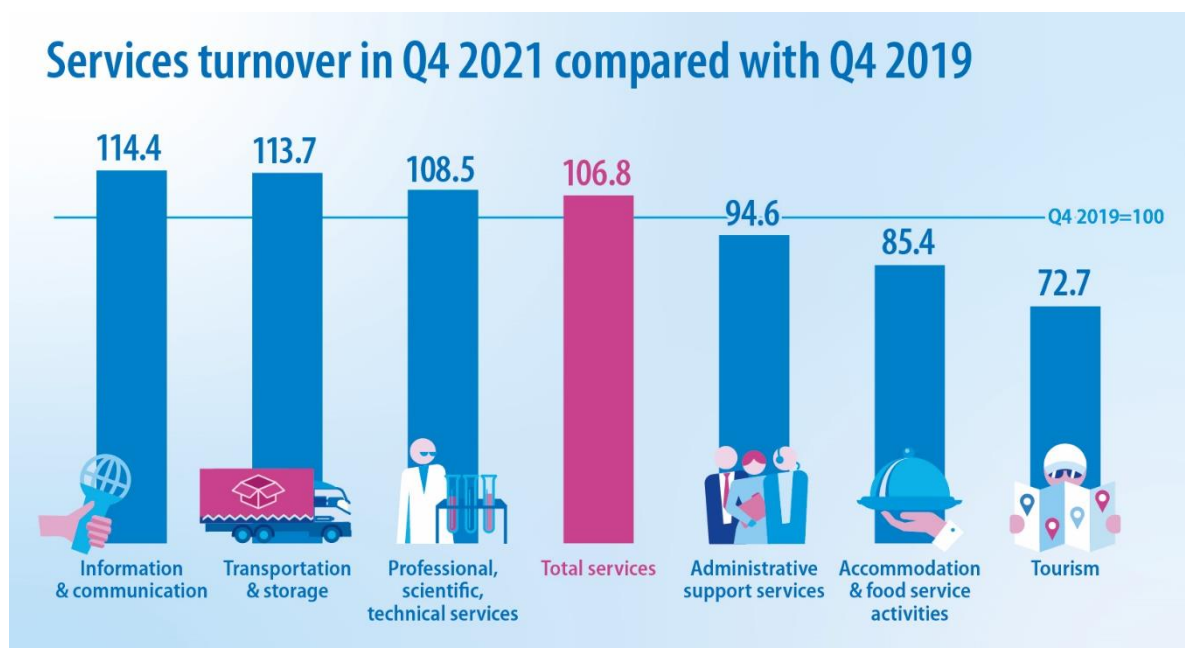
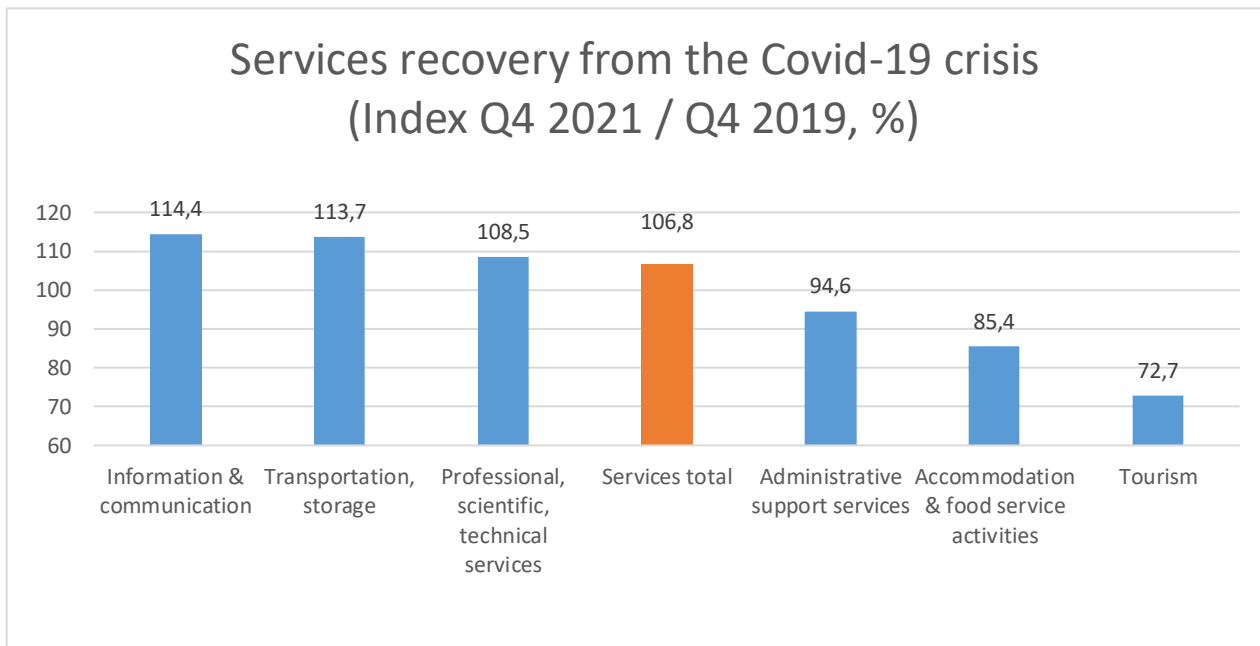


Praca domowa nr 2

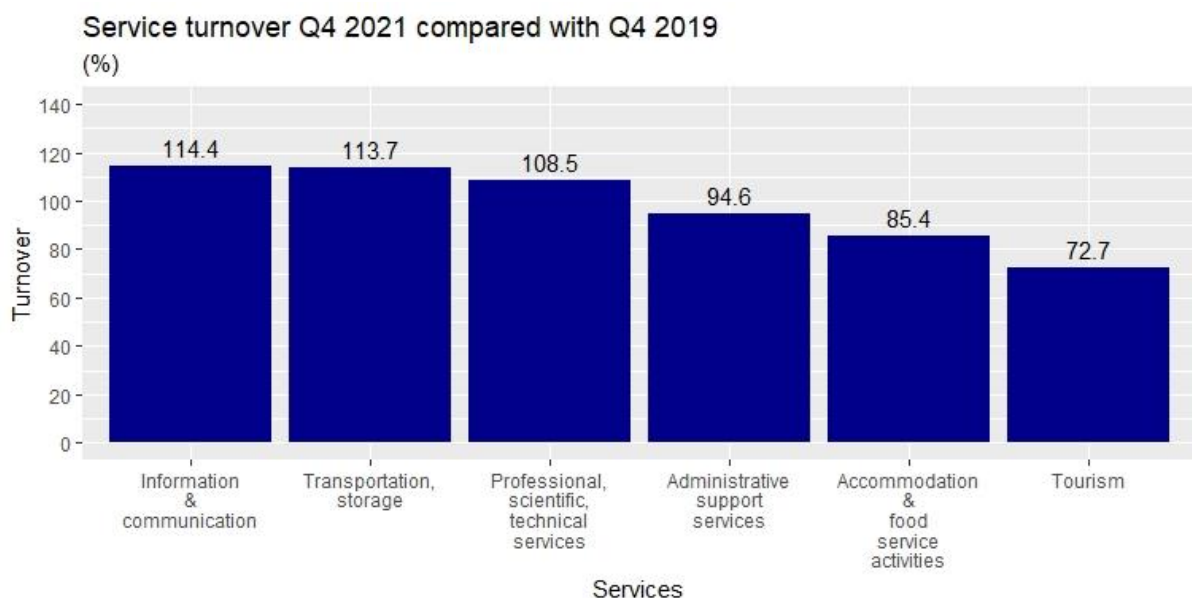
Jakub Mieczkowski

W swoim zadaniu postanowiłem wybrać tabelę ze strony Eurostat (cały artykuł pod linkiem <https://ec.europa.eu/eurostat/en/web/products-eurostat-news/-/ddn-20220328-3>). Poniżej umieściłem 2 wykresy. Pierwszy obrazuje jak wyglądał wykres w pliku excelowym z danymi, a drugi ukazuje te same dane po przerobieniu w programie graficznym. W poniższym raporcie wymienię jakie błędy zostały popełnione w obu wykresach i jakie dodatkowo niedociągnięcia posiada wykres drugi.



Pierwszym rażącym w oczy błędem jest słupek „total services”. Można się domyślić, że ta kolumna miała pokazywać stosunek obrotów usług w 4 kwartale 2021 do obrotów usług w 2019, jednakże może być mało zrozumiała i nie ma sensu porównywanie jej z pojedynczymi usługami. Tą informację lepiej przekazać w artykule. Inny błąd, który widać w szczególności na pierwszym wykresie jest skala zaczynająca się od 60 zamiast od 0. Przykładowo pierwsze wrażenie patrząc na ten wykres jest takie że „Tourism” w porównaniu do „Acomodation & food service activities” pogorszył się dwukrotnie bardziej, a jednak stosunek 85.4% do 72.7% jest zdecydowanie mniejszy niż 2. Ostatni błąd popełniony już w programie graficznym ozdabiającym wykres jest stworzenie obrazków na dole słupków. Są to dodatkowe zbędne informacje, które zamazują obraz i odciągają uwagę od danych.

Poniżej znajduje się mój wykres, który uniknął tych błędów, oraz kod który pozwolił go stworzyć:



```
install.packages("ggplot2")
library(ggplot2)
install.packages("stringi")
library(stringi)
install.packages("readxl")
library(readxl)
COVIDImpact <- read_excel("Szkola/Kuba/Studia/rstudio/COVIDImpact.xls")
View(COVIDImpact)
COVID <- COVIDImpact[87:93, 1:2]
names(COVID) <- c("Service", "Turnover")
COVID[1:7,1] <- stri_replace_all_regex(COVID$Service, " ", "\n")
COVID %>%
```

```
filter(Service != "Services total") %>%
mutate(Turnover = round(as.numeric(Turnover), 1)) %>%
ggplot(aes(x = reorder(Service, -Turnover), y = Turnover))+
geom_col(fill= "dark blue")+
labs(title = "Service turnover Q4 2021 compared with Q4 2019",
      subtitle = "(%)",
      x = "Services")+
scale_y_continuous("Turnover", seq(0,160,20), limits = c(0,140))+
geom_text(aes(label=as.numeric(Turnover)), vjust = -0.5)
```

