuje się dwa konie, które jako pierwsze prześlą metaę. Nie ma wtedy znaczenia, który z nich zwycięża. Rekordowy - klucz Winx - na 36 wyścigach, w których był favoritem, wygrał je 34 z nich.

Na co dzień typem obstawiania zakładów bukmacherskich jest zakład zwyczajny polegający na wytycaniu konia, który zwycięży goniówkę. Jest też mniej ryzykowny, zakład porządek, podczas którego typuje się dwa konie, które jako pierwsze prześlą metaę. Nie ma wtedy znaczenia, który z nich zwycięża. Rekordowy - klucz Winx - na 36 wyścigach, w których był favoritem, wygrał je 34 z nich.



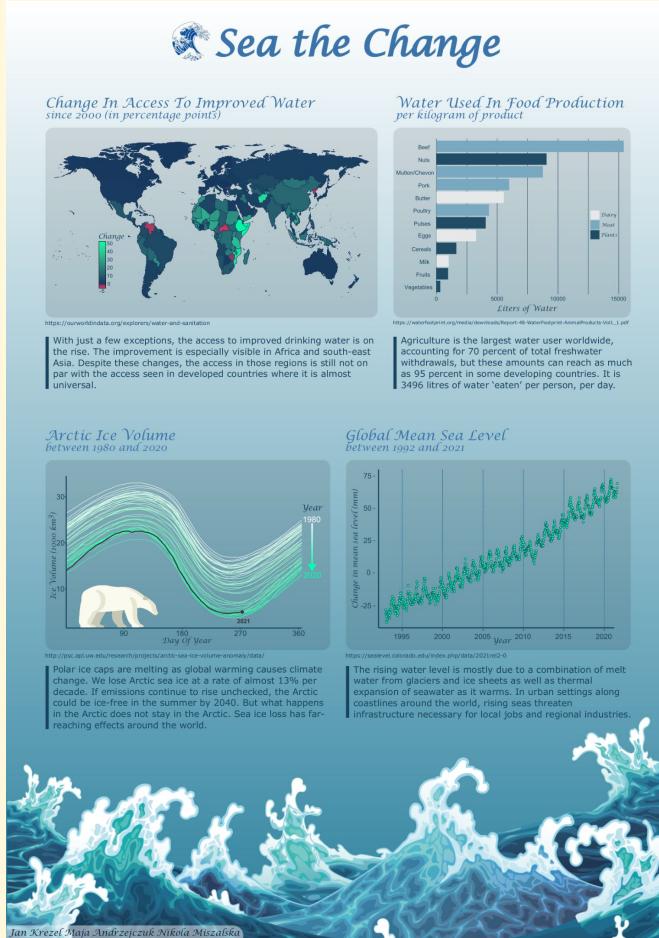
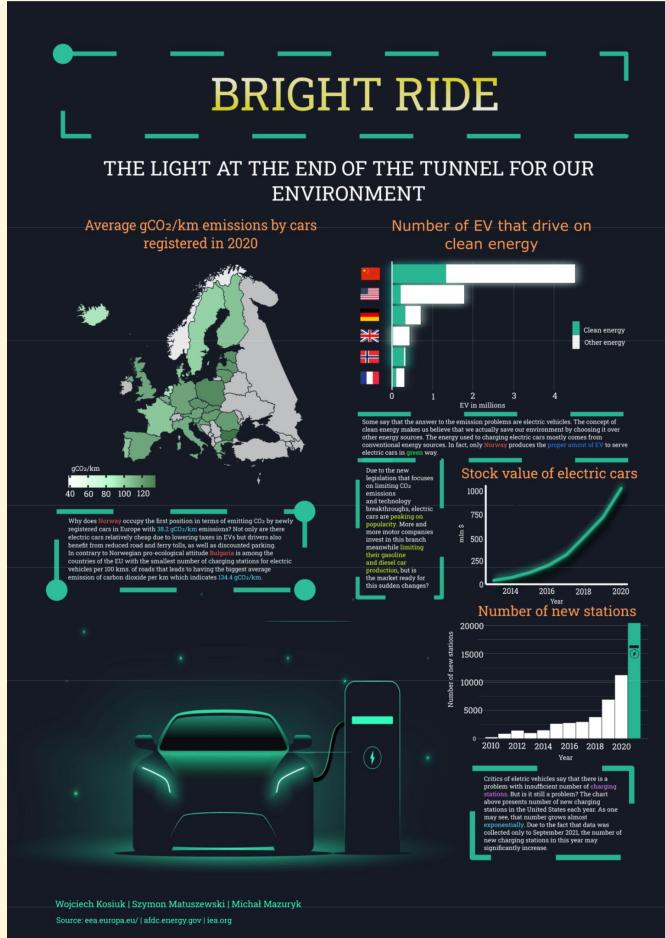
Wydział Matematyki i Nauk Informacyjnych

Autorzy:

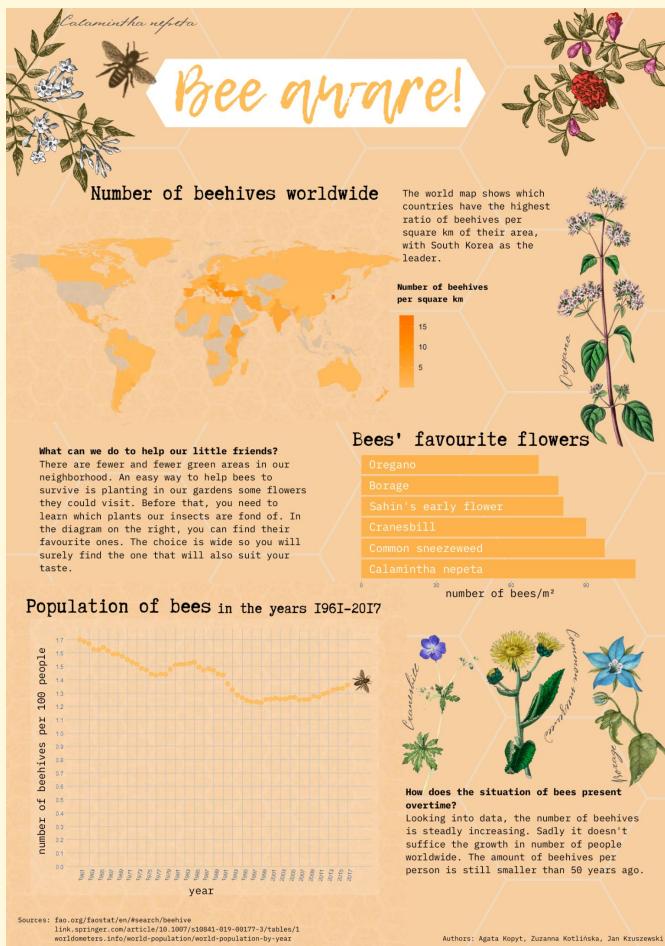
Martyna Kudłas
Barbara Jurczyk
Weronika Kyryk

Źródła:

<https://www.google.com/search?q=letharose+formula+1+world+championship+1950-2022>
<https://www.google.com/search?q=fernando+alonso+and+max+verstappen+4.737+kids+earns+1000000000>
<https://www.scmp.com/magazines/style/celebrity/article/3182964/10richest-f1-drivers-all-time-net-worths-ranked-lewis>



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

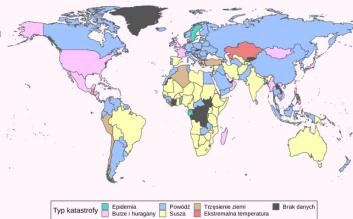


Katastrofy naturalne na mapach świata

Natura objawia swoją potęgę w wielu różnych zjawiskach. Niektóre z nich pokazują jej piękno, inne budzą grozę. My zainteresowaliśmy się katastrofami naturalnymi, które mogą być niezwykle niszczycielskie i śmiertelne. Postanowiliśmy zbadać, które kraje są najbardziej dotknięte przez katastrofy naturalne i które kataklizmy występują w konkretnych częściach świata.

Najbardziej dotkliwe katastrofy naturalne w poszczególnych krajach Na przestrzeni lat 1971-2008

Która z katastrof naturalnych dotyczy najwięcej obywateli danego państwa? Oto dane, ze wyniki są zgodne z naszymi przewidywaniami. W większości krajów europejskich najwięcej osób dotkniętych jest przez powódź, natomiast w Afryce dominują susze. W USA są to burze i huragany, a w Chile trzęsienia ziemi.



Występowanie katastrof naturalnych Na przestrzeni lat 2010-2018

Dzień występują konkretne katastrofy? Po naniesieniu pojedynczych wystąpień na mapę widać, że w dużej mierze pokrywa się z oczekiwaniemi które moglibyśmy mieć po przeanalizowaniu poprzedniej mapy. Niestety w niektórych częściach świata zbieranych jest znacznie mniej danych niż w innych. Dlatego na przykład w Europie jest znacznie więcej zaznaczonych wydarzeń niż w Brazylii czy Rosji.

Dotkliwość katastrof naturalnych Na przestrzeni lat 2010-2018

W którym kraju katastrofy naturalne są najbardziej dotkliwe dla cywilów? Aby to sprawdzić, dla każdego z nich policzyliśmy nasze własne wskaźniki. Jest to liczba osób dotkniętych w wyniku kataklizmów przez populację danego kraju, uśrednioną na przestrzeni lat. Następnie pogrupowaliśmy ten wskaźnik na 5 stopni dotkliwości i umieszciliśmy na mapie. Okazuje się, że w najdotkliwszych krajach, takich jak Filipiny, Niger czy Peru, wskaźnik osiąga wartości powyżej 1.

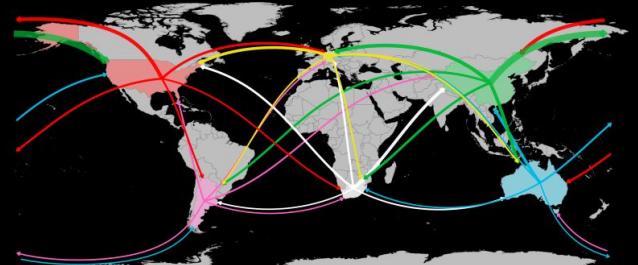


Jakub Piwko
Krzysztof Wodnicki
Łukasz Tomaszewski

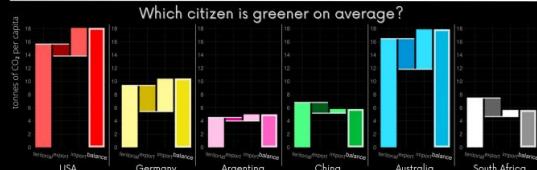
<https://medium.com/responsibleml/posters-that-change-the-perspective-on-climate-and-the-environment-c3682c0f6c39>

The hidden emissions

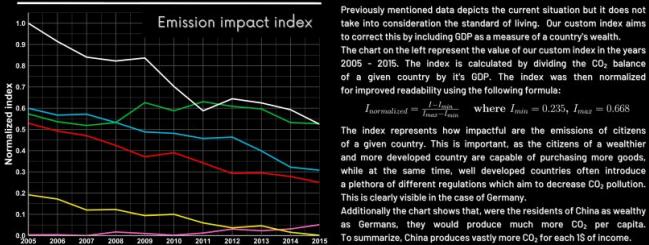
CO₂ in production and trade



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 CO₂ 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 CO₂ demand surpasses the Chinese by a whooping 10 tonnes disparity.



Faculty of Mathematics and Information Science

Patryk Rakus Kacper Trębacz Małwina Wojewoda

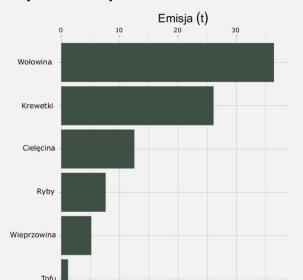
Data sources: idata.org stats.oecd.org

Produkcja żywności a ekologia

Emisja CO₂ per capita w roku 2013
w wyniku produkcji żywności

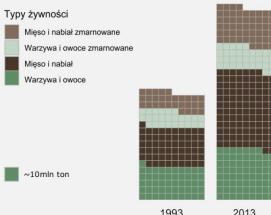


Emisja CO₂ w wyniku produkcji żywności w przeliczeniu na 1000 kcal



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

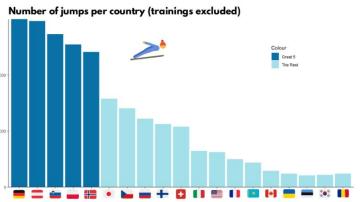


Spożycie wybranych produktów w 2013 roku
W kilogramach per capita



David Pludowski
Antoni Zajko
Grzegorz Kiersnowski
Źródła: FAOSTAT, Kaggle, OWID

See how fast they fly

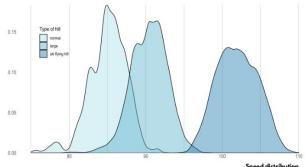


The Great 5

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.

Distribution of take-off speed by type of hill

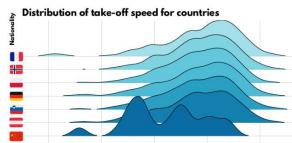


The bigger the hill, the faster you go!

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 gives 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!

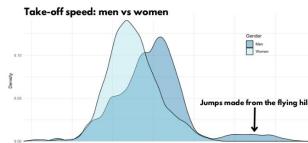
Speed by country

Due to different techniques and proficiency of their technical teams 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 for now, there are no women's competitions on flying hills.



Sources:

https://github.com/wrotki8778/Ski_jumping_data_center
<https://www.fis-ski.com/DB/general/biographies.html>

Authors:
Michał Gęraja
Rafał Pyzowski
Jakub Seliga

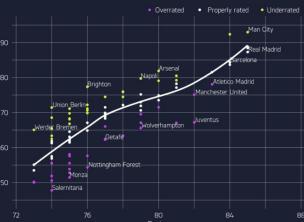


**Faculty of Mathematics
and Information Science**
WARSAW UNIVERSITY OF TECHNOLOGY

FIFA VS REALITY

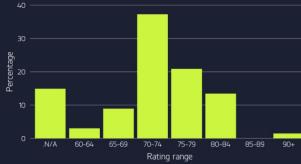
With the release of every single FIFA people argue about the ratings of players and call the creators biased. But is FIFA really that bad in terms of ratings and statistics?

FIFA TEAM RATINGS VS SPI

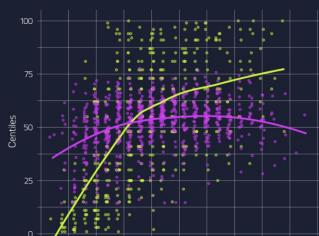


The first graph shows the dependence of the teams' overall rating in FIFA 23 on the soccer power index as of the start of the 22/23 season. SPI determines how good a team is based on match statistics from the last few months. The trend is that the higher the SPI, the higher the rating. However, at the 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



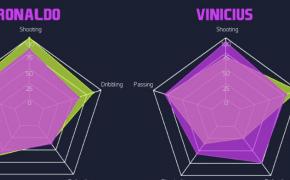
DEPENDENCE OF FORWARDS STATISTICS ON AGE



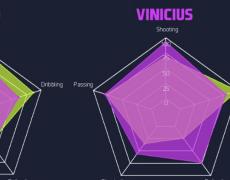
Players in FIFA are rated also on factors unrelated to their skills such as age of a player. The graph on the left compares the strikers' statistics in reality and in FIFA. According to the trend lines of these statistics, older players, despite a significant decrease in technical skills and physical 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.

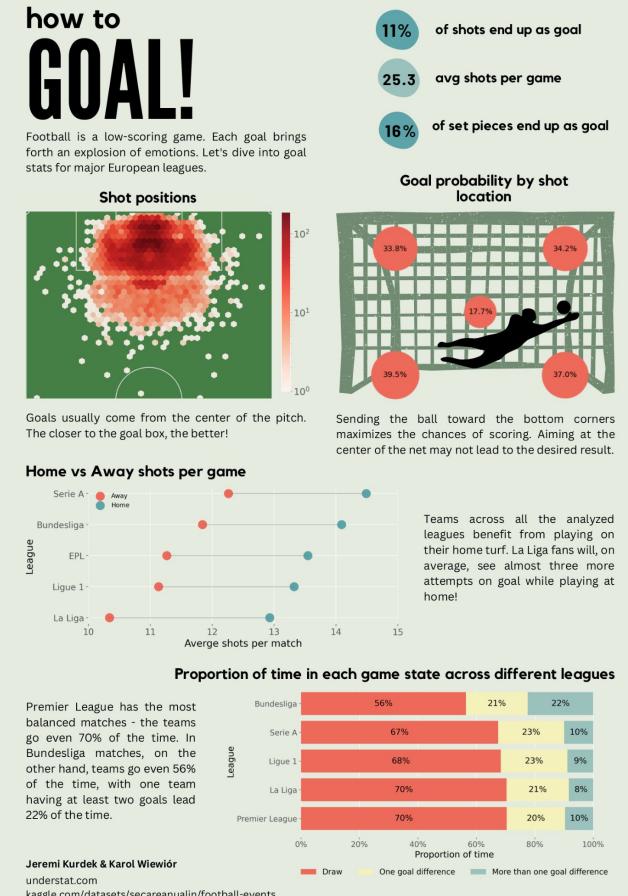
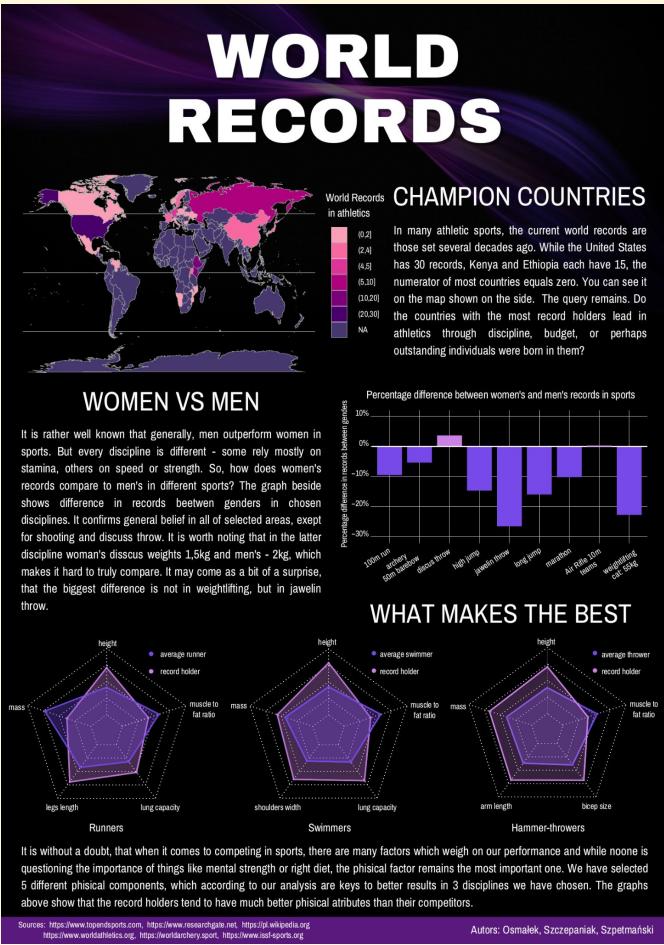
RONALDO

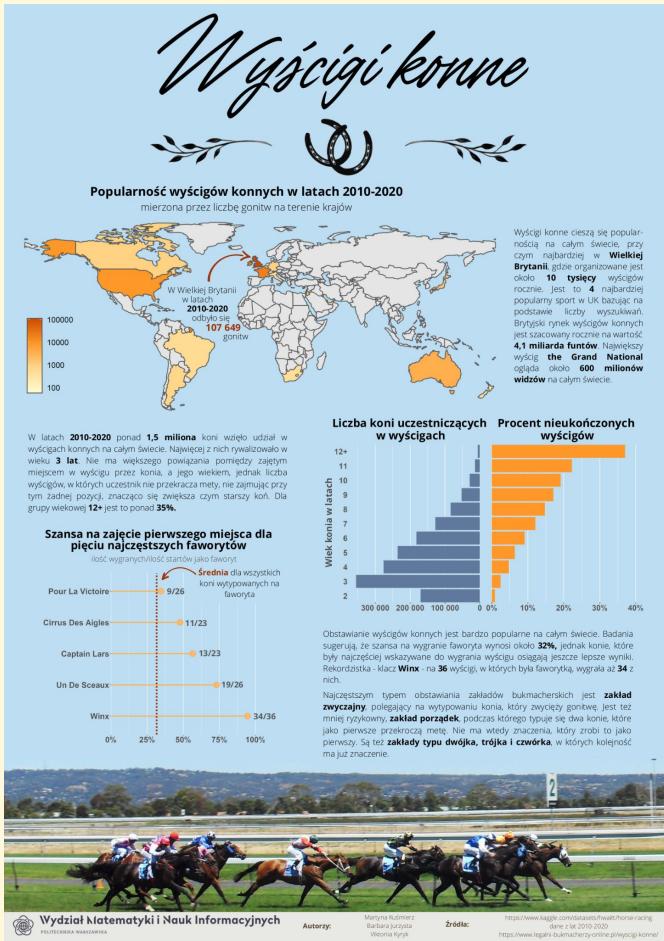


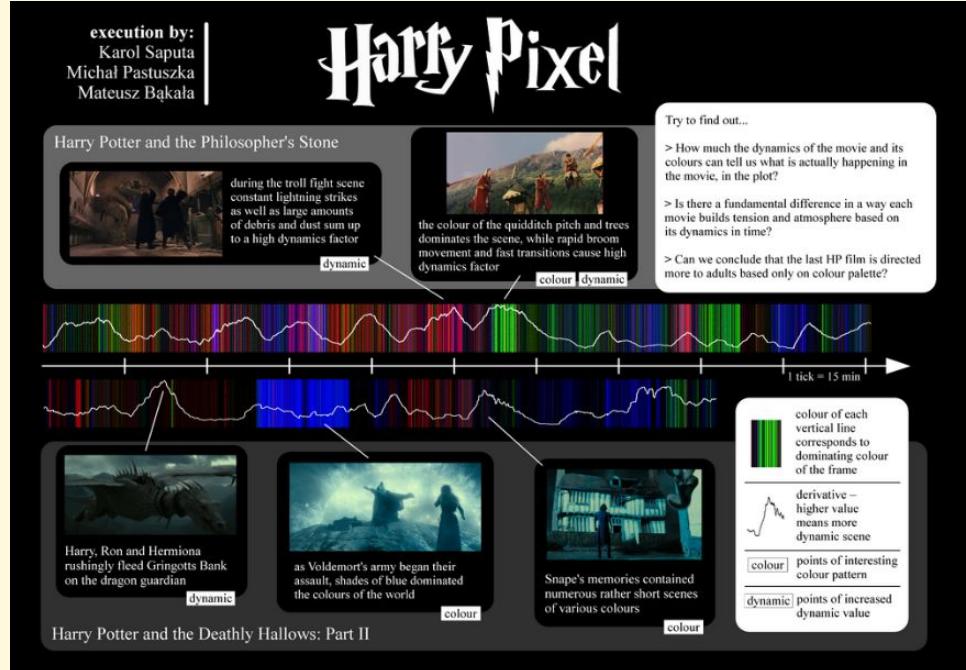
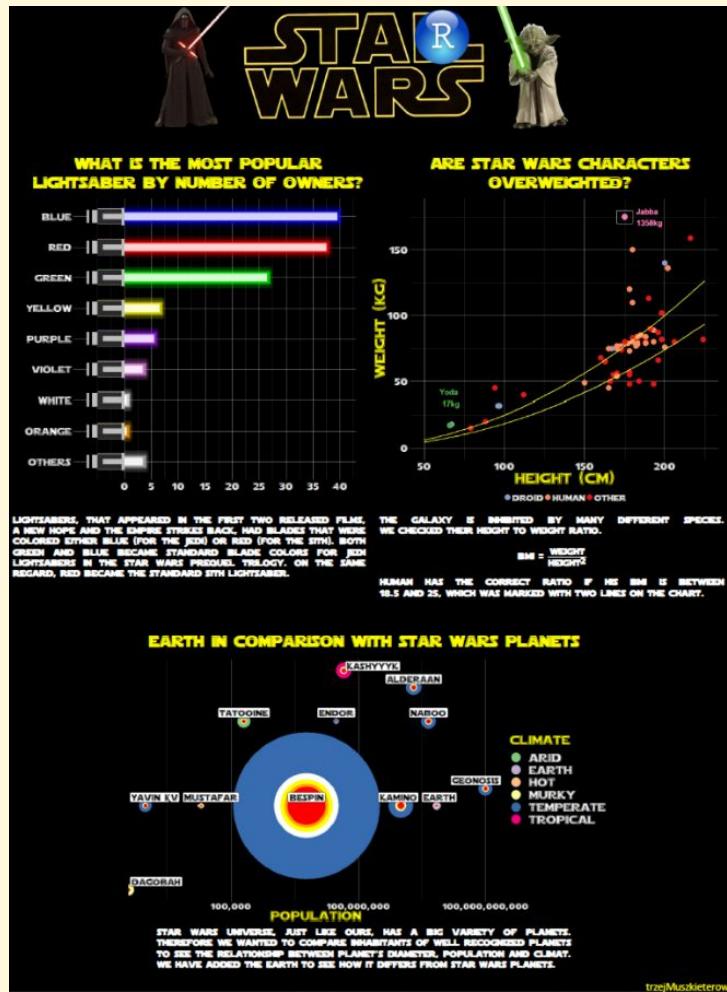
VINICIUS

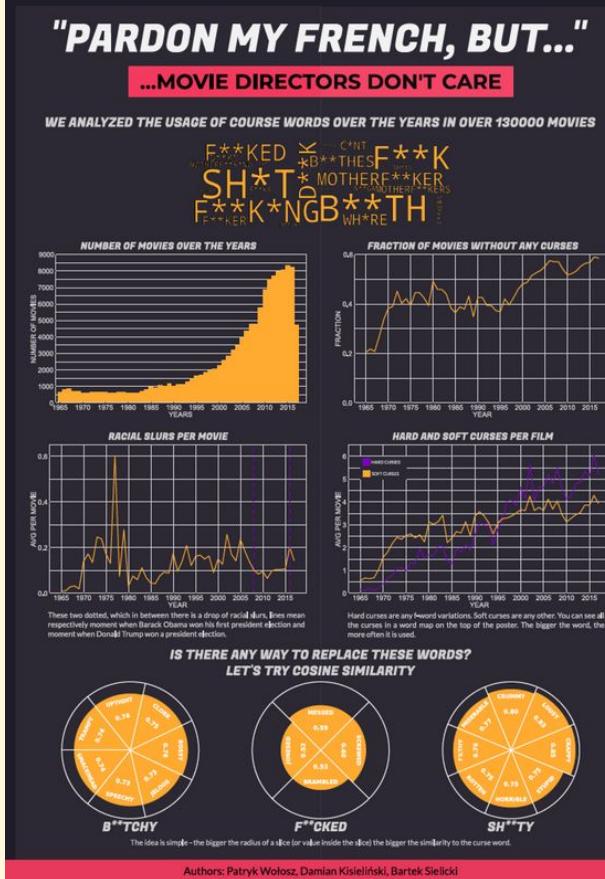
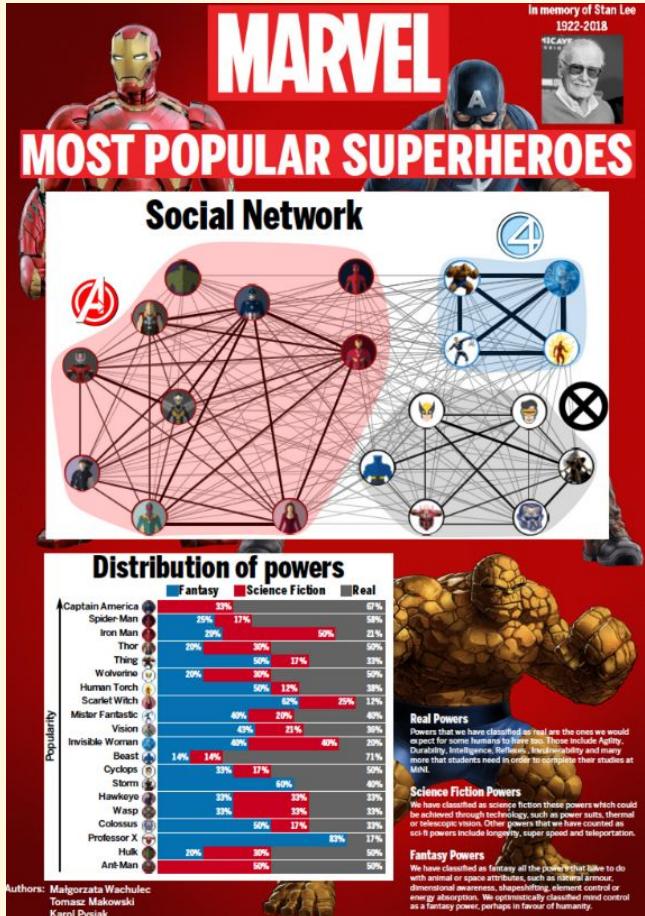


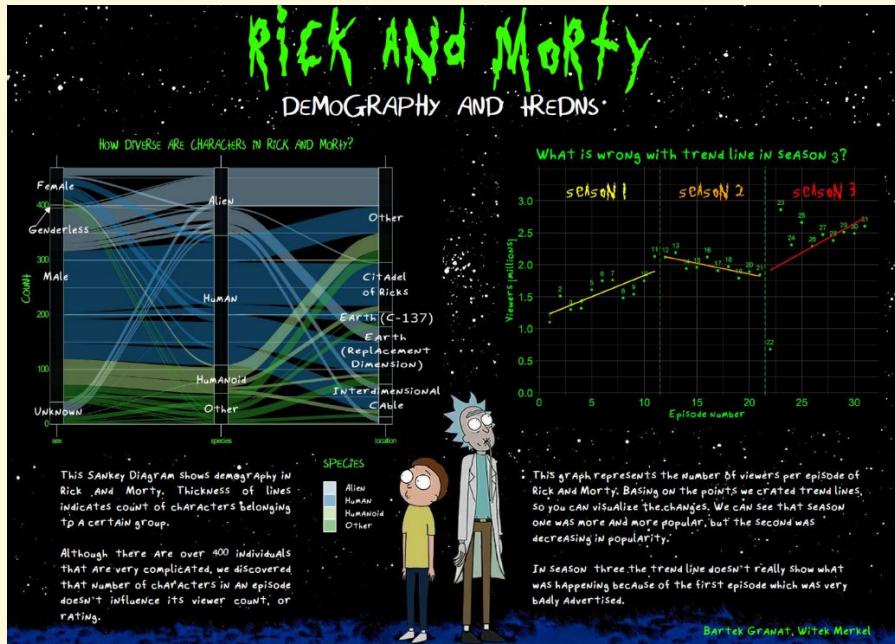
Sources: <https://www.transfermarkt.com/transfers/transfercenter/statistics/position/forward>
<https://www.fifa.com/fifaworldcup/en/2022/statistics/positions/forward>







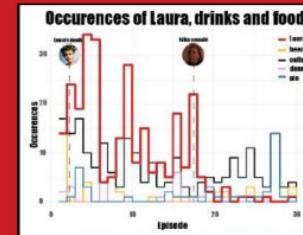




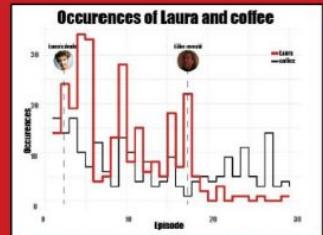
Did coffee help to solve the murder case?

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.

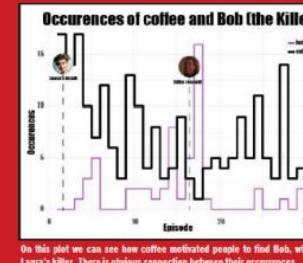
The most important thing on charts below are 2 marked moments: Laura's death and killer revelation. Apart from that, we can observe occurrences of specific words in subtitles per episode. There were 30 episodes in total in Twin Peaks series.



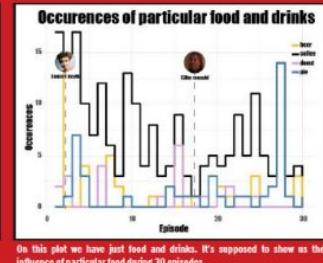
On this plot we can see all most popular food and drinks that were consumed during the story. We can observe the fact that coffee is the most consumed drink during the whole series. Which is pretty obvious - it's Twin Peaks favorite!



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.

