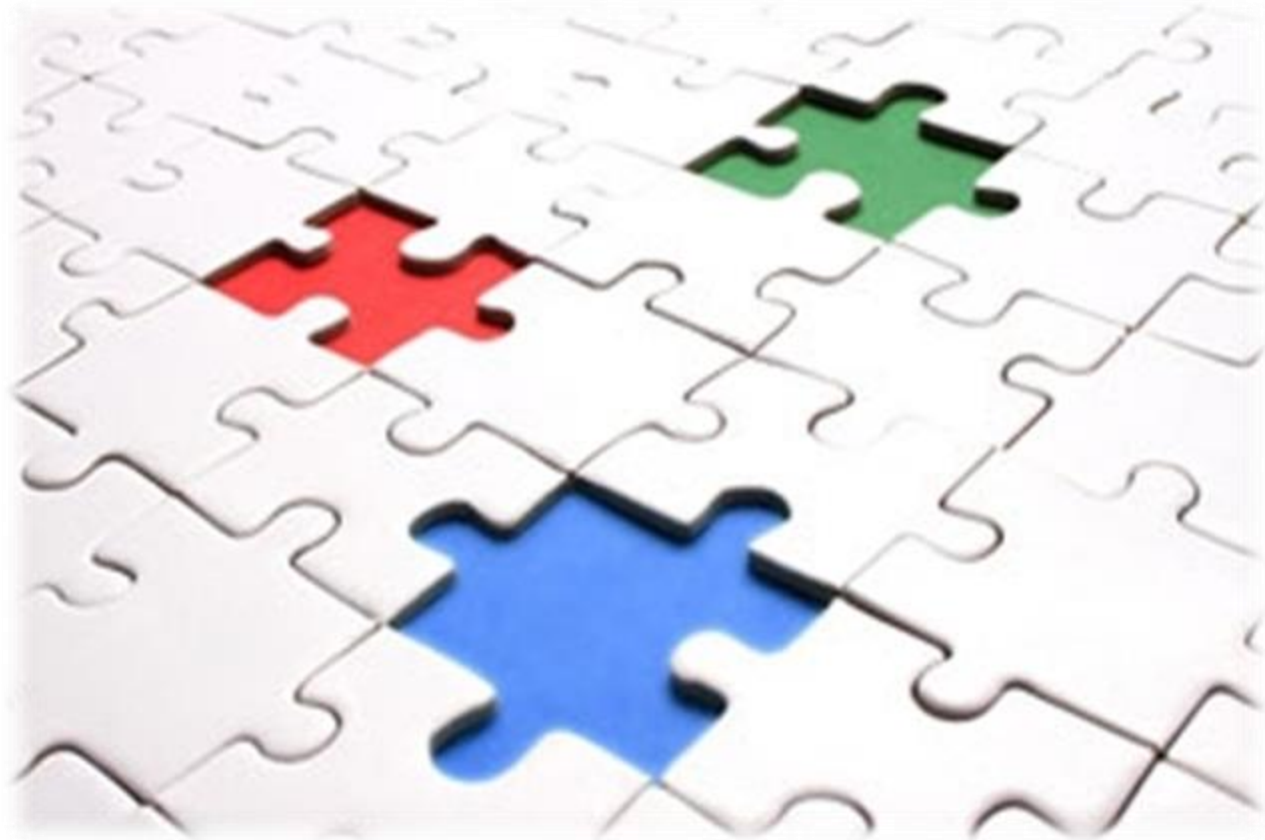


# The patterns of missingness

Evaluating Visualizations with Missing Data

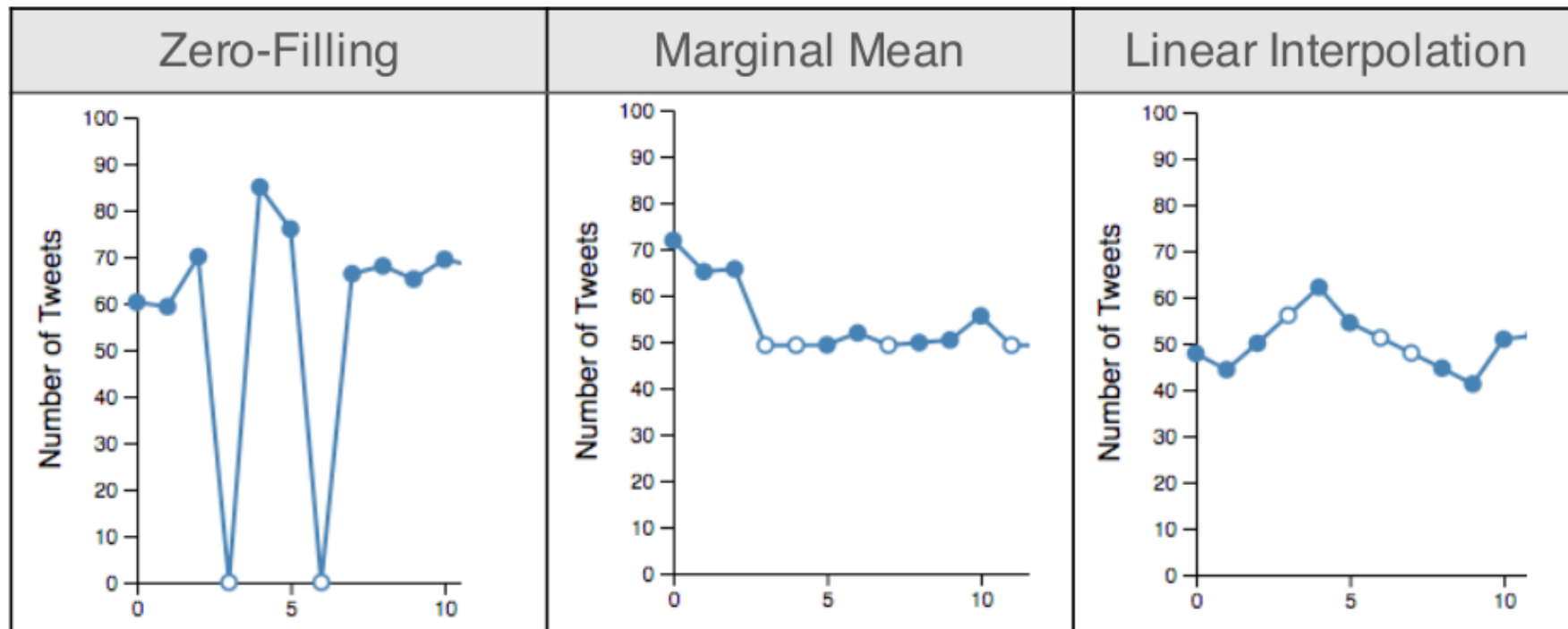
# Harsh data scientist reality



Niebezpieczeństwo  
przedstawiania  
niedokładnych danych



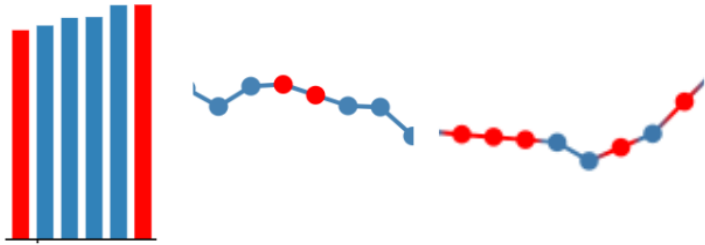
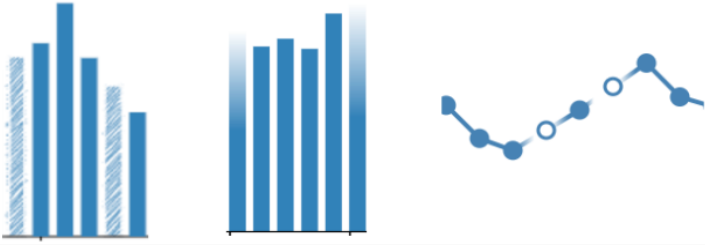
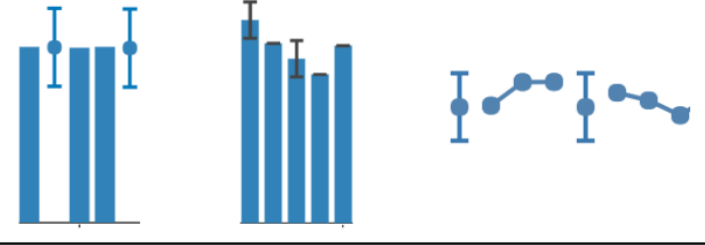

# Instead of real data - imputation



# Imputacja a wizualizacja

czyli nota o terminologii i nie tylko

# Metody wizualizacji NA

Highlight	
Downplay	
Annotation	
Information Removal	

# Results – line graphs

- Incomplete data credibility check
- Averaging
  - Linear regression the most suitable
  - Zero-filling has clearly negative impact
- Trend Detection
- Connected error bars better than disconnected



(d) Connected Error  
Bars



(e) Disconnected Error  
Bars

# Results - bar charts

-Averaging

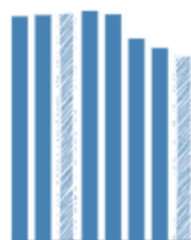
-Trend



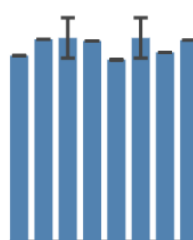
(a) Data Absent



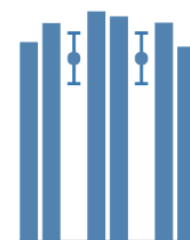
(b) Color Bars



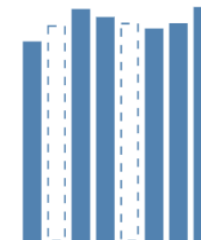
(c) Sketched Bar



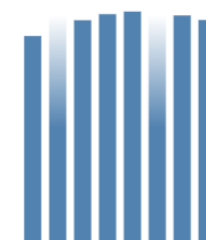
(d) Bars with Error Bars



(e) Points with Error Bars



(f) Bars with Dashed Outlines



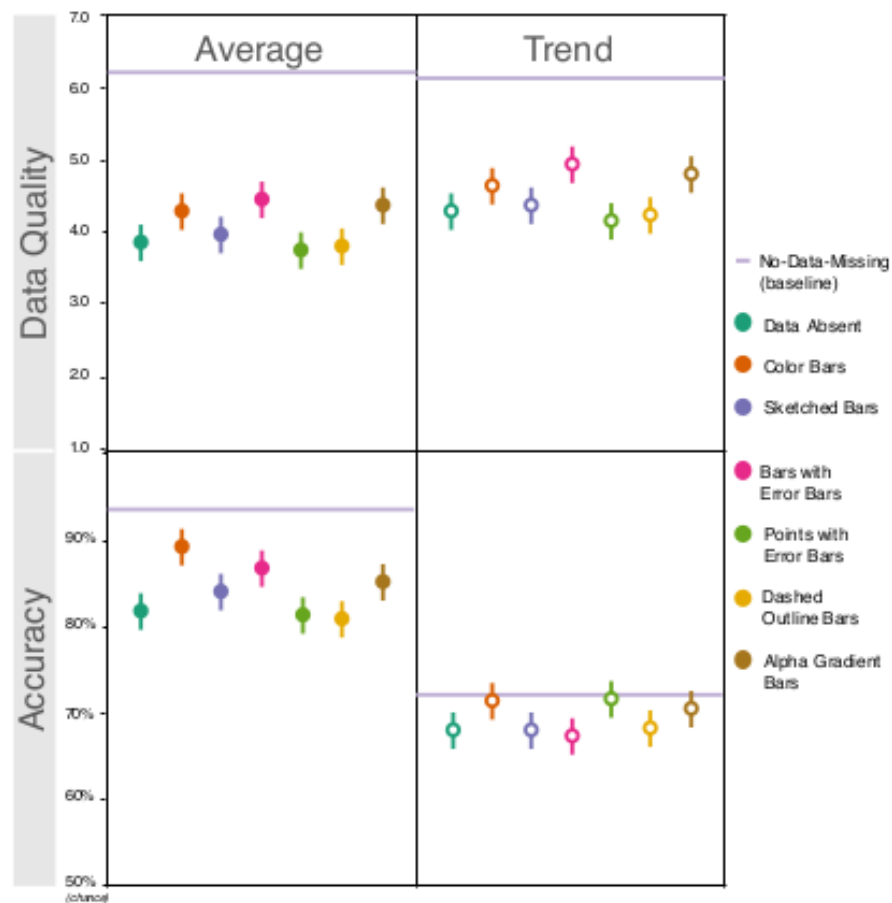
(g) Alpha Gradient Bars



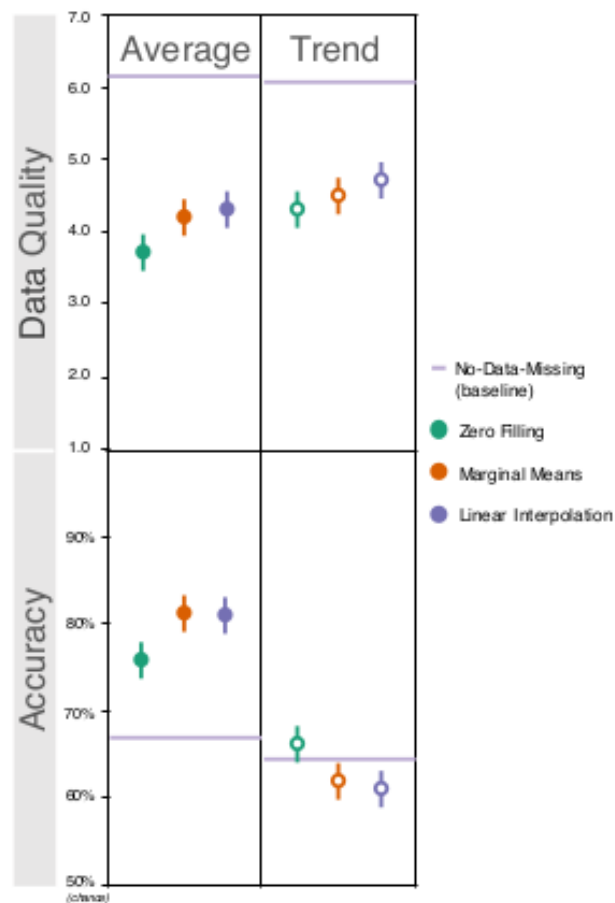
Podsumowanie

# Źródła

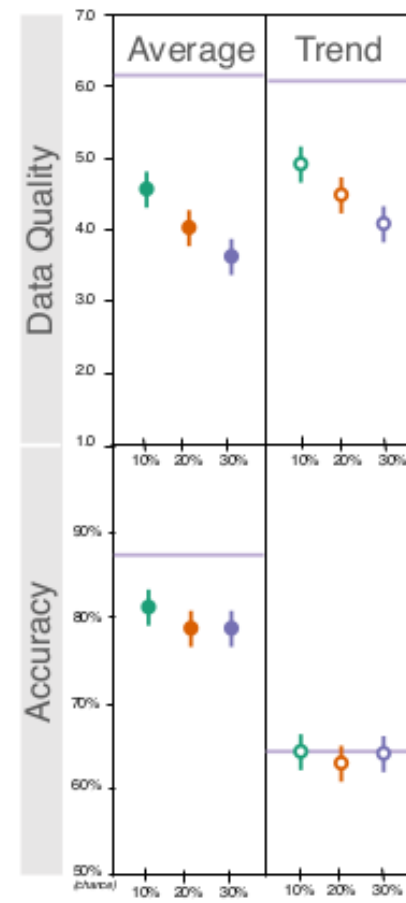
1. Where's My Data? Evaluating Visualizations with Missing Data, Hayeong Song, Danielle Albers Szafir
2. [https://en.wikipedia.org/wiki/Imputation\\_\(statistics\)](https://en.wikipedia.org/wiki/Imputation_(statistics))



(a) Visualization Methods



(b) Imputation Methods



(c) Amount Missing

Bar Charts