

Compte-Rendu du projet d'Analyse Exploratoire

Des médailles aux JO !

Duc Hau NGUYEN; Anaïs RABARY

4 Décembre 2018

Contents

Introduction	2
Contexte	2
Problématique	2
1. Première axe: l'âge qui optimise la performance	2
1.1 Étude préliminaire	2
1.2 Analyse	4
1.3 Résultats	4
2.Problématique 2	4
2.1 Présentation	4
2.2 Analyse	4
2.3 Résultats	4
3.Problématique 3	4
3.1 Présentation	4
3.2 Analyse	4
3.3 Résultats	4
Conclusion	4
R Markdown	4
Including Plots	4
References	5

```
## Parsed with column specification:
## cols(
##   ID = col_double(),
##   Name = col_character(),
##   Sex = col_character(),
##   Age = col_double(),
##   Height = col_double(),
##   Weight = col_double(),
##   Team = col_character(),
##   NOC = col_character(),
##   Games = col_character(),
##   Year = col_double(),
```

```
## Season = col_character(),  
## City = col_character(),  
## Sport = col_character(),  
## Event = col_character(),  
## Medal = col_character()  
## )
```

Introduction

Contexte

Les Jeux Olympiques (JO) sont une compétition dont nous avons hérité de la Grèce Antique. On lit souvent que seuls les citoyens les plus riches pouvaient y participer. La trêve olympique permettait à la civilisation d'oublier, le temps de 12 travaux, les guerres et troubles de l'époque. Après quelques maigres tentatives, c'est seulement depuis 1896 que les Jeux Olympiques modernes voient le jour, à Athènes.

Dans le cadre du projet d'Analyse Exploratoire à l'INSA Toulouse, avec l'accompagnement de notre professeur Gilles Tredan, nous avons récupéré sur le site Kaggle un dataset sur ces Jeux Olympiques moderne. Ce dernier recense des informations sur tous les athlètes ayant participé aux JO, les médailles qu'ils ont gagnés, leur équipe, l'année de participation, ...

Problématique

Les Jeux Olympics sont devenus le symbol de la force physique et mentale de l'être humain. C'est l'occasion pour que les élites de chaque pays se réunissent et montrent leurs spectaculaires performances. D'un autre côté, ces occasions qualifiées de spectaculaires sont un sujet intéressant pour des scientifiques : Le dernier siècle a prédit que l'humain atteindra bientôt ses limites physiques, tandis que l'on voit toujours aujourd'hui des sportifs continuer d'établir de nouveaux records.

Avec plusieurs centaines d'années d'existence des JO, l'humanité a identifié quelques facteurs optimisant la performance physique. L'historique des JO pourraient surement fournir une indication à la question: Quels sont les facteurs qui influent/optiment la performance physiques dans les JO ?

Pour établir une étude d'exploration sur ce problème, il est nécessaire de croiser nos données avec des données de la source World Wide Data, sur les PIB des pays (...) (cf. dans biblio).

Par la suite de l'étude, on analyse sur les grandes axes: caractéristique de pays, les mesures physiques de participants, leur ages, ... (TODO). Chaque axes nous mène à analyser des phénomènes particuliers ou nous demande une étude supplémentaire pour justifier.

1. Première axe: l'âge qui optimise la performance

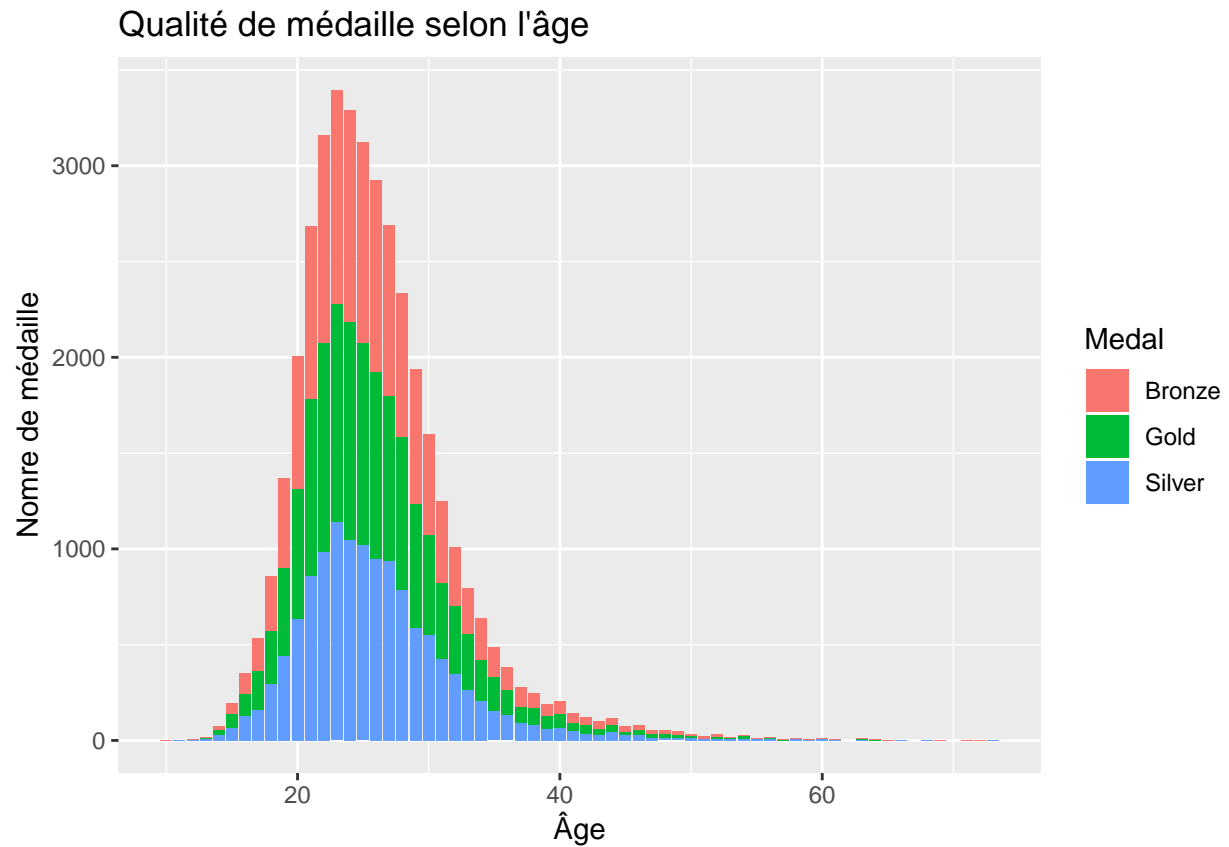
1.1 Étude préliminaire

L'approche naïve sera d'analyser le nombre de médailles que chaque âge a pu obtenir, quelque soit l'année des Jeux.

```
olympic_dataset %>%  
  count(Age, Medal) %>%  
  subset(Medal != "None") %>%
```

```
ggplot(aes(x = Age, y = n, fill = Medal)) + geom_bar(stat = "identity") +  
labs(x="Âge", y="Nomre de médaille", title = "Qualité de médaille selon l'âge")
```

Warning: Removed 3 rows containing missing values (position_stack).



1.2 Analyse

1.3 Résultats

2.Problématique 2

2.1 Présentation

2.2 Analyse

2.3 Résultats

3.Problématique 3

3.1 Présentation

3.2 Analyse

3.3 Résultats

Conclusion

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

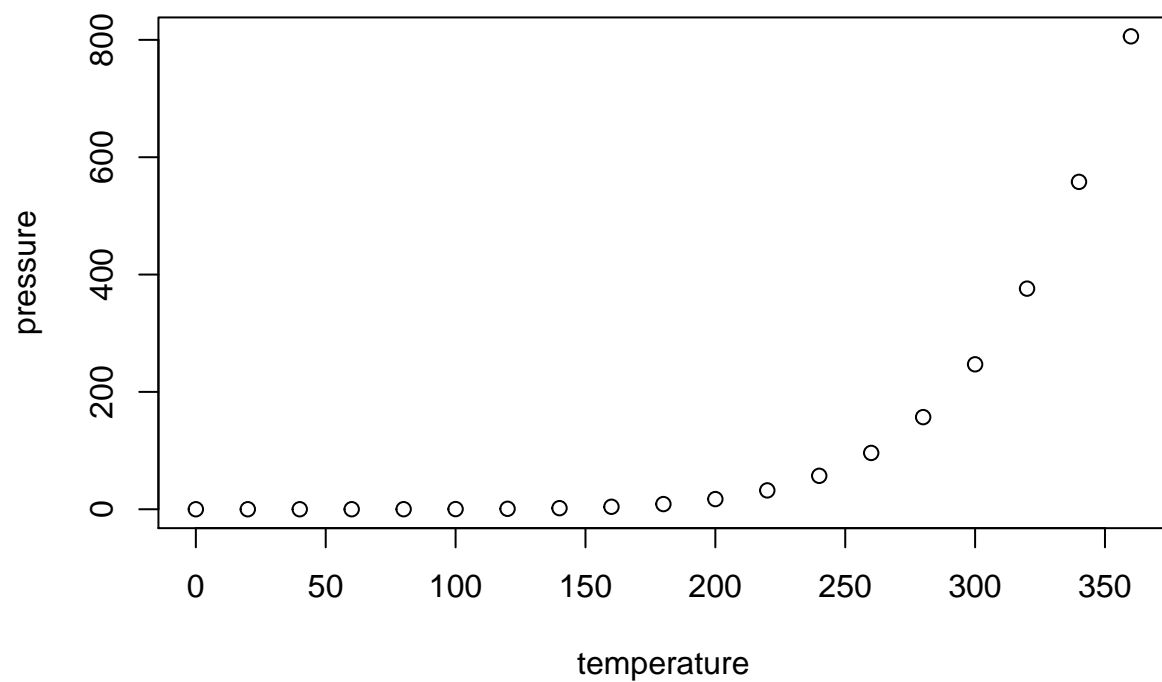
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   :  2.00
## 1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##   Mean  :15.4    Mean   : 42.98
## 3rd Qu.:19.0    3rd Qu.: 56.00
##   Max.  :25.0    Max.   :120.00
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

References