



Wstęp do wizualizacji danych

Część I

Hubert K. Kobierzewski

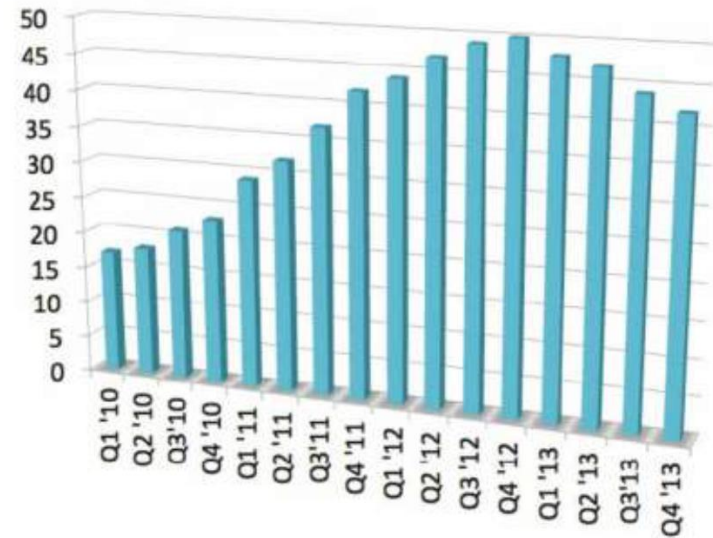
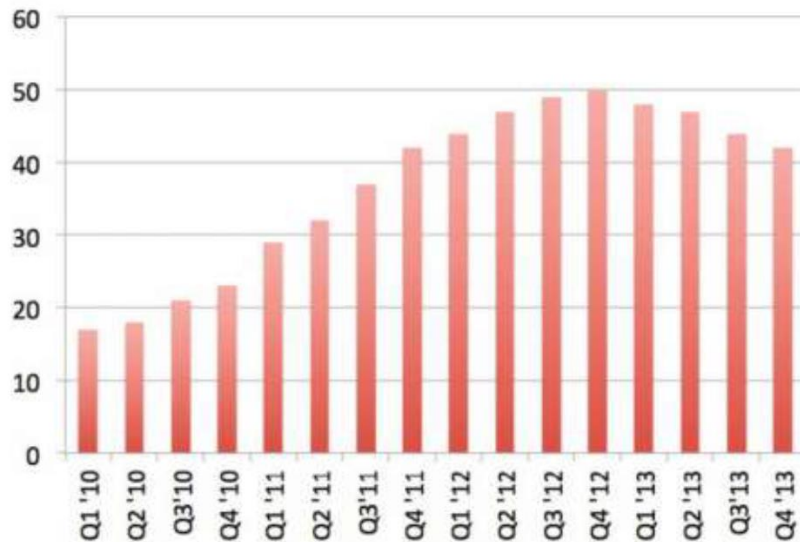
Hubert K. Kobierzewski

- BI Practice Lead w Codec (ponad 11 lat)
- Specjalizacja: Hurtownie danych, procesy ETL oraz Business Intelligence
- Ex-programista
- Certyfikowany głównie w platformie danych MSFT (MCDBA, MCTS, MCITP, MCSE – BI, MCT)
- Lider grup społecznościowych Data Community Poland oraz Warsaw Power BI User Group

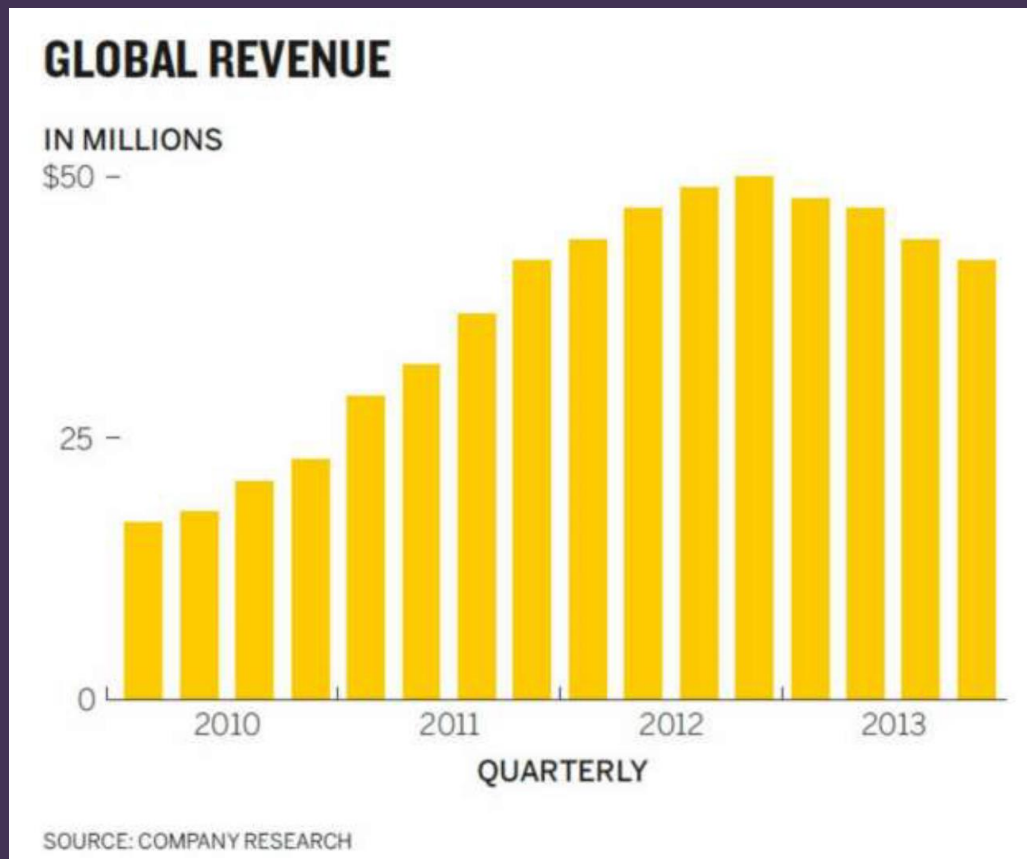


JAK DOBRZE WIZUALIZOWAĆ DANE

Bardzo łatwo można stworzyć zły wykres



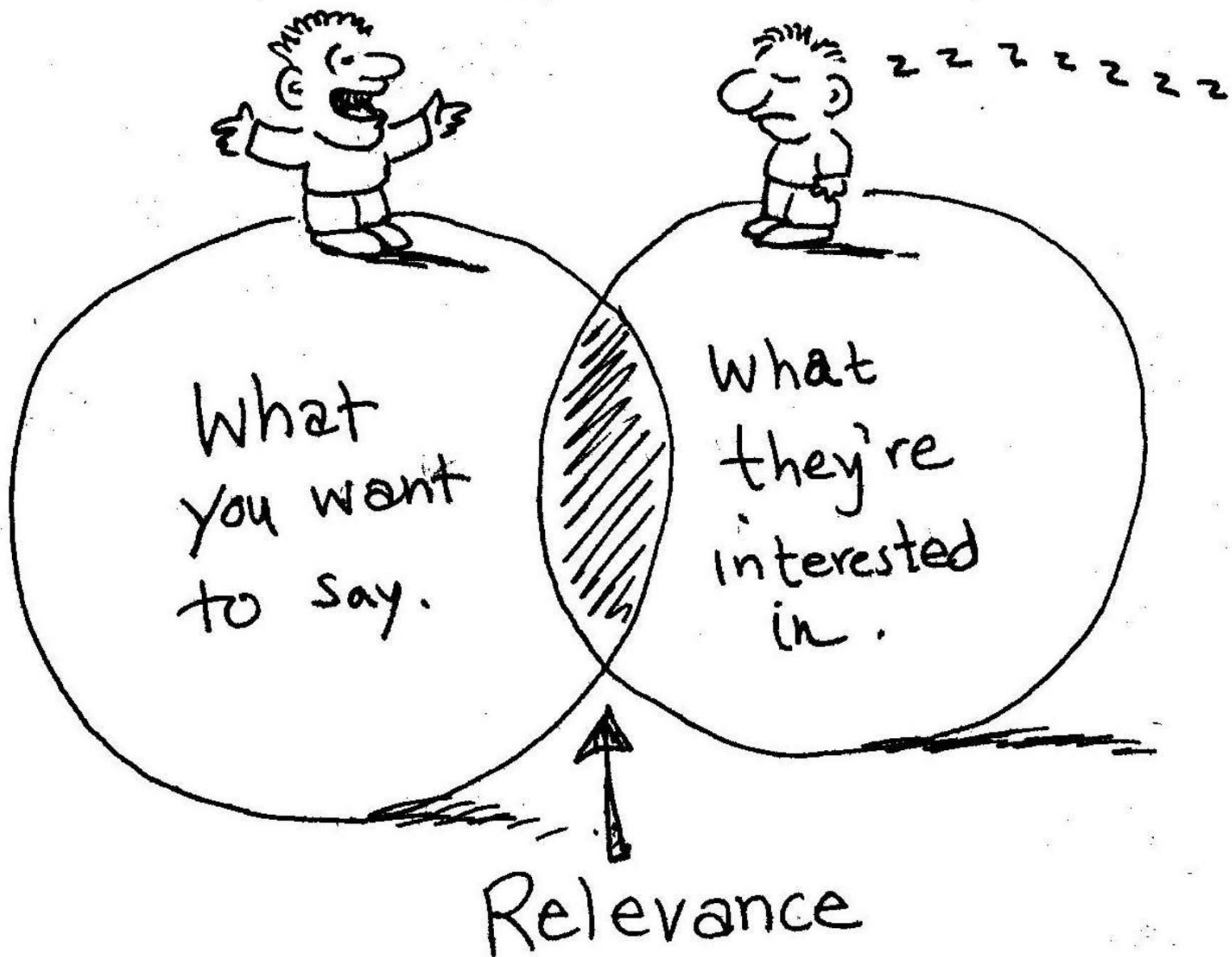
A czy ten może być lepszy?



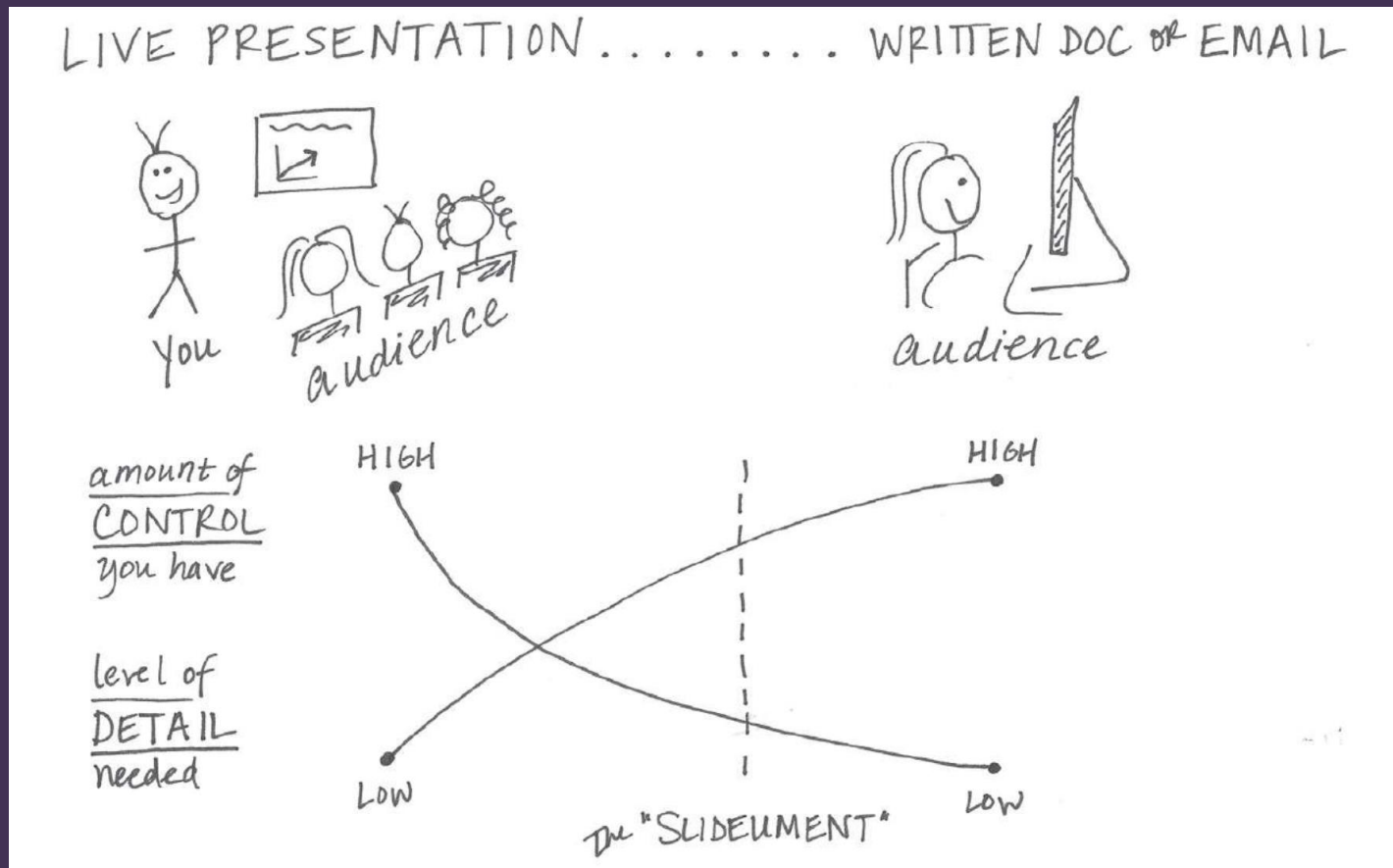
- Czy jest to dobry wykres?
- Na pewno jest on lepszy od poprzednich, ale czy jest wystarczająco dobry?
- Jakie są kryteria uznania wykresu za dobry?

Znaczenie kontekstu

- Nie sposób jest powiedzieć czy wykres jest dobry lub zły nie znając jego kontekstu
- Kryteria:
 - Kim są odbiorcy wykresu?
 - Jakie są ich potrzeby?
 - Jaką myśl chcemy zaprezentować? Co chcemy zaprezentować vs to co powinno być zaprezentowane
 - [ostatecznie] Jak to powinno być zaprezentowane?
- Jeśli zaprezentujemy ten wykres zarządowi, który zna wyniki kwartalne to może ten wykres nie będzie najlepszy. Możemy utracić uwagę odbiorców.



Poziom szczegółowości

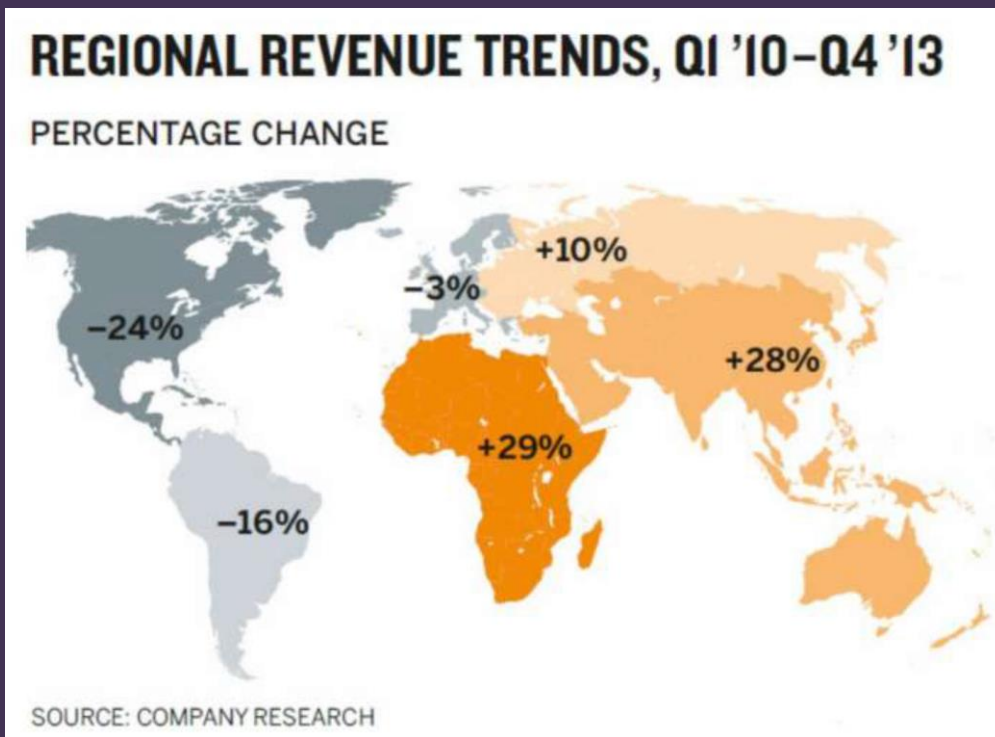


Kto jest odbiorcą naszego raportu? Co należy na nim pokazać?



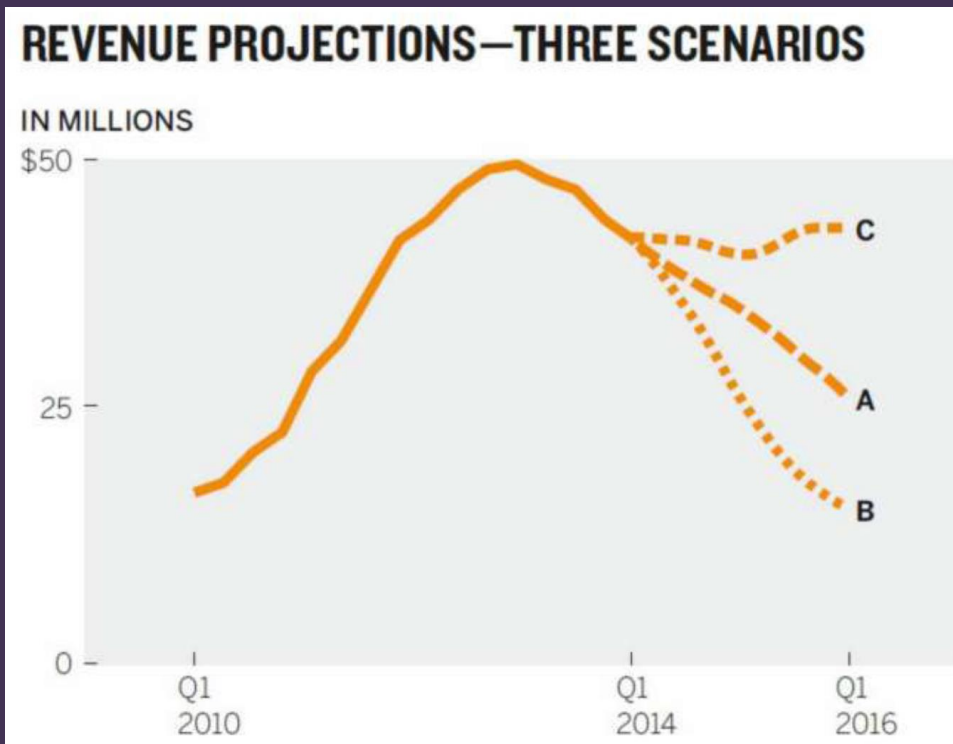
- Co jeśli zarząd rozważa gdzie inwestować w celu odwrócenia trendu spadkowego?
- Może lepiej pokazać rozbić na poszczególne regiony
- Te same dane, inny wykres

Te same rezultaty – innym sposobem



- Co lepsze: kolumny czy mapa?

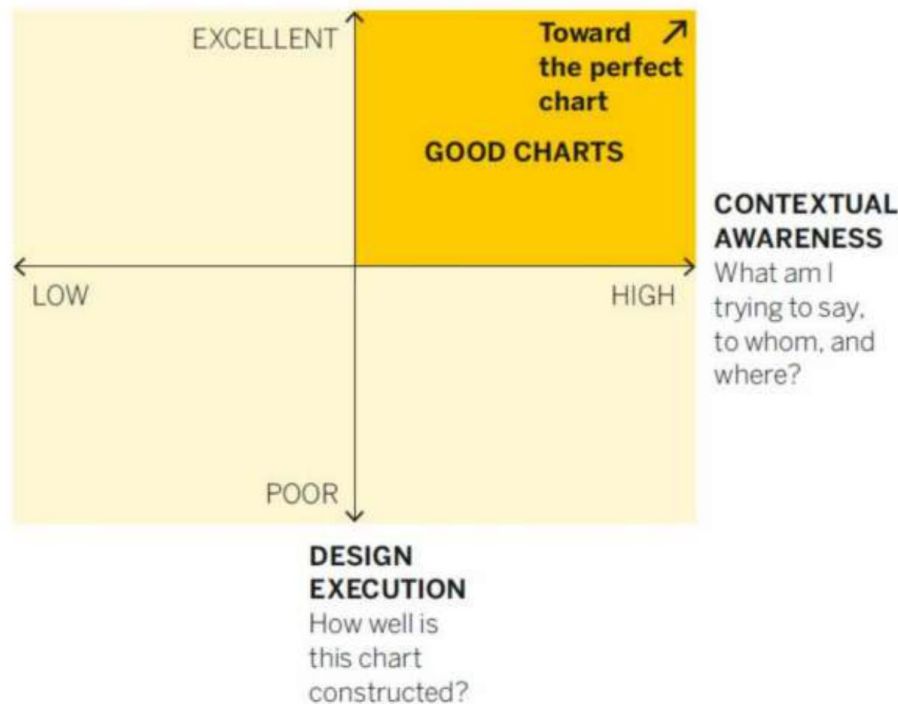
Czym jest zainteresowany nasz odbiorca?



- Jeśli w planie spotkania zarządu jest dyskusja nad strategią i scenariuszami to może lepiej jest pokazać je?
- Jaki jest cel pokazywania samej przeszłości jeśli będziemy rozmawiać o przyszłości?

Co sprawia, że wykres jest dobry?

THE GOOD CHARTS MATRIX

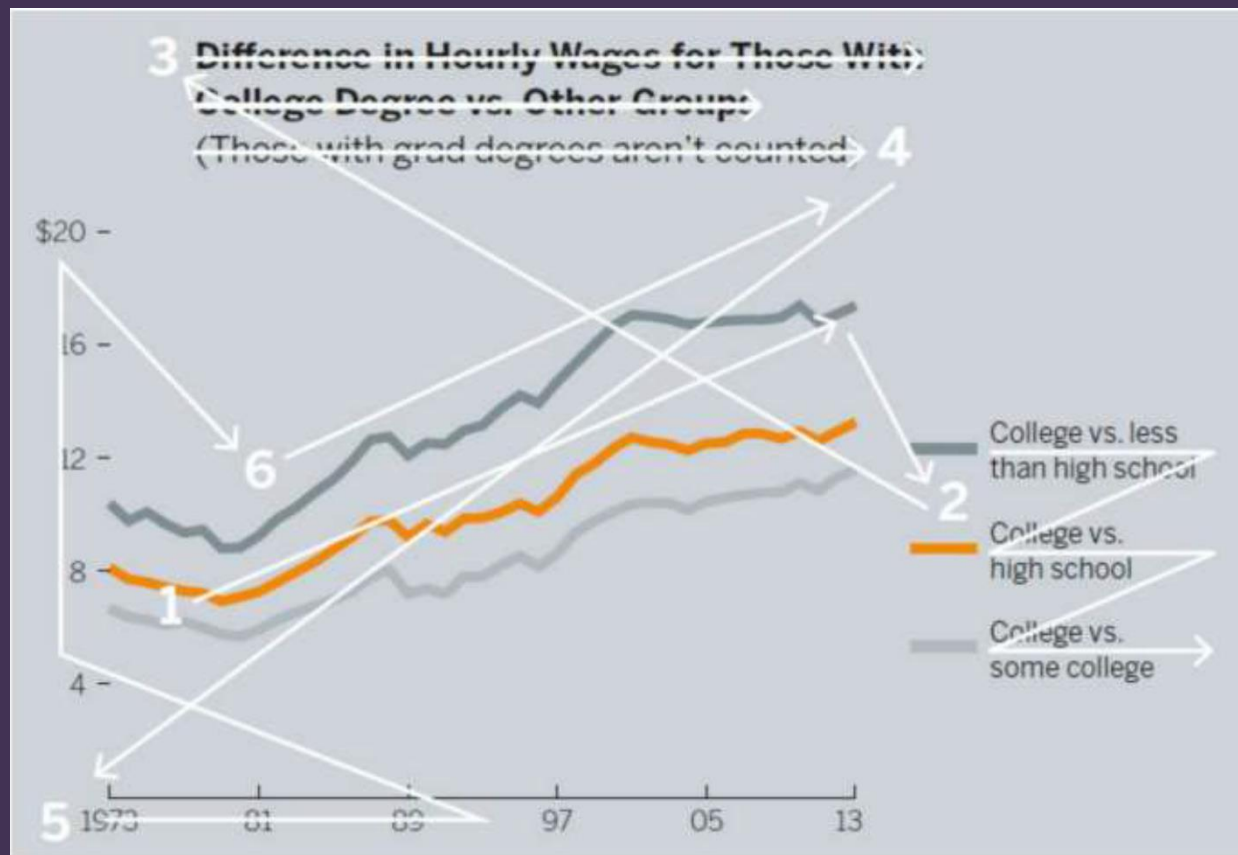


- Dobry wykres to taki, który jest prawidłowo skonstruowany i który doskonale przekazuje to, co chcemy zaprezentować (biorąc pod uwagę adresata i sposób, w jaki przekazywana jest jego główna idea)

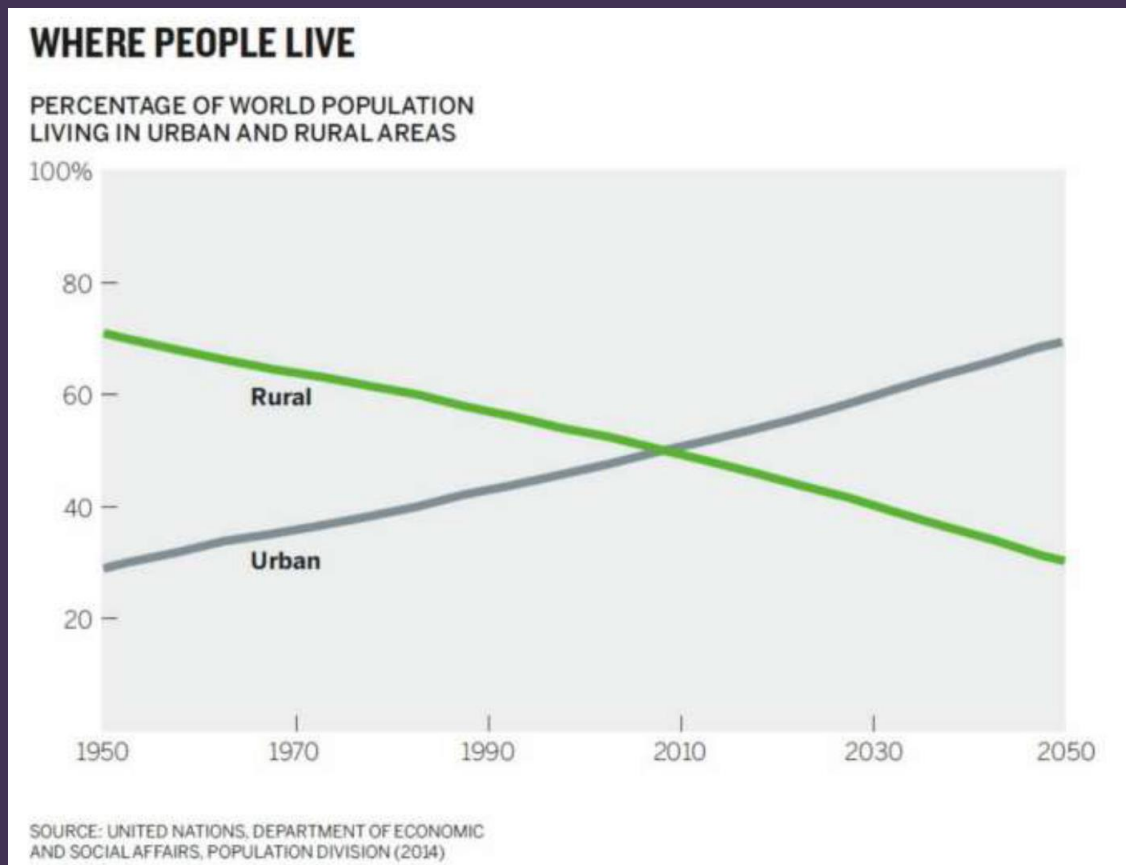
Jak ludzie czytają raporty?

- Nie ma określonej kolejności tak, jak w czytaniu tekstu
- Najpierw zwracamy uwagę na elementy wyróżnione
- Możemy skupiać uwagę tylko na kilku elementach jednocześnie
- Szukamy kontekstu/znaczenia
- Polegamy na przyzwyczajeniach

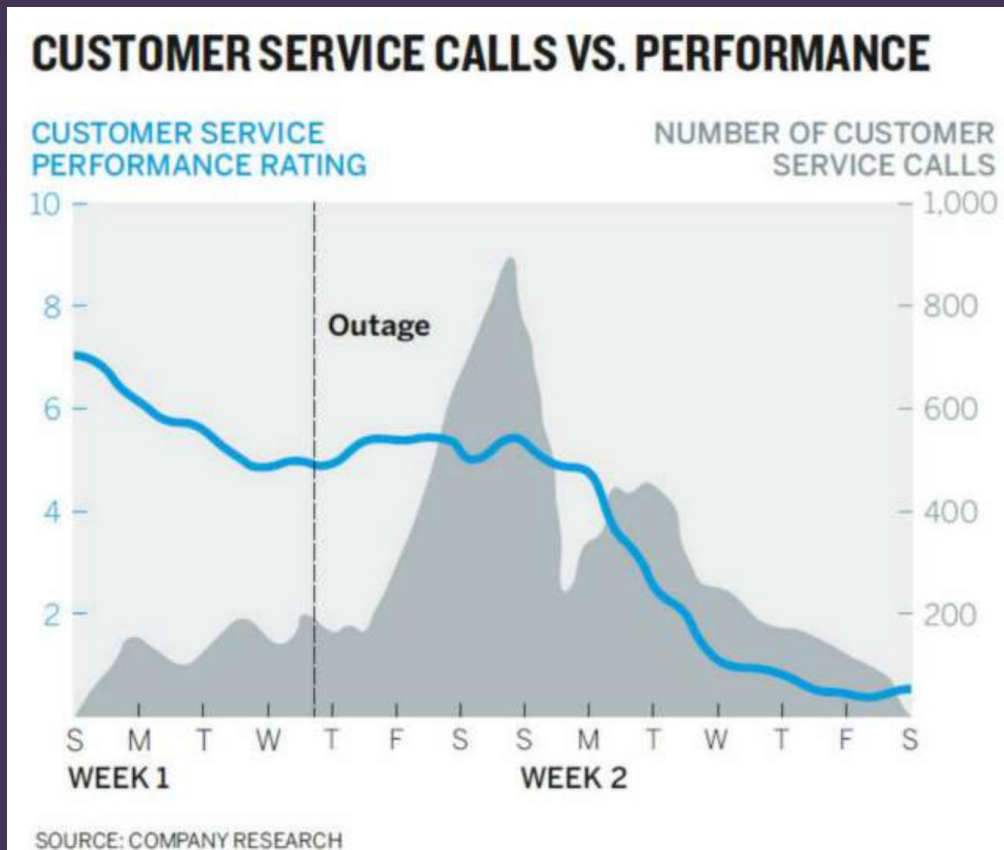
Kolejność „czytania”



Element wyróżniony

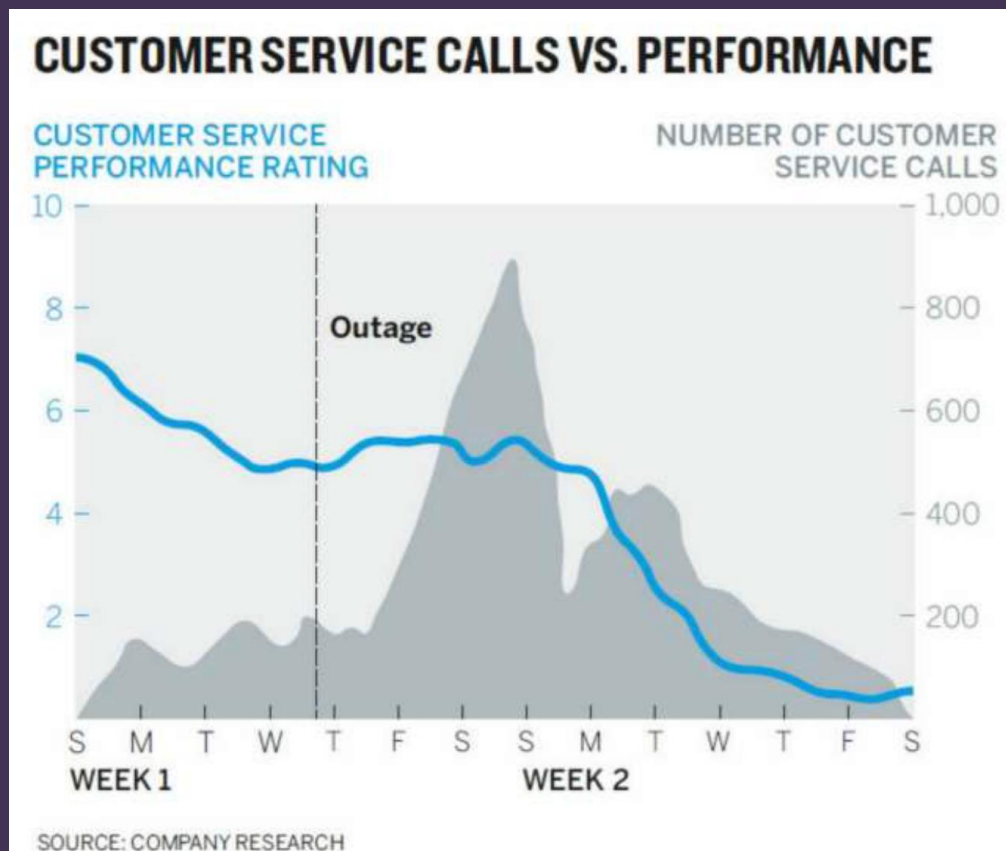


Czemu się przyglądamy?



- Jaką treść niesie ten wykres?
- Jaka historia stoi za tym wykresem?
- Co jeśli treścią jest fakt, iż nawet po spadku ilości zgłoszeń telefonicznych wydajność pracowników CC spadała?

Przekazywanie treści

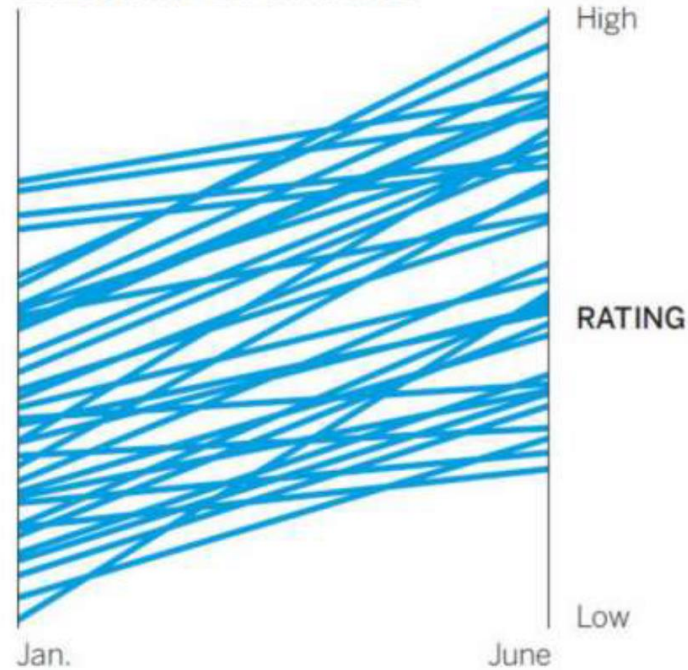


- Co jeśli treścią jest fakt, iż nawet po spadku ilości zgłoszeń telefonicznych wydajność pracowników CC spadała?
- Co więcej, zwróćmy uwagę na to kiedy zaczął się ten spadek
- Pomyśl o Twoim kierowniku...

Obserwujemy tylko kilka elementów jednocześnie

TEAM PERFORMANCE

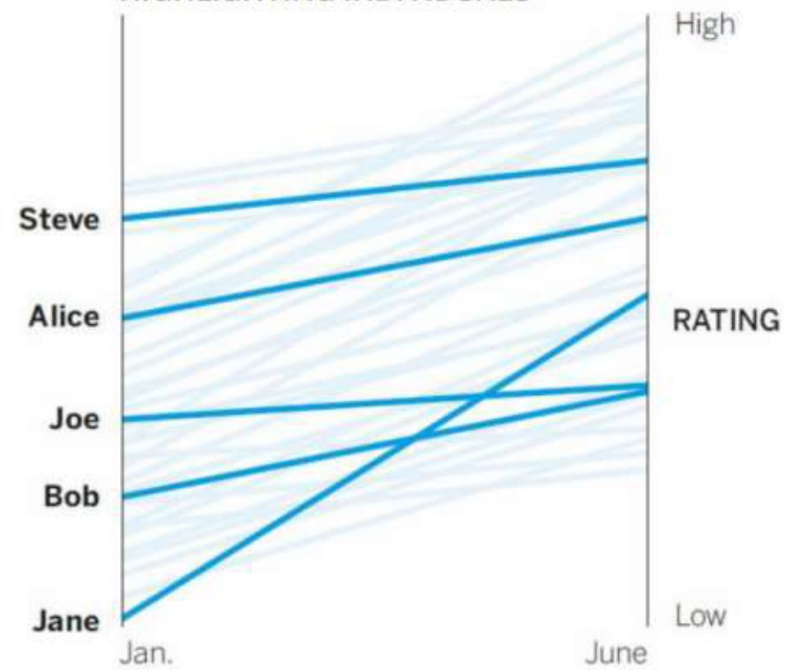
BLUE TEAM PERFORMANCE



SOURCE: COMPANY RESEARCH

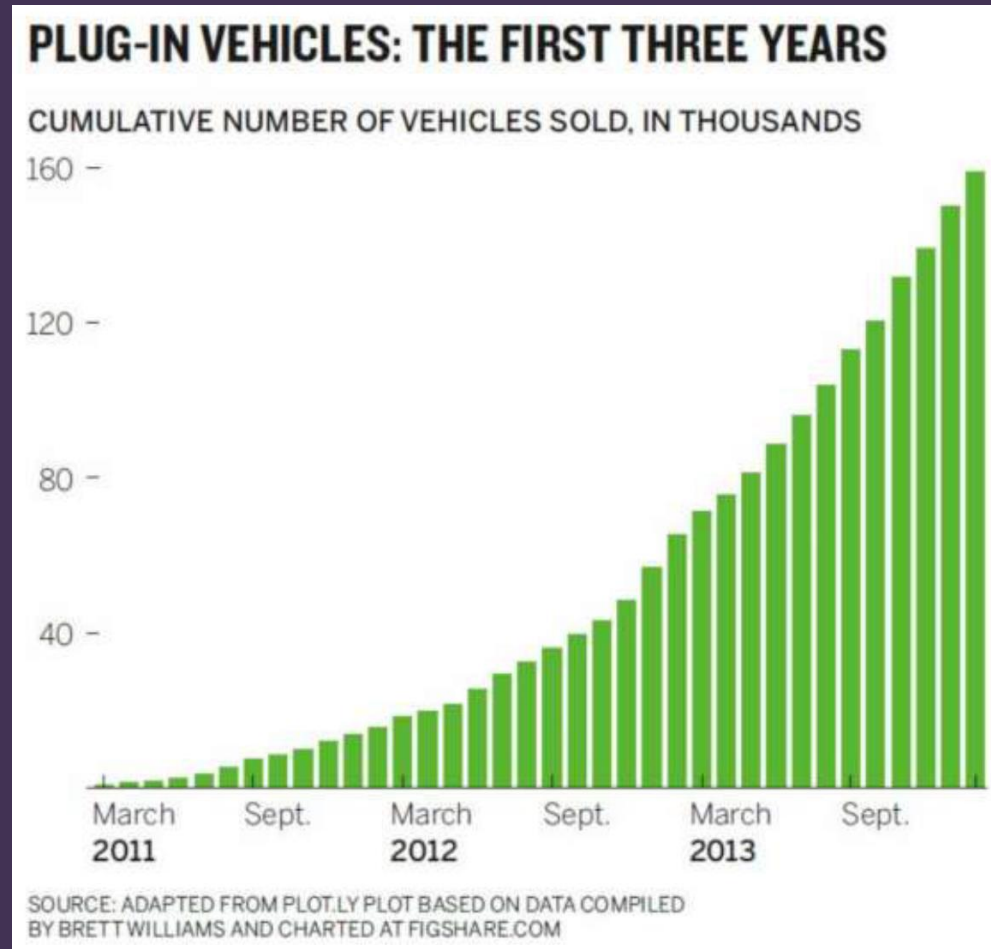
TEAM PERFORMANCE

HIGHLIGHTING INDIVIDUALS



SOURCE: COMPANY RESEARCH

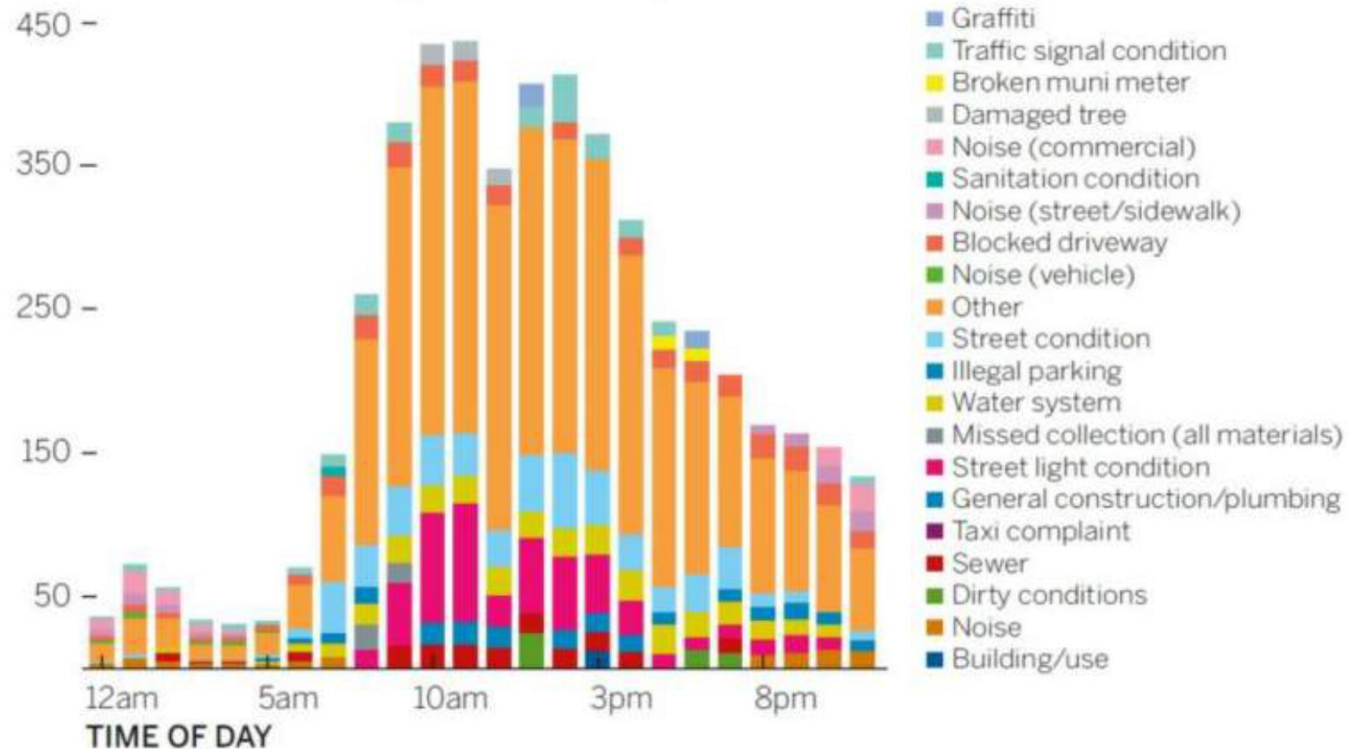
Nie skupiamy się na pojedynczych wartościach



Może trochę tego zbyt dużo...

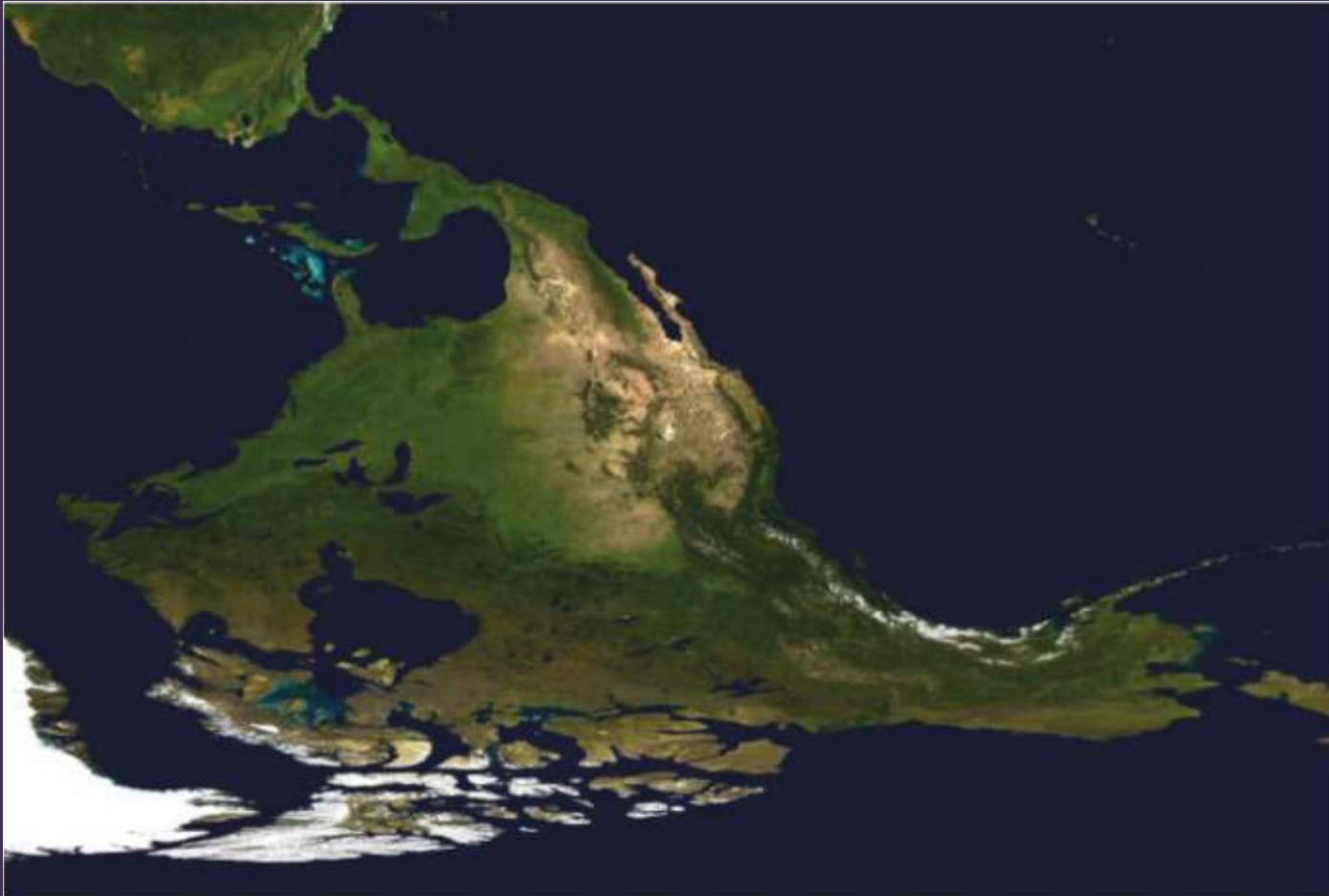
THE MOST COMMON 311 COMPLAINTS IN NYC

NUMBER OF COMPLAINTS (IN THOUSANDS)

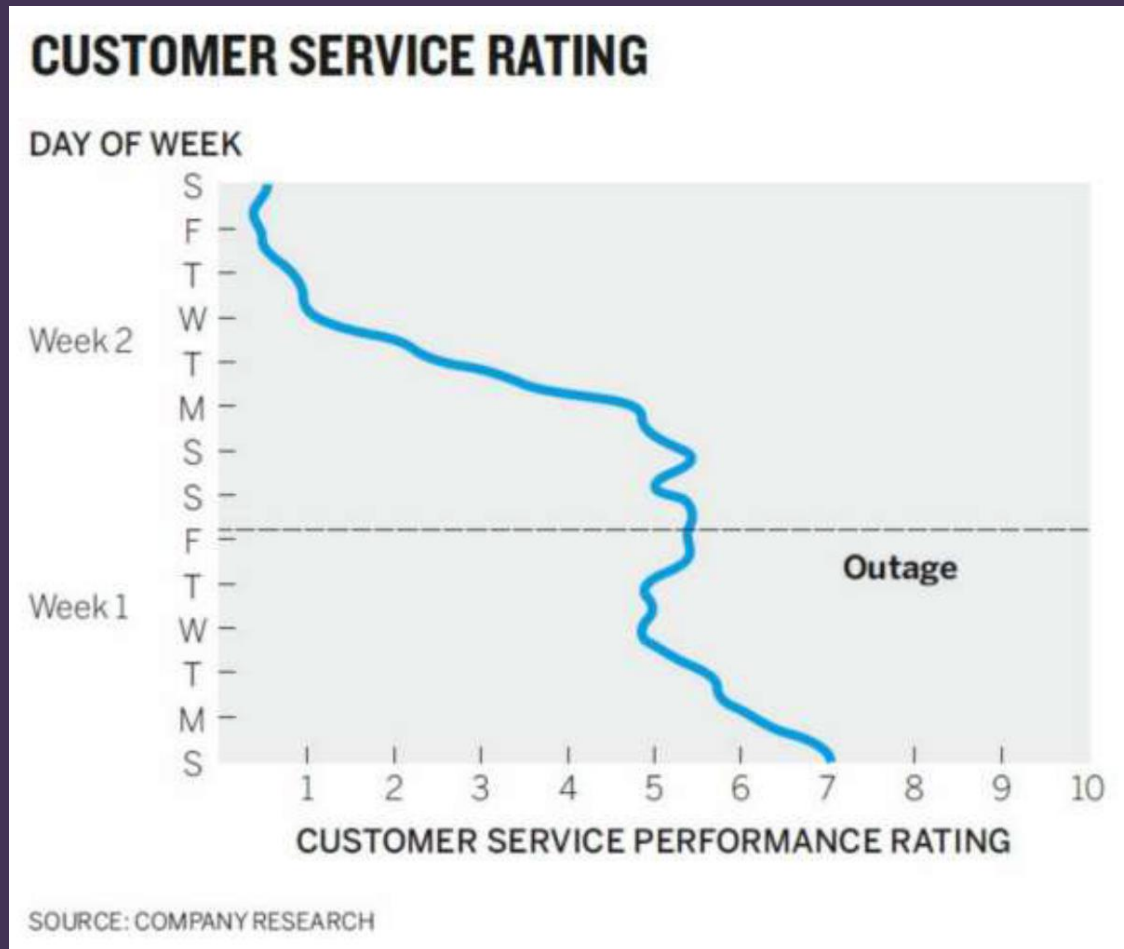


SOURCE: PLOT.LY

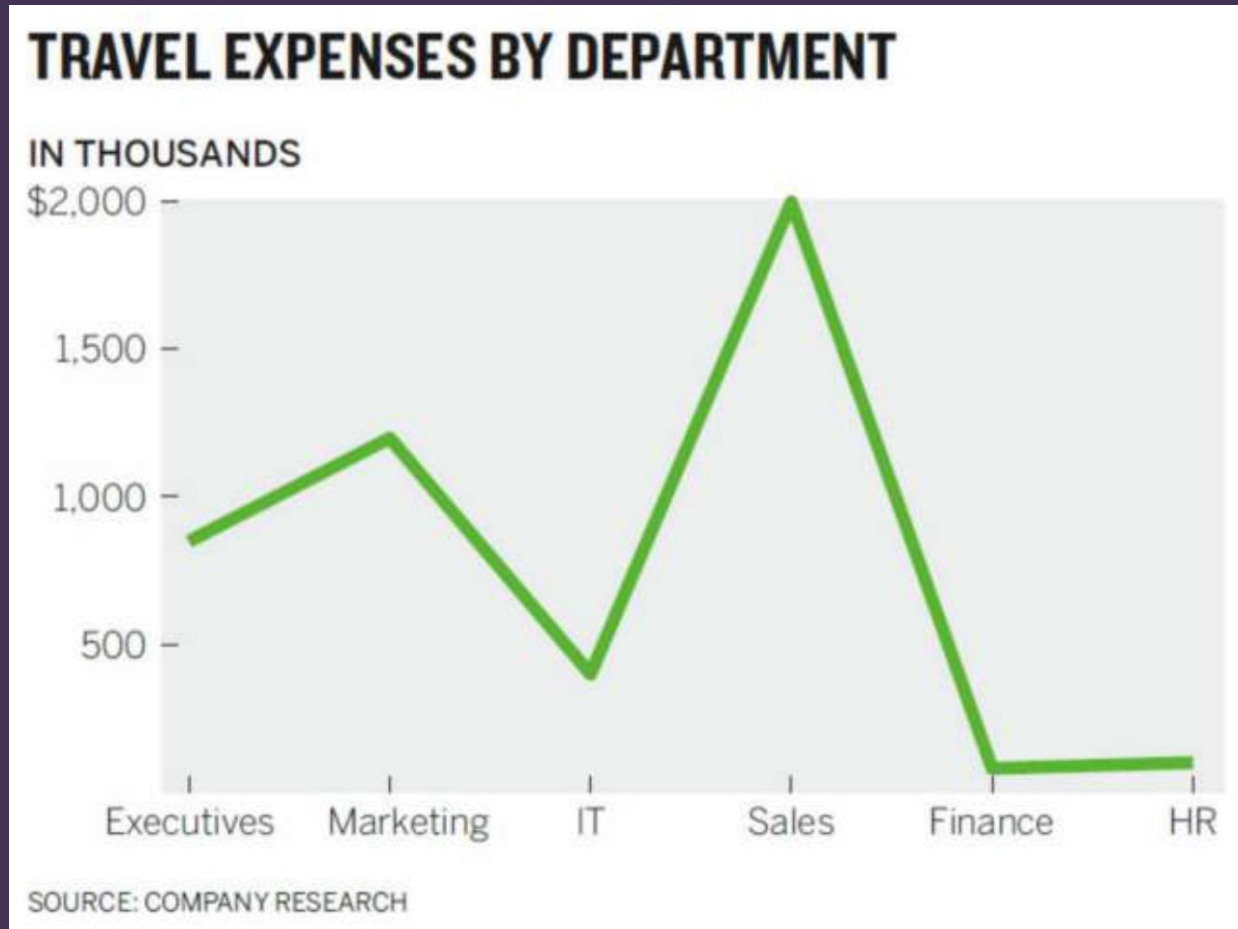
Polegamy na przyzwyczajeniach i metaforach



Polegamy na przyzwyczajeniach i metaforach

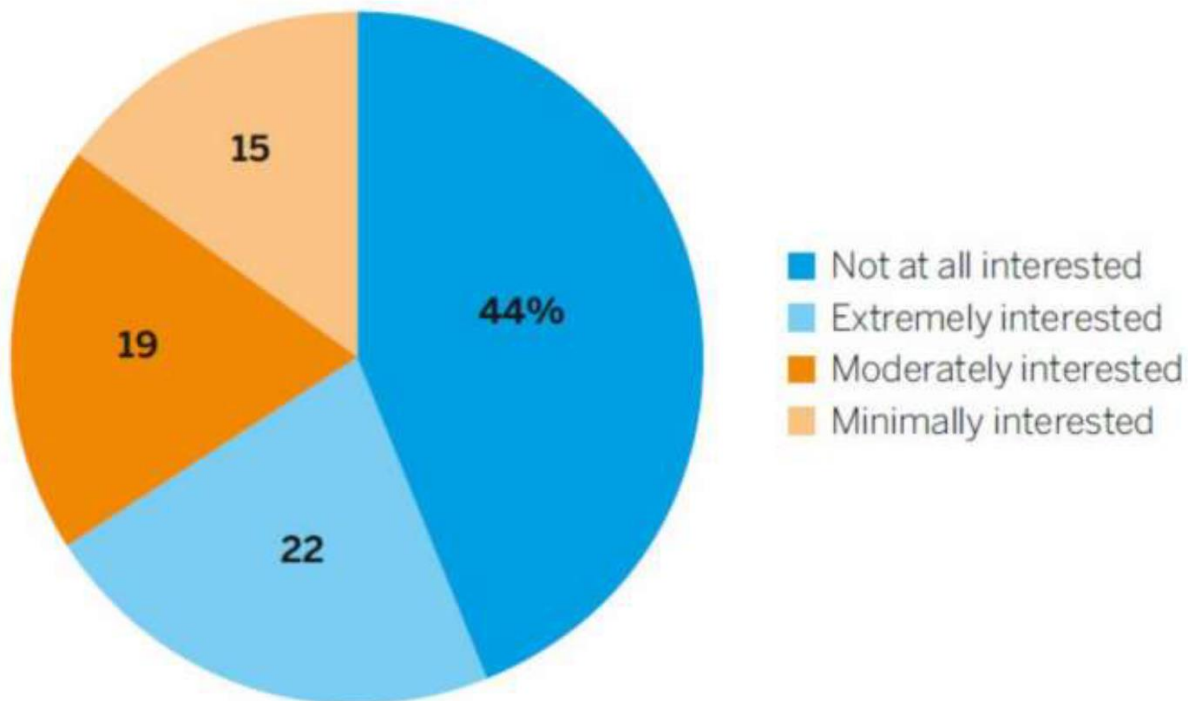


Polegamy na przyzwyczajeniach i metaforach



Siła przyzwyczajenia

HOW INTERESTED ARE YOU IN THIS PRODUCT?



SOURCE: COMPANY RESEARCH

Zasady

- Podobne kolory dla podobnych elementów
- Nasycenie koloru odpowiada wyższym i niższym wartościom – jaśniejsze odcienie przeznaczone są dla mniejszych wartości
- Kategorie powinny być posortowane

HOW INTERESTED ARE YOU IN THIS PRODUCT?



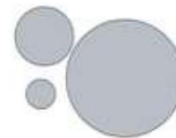
SOURCE: COMPANY RESEARCH



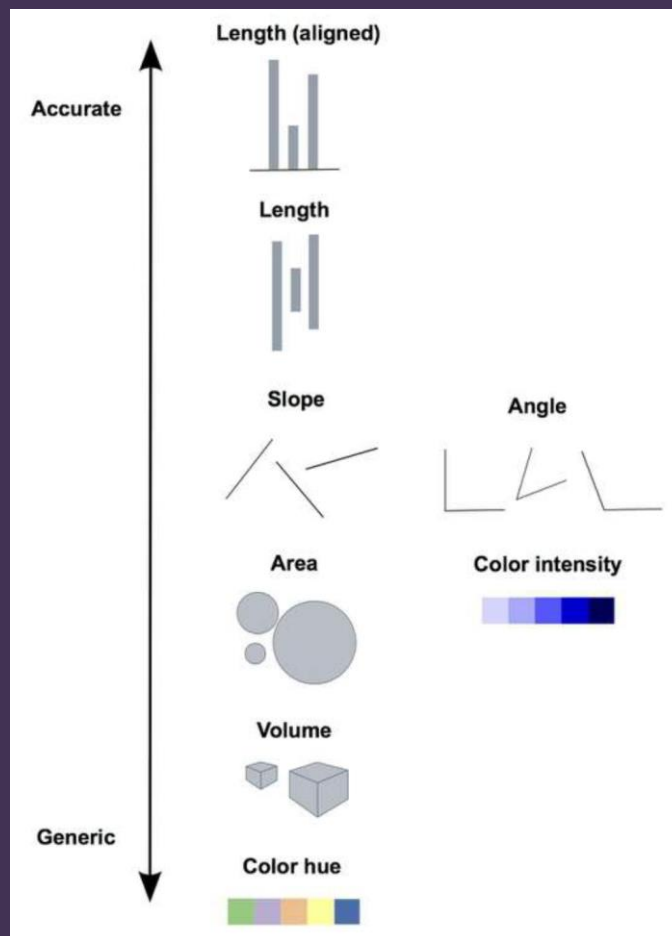
Postrzeganie elementów raportów

PERCEPCJA

Kodowanie danych wizualizacjami

Length**Slope****Angle****Length (aligned)****Color hue****Volume****Area****Color intensity**

Precyzja percepcji

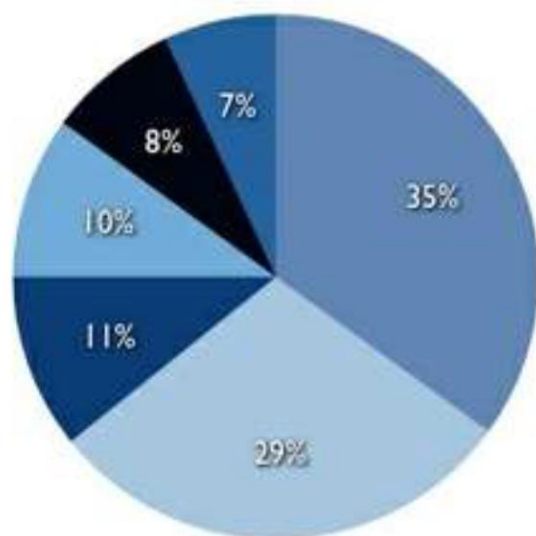


Porównaj dwa wykresy

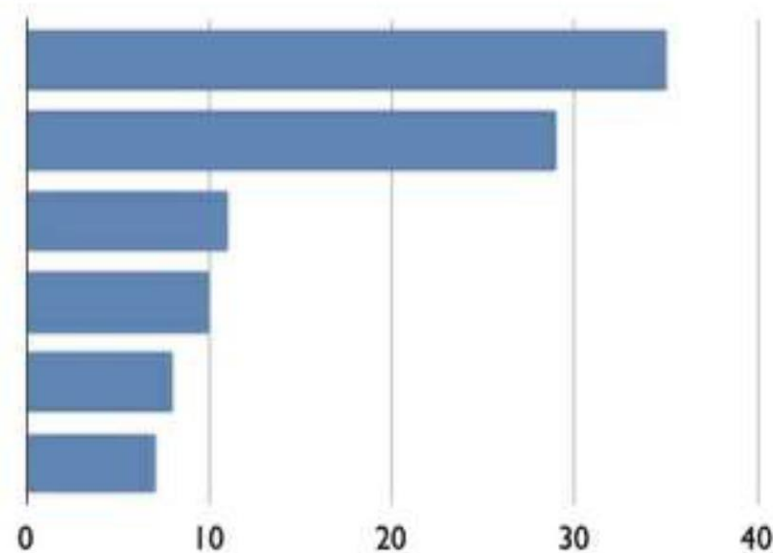


Porównaj dwa wykresy

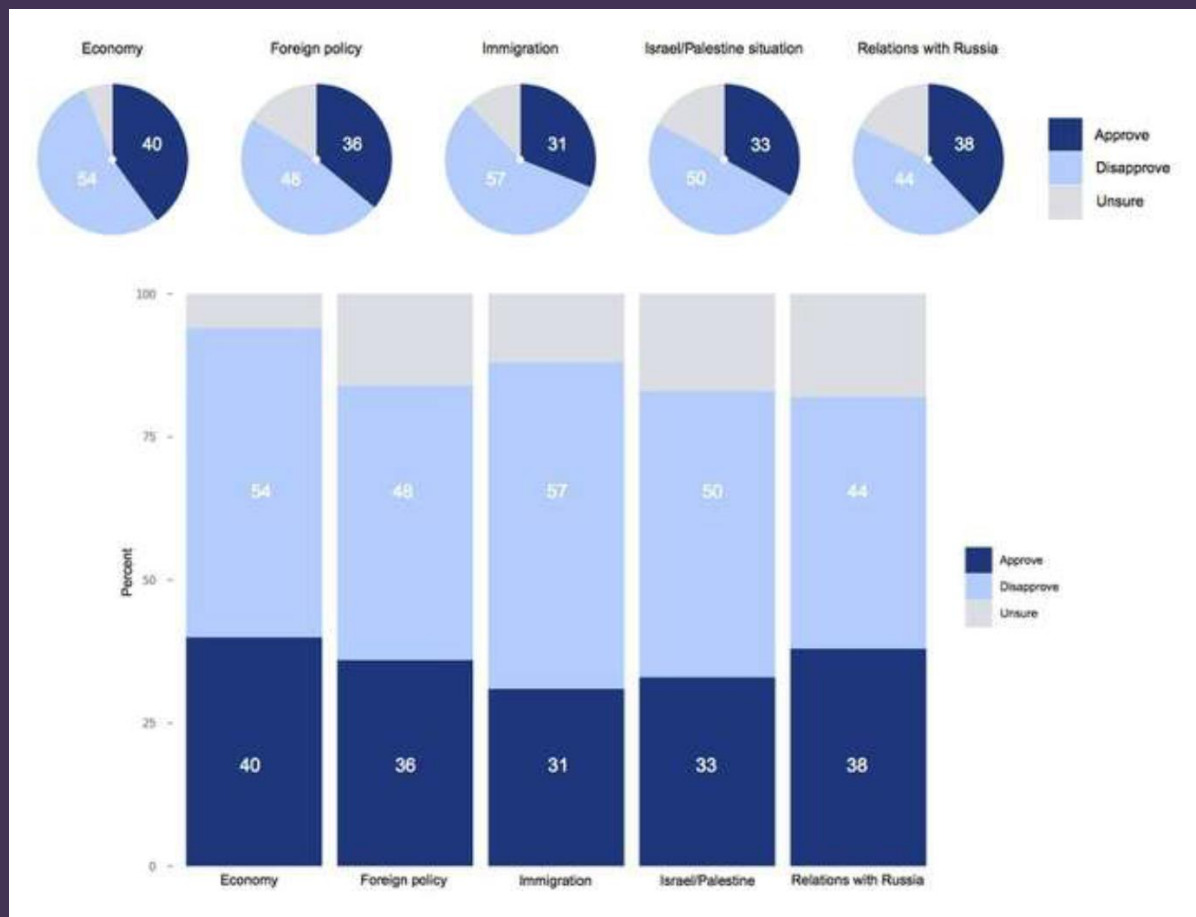
Grudge Match



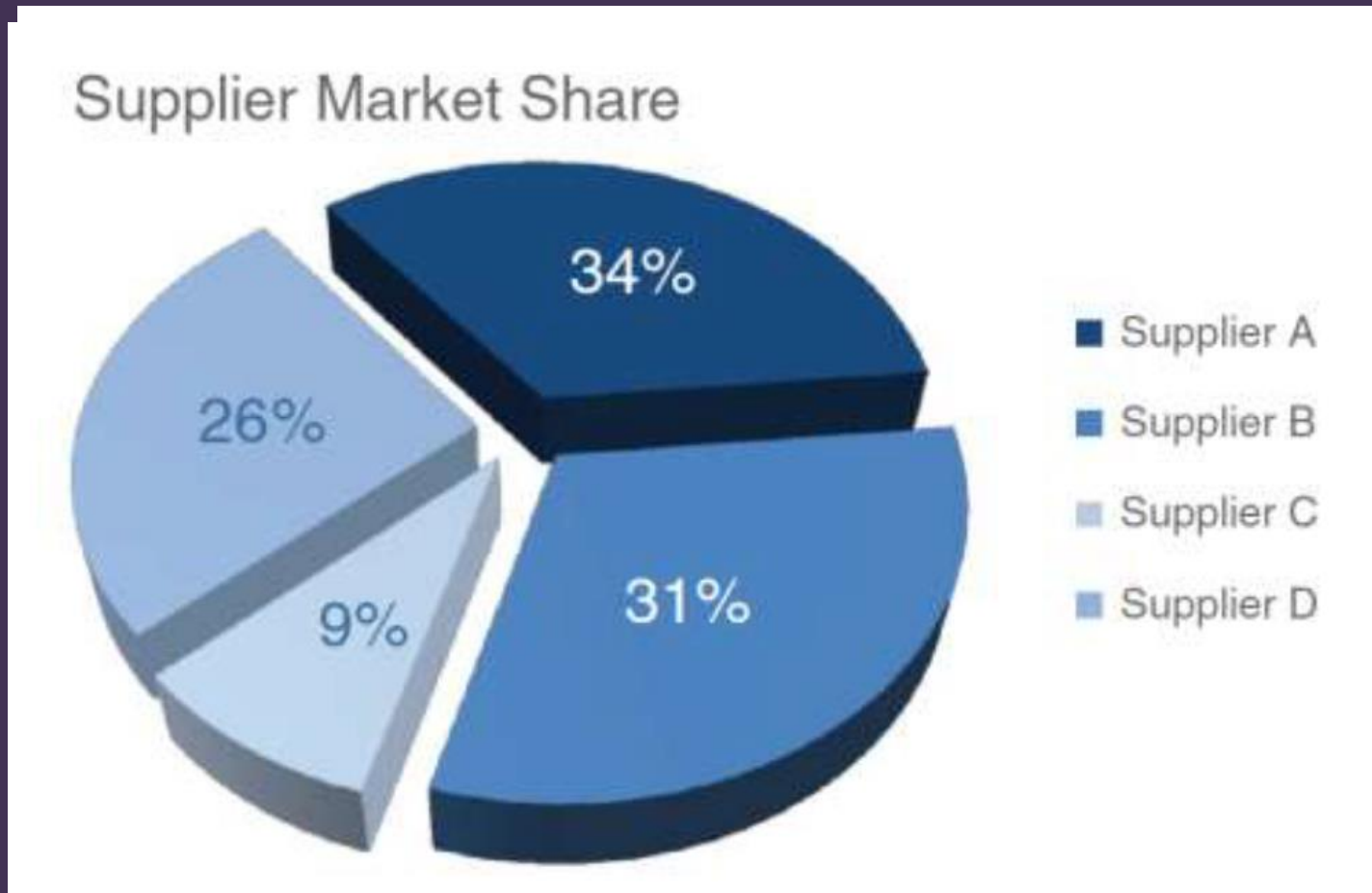
VS.



Porównaj dwa wykresy

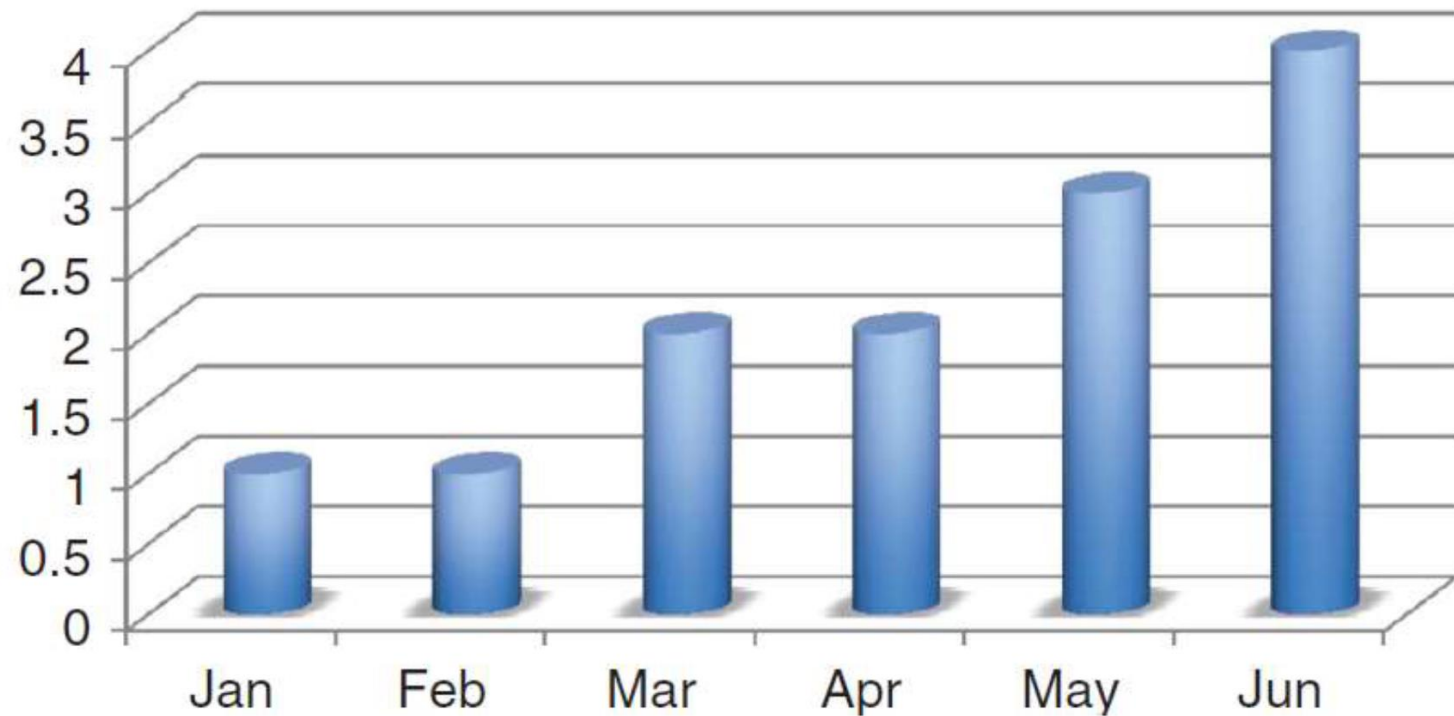


Unikaj wykresów kołowych



Unikaj 3D

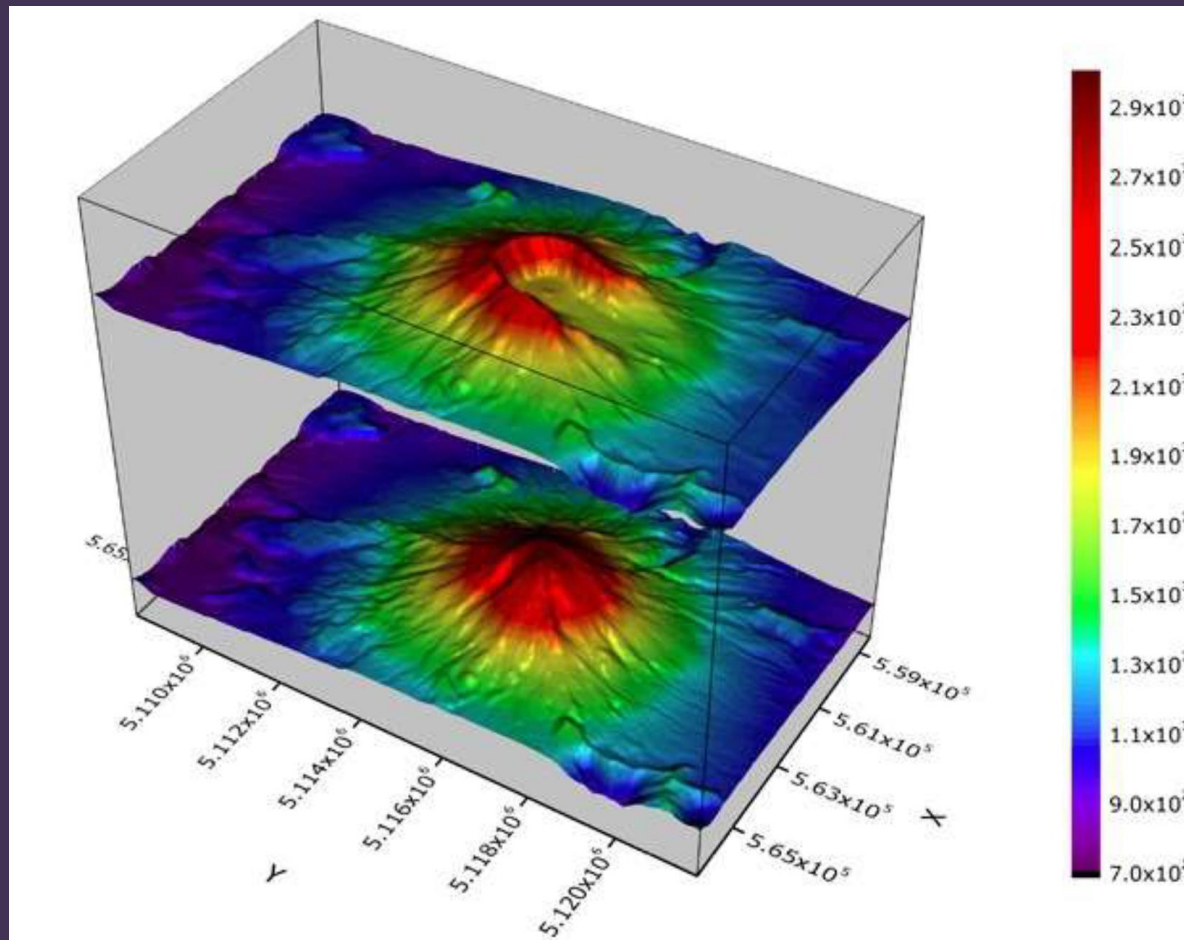
Number of issues



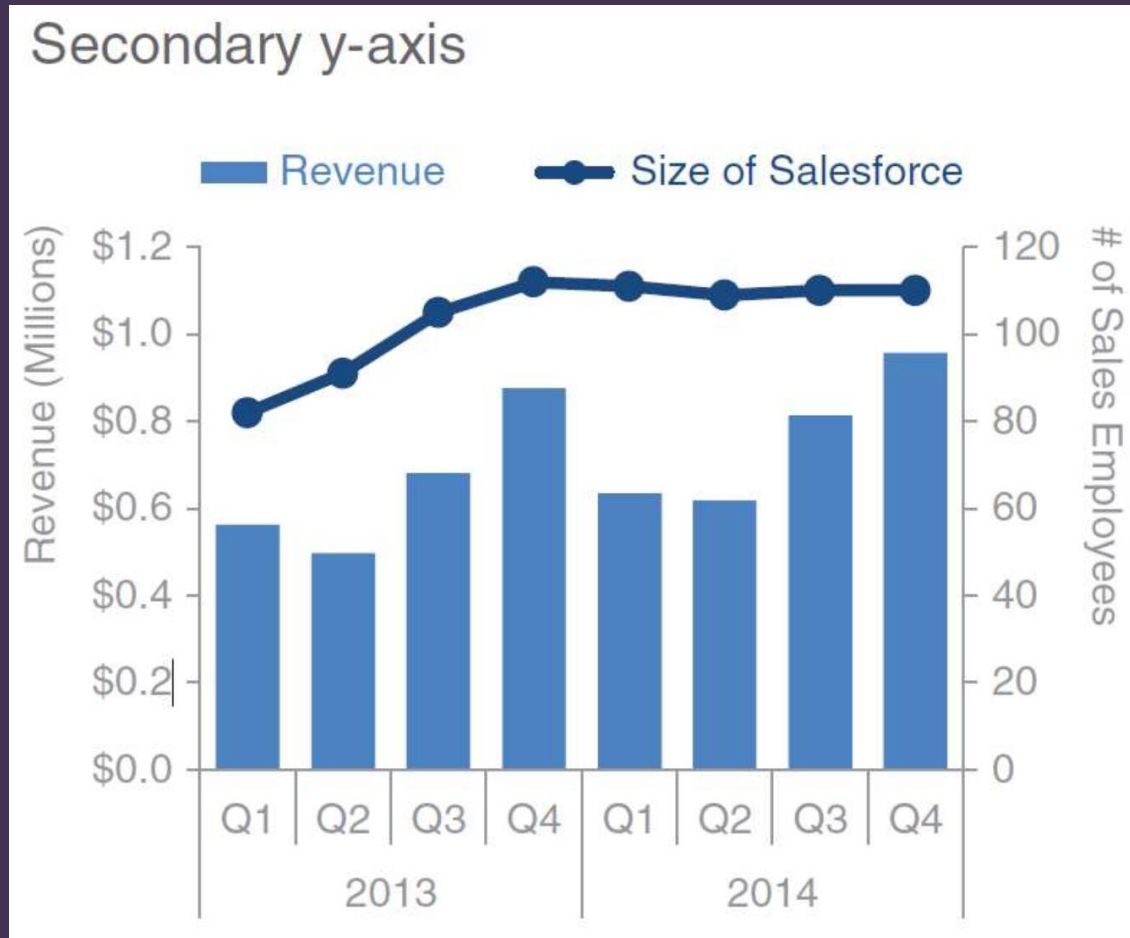
Wykresy 3D to samo zło...



...ale nie zawsze



Unikaj dodatkowych osi



Jak to poprawić?

Alternative 1: label directly



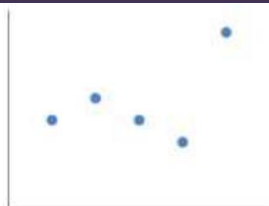
Alternative 2: pull apart vertically



Elementy wizualizacji danych

91%

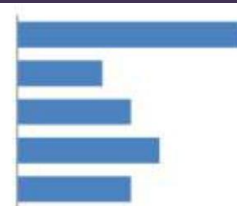
Simple text



Scatterplot



Vertical bar



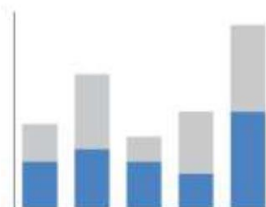
Horizontal bar

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

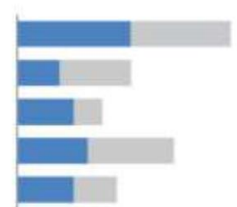
Table



Line



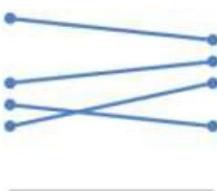
Stacked vertical bar



Stacked horizontal bar

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

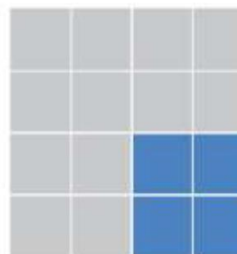
Heatmap



Slopegraph

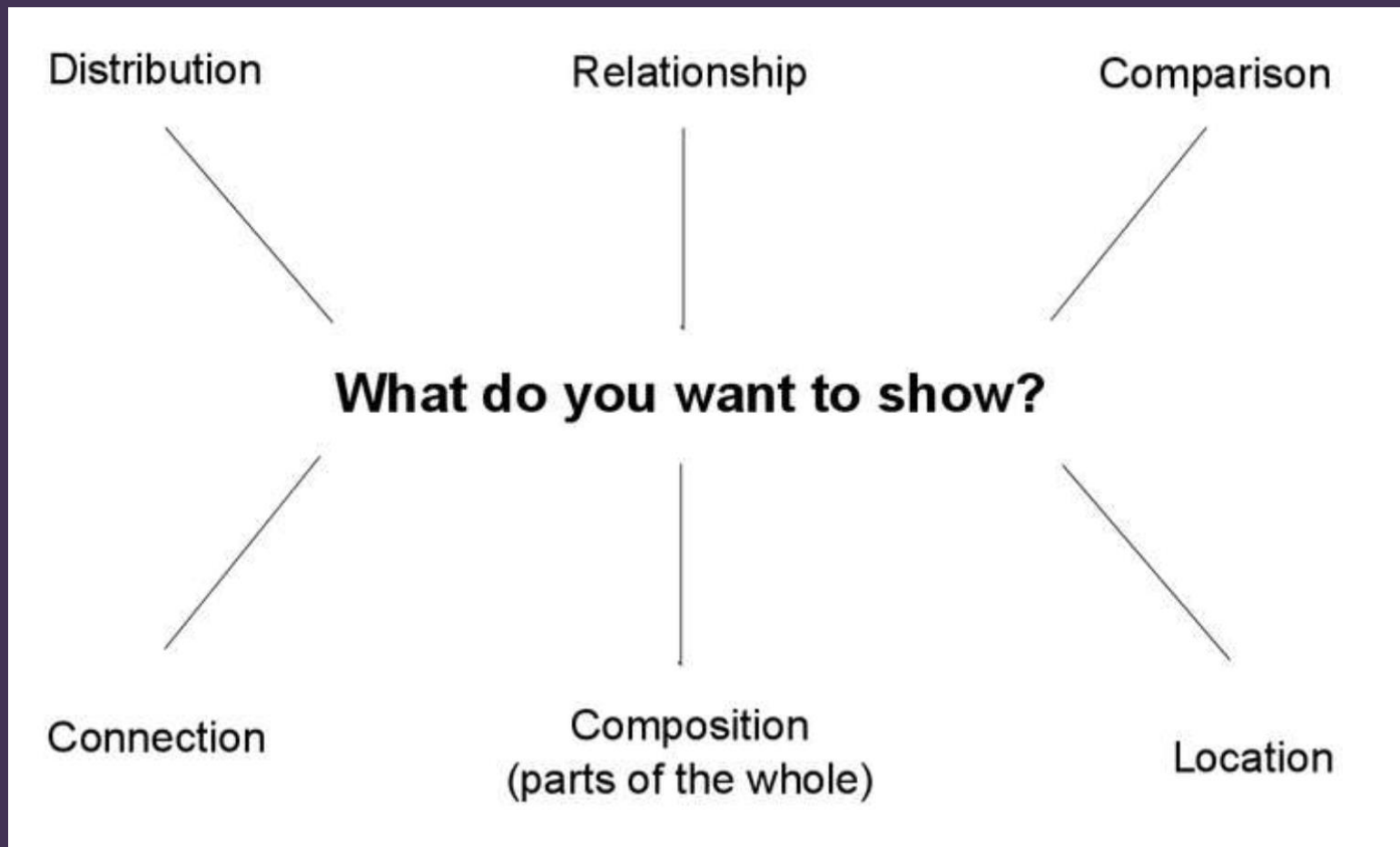


Waterfall

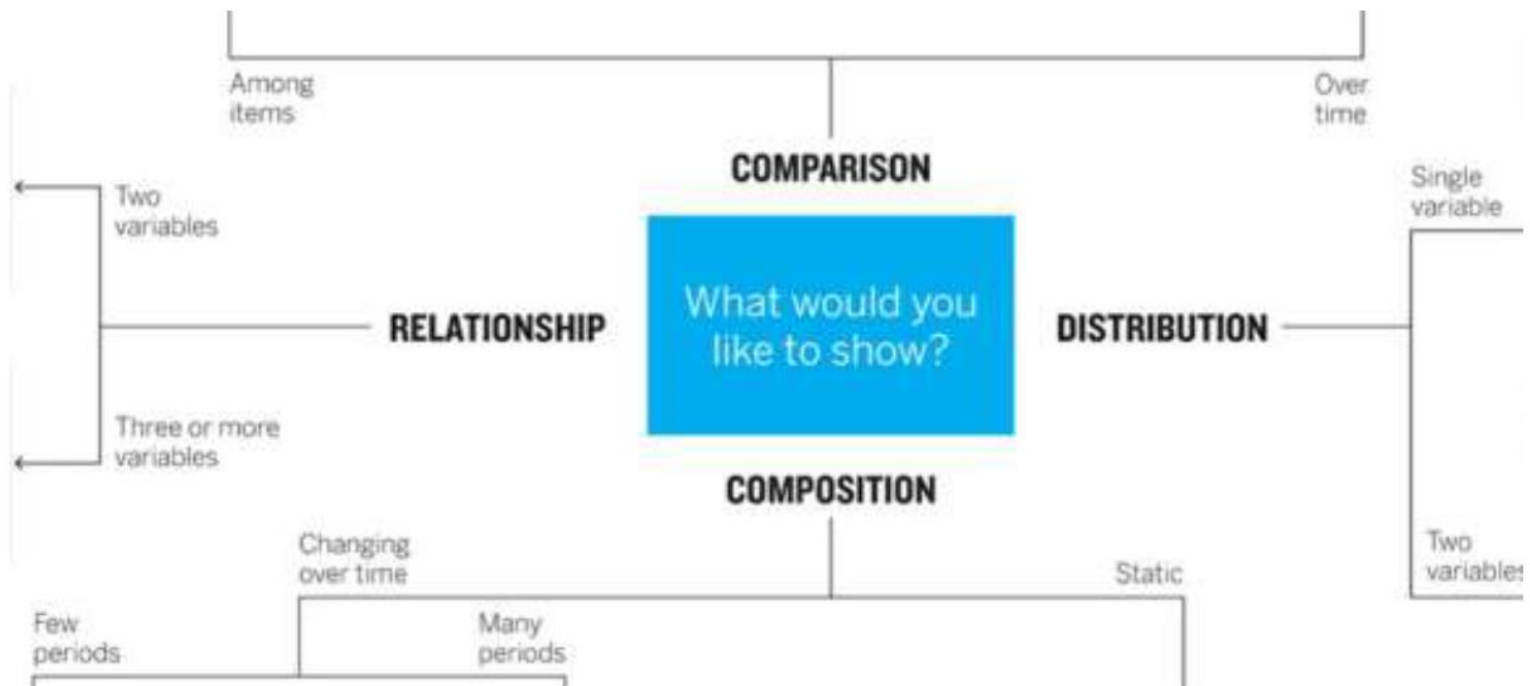


Square area

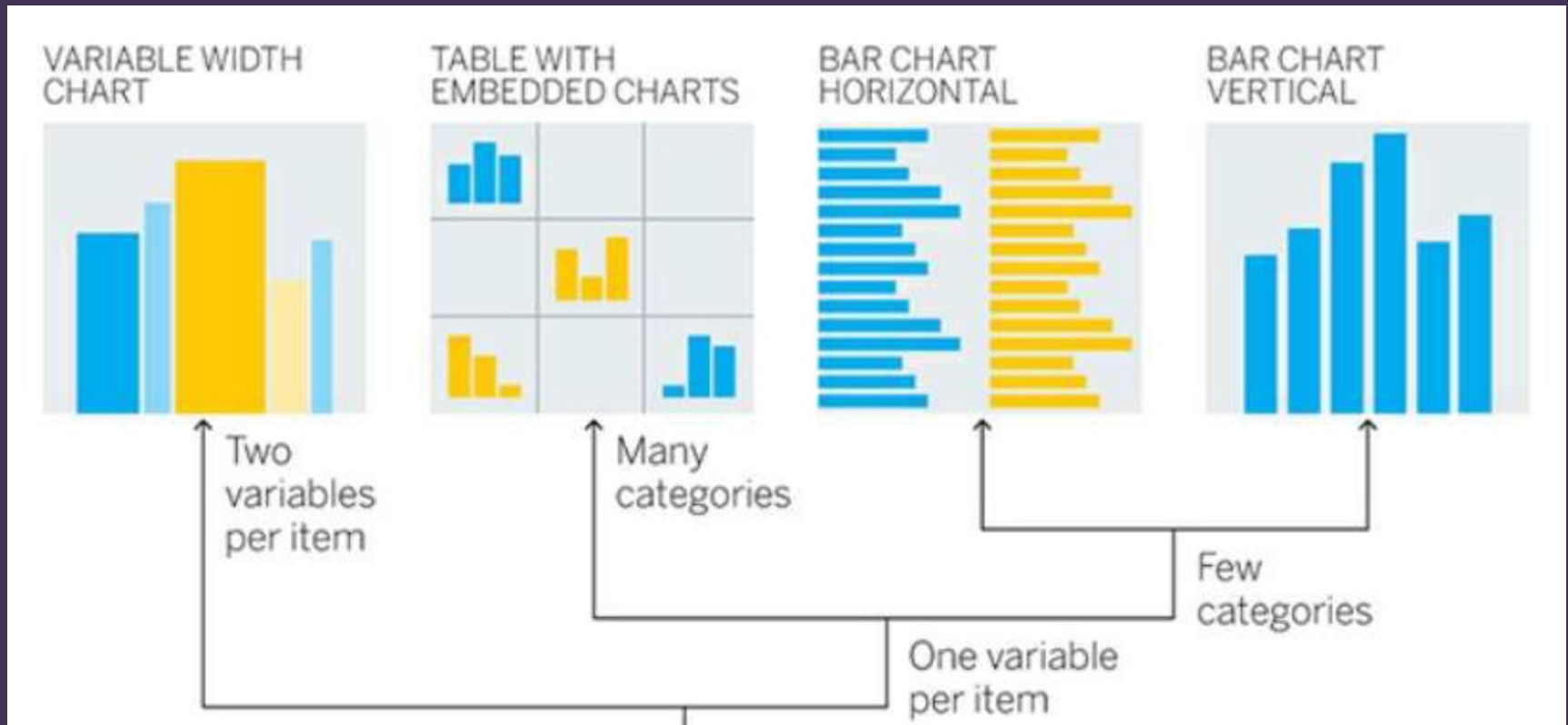
Co chcesz zaprezentować?



Co chcesz zaprezentować?

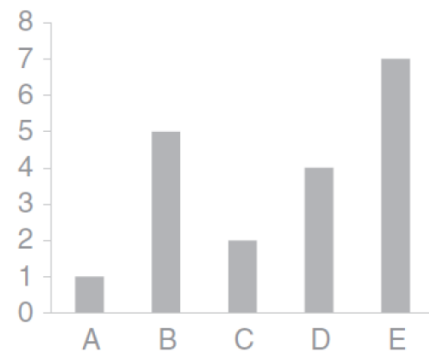


Porównanie pomiędzy elementami

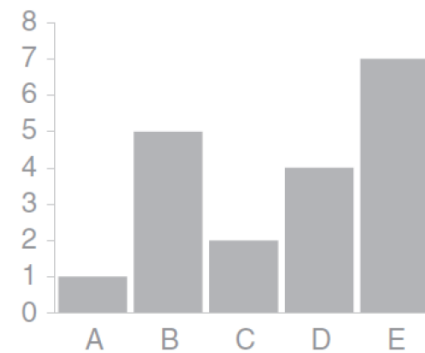


Wykresy kolumnowe

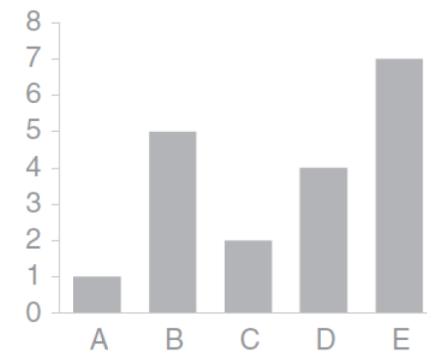
Too thin



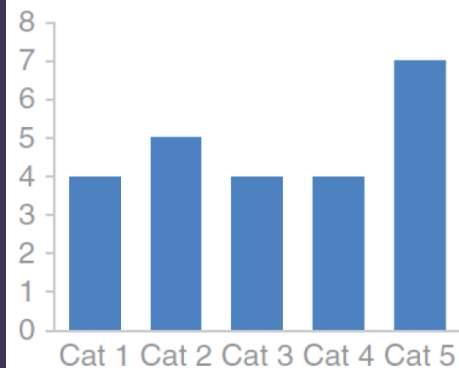
Too thick



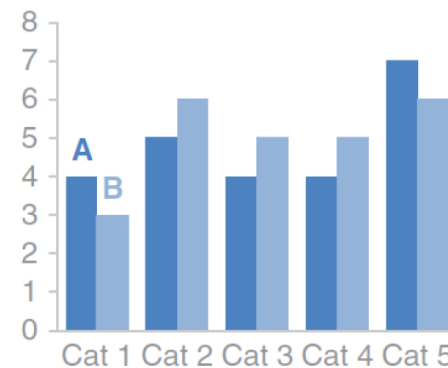
Just right



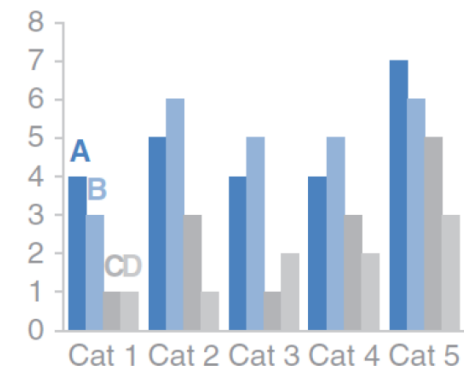
Single series



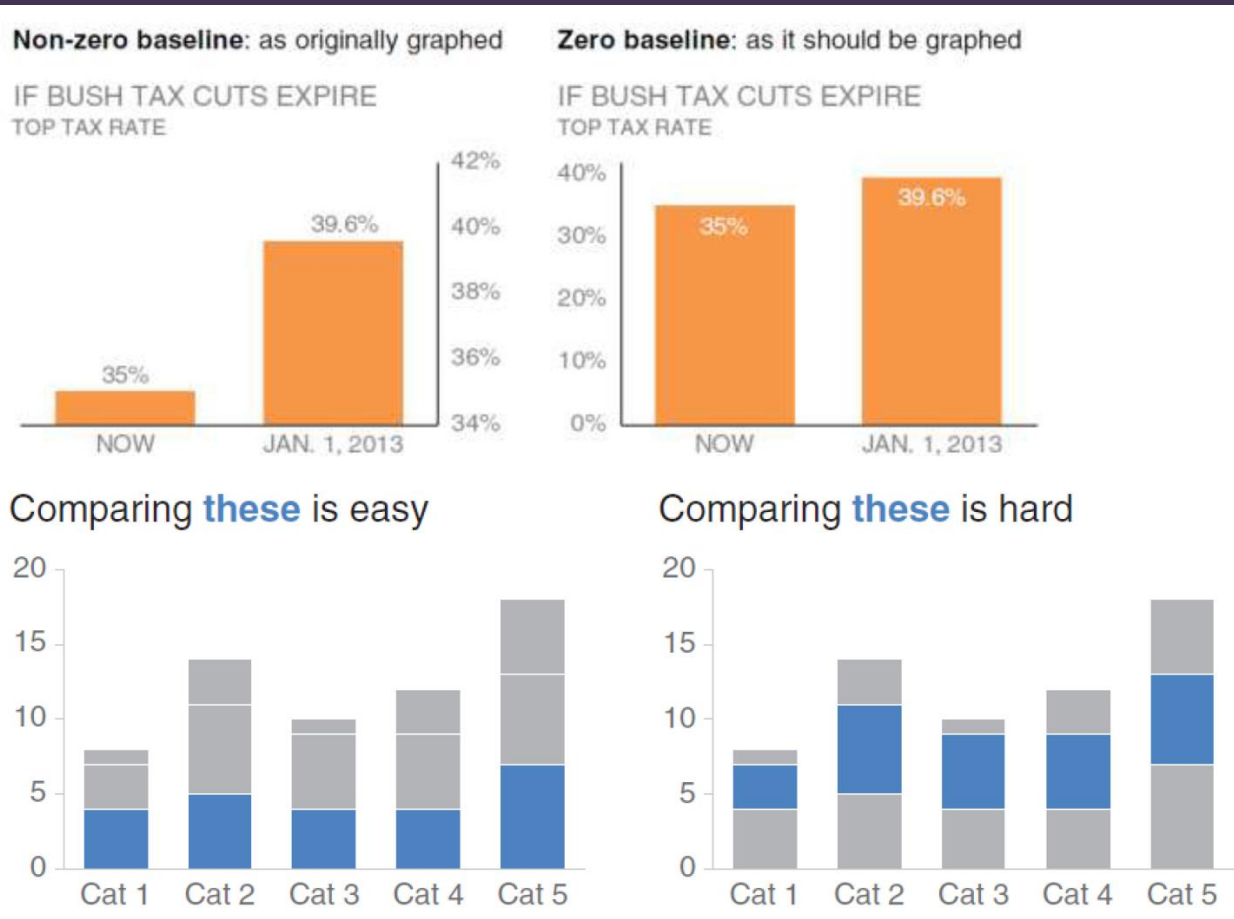
Two series



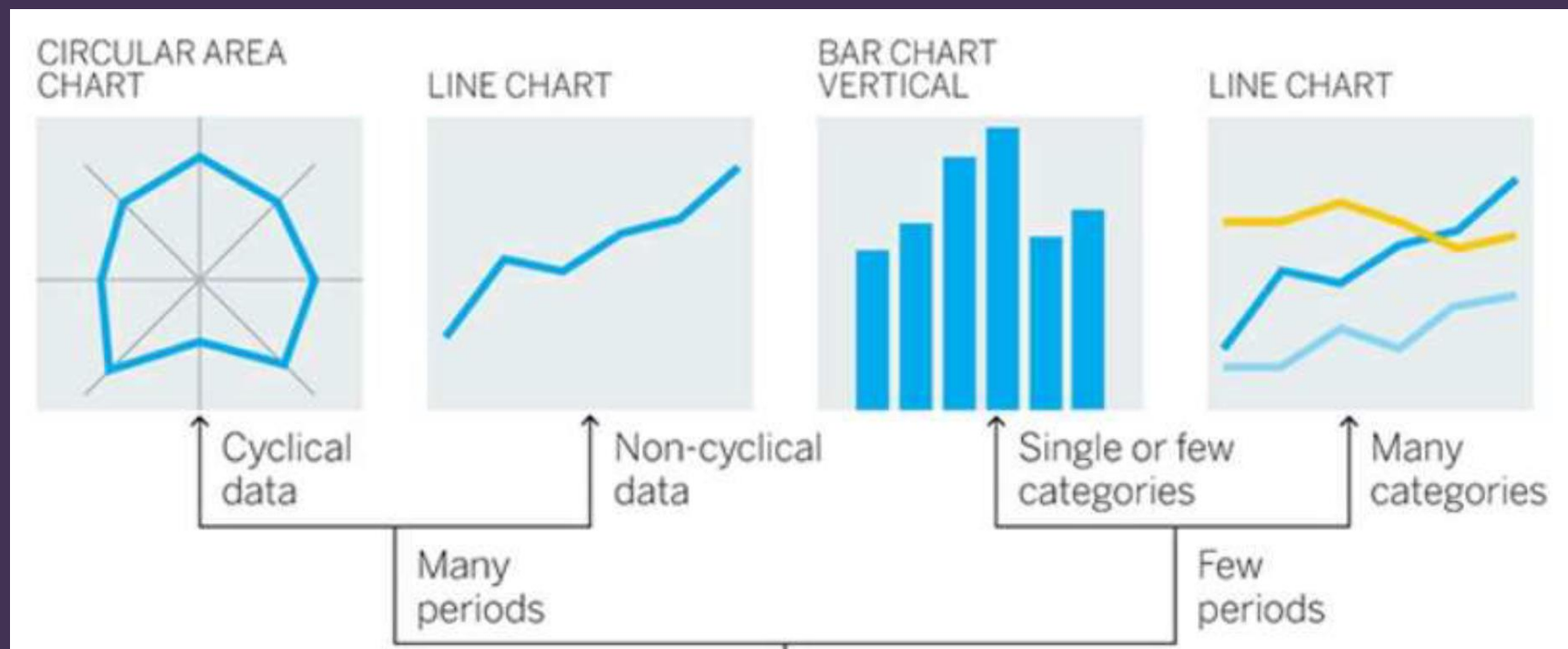
Multiple series



Wykresy kolumnowe (cd)

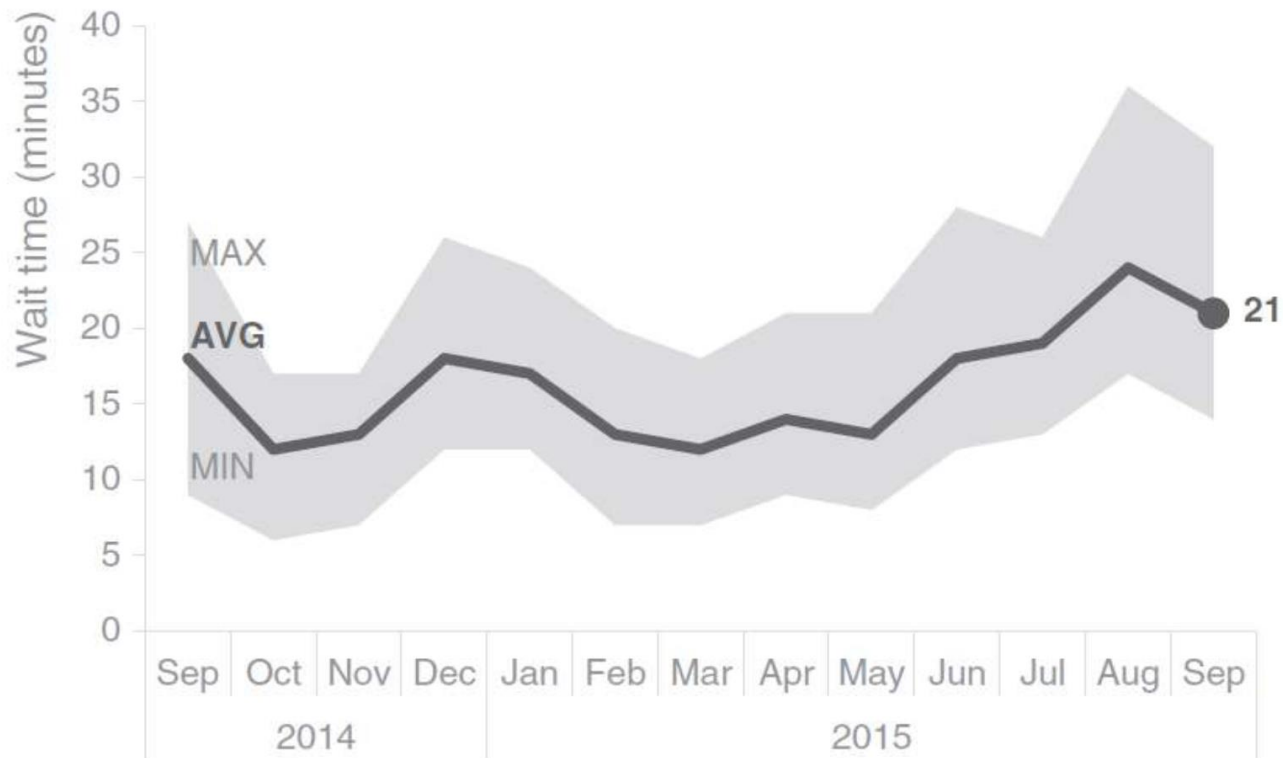


Porównanie w czasie

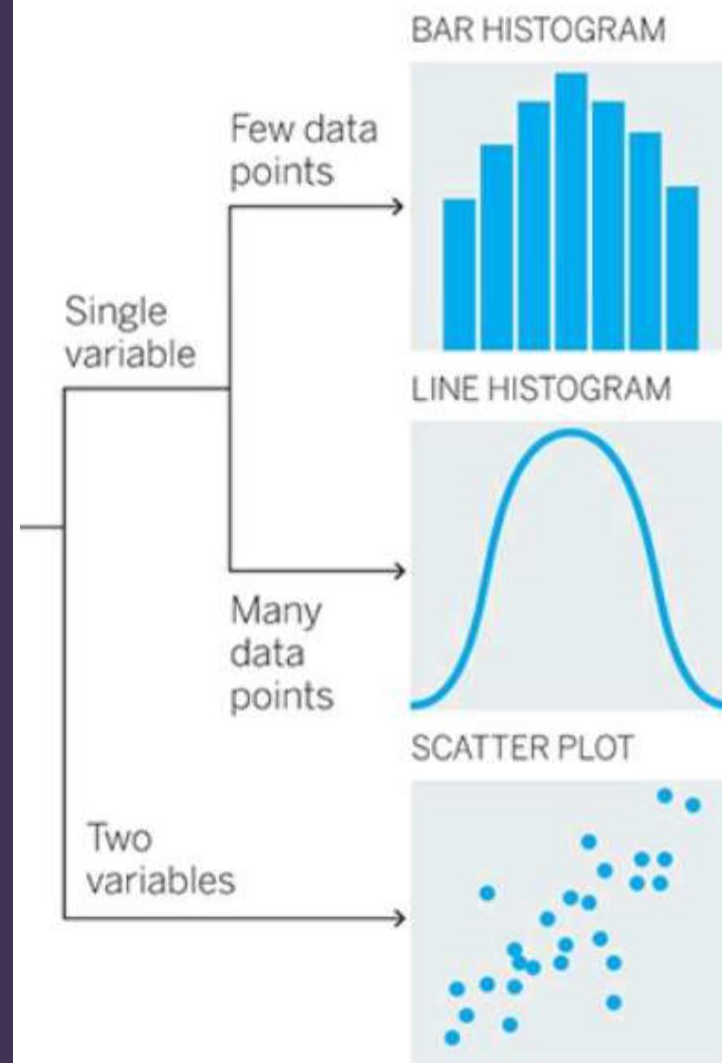


Wykres liniowy

Passport control wait time
Past 13 months



Rozkład



Zależności

SCATTER PLOT



Two
variables

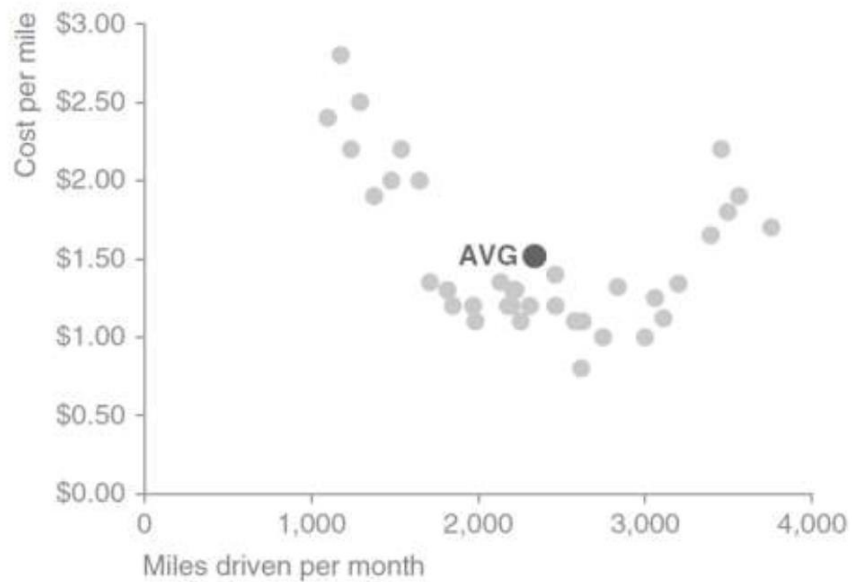
SCATTER PLOT
BUBBLE SIZE



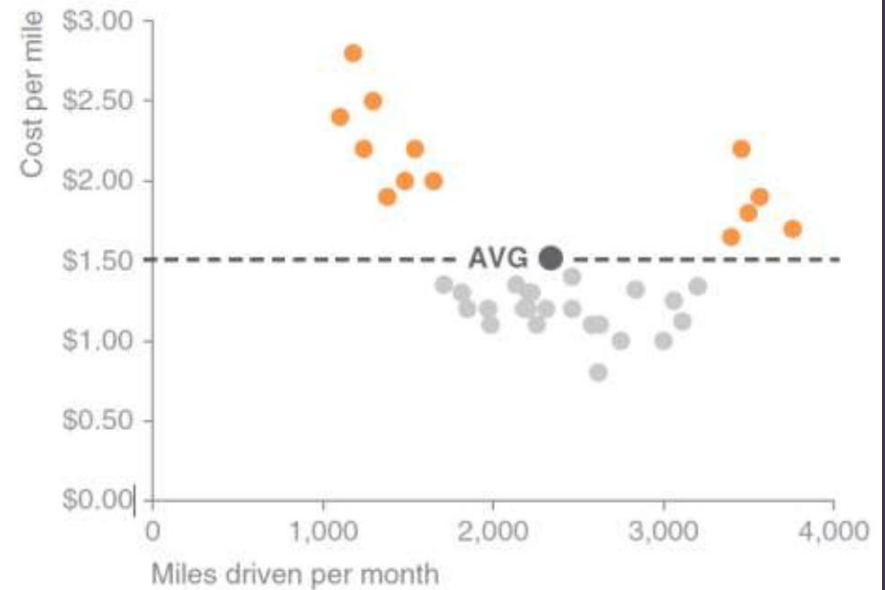
Three or more
variables

Wykres punktowy

Cost per mile by miles driven



Cost per mile by miles driven



Kompozycja (przebieg w czasie)

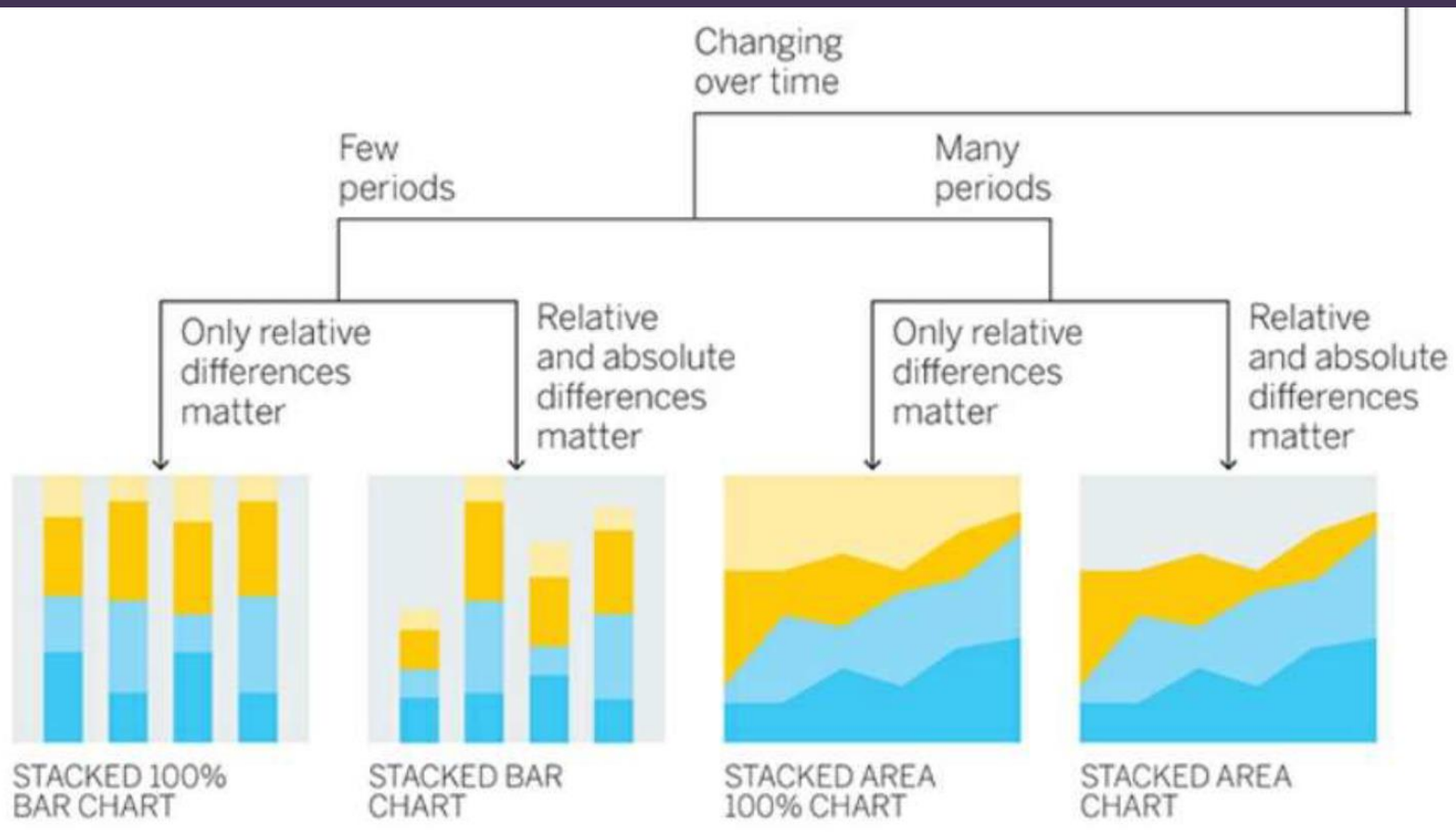
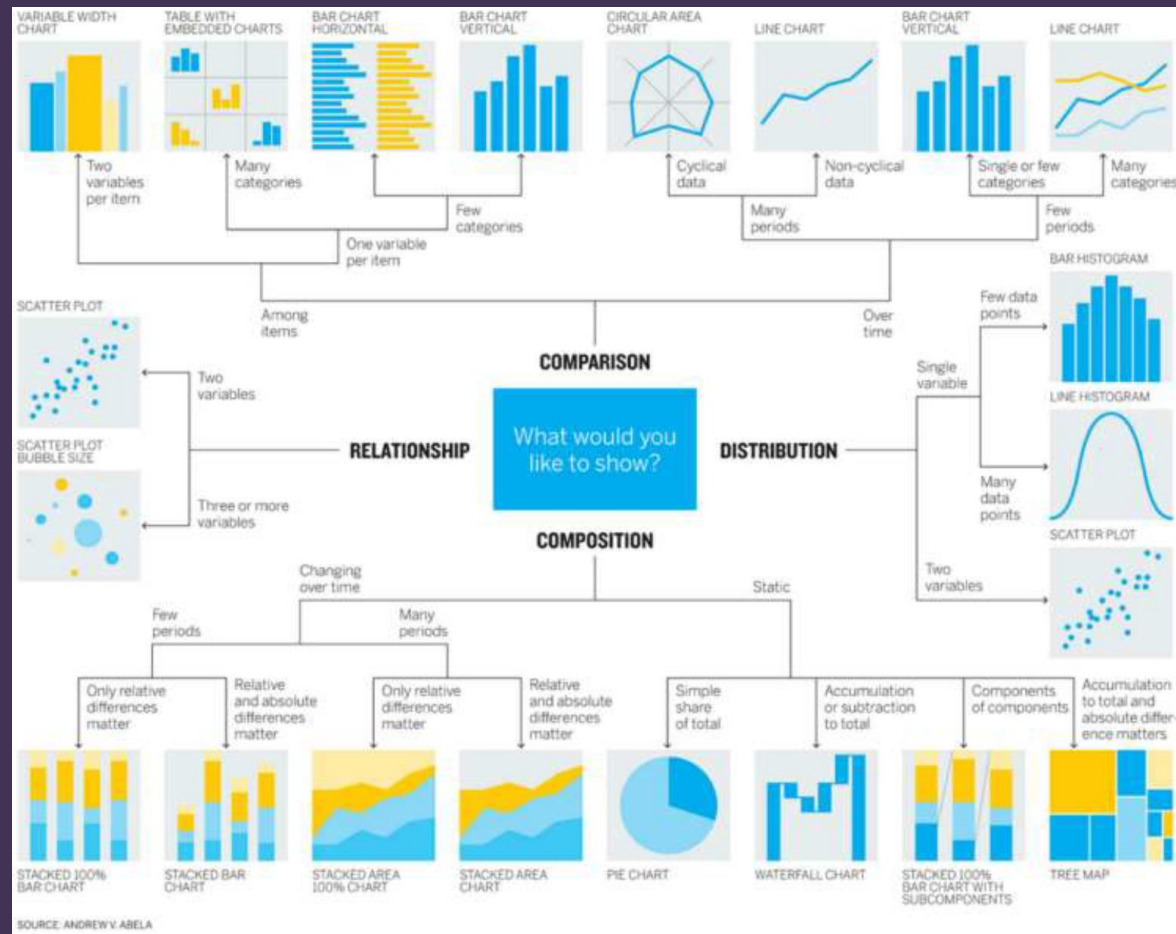


Diagram wyboru wykresu Andrew Abeli



Więcej...

- Diagram wyboru rodzaju wizualizacji, Andrew Abela
https://extremepresentation.typepad.com/blog/2006/09/choosing_a_good.html
- International Business Communication Standards <https://www.ibcs.com/>
- „Kwiatki” wizualizacyjne <https://viz.wtf/>