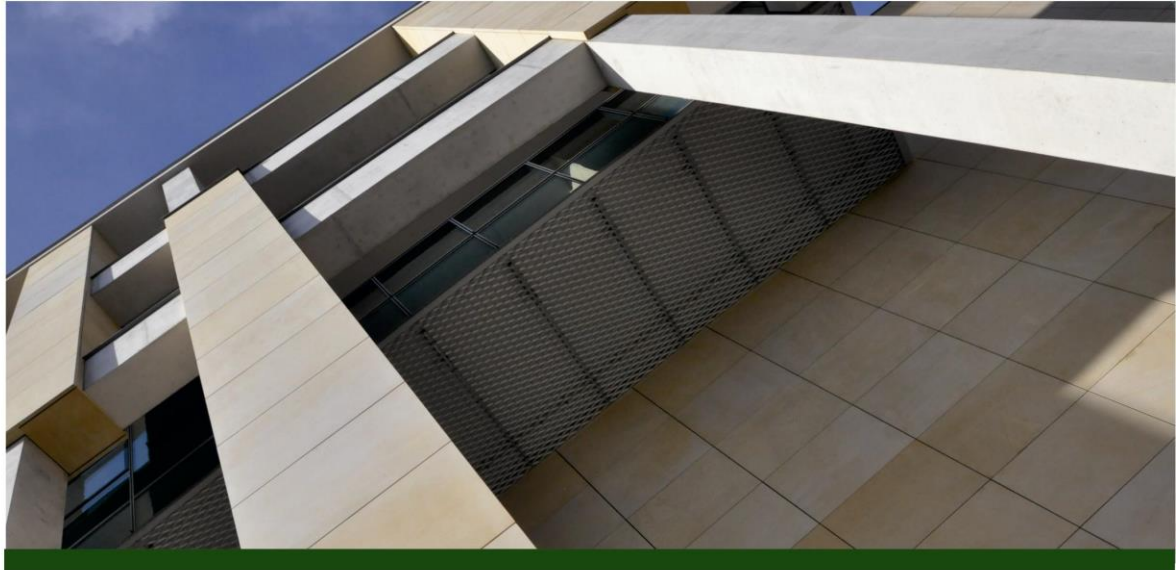




UNIWERSYTET  
EKONOMICZNY  
W POZNANIU



# Inform.narz.wizualizacji danych

II rok, studia II stopnia, rok  
akademicki 2021/2022

mgr [Piotr Kałużny](#)

# Informatyczne narzędzia wizualizacji danych

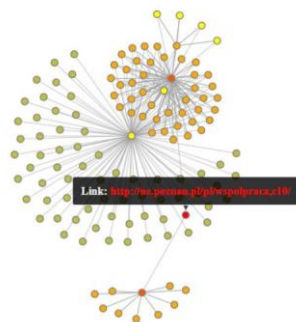
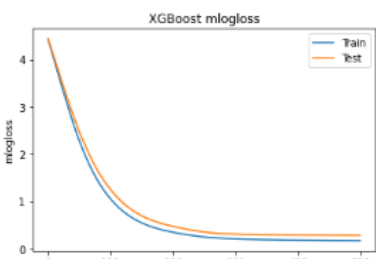
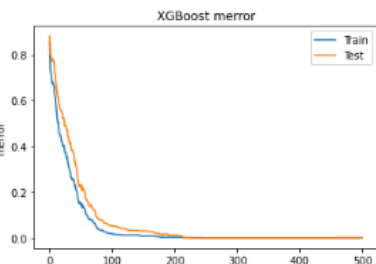
- Ćwiczenia: 15h = 15x45 min
- Tematyka: wizualizacja danych:
  - Wprowadzenie
    - Po co jest wizualizacja danych?
    - Wizualizacja jako produkt
  - Przygotowywanie danych do wizualizacji
    - Formaty danych: ilościowe, jakościowe, grafowe, szeregi czasowe, geograficzne
    - Filtrowanie, sortowanie, grupowanie, pola obliczeniowe
  - Wykresy i wizualizacje
    - Podstawowe typy wykresów, kolorystyka
    - Tworzenie wykresów, wizualizacja jako narzędzie analityczne
    - Tworzenie dashboardów i stories na podstawie danych
  - [Sylabus](#) (na aktualnym e-sylabusie jest błąd w literaturze)



# Wizualizacja - prowadzący

- Od kilku (6) lat w projektach badawczych i komercyjnych:
  - Wizualizacje procesów, map, wykresy, wyniki machine learningu, dashboardy wizualizacyjne,
  - UI design i prototypowanie rozwiązań (w tym mobilnych),
  - Grafy społecznościowe, analiza sensorów, analiza tekstu, procesów biznesowych i clickstreamów.
- Moje języki/technologie:
  - Python: pandas, numpy, scikit, pyspark, **bokeh, seaborn, ggplot, folium, plotly**, flask i inne...
  - JS: **d3.js, chart.js, mapbox, leaflet**, rickshaw
  - C#: Unity
  - R: język znam, nie czuję się specjalistą

# Może lepiej warto zobaczyć ;)



Poziom: 0 url: <http://ue.poznan.pl/wspolpraca,c10/>

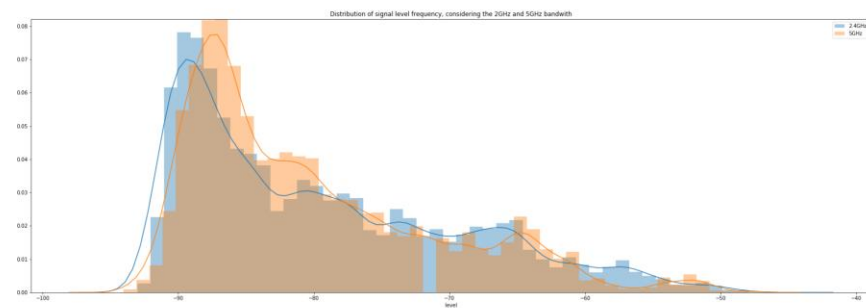
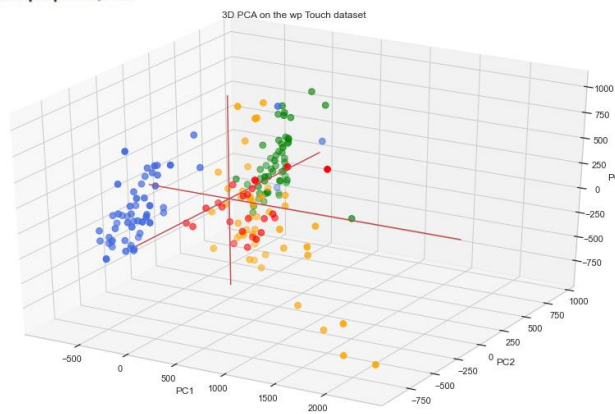
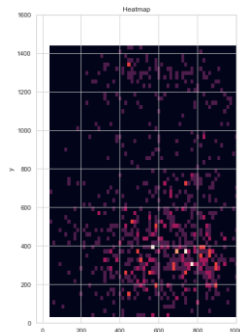
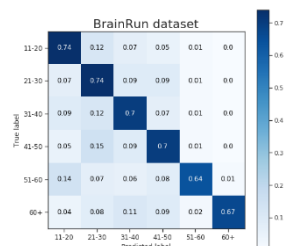


Figure 1: Distribution of signal level of observed APs with 2,4 and 5 GHz working frequency.  
Source: own data

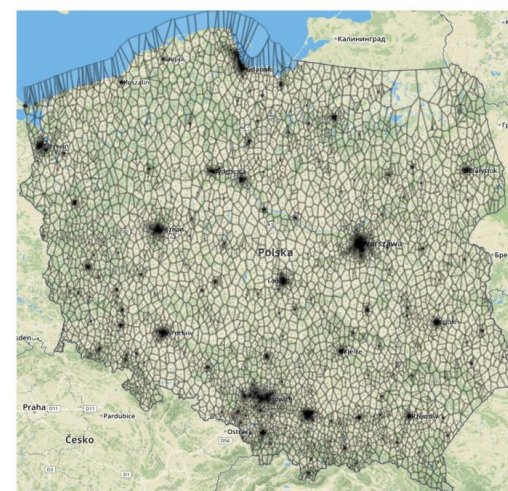
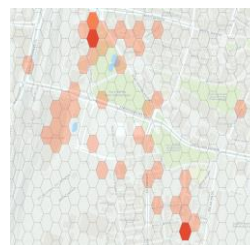
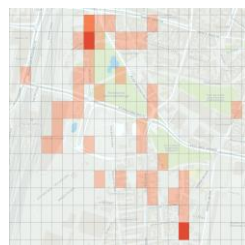
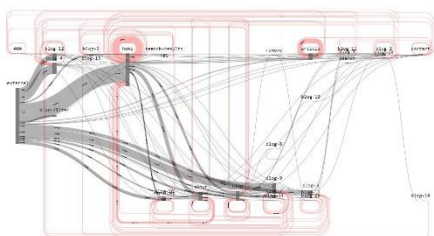


Figure 1. Visualization of Voronoi cell tessellation prepared for our study.

# Narzędzia

- **Python** jako narzędzie ponieważ:
  - narzędzia komercyjne - Qlik, Tableau, Power BI mają ograniczone zastosowania.
  - Python to wraz z R najłatwiejsze języki do wizualizacji
  - łatwo można dołożyć JS i stworzyć dynamiczną wizualizację albo dodać Flask i stworzyć usługę.
  - kierunek obejmuje przedmioty w głównym toku: ASI, Uczenie Maszynowe, Big Data – gdzie Python to narzędzie do przekazania treści, jest więc spójnie.
- Inne ciekawe podejście to wizualizacje robione przez JS i biblioteki wspierające (d3, chart.js, DataTables).

# Treści

1. Wprowadzenie i prezentacja „spodziewanych efektów”.
  2. Percepcja, przetwarzanie i reprezentacja wizualna danych
  3. Środowisko – Jupyter notebook / Google Colaboratory i Python – matplotlib I
  4. Pandas i matplotlib II
  5. Seaborn i ggplot
  6. Typy wykresów
  7. Kolorystyka i percepcja, palety kolorów
  8. Grupowanie i przetwarzanie danych, obliczenia
  9. Dane geograficzne – formaty danych i projekcje, basemap, folium
  10. Interaktywne elementy wykresów
  11. \*Grafy / \* Interaktywne elementy wykresów II - Dashboardy
  12. \*Stories z danych – Use Case do zaliczenia
  13. Prezentacja projektów I
  14. Prezentacja projektów II
  15. Prezentacja projektów III
- \*Może ulec zmianie ze względu na zainteresowania i problemy z projektami





# Zaliczenie

- Bez względu na sytuację epidemiologiczną:
  - Zaliczenie projektem na koniec zajęć
- Projekt opisujący wizualizację wybranego zbioru danych na ocenę
- Dodatkowo są możliwe pkt za aktywność, może będą jakieś dodatkowe zadania ale nie obiecuje – wolałbym, żebyście się skupili na projektach
- Projekt może korzystać z danych z pracy magisterskiej – jeżeli takowe wykorzystujecie

# Zaliczenie - kryteria

Kryterium	Zakres pkt	Opis
Oryginalność problemu	0-5 pkt	<b>Ciekawy problem badawczy</b> , zbiór nie przedstawiany w tutorialach i możliwy do analizy.
Prezentacja	0-5 pkt	Zainteresowanie tematem, ciąg logiczny, składnia, długość!
Story / Prezentacja i opis zbioru i badanych problemów	0-5 pkt	Opis samego zbioru danych, etapów postępowania przy analizie, cel poszczególnych wizualizacji, opisanie przebiegu i logiczny ciąg całości.
Wykorzystanie różnych wizualizacji	0-10 pkt	Standardowe wizualizacje - 5pkt (wykres kolumnowy, gęstości). Żeby mieć 10 punktów przynajmniej 4 sensowne <b>typy</b> wykresów (+2 na każdą osobę w grupie).
Filtrowanie, grupowanie	0-10 pkt	Przekształcenia danych, nie korzystanie tylko z danych źródłowych, grupowanie, usuwanie outlierów, facetGrid. Ciekawe przekształcenia (nie korzystanie z danych surowych) i transformacje formatu danych.
Dobór wizualizacji do problemu	0-15 pkt	Odpowiedni dobór wykresu do przedstawianej zależności, czytelność, kolorystyka. Za każdy błąd od -2 do -4pkt.
Dodatkowe punkty	0-10 pkt	Inna ciekawa technologia, nieomawiana biblioteka, usługa zamiast wizualizacji, nietypowy typ lub źródło danych (grafowe z przekształceniami albo społecznościowe, procesy, sensory, scrapping)
RAZEM	50 pkt + max 10 dodatkowych	



# Ogólnie

- Samodzielnie albo do 3 os – będę wymagał szerszej wizualizacji od 3 os. niż od 1, każda musi pokazać swój wkład.
- Projekt na 3:
  - Prosty zbiór danych (np. sprzedaż), podstawowe transformacje: rozkład zmiennych, wykresy kolumnowe, kolory bez palet.
- Projekt na 4:
  - Zbiór danych nie prosto z pierwszego tutoriala ML, ciekawe wykresy (boxplot, geografia, grafy), brak błędów i dobre dobranie kolorów, grupowanie danych.
- Projekt na 5:
  - J.w. + nietypowy zbiór danych (geografia, grafy, procesy), **albo** nietypowa technologia **albo** dobra i rzetelna analiza – wyciągnięcie wniosków z wizualizacji, przekształcenia i grupowania (analiza średnich w grupach, outlierzy, korelacja zmiennych itd.).
- Wymagania pomiędzy 1 os. a 3 os. zespołem nie różnią się znacząco, głównie oczekiwaną liczbą typów wykresów i wyciągniętych wniosków. Kilkuosobowe zespoły muszą wskazać jednoznacznie swój wkład w projekt. Cięższy do zebrania/analizy zbiór to dobre wyzwanie dla 3 os.



# Szczegóły oceniania

- Forma oddania każdego projektu ma być jednym dokumentem w **PDF/HTML** (w przypadku Collaba/Jupytera to proste), w przypadku innych narzędzi musicie doprowadzić projekt także do jednego dokumentu – raportu z opisem całości. Obowiązuje was także prezentacja wyników, która może się odbywać po prostu jako odczytywanie jupitera.
- Dodatkowe punkty nie są obowiązkowe... mogą nie dać żadnych ;). Możecie się natomiast dopytać wcześniej czy coś się kwalifikuje na nie.
- **W przypadku wątpliwości co do samodzielności/oceny mogą pytać o to co robi konkretny kod, czemu dany wykres został dobrany do pokazania prawidłowości bądź jaki był wkład poszczególnych osób w dane wykresy – możecie każdy wykres oznaczyć nazwiskiem w () aby uniknąć sporów.**

# Prezentacja

- Prezentacja między 5 a 10min, dłuższe prezentacje będą skutkowały odejmowaniem pkt – wyobraźcie sobie, że macie niecierpliwego szefa przed sobą ;)
- Co ma być w prezentacji:
  - Założenia - co to za zbiór danych: źródło, co bada
  - Cele - co chcieliście zbadać, przedstawić za pomocą wizualizacji
  - Opis zastosowanych technik (wykresy, kolory, przekształcenia), najciekawsze wykresy.
  - Opis wniosków



# Teams

- Zajęcia w tym semestrze są na na żywo. Musicie być obecni na ćwiczeniach/labolatoriach – zgodnie z regulaminem studiów.
- Gdyby coś się zmieniło – grupy na Teams są gotowe.

# Obecność na zajęciach

- Obecność na ćwiczeniach jest obowiązkowa
- Każdy student jest **zobowiązany do potwierdzenia swojej obecności na moodle w module frekwencja**
- Formalne usprawiedliwienie należy dostarczyć w terminie 4 dni roboczych od daty zaliczenia (osobiście lub mailem)
- Usprawiedliwiona nieobecność nie zwalnia studenta ze znajomości materiału realizowanego na zajęciach w tym czasie

# Oceny

Zakres PKT	Zakres %	Ocena
<30	(0-60)	2
30-34	<60-70)	3
35-39	<70-80)	3,5
40-44	<80-90)	4
45-47,5)	<90-95)	4,5
<47,5-50+	<95-100)	5

# Poprawka

- Przysługują dwie poprawki w trybie regulaminowym
- Poprawka gdy suma punktów zdobytych podczas całego semestru  $< 60\%$
- Poprawka w formie projektu zaliczanego live na Teams (będę prosił o przygotowanie analizy i **będę się pytał o kod i wyjaśnienie go**)
- Termin poprawki zostanie wskazany później



# Reguły akademickiej uczciwości

1. Kopiowanie lub przeformułowywanie słów, pomysłów lub pracy innych osób i przedstawianie ich jako własnych oraz nie zapewnianie stosownych odwołań.
2. Wykorzystywanie lub udostępnianie w trakcie sprawdzianu jakichkolwiek niezatwierdzonych materiałów z nim związanych, wewnątrz lub na zewnątrz pomieszczenia, w którym się odbywa.
3. Komunikowanie się w trakcie sprawdzianu w dowolny sposób z jakąkolwiek osobą, która nie jest do tego upoważniona.
4. Udzielanie lub pobieranie pomocy od innych w trakcie sprawdzianu.
5. Zabieranie, przekształcanie, ukrywanie, uszkodzanie, niszczenie jakiegokolwiek rzeczy związanej z przygotowaniem projektów, badań lub sprawdzianów.

# Materiały do zajęć

- <http://moodle.ue.poznan.pl>
- Kurs:
  - Informatyczne narzędzia wizualizacji danych -  
stacj. II st. - (ćw.) - mgr Piotr Kałużny
- <https://moodle.ue.poznan.pl/course/view.php?id=14201>
- Hasło: inwd2021gc0X  
*gdzie X to nr grupy*

# Kontakt

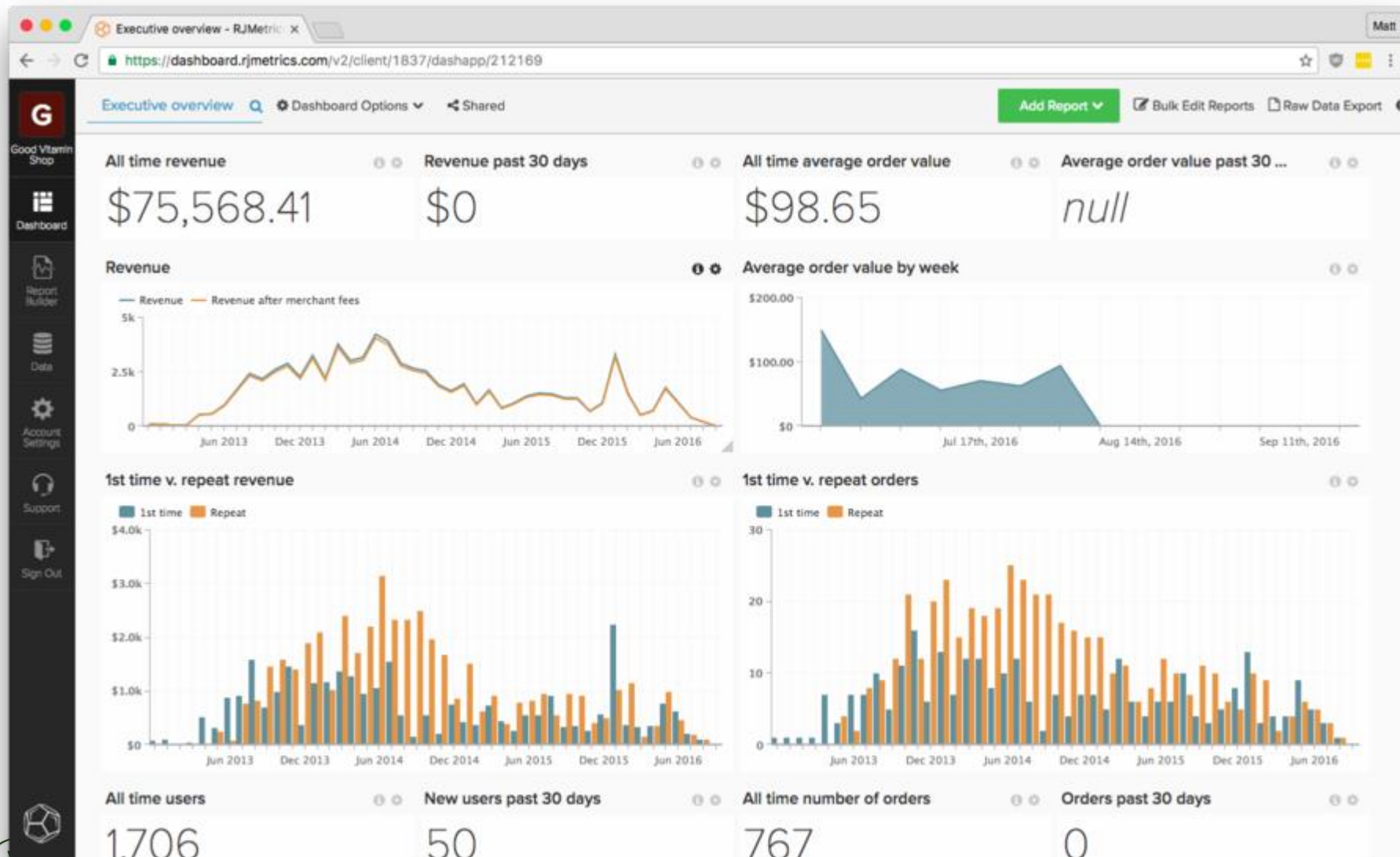
- [piotr.kaluzny@ue.poznan.pl](mailto:piotr.kaluzny@ue.poznan.pl)
- **W ważnych sprawach ten i tylko ten mail, nie Teams, nie chat moodle, nie USOS – odpowiadam tylko na tego maila.**
- W tytule [INWD]
- Dyżury – Oficjalnie 411 C
- **UWAGA! Na UEP pracujemy w 95% zdalnie, więc kontakt via mail:**
  - Wtorek 13:00-14:00
  - Środa 16:15-17:15
  - W razie potrzeby dyżuru proszę o maila wcześniej – trzeba to wszystko skonfigurować

# Pytania? Wątpliwości?

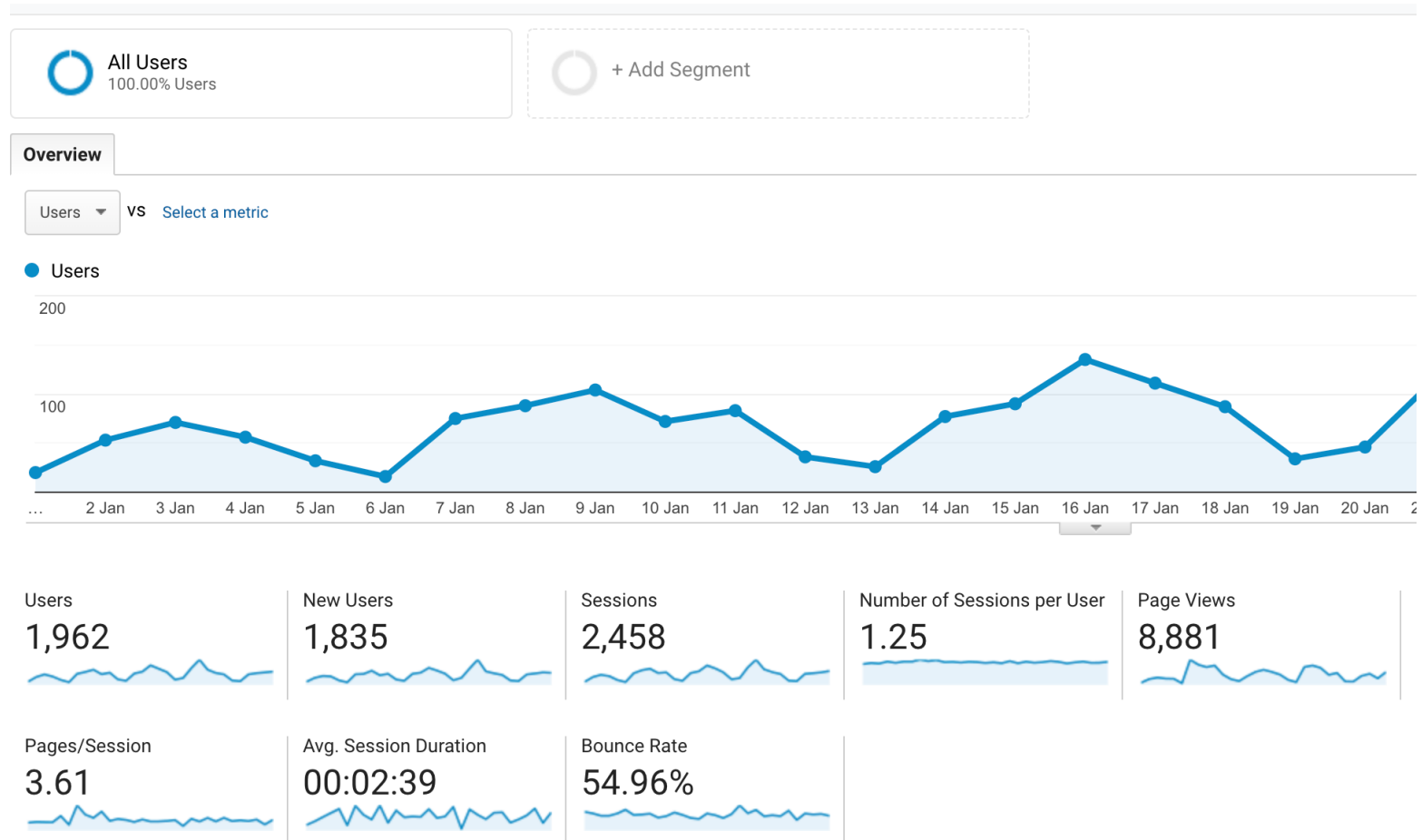


UNIwersYTET  
EKONOMICZNY  
W POZNANIU

# Wizualizacja jako produkt

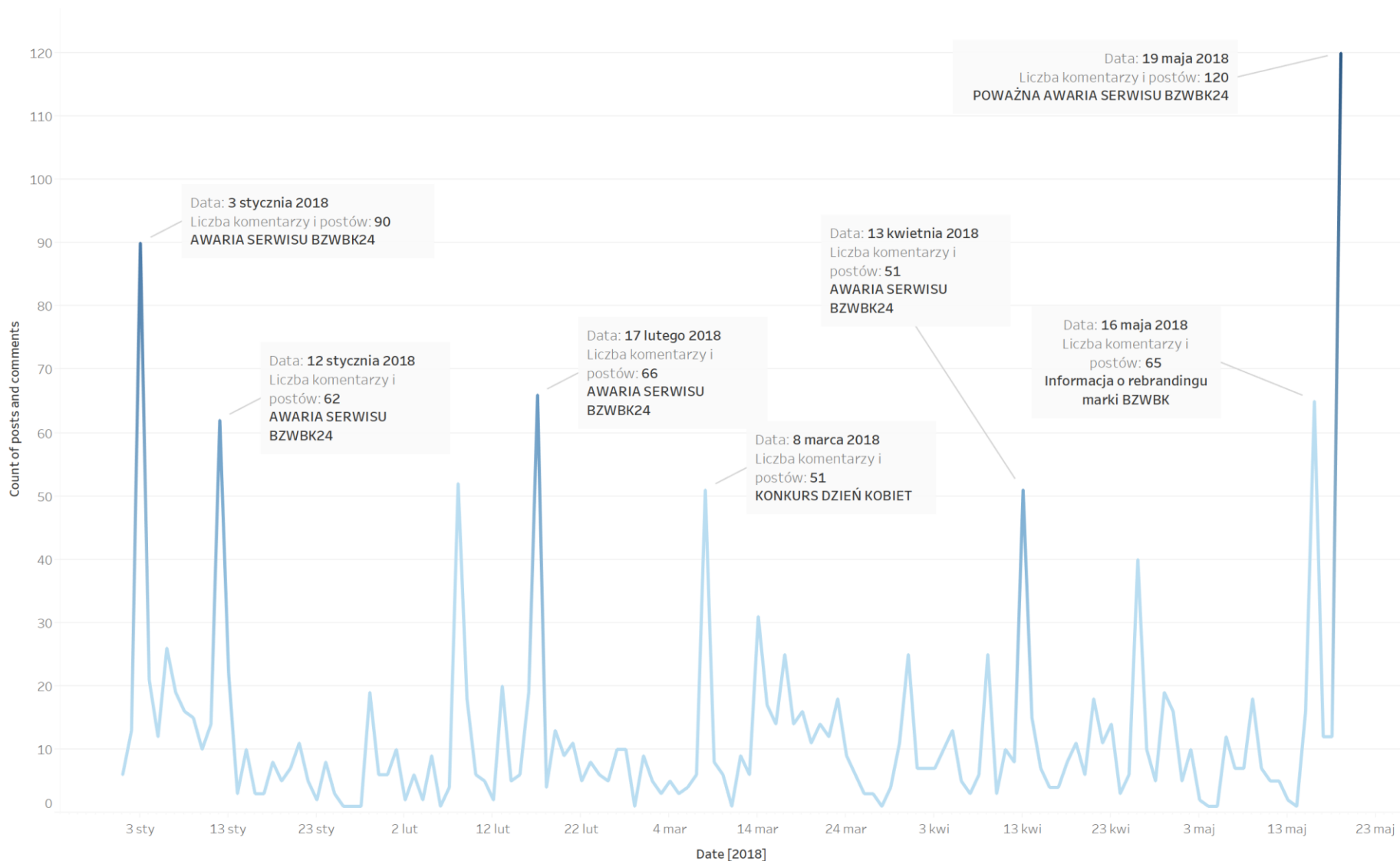


# Wizualizacja jako produkt



# Wizualizacja jako produkt – case BZWBK

Aktywność klientów na portalu Facebook





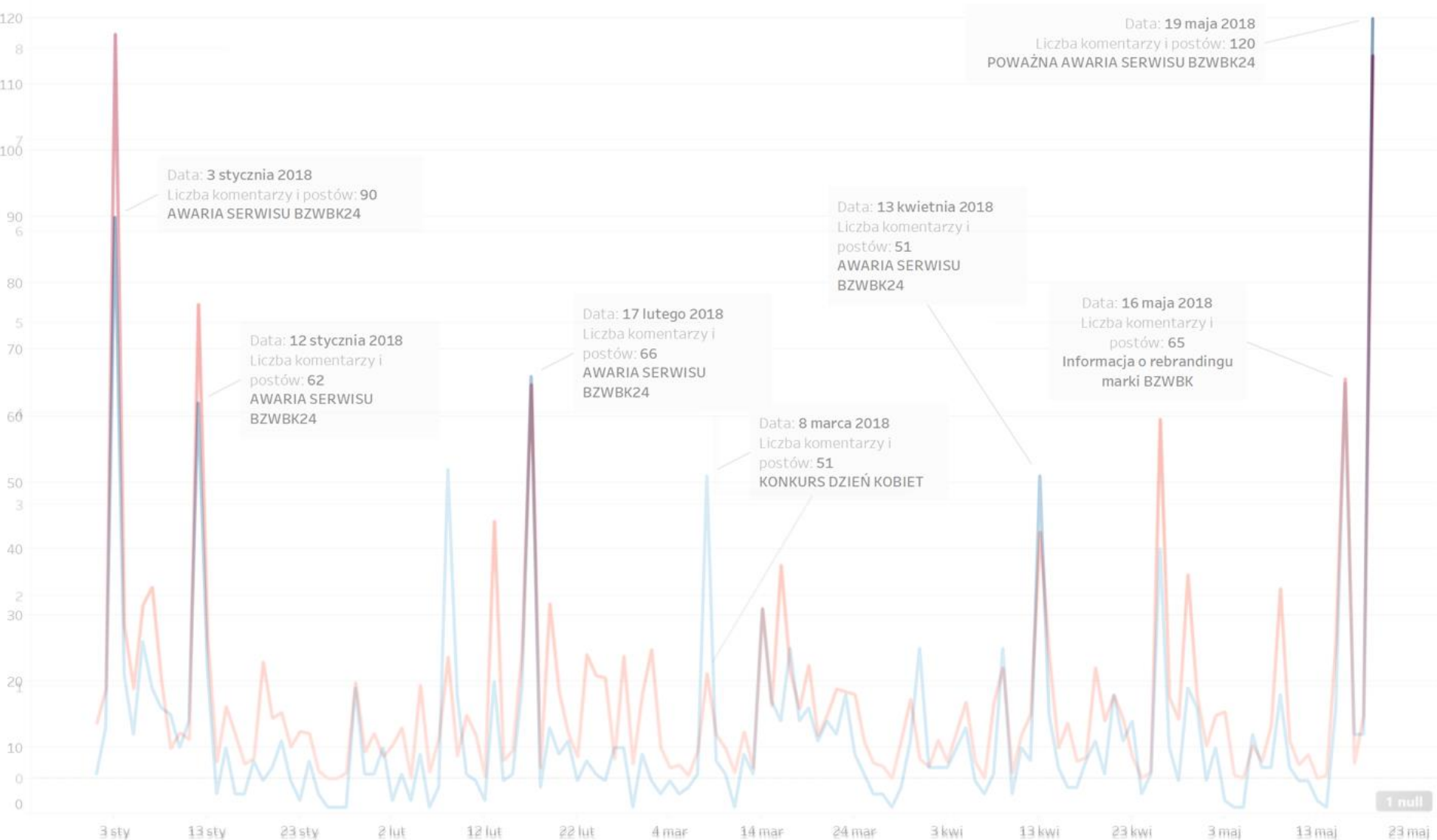
# Wizualizacja jako produkt – case BZWBK

Analiza sentymentu komentarzy - negatywne nasycenie

## Wydźwięk negatywny

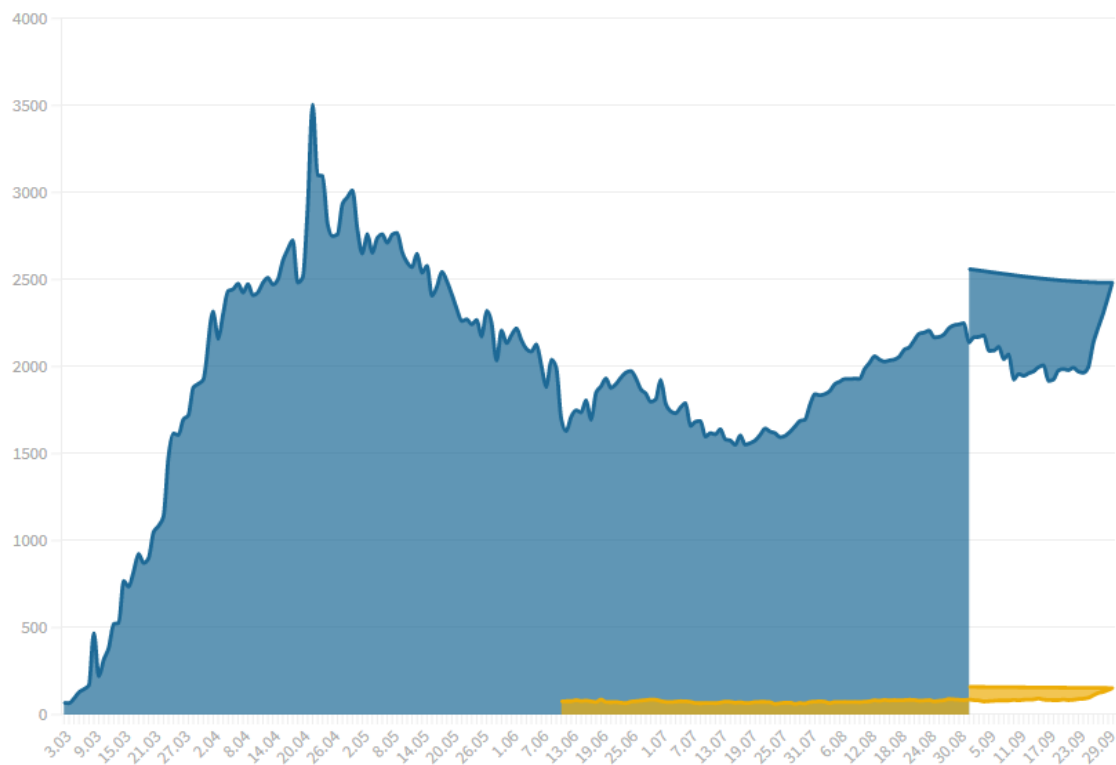


# A może oba?



## Liczba osób hospitalizowanych na COVID-19 i zajętych respiratorów

■ Osoby hospitalizowane ■ Zajęte respiratory



Źródło: Ministerstwo Zdrowia

✿ A Flourish chart

**Autor:** ads\mtom

**Źródło:** tvn24.pl

**Źródło zdjęcia głównego:** Leszek Szymański/PAP

**PODZIEL SIĘ**



UNIWERSYTET  
EKONOMICZNY  
W POZNANIU

# A wy czego oczekujecie od tego przedmiotu?



UNIwersYTET  
EKONOMICZNY  
W POZNANIU