## **DM Data visualisation**

RAKOTOBE Muriella CHOLLET Antoine

Le graphique nous indique dans le titre « Perte d'emploi par trimestre » : le titre est mal posé car il s'agit en fait du nombre de demandeurs d'emploi aux Etats-Unis et non la perte d'emploi. De plus, il n'est pas précisé dans le titre que ces données se basent sur les Etats-Unis. Nous le supposons au vu des chiffres dans le graphique et de où a été publié ce graphique, sur America's newsroom, mais il était tout à fait possible qu'il s'agisse seulement que de certaines régions.

Nous avons une droite sur le graphique : cela nous amène à penser que le nombre de chômeurs aux Etats-Unis augmente de manière linéaire alors que ce n'est pas le cas. Cela est lié au choix des mois qui sont représentés, donc il est d'autant plus important de représenter tous les mois, ou de manière générale avoir une fréquence régulière de mois, par exemple tous les 2 ou 3 mois, pour mieux voir l'évolution du nombre de chômeurs aux Etats-Unis et notamment constater que l'évolution n'est pas linéaire.



Le graphique ne permet pas de bien voir l'évolution du nombre de chômeurs aux Etats-Unis. L'auteur a choisi de représenter cette évolution en choisissant 4 mois : Décembre 2007, Septembre 2008, Mars 2009 et Juin 2010. Le choix de ces 4 mois nous paraît ambigü : il aurait fallu mettre des résultats pour chaque mois, par exemple Décembre 2007, Janvier 2008, Février 2008, etc... Jusqu'à Juin 2010.

Ces données sont issues du BLS (Bureau of labour statistics). Cependant, la source reste incomplète car il aurait fallu que l'auteur précise également la nomenclature sur laquelle se base ces données. Par exemple, il est souvent précisé le niveau NUTS lorsque les données viennent d'Eurostat.

A partir des critiques faites sur le graphique qui nous a été donné, nous proposons une visualisation alternative des données via le graphique suivant :



Nous sommes allés sur le site du BLS afin de reprendre les données concernant le nombre de chômeurs aux Etats-Unis pour représenter tous les 4 mois son évolution au cours du temps. Nous avons choisi de conserver la visualisation sous forme de courbe car, de notre point de vue, c'est ce qui permet au mieux de voir à l'oeil cette évolution au cours du temps. Nous avons mis un titre plus explicatif et complet, en rajoutant clairement que l'étude se base sur les Etats-Unis, et nous avons également ajouté un sous-titre explicitant qu'il s'agit de l'évolution du nombre de chômeurs de 2007 à 2012. Concernant la source, nous l'avons complété en mettant le nom de l'étude pour avoir plus de précision sur la provenance des données.

Nous constatons que l'ancien graphique était trompeur car nous voyons une forte diminution du nombre de chômeurs aux Etats-Unis.

```
File Edit Code View Plots Session Build Debug Profile Tools Help
     ▼ | 😭 ▼ | 🔒 | 📄 | 🍎 Go to file/function
                                                                     ■ • Addins •
  dmcode.R × unempl ×
                        Source on Save
                                                                                                                                                              → Run → Source
          library(ggpmisc)
          library(ggpubr)
library(scales)
          unempl <- read.csv2("P:/dataviz/unempl.csv",header = T) #
row.names(unempl) <- unempl$Year #changement des index des
unempl<- unempl[-61,] #suppression d'une ligne parasite</pre>
    13
    14
   15
16
17
18
19
          unempl$Year <- mdy(unempl$Year) #conversion au format date mois/jours/annee</pre>
          options(scipen=999) # desactivation de la notation scientifique
    20
   21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
             <- ggplot(unempl, aes(x = Year, y = Unemployment,group=1))+
geom_point(aes(y=Unemployment))+ #affichage des points
geom_line(aes(y=Unemployment))+ #affichage de la courbe associee aux points
theme_minial()+ #modification du theme
ylim(8000,17000)+ #initialisation de l'echelle d'un nombre de chomeurs pour eviter une vision</pre>
             #Trop grande des variation

#Trop grande des variation

scale_x_date(date_breaks = "4 month", labels=date_format("%m-%Y"))+ #modification de l'echelle des mois

#pour afficher tout les 4 mois au format mois-annee
            print(a) #affichage
```