

# Prognozowanie i symulacje

## Wykład 1:

### Informacje organizacyjne Podstawowe pojęcia związane z prognozowaniem

**WYKŁAD:** Marek Sobolewski

**LABORATORIA:** Roman Szostek, Marek Sobolewski

**STRONA INTERNETOWA:** [mareksobolewski.v.prz.edu.pl](http://mareksobolewski.v.prz.edu.pl)

# Po co nam analiza danych, umiejętność prognozowania?

1. Żyjemy w czasach, kiedy informacja stała się towarem...

Dobrze jest umieć analizować dane, po to żeby lepiej wykorzystywać MOŻLIWOŚCI jakie nam stwarza nowoczesność XXI wieku.

2. Żyjemy w czasach, kiedy informacja stała się narzędziem...

Dobrze jest wiedzieć, jak inni mogą wykorzystywać to narzędzie i umieć ochronić się przed ZAGROŻENIAMI, jakie stwarza nam nowoczesność XXI wieku.

# Najsłynniejsze powiedzenie o statystyce

*Są trzy rodzaje kłamstwa:  
normalne, diabelskie i statystyka*

[Autor niejednoznaczny: M. Twain, B. Disraeli, G.B. Shaw]

**I dopowiedzenie...**

*Ale to nie statystyka kłamie,  
tylko ludzie, którzy się nią posługują!*

# Organizacja zajęć

Wykłady: 30 godz. (15\*)

Laboratoria: 30 godz. (15\*)

**\* na studiach niestacjonarnych**

## Warunki zaliczenia przedmiotu:

- 1. Obecność na zajęciach (dotyczy laboratoriów i wykładów!).**
- 2. Uzyskanie pozytywnej oceny z laboratoriów (kolokwium, wejściówki i inne formy sprawdzania wiadomości).**
- 3. Uzyskanie pozytywnej oceny z EGZAMINU.**
- 4. Końcowa ocena będzie ważoną średnią ocen z zaliczenia laboratoriów i egzaminu.**

# **EGZAMIN – w formie praktycznej!!!**

**Czym się będzie różnił egzamin od kolokwiów na zajęciach laboratoryjnych:**

- będzie obejmował całość materiału;
- będzie wymagał sprawnego posługiwania się narzędziami analitycznymi w ograniczonym czasie;
- będzie na nim wymagana znajomość metod poznawanych na laboratoriach, ale w rozszerzonym zakresie – niemal na każdym wykładzie będą wskazywane pewne dodatkowe umiejętności, które będą także weryfikowane na egzaminie.

## Cel zajęć

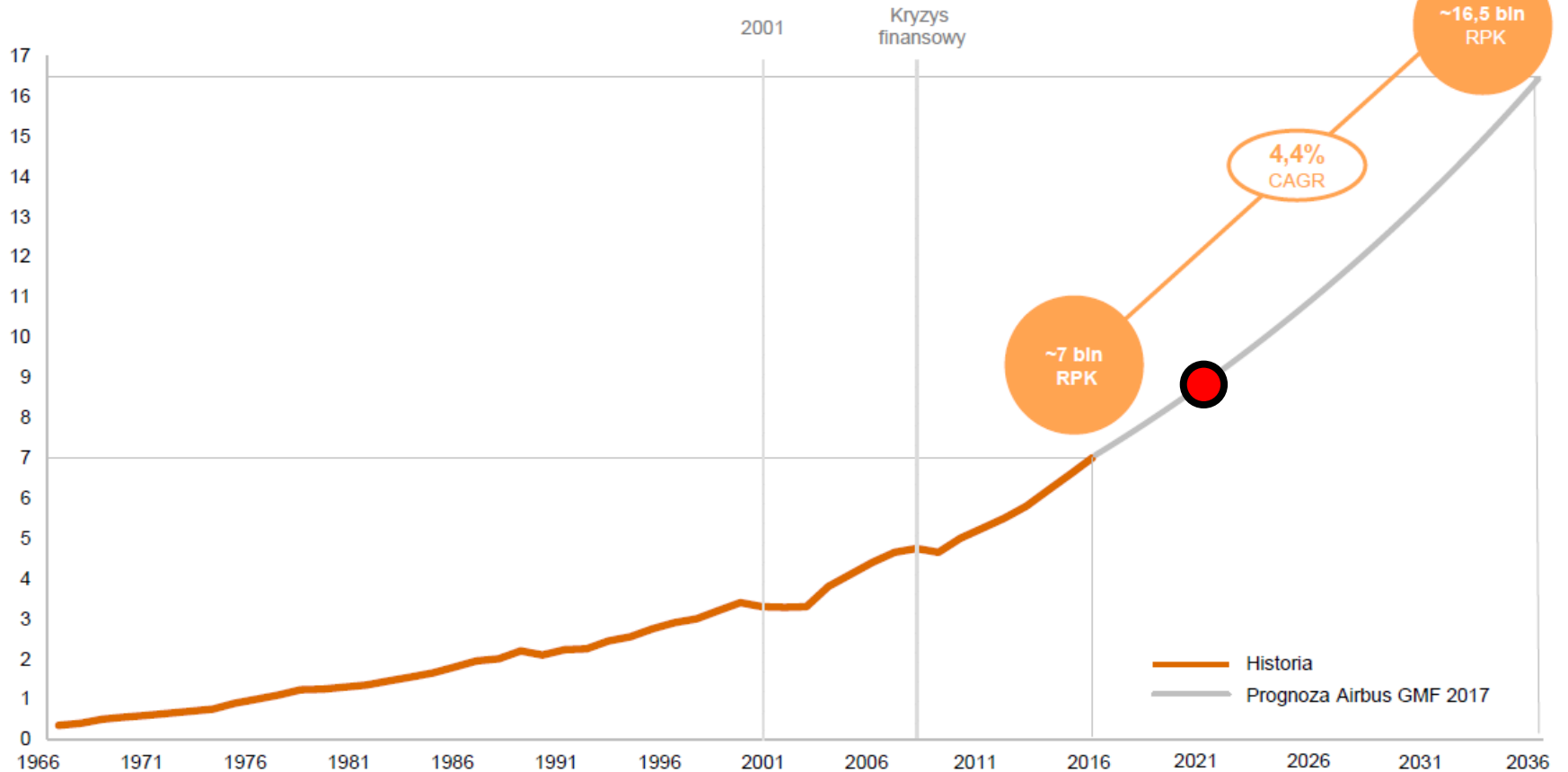
Celem głównym, co jest dość oczywiste, jest nauka podstawowych metod prognozowania.

Nie każdy jednak będzie w swojej pracy sam dokonywał prognoz, dlatego równie ważne są pewne cele poboczne:

- wiedzieć, jakie są możliwości, ale i ograniczenia prognozowania za pomocą metod statystycznych – umieć zlecić i wyegzekwować wykonanie prognoz od zespołu współpracowników;
- wiedzieć, jak ostrożnie należy podchodzić do prognoz opracowanych przez różne instytucje (zaraz będzie przykład).

# Przykłady prognoz – udanych i nieudanych

Ruch lotniczy na świecie (bln RPK\*)

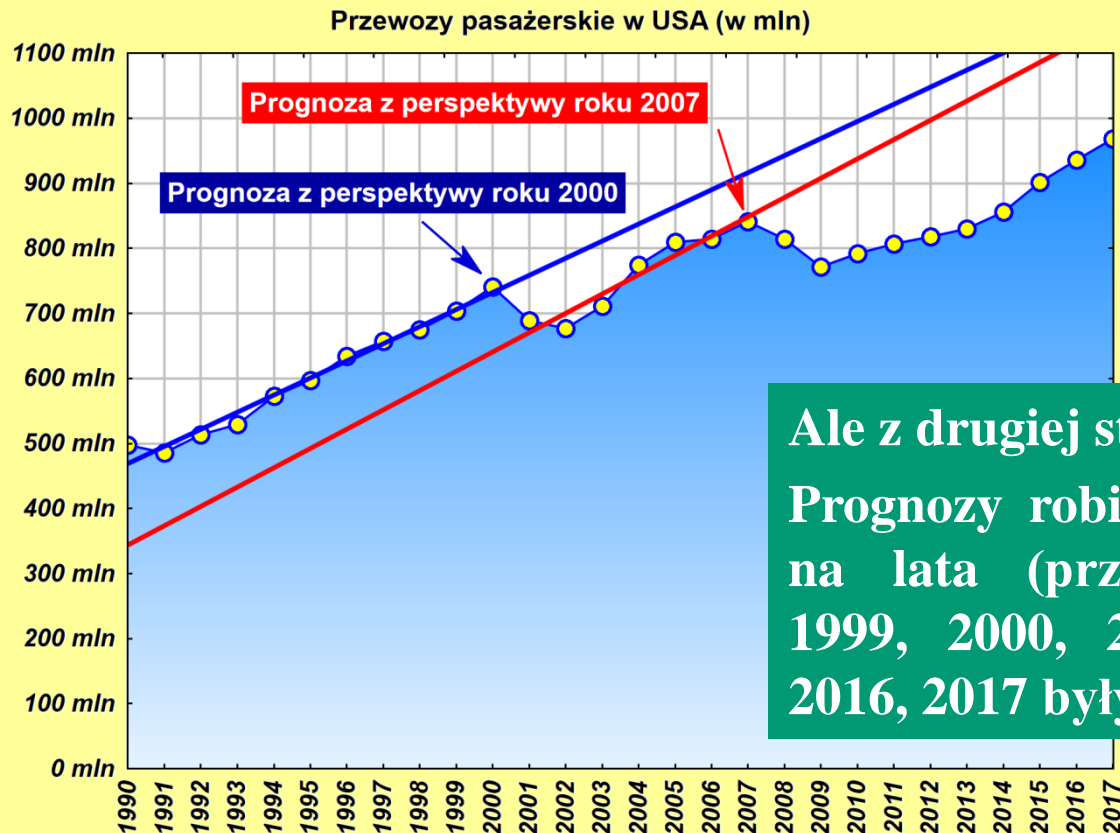


\* RPK (Revenue Passenger Kilometer) – liczba pasażero-kilometrów

CAGR - (Compound Annual Growth Rate) - skumulowany roczny wskaźnik wzrostu

# Przykłady prognoz – udanych i nieudanych

Poniżej przedstawiono prognozy przewozów pasażerskich w USA z perspektywy 2000 oraz 2007 roku. Jak widać, prognozy te były dalekie od rzeczywistego przebiegu badanego zjawiska, co każe ostrożnie podejść do „łatwego” przewidywania przyszłości na podstawie ekstrapolacji trendów.



Ale z drugiej strony...

Prognozy robione w ten sposób na lata (przykładowo): 1998, 1999, 2000, 2006, 2007, 2015, 2016, 2017 byłyby bardzo trafne.

Źródło: Opracowanie własne na podstawie danych  
United States Department of Transportation



# Uwagi dodane w roku 2021 (wiosną!)

Nie sposób nie wspomnieć o ogromnym znaczeniu prognozowania w okresie kryzysu „epidemicznego”, który rozpoczął się niemal dokładnie rok temu.

Niestety, sposób analizy danych o liczbie zakażeń wirusem SARS-Cov-2 budzi wiele zastrzeżeń. Oto niektóre z nich:

- 1) Prezentowanie danych skumulowanych** – i to od początku trwania zjawiska – nie ma żadnego sensu, tworzy złudzenie przewidywalności tego zjawiska i ukierunkowuje sposób interpretacji.
- 2) Brak ścisłego określenia co się analizuje** – liczbę dodatnich wyników testów diagnostycznych, liczbę zakażeń, liczbę zachorowań?
- 3) Brak kontekstu w prezentowaniu zakażeń wirusem SARS-Cov-2** – stąd wiele oczywistości i banalnych stwierdzeń, których nie powinno się uznawać za „sukces” modeli prognostycznych, tylko efekt corocznej sezonowości chorób wirusowych.
- 4) Brak świadomości braku stabilności badanego zjawiska** – chodzi tu o błędne koło samosprawdzalności prognoz – „złe prognozy powodują wzmożone testowanie, co prowadzi do wykrycia większej liczby zakażeń”.

# Tematyka zajęć

- 1. Opis danych czasowych (wykresy, indeksy dynamiki, analiza sezonowości).**
- 2. Prognozowanie na podstawie modeli tendencji rozwojowych (trendy liniowe i nieliniowe – wykresy i analiza regresji) bez i z efektami sezonowymi.**
- 3. Prognozowanie na podstawie modeli adaptacyjnych (średnie ruchome, wyrównywanie wykładnicze).**
- 4. Symulacje przebiegu zjawisk ludnościowych – prognoza demograficzna.**
- 5. Prognozowanie zmiennej dychotomicznej – regresja logistyczna.**
- 6. Inne metody, rozszerzenia, uzupełnienia.**

# Program *STATISTICA*

Większość obliczeń i prezentacji graficznych wykonywać będziemy za pomocą programu *STATISTICA-13* – specjalistycznego narzędzia analizy danych.

Program *STATISTICA* jest zainstalowany w pracowniach, ponadto licencja, którą posiada Politechnika Rzeszowska, uprawnia studentów do posługiwania się programem na komputerach domowych.

Informacje o sposobie instalacji i wersja instalacyjna programu *STATISTICA* znajdują się na stronie:

<http://ci.prz.edu.pl/licencje-oprogramowania>

Część obliczeń i analiz wykonywać będziemy także za pomocą arkusza kalkulacyjnego *Excel*.

# Analitik danych na rynku pracy

Ranking najlepszych zawodów wg raportu „Jobs Rated Almanac” 2015 przygotowanego przez serwis CareerCast \*

1. aktuariusz – ok. 94 tys. dol. rocznie;
2. audiolog – 71 tys. dol.;
3. matematyk (ale taki od biznesu i przemysłu) – 102 tys. dol.;
4. statystyk analityk (prowadzi badania i analizuje dane) – 79 tys. dol.;
5. inżynier biomedyczny – 89 tys. dol.;
6. naukowiec (chodzi o kogoś, kto jest jednocześnie informatykiem, statystyko-analitikiem i jeszcze ma wiedzę z konkretnej dziedziny) – 89 tys. dol.;
7. higienistka stomatologiczna – 71 tys. dol.;
8. software inżynier – 93 tys. dol.;
9. terapeuta (zarówno od ciała, jak i od duszy) – 77 tys. dol.;
10. analityk systemów komputerowych – 81 tys. dol.;
11. patolog mowy (coś takiego jak logopeda, bo zajmuje się zaburzeniami mowy i słuchu, czyli osobami z problemami z komunikacją) – 71 tys. dol.;
12. dietetyk – 56 tys. dol.

\*Przy tworzeniu zestawienia brano pod uwagę nie tylko zarobki, lecz także stopień niebezpieczeństwa profesji czy wysiłek konieczny do wykonania obowiązków, stres towarzyszący wykonywaniu pracy oraz łatwość znalezienia zatrudnienia.

# Czy korzystamy z prognoz?

**W codziennym życiu korzystamy (czasem bezwiednie) z różnych przewidywań dotyczących przyszłości:**

- **prognozuje się liczbę studentów, żeby zaplanować budżet uczelni, inwestycje;**
- **prognozuje się kursy akcji;**
- **prognozuje się popyt na samochody, żeby prognozować wpływy z podatku od sprzedaży nowych samochodów;**
- **...a producenci prognozują popyt na samochody, żeby wiedzieć, ile ich produkować.**
- **prognozuje się prawdopodobieństwo wypadku, żeby ustalić koszty ubezpieczenia samochodu;**
- **prognozuje się wynik wyborów – badanie opinii publicznej to też prognoza;**
- **...**
- **i wreszcie na końcu wymieńmy – choć to wcale nie najmniej ważna prognoza – przewidywanie pogody.**

# Czy sami prognozujemy?

**Wiele nieformalnym prognoz dokonujemy każdego dnia:**

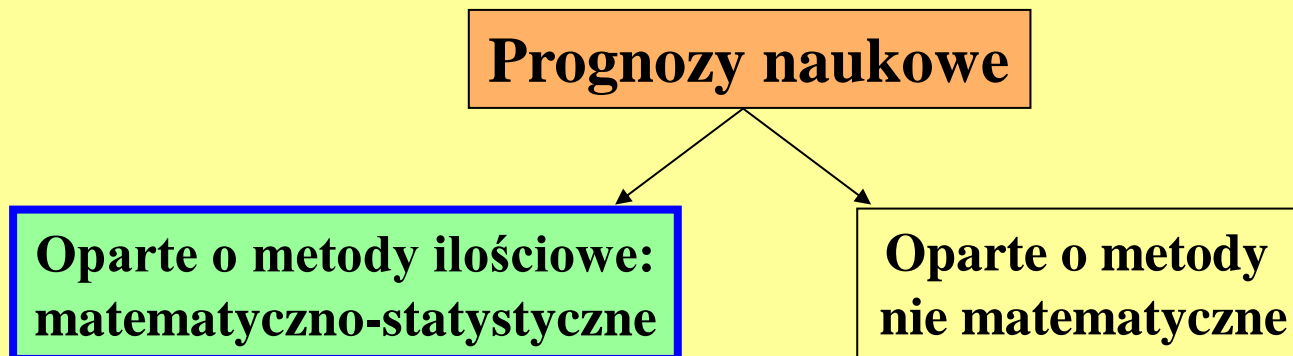
- **czy odbędzie się dzisiejszy wykład z *Prognozowania*?**
- **czy uda mi się za pierwszym razem zdać egzamin?**
- **czy w najbliższy weekend będzie pogoda, która pozwoli na wyjazd na narty?**
- **czy kupić samochód z silnikiem benzynowym, dieslem czy na gaz – tu skorzystamy z prognozy dotyczącej zużycia paliwa oraz prognozowanych cen paliw?**
- **... ..**

# W jaki sposób możemy prognozować?

- Przewidywania racjonalne – są oparte na potocznie rozumianym doświadczeniu, lecz nie używa się w nich metod naukowych.

- Przewidywania naukowe – oparte na regułach naukowych, które charakteryzują się ścisłością, precyzją i powtarzalnością.

- Przewidywania nieracjonalne – wróżby, prorocтва, przepowiednie.



# **Modelowanie (symulacja) w prognozowaniu**

**Każda sformalizowana próba wnioskowania o nieznanej rzeczywistości (przyszłości) będzie oparta na pewnym uproszczonym (co nie znaczy, że banalnym) modelu opisującym najważniejsze aspekty badanego problemu.**

**W wielu sytuacjach nie znamy tego modelu i dobieramy go na podstawie danych historycznych.**

**Czasem „mechanizm” rozwoju danego zjawiska jest znany, choć jego dokładny przebieg zależy od przyjętych założeń i przypadku – wtedy dokonujemy symulacji przebiegu tego zjawiska w przyszłości.**



# Ciekawostki

Prognozy mogą dostarczać informacji – ukierunkowywać pewne działania, mogą też pełnić funkcje aktywizujące – ostrzegawcze.

Nie zawsze chcemy, żeby prognozy się spełniły.

Prognozy mogą mieć właściwość „samospełnienia się” (mit o Edypie). Prognoza ekonomiczna, która będzie zapowiadała wzrost cen jakiegoś towaru może do niego doprowadzić z powodu znaczącego wzrostu popytu, generowanego przez klientów kupujących „na zapas”. I w ten sposób wygenerować wzrost cen.

Często używa się klasyfikacji prognoz, w których niewiedza jest spowodowana jest czynnikiem czasowym, na krótko-, średnio- i długoterminowe. Należy mieć świadomość, iż w przypadku zjawisk demograficznych prognoza pięcioletnia będzie miała charakter krótkoterminowej, podczas gdy ten sam horyzont czasowy w odniesieniu do kursów akcji wydaje się niewyobrażalnie odległy.

# Prognozowanie – potrzeba czy moda?

Prognozowanie stało się konieczne i możliwe dopiero z nadejściem epoki cywilizacji informacyjnej. Bezprecedensowe przyspieszenie tempa zmian, z jakim mamy do czynienia w ostatnich kilkudziesięciu (a szczególnie kilkunastu latach) zmusiło ludzi do podjęcia prób przewidywania przyszłości, w kontekście konsekwencji jakie te zmiany mogą wywołać.

*K. Boulding: „Świat dzisiejszy... jest tak różny od świata, w którym ja się urodziłem, jak tamten był różny od czasów Juliusza Cezara. Mówiąc z grubsza, urodziłem się w środkowym okresie historii człowieka. Prawie tyle samo wydarzyło się, odkąd ja się urodziłem, co na przestrzeni całego okresu poprzedzającego moment moich urodzin.”*

# Szybkość zmian...

A. Toffler: „...gdyby ostatnie **50 tysięcy lat** istnienia człowieka podzielić na odcinki życia wynoszące około **62 lata**, to okazałoby się, że mieliśmy już **800 takich odcinków**. Z tych ośmiuset pełne **650 przeżyliśmy w jaskiniach**. Zaledwie w ostatnich **70** takich odcinkach udało się człowiekowi nawiązać faktyczną łączność pomiędzy jednym a drugim okresem życia – co stało się możliwe dzięki pismu. Zaledwie w ostatnich **6** odcinkach po raz pierwszy udostępniono człowiekowi słowo drukowane w skali masowej. Zaledwie w ostatnich **4** – możliwym stało się odmierzanie czasu z taką dokładnością. Zaledwie w ciągu ostatnich **2** – ktoś gdzieś posłużył się silnikiem elektrycznym.”

Telefon komórkowy w powszechnym użytku, to **0,2** takiego odcinka czasowego, niewiele krócej lub dłużej korzystamy z tak ogromną intensywnością z komputerów osobistych, Internetu, aparatów cyfrowych.

# Jak ocenić wiarygodność prognozy?

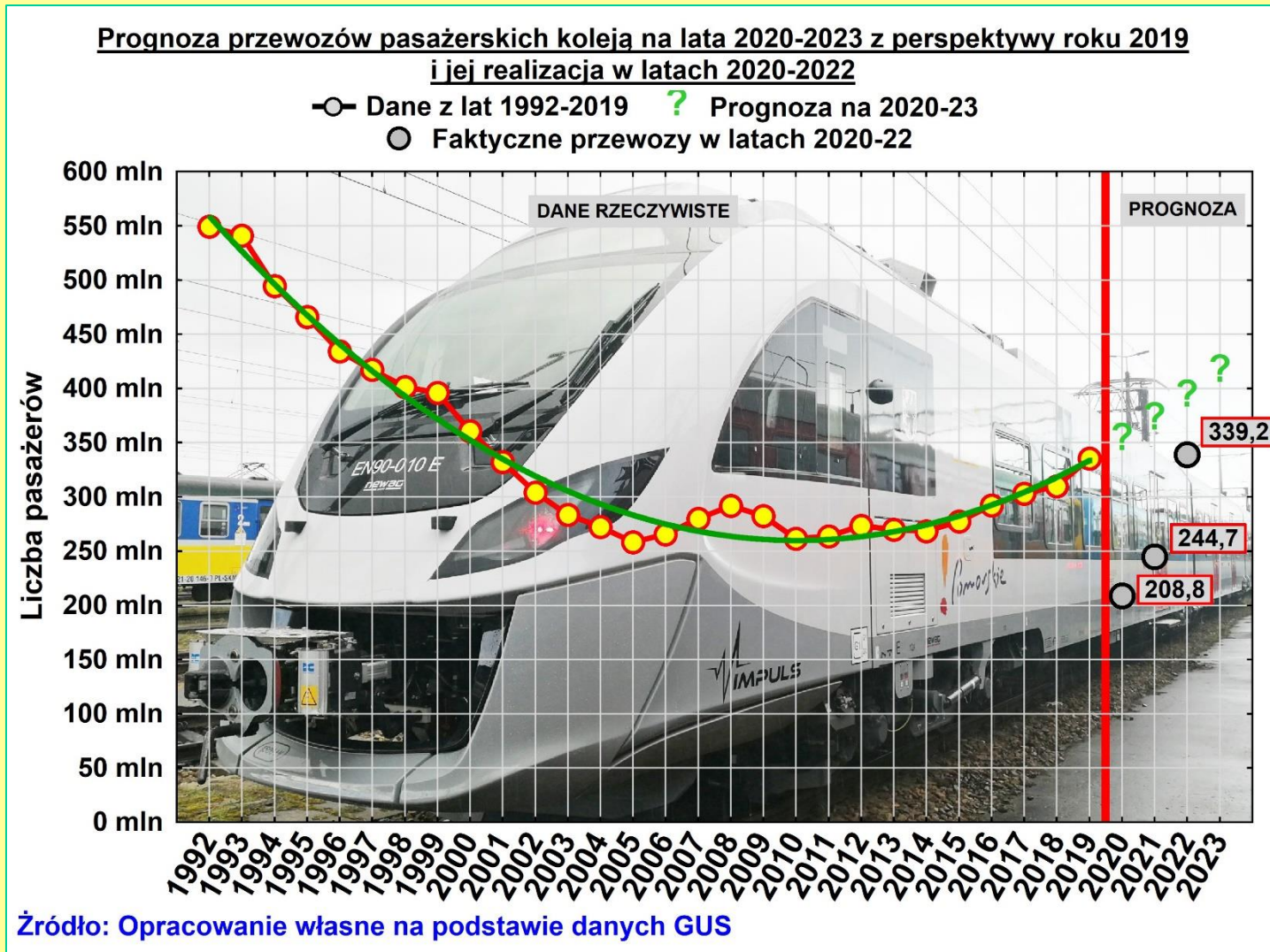
Podstawowym warunkiem skuteczności prognoz jest pewna **stabilność modelowanego zjawiska w czasie i przestrzeni**

Wydaje się, iż można wysnuć tezę o większej stabilności zjawisk naturalnych i co za tym idzie, można spodziewać się dużej sprawdzalności prognoz w medycynie, biologii, demografii a nieco mniejszej w socjologii. Zjawiska *par excellence* ekonomiczne (często wręcz „wirtualne”) jak na przykład kursy akcji, walut są bardzo niestabilne i jako takie stwarzają duże problemy podczas tworzenia prognoz.

Niestety, błędy prognozy możemy wyliczyć dopiero po jej zrealizowaniu, a więc wtedy, gdy staje się już niepotrzebna. Warto to jednak robić, bo przecież uczymy się na błędach.

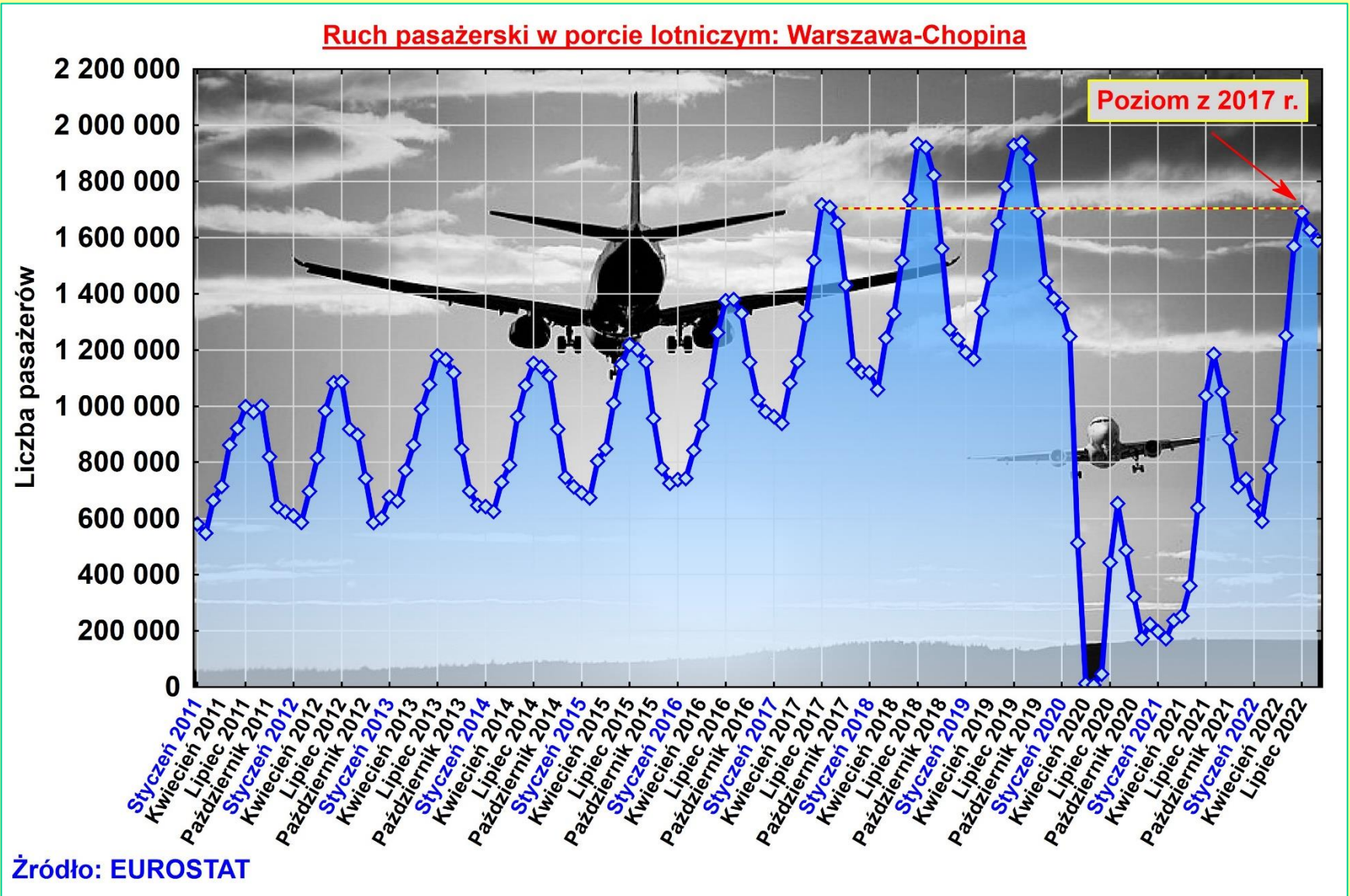
# A z tą stabilnością nie jest za dobrze... (1)

Ostatnie 2-3 lata to nie jest dobry czas dla poważnych analityków danych...



# A z tą stabilnością nie jest za dobrze... (2)

Ostatnie 2-3 lata to nie jest dobry czas dla poważnych analityków danych...



# Źródła wiedzy

## Podstawowe proponowane podręczniki:

**Hydzik P., Sobolewski M., *Komputerowa analiza danych społeczno-gospodarczych*, Oficyna Wydawnicza Politechniki Rzeszowskiej, Rzeszów 2007.**

***Prognozowanie gospodarcze. Metody i zastosowania*, pod red. nauk. Cieślak M., PWN, Warszawa 1997.**

**Dittmann P., *Metody prognozowania sprzedaży w przedsiębiorstwie*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, 1998.**

**Zeliaś A., *Teoria prognozy*, PWE, Warszawa 1997.**

**Aczel D.A., *Statystyka w zarządzaniu*, PWN, Warszawa 2000.**

**Koronacki J., Mielniczuk J., *Statystyka dla studentów kierunków technicznych i przyrodniczych*, Wydawnictwo Naukowo-Techniczne, Warszawa 2000.**

**Dobosz M., *Wspomagana komputerowo analiza danych*, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2001.**

**Stanisz A., *Przystępny kurs statystyki z wykorzystaniem pakietu Statistica PL na przykładzie danych z medycyny. Tom I i II*. StatSoft Polska, Kraków 2000.**

# Źródła wiedzy

## Inne materiały

(dla osób zainteresowanych problematyką –  
w szczególności „filozofią” prognozowania):

**Z. Czerwiński, *Jeszcze raz o prognozach*, Przestrzenno-czasowe modelowanie zjawisk społeczno-gospodarczych, Wyd. AE Kraków, 2000.**

***Świat i Polska 2005. Prognozy, analizy, opinie.* The Economist/Polityka, grudzień-luty 2005.**

**Kinsella K., Velkoff V.A., *An Aging World: 2001*, US. Census Bureau, Washington 2002.**

**Toffler A., *Szok przyszłości*, Wydawnictwo Zyska i S-ka, Poznań 2004.**

**Fukuyama F., *Koniec człowieka*, Wydawnictwo Znak, Kraków 2004.**