

# Poprawa jakości wykresu opublikowanego w internecie z wykorzystaniem pakietu *plotly*

Hubert Kowalski

December 9, 2023

## 1 Wizualizacja źródłowa

Wizualizacja źródłowa została opublikowana przez kanadyjskiego nadawcę radiowo-telewizyjnego *CBC* w corocznym raporcie z działalności. W ich raporcie z lat 2017 - 2018, opublikowanym 31 marca 2018 roku, można znaleźć wiele wykresów wymagających poprawy, jednak na potrzeby tego zadania skupimy się na wykresie dotyczącym źródeł finansowania stacji.

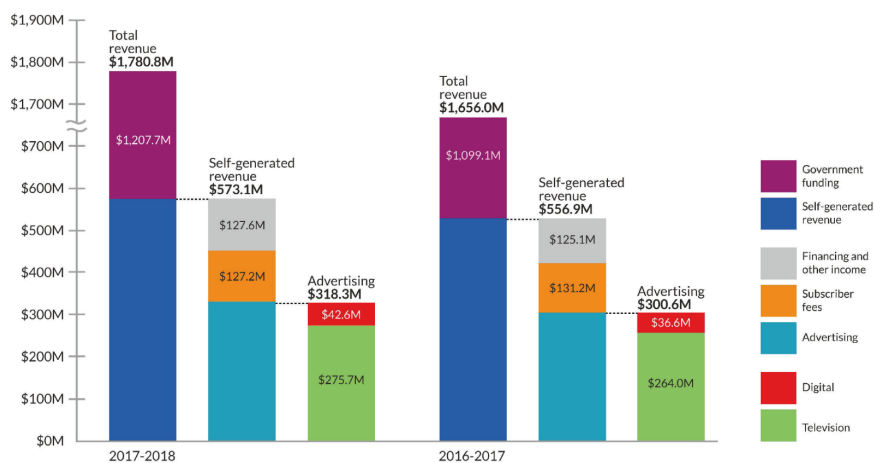


Figure 1: Wykres: [wizualizacja źródłowa](#) Pełny raport: [raport](#)

Na wykresie 1 przedstawione są wartości przychodów (wyrażone w milionach dolarów) z danej kategorii. Na pierwszy rzut oka może wydawać się, że wykres jest poprawny, jednak po głębszej analizie można zauważyć następujące nieprawidłowości. Na osi Y znajduje się przerwa, przez co słupek odpowiadający finansowaniu przez rząd wydaje się mniejszy niż ten reprezentujący fundusze własne, podczas gdy w rzeczywistości *CBC* finansowane jest w większości ze środków publicznych. Dodatkowo, słupki reprezentujące finansowanie własne oraz finansowanie przez reklamy nie powinny stanowić odrębnej części na wykresie. Przykładowo, na wykresie słupek czerwono-zielony jest jedynie podpodziałem słupka turkusowego.

## 2 Nowa wizualizacja

Na początku konieczne było przygotowanie danych potrzebnych do stworzenia wykresu.

```
1 government <- c(1099.1, 1207.7)
2 financing <- c(125.1, 127.6)
3 subs <- c(131.2, 127.2)
4 digital <- c(36.6, 42.6)
5 television <- c(264.0, 275.7)
6 years <- c("2016-2017", "2017-2018")
7 sum <- government + financing + subs + digital + television
8
9 governmentF <- paste0("<b>$", government, "M</b>")
10 financingF <- paste0("<b>$", financing, "M</b>")
11 subsF <- paste0("<b>$", subs, "M</b>")
12 digitalF <- paste0("<b>$", digital, "M</b>")
13 televisionF <- paste0("<b>$", television, "M</b>")
14 sumF <- paste0("<b>$", sum, "M</b>")
15
16 colnames <- c("Government funding", "Financing and other income",
17             "Subscriber fees", "Digital advertising",
18             "Television advertising", "Year")
19
20 df <- data.frame(government, financing, subs, digital, television,
21                 years)
22 df <- df %>%
23   mutate(years = factor(years, levels = c("2016-2017",
24     "2017-2018")))
24 colnames(df) <- colnames
```

Zmienne typu *government*, *financing*... zawierają dane liczbowe, a zmienne *governmentF*, *financingF*... - sformatowane napisy do wyświetlenia na wykresie. Następnie tworzona jest ramka danych z odpowiednimi nazwami kolumn.

Następnie przygotowywane są etykiety wyświetlane podczas najeżdżania myszką na odpowiednią część wykresu.

```
1 labels <- c("Government funding", "Financing and other income",
2             "Subscriber fees", "Digital advertising", "Television
3             advertising")
4 categories <- c("Government", "Self-generated revenue",
5                 "Self-generated revenue", "Advertising", "Advertising")
6 full_labels <- paste0("<b>Source: </b>", labels, "<br><b>Category:
7             </b>", categories)
```

Na koniec tworzony jest wykres interaktywny:

```
1 plot_ly(df, x = ~years, y = ~television, type = 'bar', name =  
  'Television advertising', text = televisionF, textfont = t,  
  textposition = 'auto', hoverinfo = 'text', hovertext =  
  paste0(full_labels[5], "<br><b>Value: </b>", televisionF ),  
2      marker = list(color = "#0066CC", line = list(color =  
        'rgb(8,48,107)', width = 3)) ) %>%  
3 add_trace(y = ~digital, name = 'Digital advertising', text =  
  digitalF, hoverinfo = 'text', hovertext =  
  paste0(full_labels[4], "<br><b>Value: </b>", digitalF ),  
4      marker = list(color = "#A00080")) %>%  
5 add_trace(y = ~subs, name = 'Subscriber fees', text =  
  subsF, hoverinfo = 'text', hovertext = paste0(full_labels[3],  
  "<br><b>Value: </b>", subsF),  
6      marker = list(color = "#F37295")) %>%  
7 add_trace(y = ~financing, name = 'Financing and other income',  
  text = financingF, hoverinfo = 'text', hovertext =  
  paste0(full_labels[2], "<br><b>Value: </b>", financingF ),  
8      marker = list(color = "cornsilk")) %>%  
9 add_trace(y = ~government, name = 'Government funding', text =  
  governmentF, hoverinfo = 'text', hovertext =  
  paste0(full_labels[1], "<br><b>Value: </b>", governmentF ),  
10     marker = list(color = "#CD2626")) %>%  
11 add_annotations(x = ~years, y = ~sum+50, text = paste0("Total  
  revenue: ", sumF), showarrow = FALSE, font = list(size = 16))  
  %>%  
12 layout(title = '<b>Revenue and other sources of funds for the  
  <i>Canadian National Broadcasting Company (CBC)</i></b>',  
13     font = list(size = 16),  
14     xaxis = list(title = 'Year'),  
15     yaxis = list(title = list(text = "Value (in dollars) for  
        each source of funds", standoff = 10), gridcolor =  
        "darkgrey"),  
16     bargap = 0.6,  
17     margin = m,  
18     legend = list(x = 100, y = 0.5, title =  
        list(text = "<b>Type of source</b>"),  
19     bargap = 0.6,  
20     hoverlabel = list(font = list(size = 16))  
21     ) %>%  
22 config(displayModeBar = FALSE)
```

Rezultat:

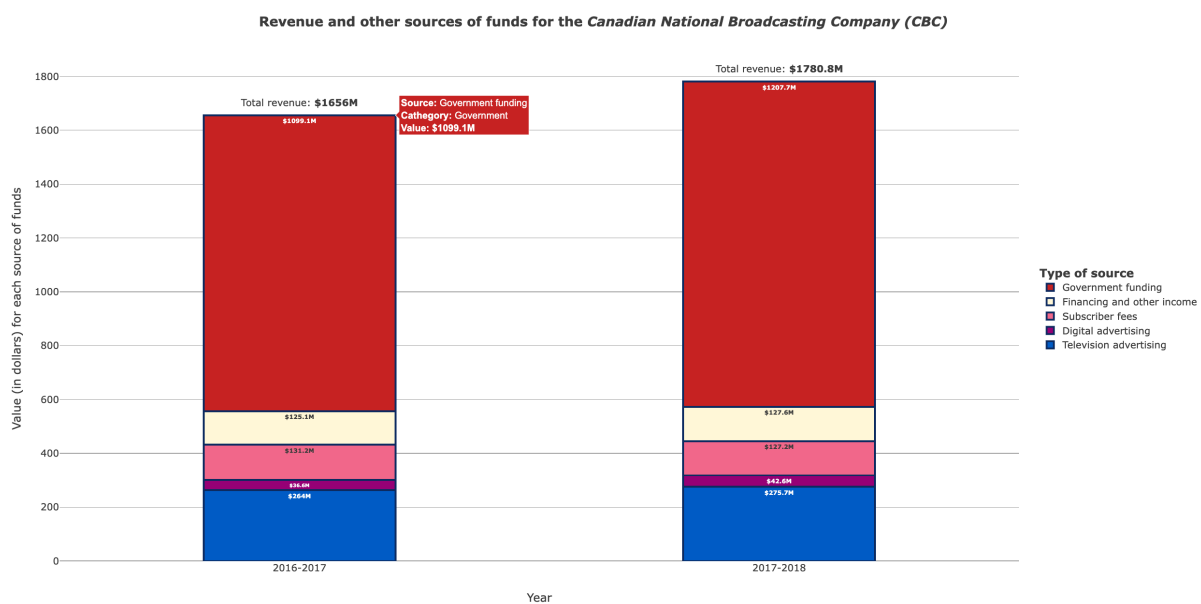


Figure 2: Poglądowy widok poprawionego wykresu. Aby skorzystać z interaktywności należy pobrać plik *Wykres.html* dostępny w repozytorium.

### 3 Uzasadnienie

Stworzony wykres posiada tytuł i opisy osi. Oś Y nie jest przerywana. Podpodziały źródeł finansowania nie stanowią już odrębnych słupków, a informacja o kategorii do której należy dane źródło finansowania jest zawarta w etykiecie po najechaniu na daną część wykresu.