

Visualisation

Serge-Étienne Parent

01 février 2019

Plan

1. Les graphiques: avant de publier
2. Choisir le bon type de graphique
3. L'approche impérative et l'approche déclarative
4. Visualisation en R
5. Fermeture

Objectifs spécifiques

À la fin de ce chapitre, vous

- comprendrez l'importance de l'exploration des données
- comprendrez les guides généraux pour créer un graphique approprié
- comprendrez la différence entre les modes impératifs et déclaratifs pour la création de graphique
- serez en mesure de créer des nuages de points, lignes, histogrammes, diagrammes en barres et boxplots en R
- saurez exporter un graphique en vue d'une publication

1. Les graphiques: avant de publier

Un outil d'exploration

Animation montrant la progression du jeu de données *Datasaurus* pour toutes les formes visées. Source: Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing.

Cinq qualités d'un bon graphique

1. **Elle est véritable**, puisqu'elle est basée sur une recherche exhaustive et honnête.
2. **Elle est fonctionnelle**, puisqu'elle constitue une représentation précise des données, et qu'elle est construite de manière à laisser les observateurs.trices prendre des initiatives conséquentes.
3. **Elle est attrayante** et intrigante, et même esthétiquement plaisante pour l'audience visée - les scientifiques d'abord, mais aussi le public en général.
4. **Elle est pertinente**, puisqu'elle révèle des évidences scientifiques autrement difficilement accessibles.
5. **Elle est instructive**, parce que si l'on saisit et accepte les évidences scientifiques qu'elle décrit, cela changera notre perception pour le mieux.

Qualité 1. Honnêteté

```
library("tidyverse")
```

```
## -- Attaching packages -----

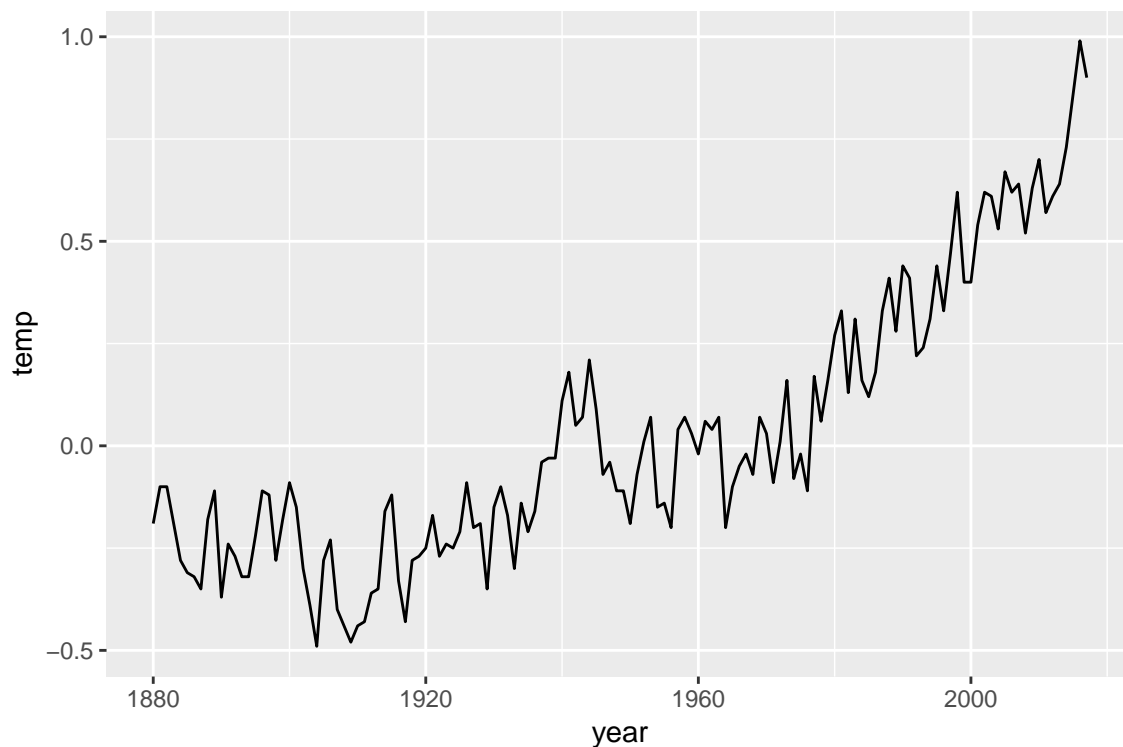
## v ggplot2 3.1.0      v purrr  0.3.0
## v tibble  2.0.1      v dplyr  0.7.8
## v tidyr   0.8.2      v stringr 1.3.1
## v readr   1.3.1      v forcats 0.3.0

## -- Conflicts -----
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

nasa_temp <- read_tsv("https://climate.nasa.gov/system/internal_resources/details/original/647_Global_T
                      col_names = c('year', 'temp', 'loess'))

## Parsed with column specification:
## cols(
##   year = col_double(),
##   temp = col_double(),
##   loess = col_double()
## )
```

```
nasa_temp %>%
  #filter(year >= 1998 & year <= 2008) %>%
  ggplot(mapping = aes(x = year, y = temp)) +
  geom_line()
```



Qualité 2. Fonctionnelle

About Company

3d Pie Chart

This slide is perfect for long text descriptions

16%

Creative Design

Lorem Ipsum is simply dummy text of the printing and typesetting industry.

24%

Desktop Application

Lorem Ipsum is simply dummy text of the printing and typesetting industry.

50%

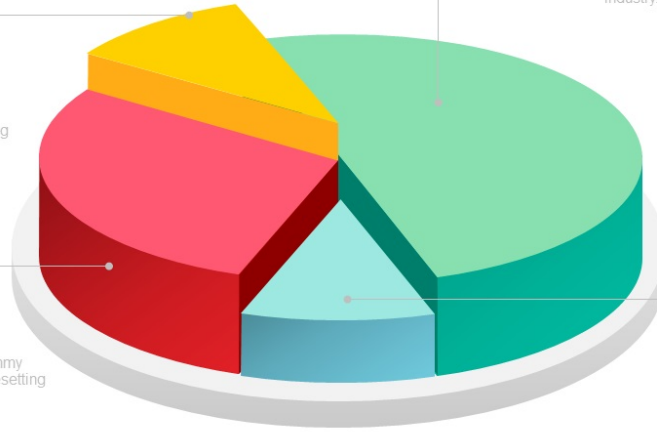
Mobile App Design

Lorem Ipsum is simply dummy text of the printing and typesetting industry.

10%

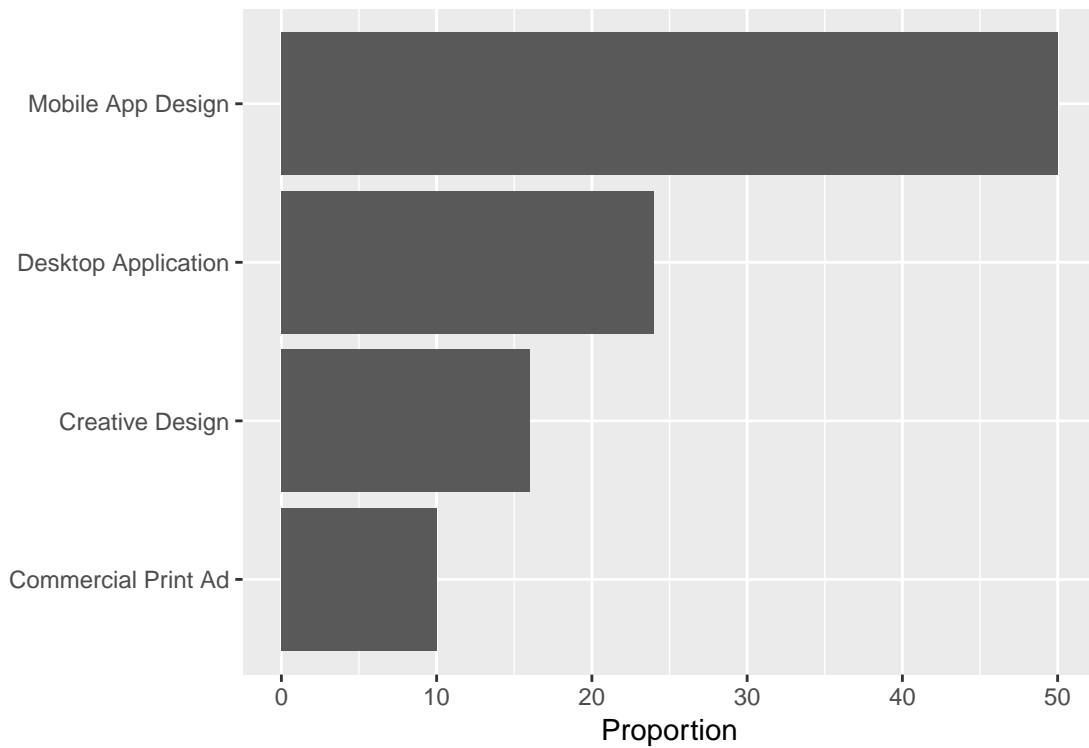
Commercial Print Ad

Lorem Ipsum is simply dummy text of the printing and typesetting industry.



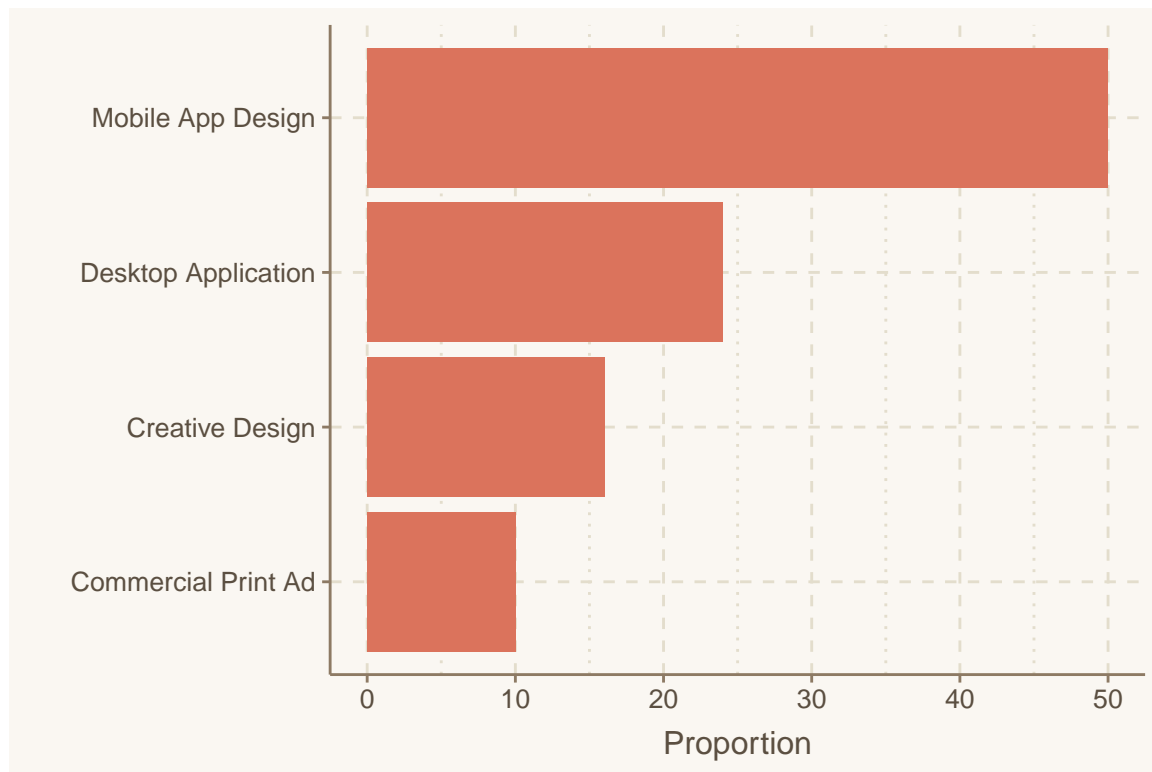
Source: Slidebazaar

```
market <- tibble(category = c("Creative Design", "Desktop Application", "Commercial Print Ad", "Mobile App Design"))
market %>%
  ggplot(mapping = aes(x = fct_reorder(category, proportion), y = proportion)) + geom_col() + labs(x = "Category", y = "Proportion")
```



Qualité 3. Visuel

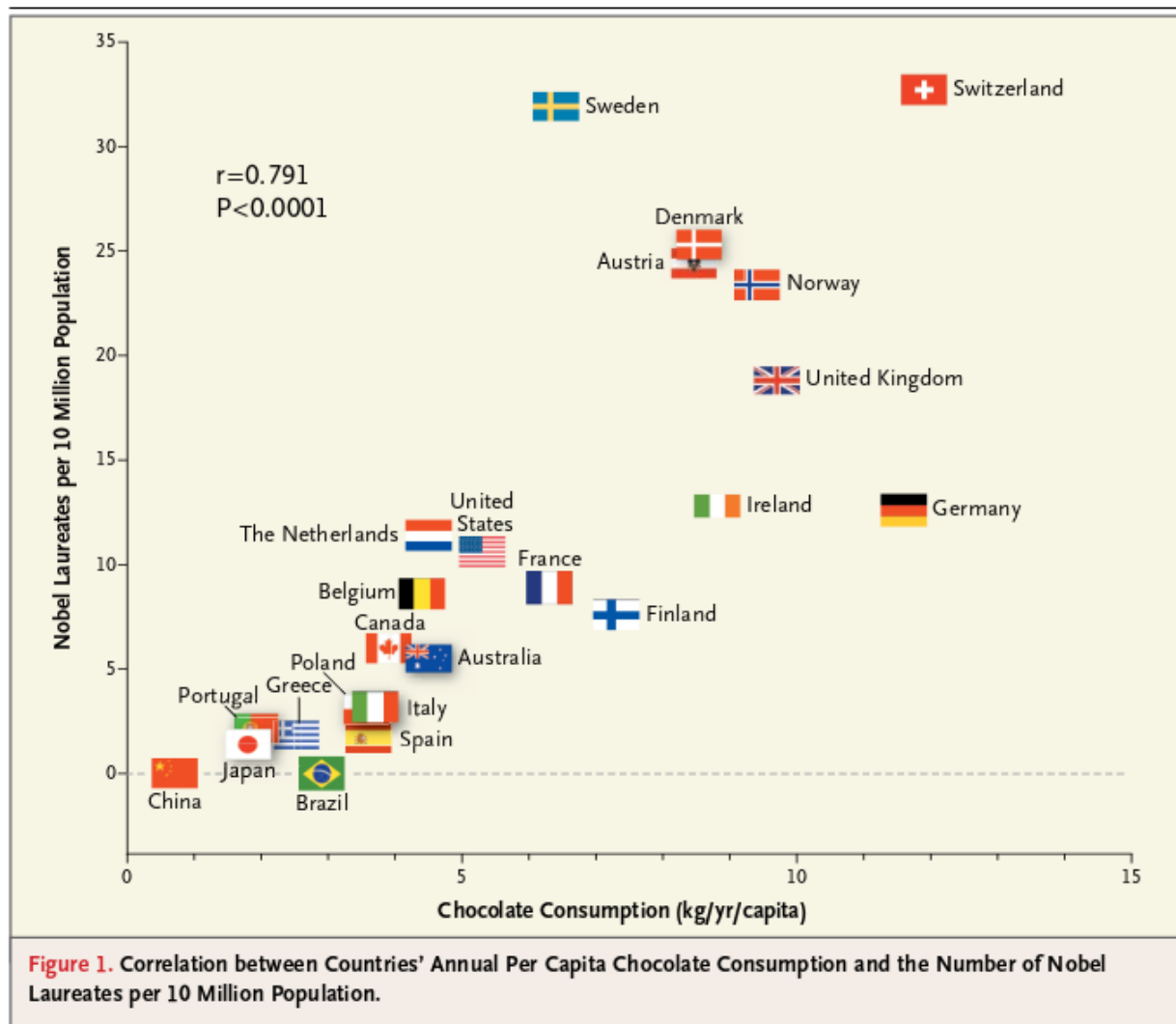
```
library("ggthemr")
ggthemr(palette = "dust", layout = "scientific", type = "outer")
market %>%
  ggplot(mapping = aes(x = fct_reorder(category, proportion), y = proportion)) + geom_col() + labs(x =
```



Qualité 4. Pertinence

Source: GIEC, Bilan 2001 des changements climatiques : Les éléments scientifiques

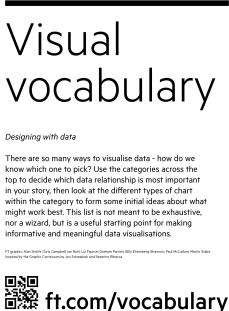
Qualité 5. Instructive



Source: Messerli, (2012)

3. Choisir le bon type de graphique

1. Mettez de l'ordre dans vos données.
2. Réfléchissez au message que vous désirez transmettre.
3. Essayez différentes représentations.
4. Testez le résultat.



ggplot2

1. **Les données.** Votre tableau est bien sûr un argument nécessaire pour générer le graphique.
2. **Les marqueurs.** Un terme abstrait pour désigner les points, les lignes, les polygones, les barres, les flèches, etc.
3. **Les attributs encodés.** La position, la dimension, la couleur ou la forme que prendront les géométries. En ggplot2, on les nomme les *aesthetics*.
4. **Les attributs globaux.** Les attributs sont globaux lorsqu'ils sont constant (ils ne dépendent pas d'une variable). Les valeurs par défaut conviennent généralement, mais certains attributs peuvent être spécifiés: par exemple la forme ou la couleur des points, le type de ligne.
5. **Les thèmes.** Le thème du graphique peut être spécifié dans son ensemble, c'est-à-dire en utilisant un thème prédéfini, mais l'on peut modifier certains détails.

ggplot2

-> basculer vers les notes de cours.

6. Fermeture