



AMAINSTITUTE.PL





V Edycja Konferencji "Modelowanie dla Biznesu"

16 stycznia 2018 • SAS Innovation Hub • Warszawa

Jak żyć w epoce AI & Machine-Learning

Karol Przanowski,
AMA Institute
Szkoła Główna Handlowa - SGH

- Znaki czasu
- Rekomendacje:
 1. Zmień myślenie – analityka w centrum
 2. Zmieniaj język analityczny w biznesowy
 3. Bądź proaktywny, twórz kulturę data driven
 4. Zaakceptuj niską jakość danych
 5. W wielu procesach znajdź zastosowanie dla modeli predykcyjnych
 6. Rozwijaj myślenie statystyczne
 7. Naucz się używać oprogramowania: SAS, R i Python
 8. Zawsze zaczynaj od zrozumienia procesu
 9. Masz problemy? Przyjdź do nas



Świat się zmienia i możemy tylko zdążyć wsiąść do kolejnego pociągu

Rynek zmienia się szybciej niż wiedza ekspercka

Dane i analityka stają się cennymi aktywami każdego przedsiębiorstwa

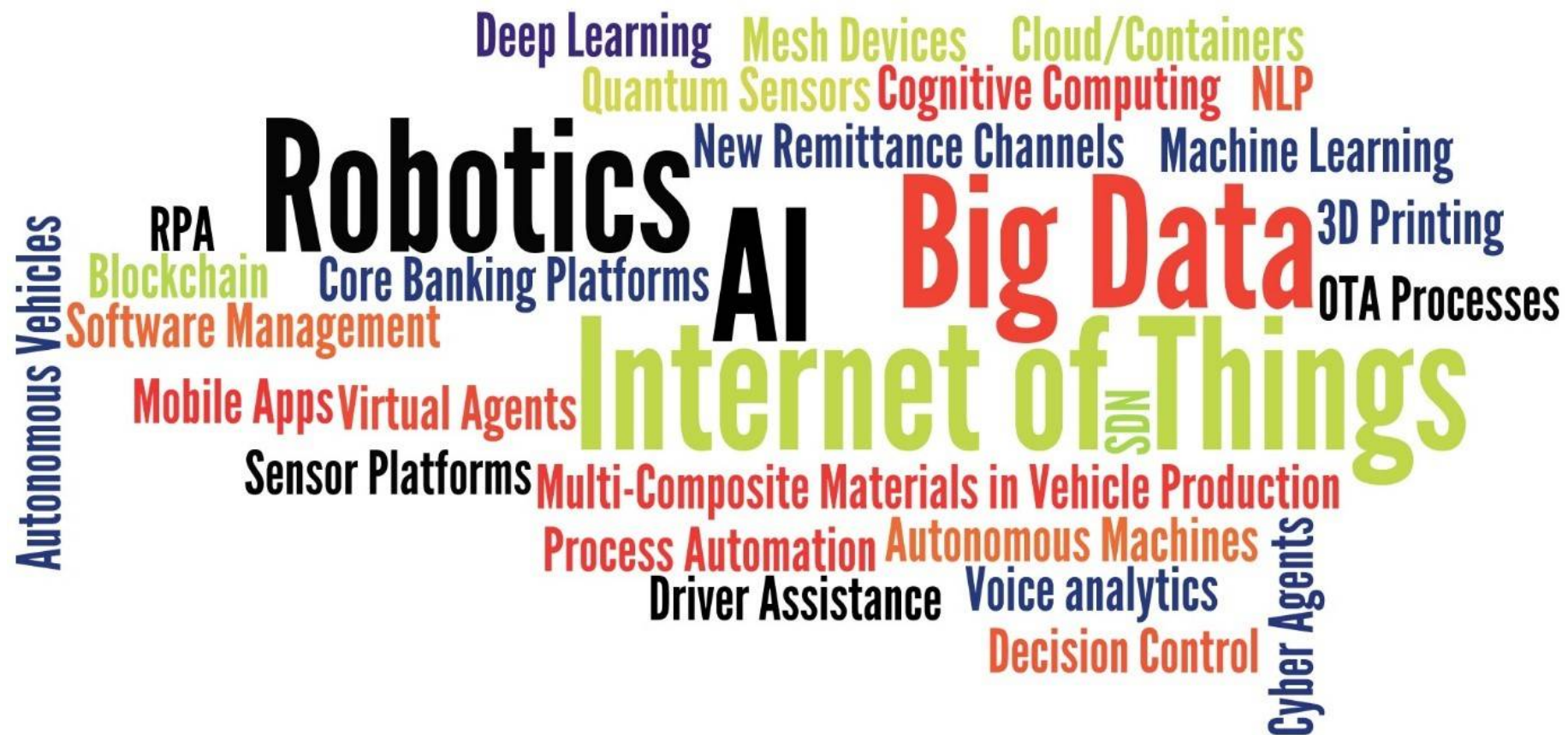
Digitalizacja wszystkiego i wszystkich

Big Data napotkało na pierwsze sukcesy i porażki

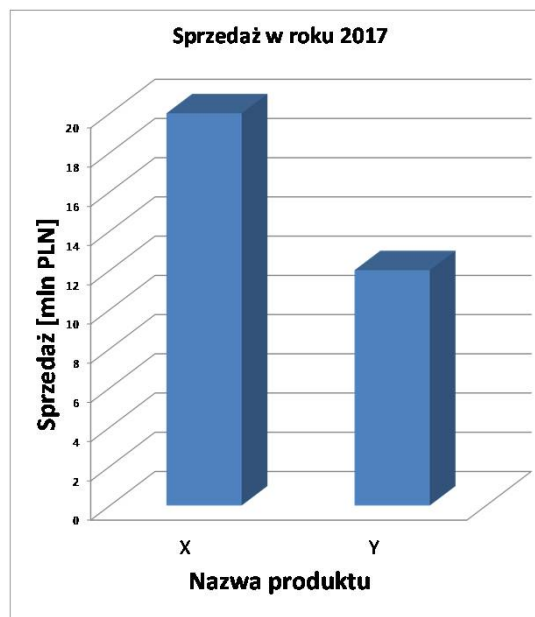
Nawet sprzedaż butów zwiększa się dzięki analityce



Nie bój się – to tylko synonimy analityki



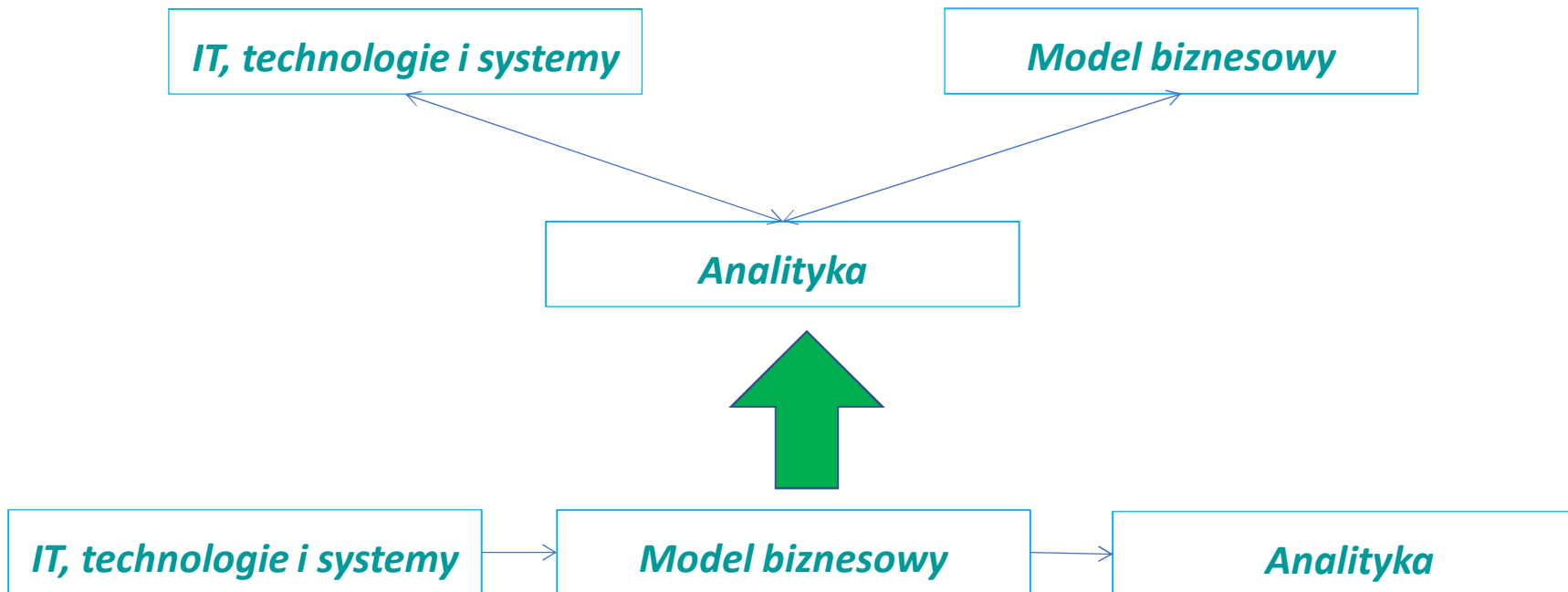
A po studiach trafiaamy do naszej pierwszej pracy!!!



Nie przerażaj się, nie musisz zderzać się ze ścianą

Zmień myślenie – analityka w centrum

Wszystko da się skopiować z wyjątkiem wewnętrznych danych i analityki



Najczęściej zaczyna się od technologii, potem buduje się model biznesowy, a na końcu analitykę

Zmieniaj język analityczny w biznesowy

Musisz przekonywać językiem zrozumiałym przez biznes

Nowy model predykcyjny zwiększy sprzedaż o 20%

Nowe reguły analityczne zmniejszą koszty o 30%

Nowe narzędzie analityczne zmniejszy odchodzenie klientów o 30%

Dzięki lepszym danym i lepszym kryteriom selekcji zwiększymy zysk o 15%



ZAŁOŻENIA:

- 20 tys. Klientów
- 348 kampanii marketingowych rocznie
- ~7 mln decyzji – wysłać czy nie?
- Koszt jednostkowy: 5
- Zarobek przy zakupie: 800
- Średnia szansa zakupu: 0,5%



Wysyłamy wszystkim

- Przychody: 28 000 000
- Koszty: 35 000 000
- Zysk: -7 000 000

Całkowicie nieopłacalne



Reguły eksperckie

- Przychody: 15 895 139
- Koszty: 12 250 000
- Zysk: 3 645 139

Zauważalne zyski



Wyniki Finansowe

Występują zauważalne zyski, ale czy można je poprawić?

Case Study (2) – możliwe warianty rozwoju

Model_A – dane bazowe

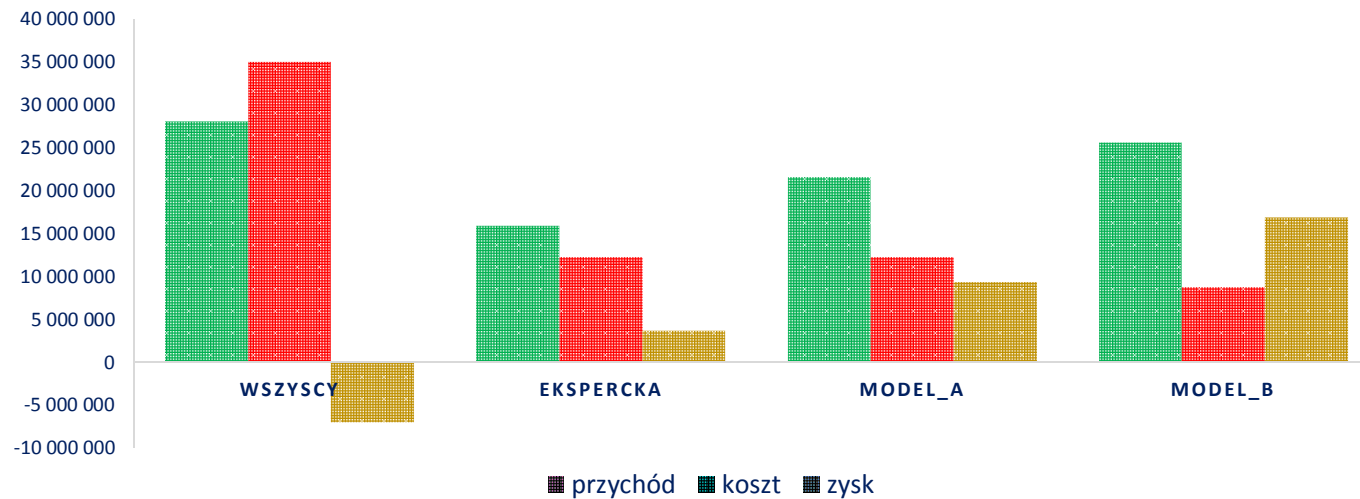
- Przychody: 21 531 163
- Koszty: 12 250 000
- Zysk: 9 281 163

Wzrost zysków o ponad 5,5 mln vs reguły eksperckie

Model_B – specjalistycznie rozbudowany zbiór danych

- Przychody: 25 599 340
- Koszty: 8 750 000
- Zysk: 16 849 340

Jeszcze większy zysk oraz obniżenie kosztów wstępnych



Wyniki Finansowe

Case Study (3) – Świat analityka

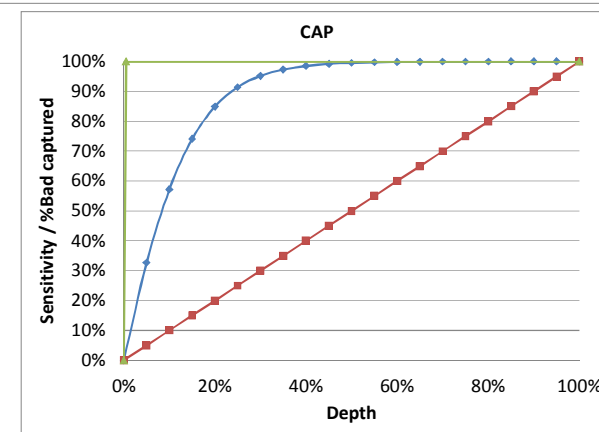
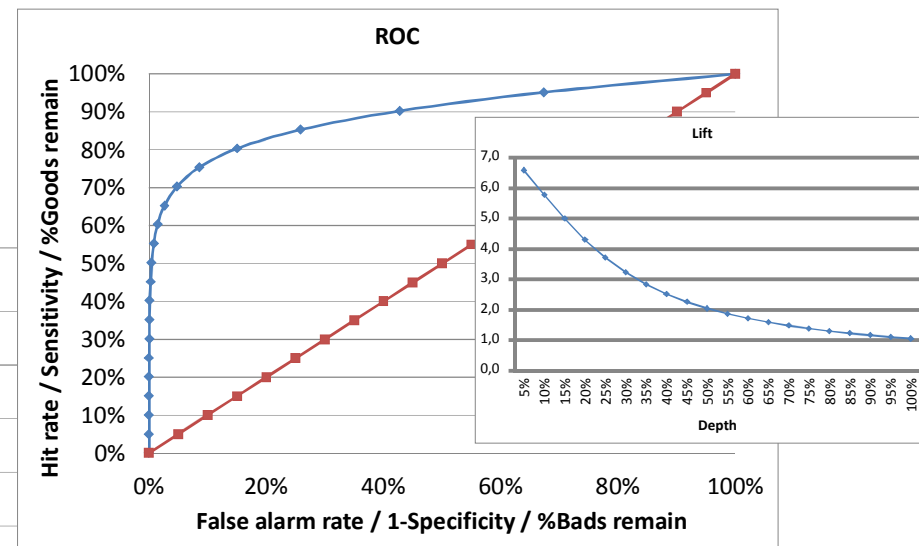
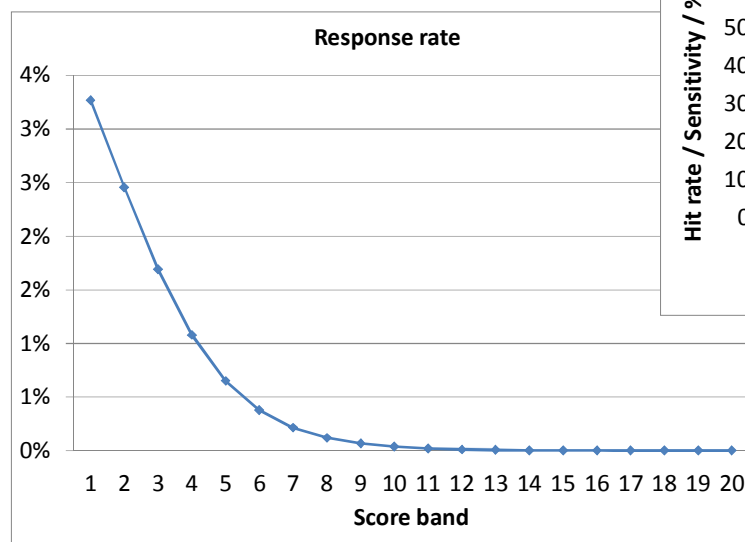
Liczba decyzji	7 000 000
Średni przychód ze sprzedaży	800
Średni koszt kontaktu	5

Globalna moc predykcyjna - Gini	78,36%
---------------------------------	--------

Globalna konwersja	0,5%
Konwersja na grupie docelowej	1,83%
Udział grupy docelowej	25,00%
Statystyka Lift na grupie docelowej	3,66
Statystyka Gains na grupie docelowej	91,43%
Koszt globalny	35 000 000
Przychód globalny	28 000 000
Wynik (zysk) globalny	-7 000 000

Koszt na grupie docelowej	8 750 000
Przychód na grupie docelowej	25 599 340
Zysk na grupie docelowej	16 849 340
Liczba ofert, grupa docelowa	1 750 000
Liczba respondentów w grupie docelowej	31 999

Liczba kampanii	29
Liczba miesięcy	12
Liczba klientów	20 115
Liczba decyzji	7 000 000

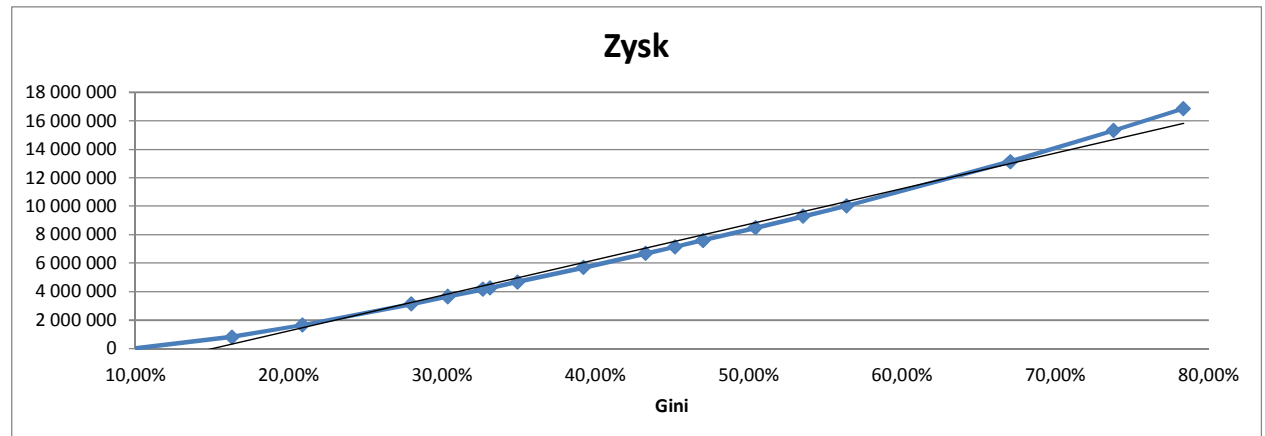


Studium przypadku w Excelu:

<http://administracja.sgh.waw.pl/pl/OW/publikacje/Strony/2015.aspx>

Zmienna: Liczba wysyłanych emaili w ciągu ostatnich 2 miesięcy			
Numer kategorii	Kategoria - opis profilu	Konwersja	Udział w populacji
1	1 < VAR_N_EMAIL_2M	5,69%	19,84%
2	0 < VAR_N_EMAIL_2M <= 1	5,17%	18,84%
3	not missing(VAR_N_EMAIL_2M) and VAR_N_EMAIL_2M <= 0	4,27%	61,31%
			100,00%

Delta Gini	Delta Zysku
1%	272 569
5%	1 362 844
10%	2 725 687



Bądź proaktywny, twórz kulturę *data driven*

Szukaj miejsc pracy, gdzie analityka jest w centrum, a jeśli jej nie ma, to pomóż to osiągnąć

Nie rozstawaj się ze swoimi raportami i narzędziami szybkiej eksploracji danych

Aktywnie zabieraj głos na spotkaniach, obalaj mity pokazując raporty

Naucz się biznesu a swojego managera analityki

Zawsze bądź gotowy do przemowy w windzie – *elevator speech*

Pytaj o szczegóły procesu

Gdy robisz raport, pytaj do czego jest potrzebny



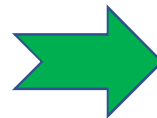
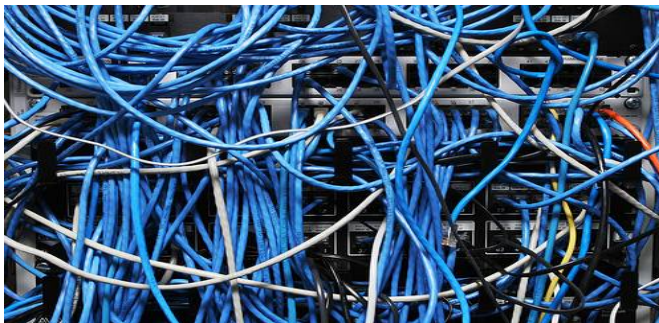
Zaakceptuj niską jakość danych,
możesz tylko zrobić lepszy krok a nie rewolucję

Big Data wcale nie oznacza duże dane – po oczyszczeniu pozostaje mało

Nie używaj niskiej jakości danych jako powodu Twoich niepowodzeń

Jeśli identyfikujesz błędy w danych, to Ty od tej chwili jesteś za nie odpowiedzialny!!!

Proponuj procesy, które od razu zbierają lepsze dane – pola słownikowe



NOWE PARADYGMATY

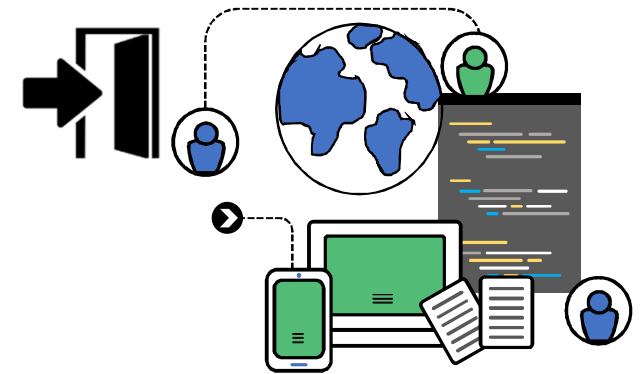
- DWH (Hurtownia danych):
 - Najpierw oczyścić potem załadować (stare)
 - Załadować i potem się martwić (nowe)
- Modelowanie (prognozowanie):
 - Znajdź przyczynę i skutek, obserwuj istotne zmienne, czynniki (stare)
 - Sprawdź, które ze zgromadzonych danych wpływają na modelowane zdarzenie, zależności od zmiennych pochodnych (nowe)
 - Podczas jednego projektu zbiera się dane a podczas drugiego się je wykorzystuje



GŁÓWNE IDEE

- Pozwólmy mówić danym
- Nie ważne dlaczego, ważne, że działa!
- (*not why, but what*)
- Najważniejsze, że da się zarabiać!!!
- Gromadźmy dane bez ich pierwotnego przeznaczenia

DATA
ENTRANCE



BIG DATA

Mayer-Schonberger V i Cukier K (2013). *Big Data: A revolution that will transform how we live, work and think*. An Eamon Dolan Book, Houghton Mifflin Harcourt, Boston, New York.

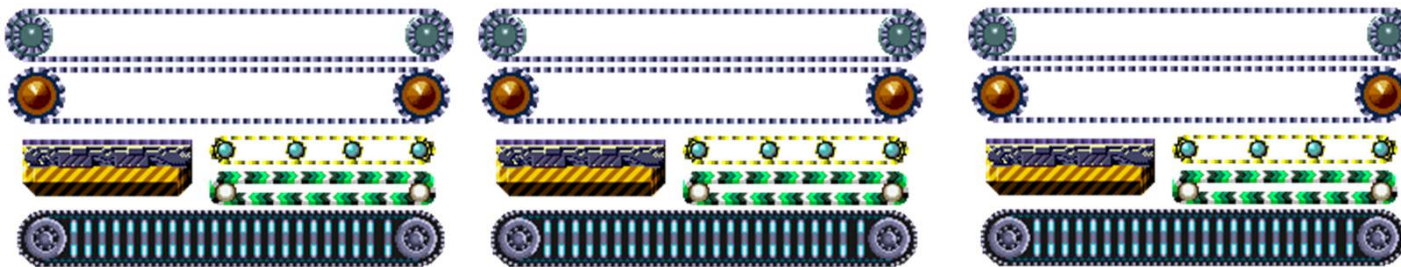
W wielu procesach znajdź zastosowanie dla modeli predykcyjnych,
stosuj procesy z automatycznym podejmowaniem decyzji

Jeśli się nad tym zastanawiasz, to konkurencja już dawno to wdrożyła

Tylko automatyczne procesy pozwolą przetrwać Tobie i Twojej firmie

Zastosowania modeli predykcyjnych ograniczone są tylko przez Twoją wyobraźnię

Prawie każdy proces masowy z decyzją (tak, nie) daje się zautomatyzować



Dlaczego modele predykcyjne nie są stosowane na szeroką skalę?

WYSOKA CENA RYNKOWA
OPROGRAMOWANIA

DŁUGI CZAS
WDROŻENIA

KONIECZNOŚĆ POSIADANIA
SPECJALISTYCZNYCH ZASOBÓW

BRAK ZROZUMIENIA WARTOŚCI
DODANEJ ANALIZY DANYCH

SIŁOSOWOŚĆ DZIAŁAŃ W
ORGANIZACJI

SKOMPLIKOWANA
TECHNOLOGIA



Trzeba przełamywać bariery, obalić *status quo* i zlikwidować silosy!!!

NIEPOWODZENIA BIG DATA

- Brak dobrych *business case*
- Dane są zbierane, ale zbyt mało uzasadnia się ich przydatność
- Lekceważone problemy jakości danych
- Umniejszanie problemu wnioskowania na podstawie obciążonej próby
- Zbyt duży nacisk położony na technologię
- Złudne nadzieje szybkiego „klikania modelu”
- Brak inwestycji w przygotowanie i wykształcenie inżyniera danych
- Brak publicznych danych, dostępnych i przykładowych



PRZYKŁADY BRANŻ, W KTÓRYCH MODELE PREDYKCYJNE MOGĄ BYĆ
WYKORZYSTYWANE

FINANSE

- Fundusze inwestycyjne i gwarancyjne
- Ubezpieczenia
- Kredyty / Leasing/ Faktoring
- Windykacja
- Ochrona przed nadużyciami

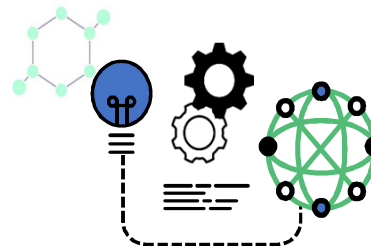
MARKETING

- Częstotliwości i rodzaj kontaktu z klientem
- Programy lojalnościowe
- Retencja w usługach abonamentowych
- Promocje cenowe
- Sprzedaż internetowa

INNE

- Centra usług wspólnych
- Punkty masowej obsługi klienta
- Domy wysyłkowe
- Logistyka
- Firmy windykacyjne

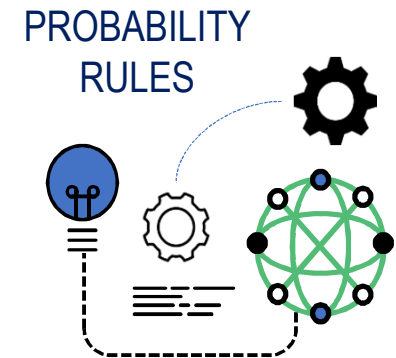
NAUKA



Rozwijaj myślenie statystyczne,

identyfikuj zjawiska masowe i wykrywaj główne prawidłowości

- Myślenie biznesowe:
 - Ile zarobimy, ile stracimy? Ważniejsze są plany bliższe i pieniądze w zasięgu ręki! Szybki sukces się liczy!
- Myślenie operacyjne:
 - Jak utrzymać proces, *status quo*? Nie zwiększajmy kosztów, róbmy efektywniej!
- Myślenie zarządcze:
 - Wystarczy 20% trudu, bo daje 80% efektów. Pareto działa!
- Myślenie statystyczne:
 - Estymować liczby ukryte, trudne do zaobserwowania
 - Pareto nie działa w przypadku analiz, każda analiza daje większe efekty i zrozumienie procesu
 - Znaleźć proporcje między udziałem zyskownych klientów a klientami odrzuconymi i zniechęconymi



Naucz się używać oprogramowania: SAS, R i Python,
w dzisiejszych czasach musisz znać płatne i bezpłatne rozwiązania

Trudniej znaleźć „mistrza” od opensource, SAS może wytyczać kierunki

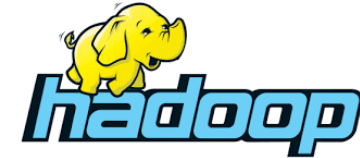
Zbyt duża liczba pakietów bezpłatnych – który wybrać?

Trudności bezpłatnych rozwiązań niweluj siecią kontaktów

Rozwijaj środowisko analityków

Rozwijaj techniki programistyczne, nie tylko przez metodę prób i błędów





PostgreSQL



INŻYNIER
DANYCH



Co jest ważniejsze?

- Wygodny interfejs - UFI
- Tanie oprogramowanie
- Hard code
- SLA i dobry HELPDESK
- **Kompetencje zespołów inżynierów danych**

INŻYNIER
DANYCH



Zawsze zaczynaj od zrozumienia procesu,
pamiętaj, że stary proces może zaburzać interpretację danych

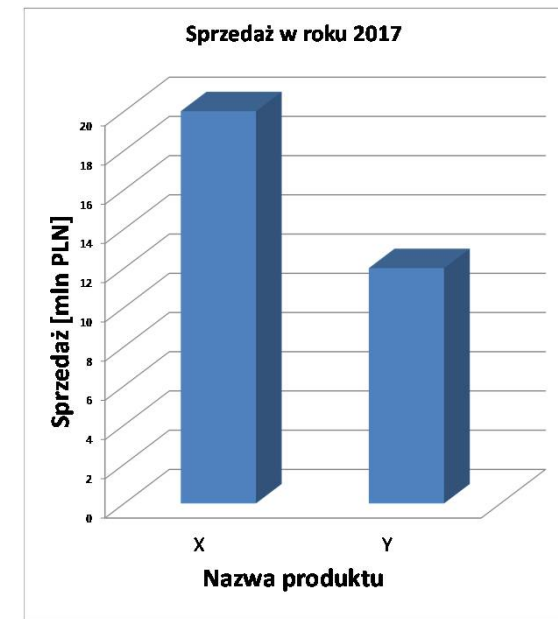
Dlaczego produkt X posiada większą sprzedaż?

Bo klienci go preferowali?

Bo pracownicy sprzedaży mieli system motywacyjny?

Bo produkt był tańszy?

Bo produkt szybko można było dostarczyć?



Masz problemy? Przyjdź do nas,

pomożemy, razem zrobimy prezentację, przekonamy Twoich przełożonych

1. Dostęp do zespołu z dużą wiedzą o lokalnym rynku
2. Elastyczne podejście do klienta, zamiast wykorzystywania generycznych schematów
3. Połączenie dużego doświadczenia biznesowego oraz IT z doświadczeniem naukowym, zachowując zdrową krytykę zarówno narzędzi analitycznych, jak i IT
 - zrozumienie biznesu i ograniczeń IT
4. Rozwiązania szyte na miarę
5. Szybki zwrot poniesionych kosztów
 - iteracyjne podejście zapewniające nacisk na otrzymanie zwrotu dla klienta w krótkim okresie