

Comptes et proportions

8INF416

Automne 2022

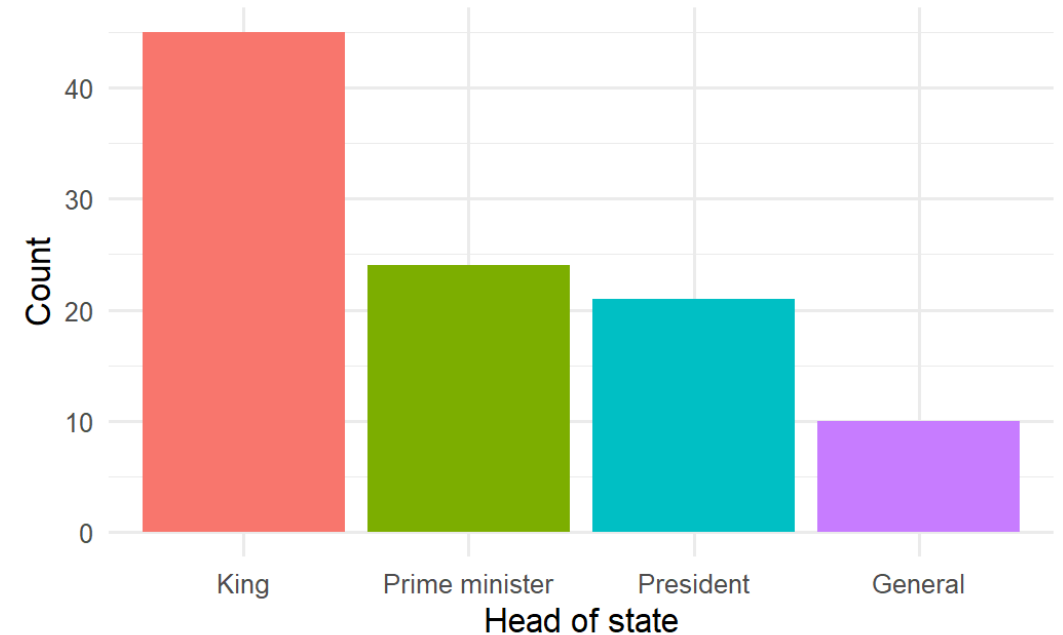
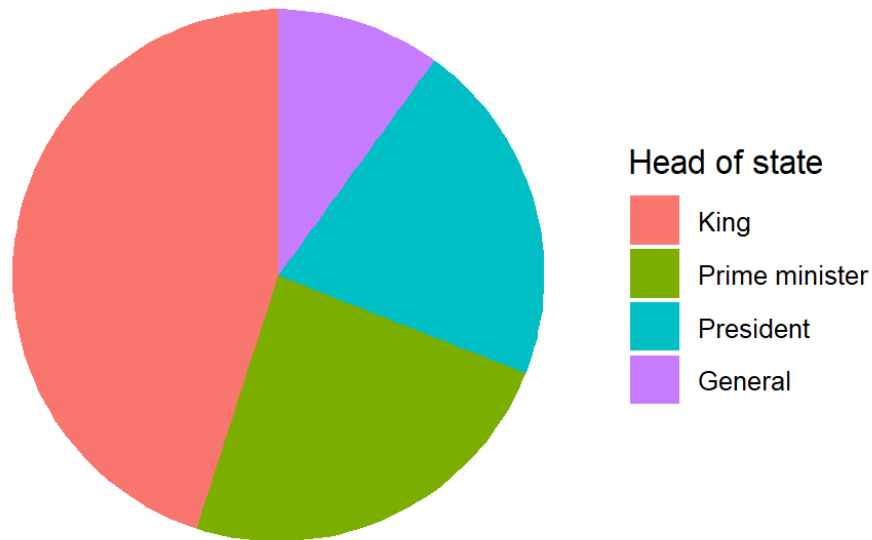
Plan pour aujourd'hui

Comptes

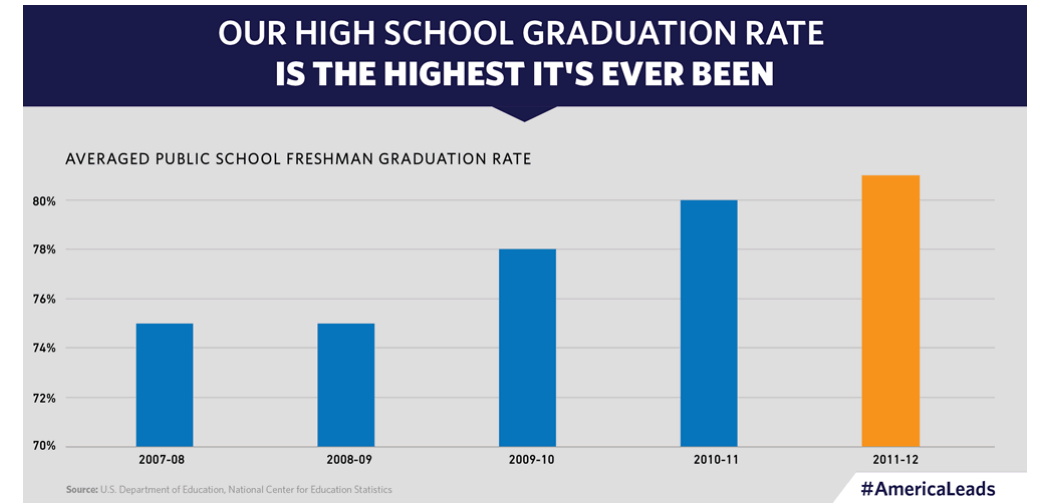
Proportions

Yay bar plots !

Nous sommes bien meilleurs pour visualiser la longueur des lignes que les angles et les surfaces



Oh pas de bar plots !



<https://blog.ed.gov/2014/10/progress-on-education-is-helping-fuel-our-economys-growth/> et <https://www.mediamatters.org/fox-news/dishonest-fox-charts-obamacare-enrollment-edition> et <https://twitter.com/ObamaWhiteHouse/status/517743415375974401>

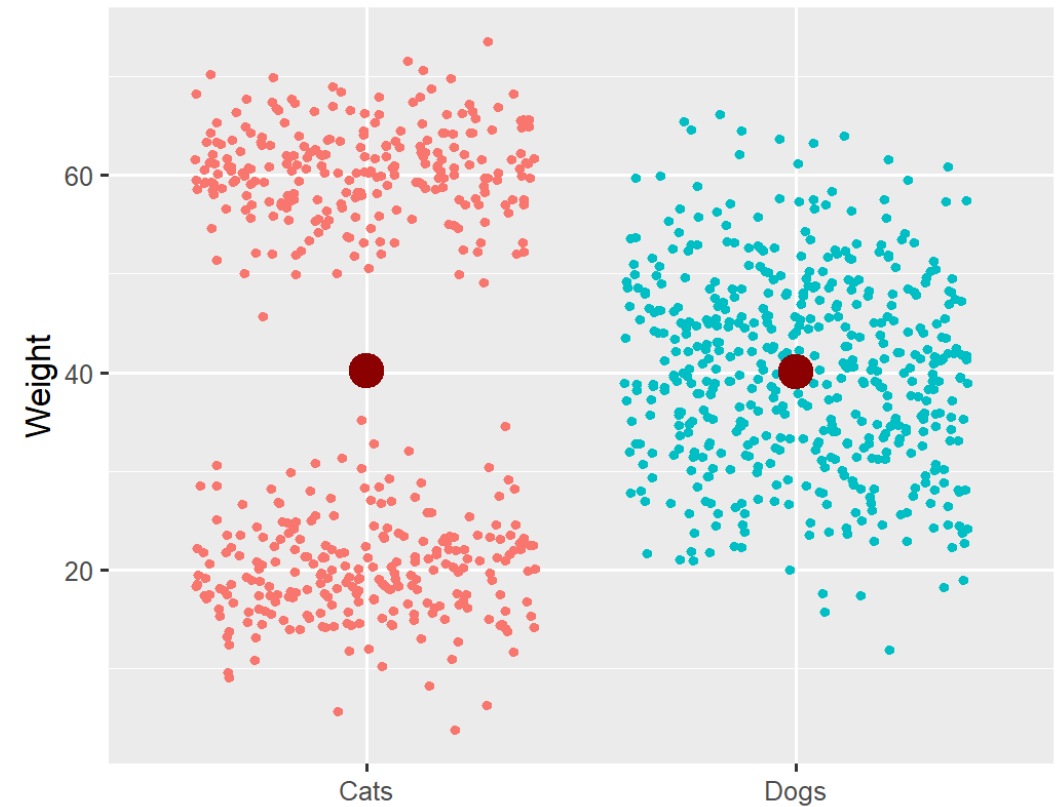
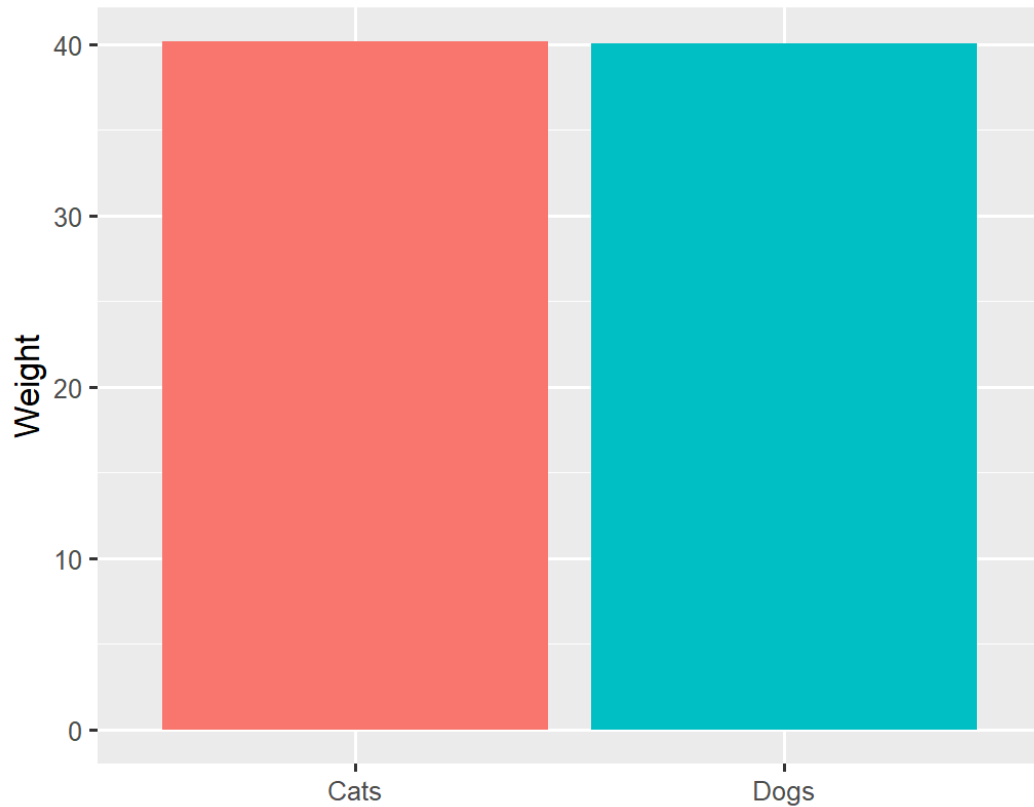
Commencer à zéro

**La longueur totale de la ligne compte,
alors ne la tronquez pas !**

Toujours commencer à 0

(Ou n'utilisez pas de barres)

Barres et stats descriptives



Afficher plus avec les graphiques en

```
ggplot(animals,  
      aes(x = animal_type,  
          y = weight,  
          color = animal_type)) +  
  geom_point(position = position_jitter(height = 1),  
            size = 1) +  
  labs(x = NULL, y = "Weight") +  
  guides(color = "none")
```

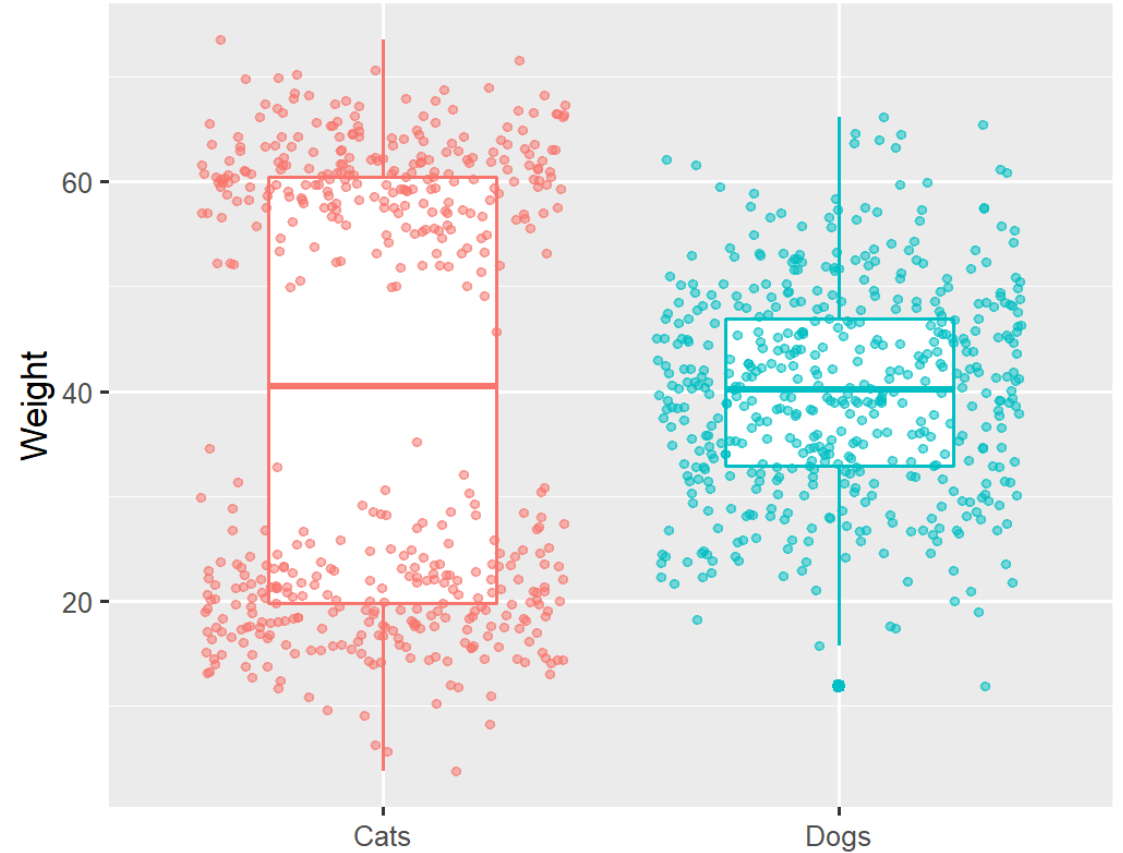
Afficher plus avec les graphiques

```
library(ggbeeswarm)

ggplot(animals, aes(x = animal_type,
                    y = weight,
                    color = animal_type)) +
  geom_beeswarm(size = 1) +
  # Or try this too:
  # geom_quasirandom() +
  labs(x = NULL, y = "Weight") +
  guides(color = "none")
```

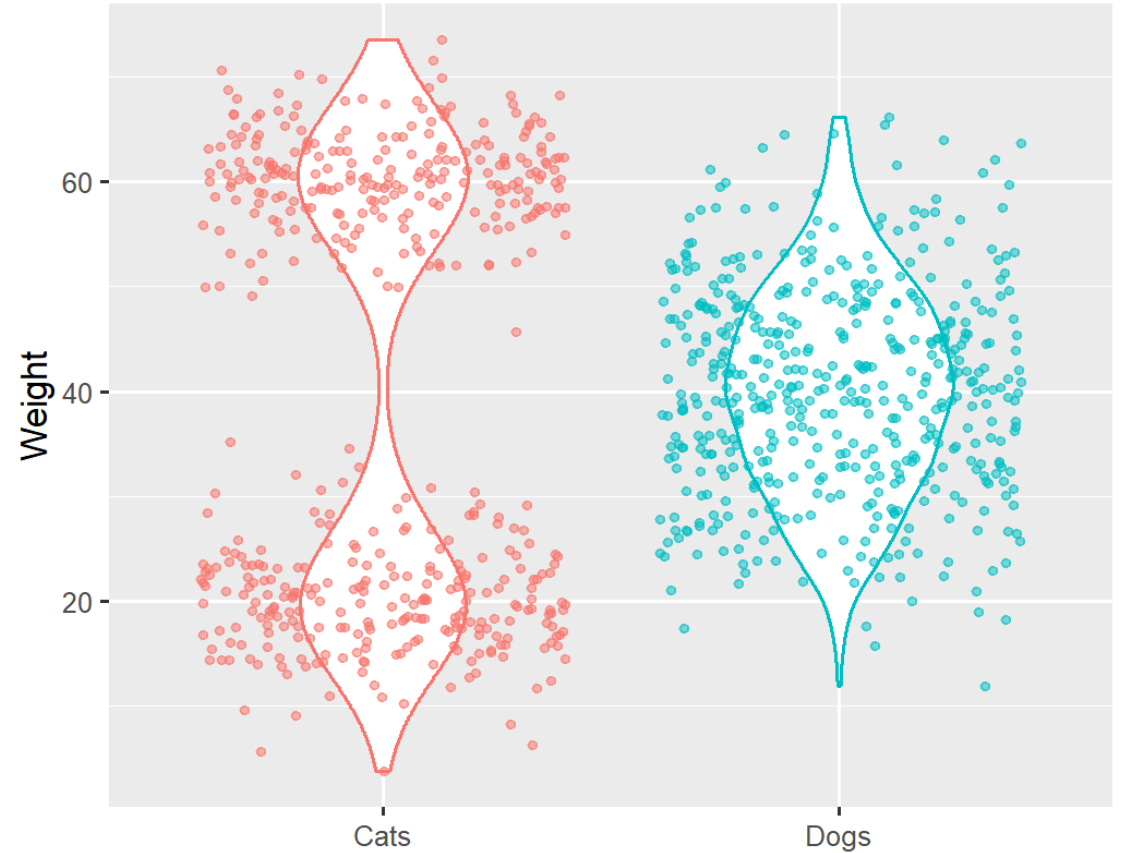

Combiner des boxplots avec des points

```
ggplot(animals, aes(x = animal_type,  
                    y = weight,  
                    color = animal_type)) +  
  geom_boxplot(width = 0.5) +  
  geom_point(position = position_jitter(height = 0.5),  
            size = 1, alpha = 0.5) +  
  labs(x = NULL, y = "Weight") +  
  guides(color = "none")
```



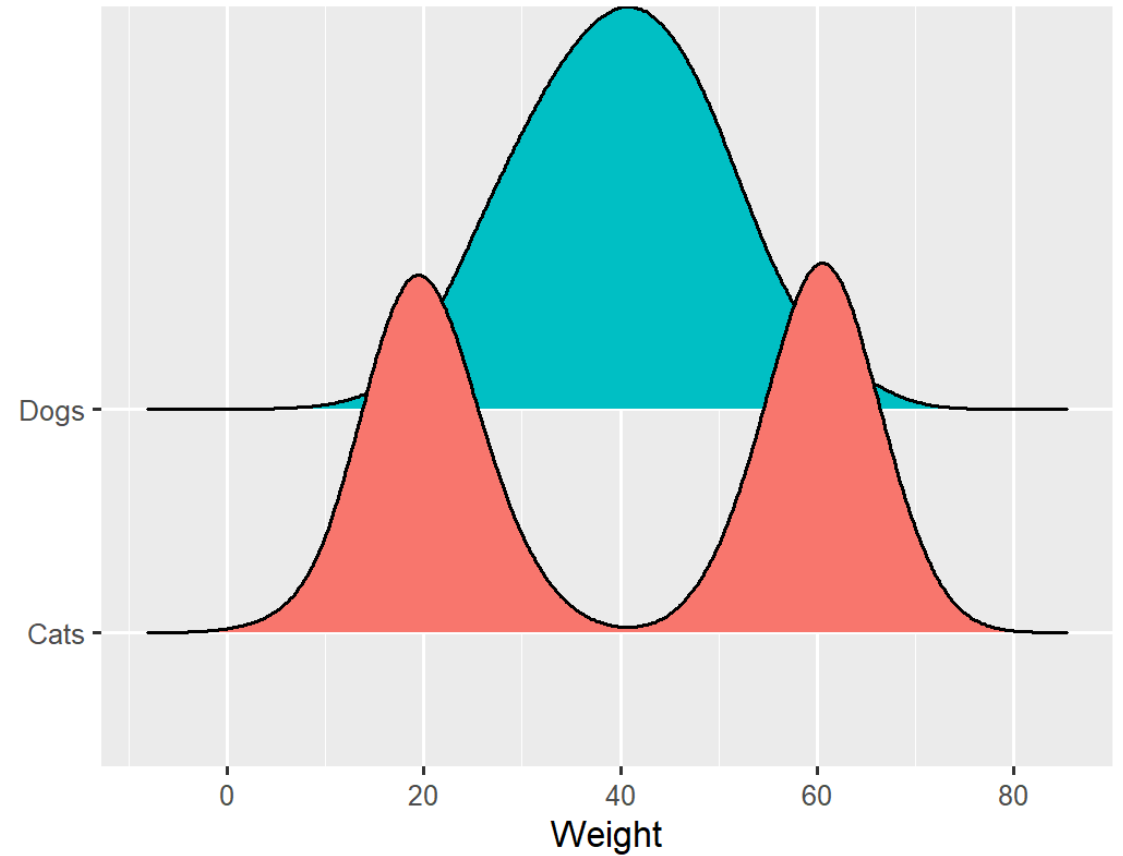
Combinez des violons avec des points

```
ggplot(animals, aes(x = animal_type,  
                    y = weight,  
                    color = animal_type)) +  
  geom_violin(width = 0.5) +  
  geom_point(position = position_jitter(height = 1), size = 1, alpha = 0.5) +  
  labs(x = NULL, y = "Weight") +  
  guides(color = "none")
```



Chevauchement des ridgeplots

```
library(ggribes)\n\n  ggplot(animals, aes(x = weight,\n                      y = animal_type,\n                      fill = animal_type)) +\n    geom_density_ridges() +\n    labs(x = "Weight", y = NULL) +\n    guides(fill = "none")
```



Règles generale

Les graphiques à barres commencent toujours à zéro

N'utilisez pas de barres pour les statistiques descriptives
Vous perdez trop d'informations.

La fin de la barre est souvent tout ce qui compte

Beaucoup d'alternatives

Nous utiliserons une version résumée de l'ensemble de données gapminder comme exemple

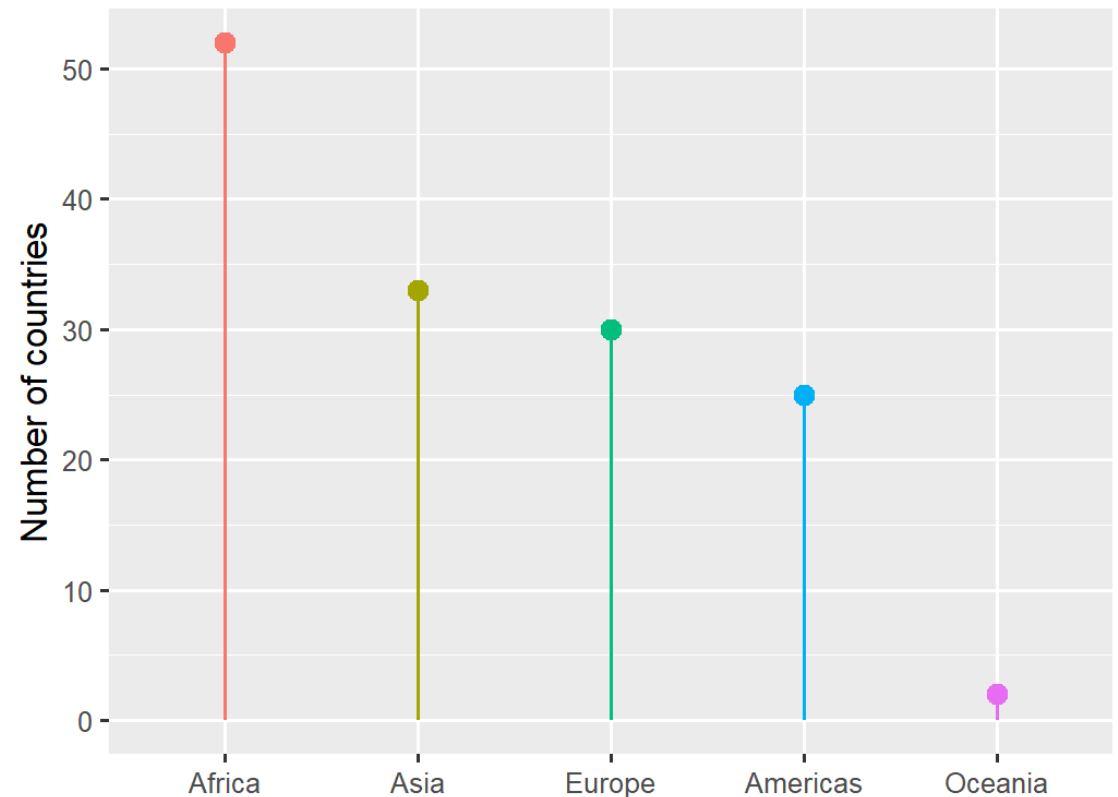
```
library(gapminder)
gapminder_continents <- gapminder %>%
  filter(year == 2007) %>% # juste 2007
  count(continent) %>% # On compte les continents
  arrange(desc(n)) %>% # Trier par nombre de pays
  # Faire du continent un facteur ordonné
  mutate(continent = fct_inorder(continent))

ggplot(gapminder_continents,
       aes(x = continent, y = n, fill = continent)) +
  geom_col() +
  guides(fill = "none") +
  labs(x = NULL, y = "Number of countries")
```

Alternatives : graphiques Lollipop

Puisque la fin de la barre est importante, mettez-la en valeur

```
ggplot(gapminder_continents,  
       aes(x = continent, y = n,  
           color = continent)) +  
  geom_pointrange(aes(ymin = 0, ymax = n)) +  
  guides(color = "none") +  
  labs(x = NULL, y = "Number of countries")
```

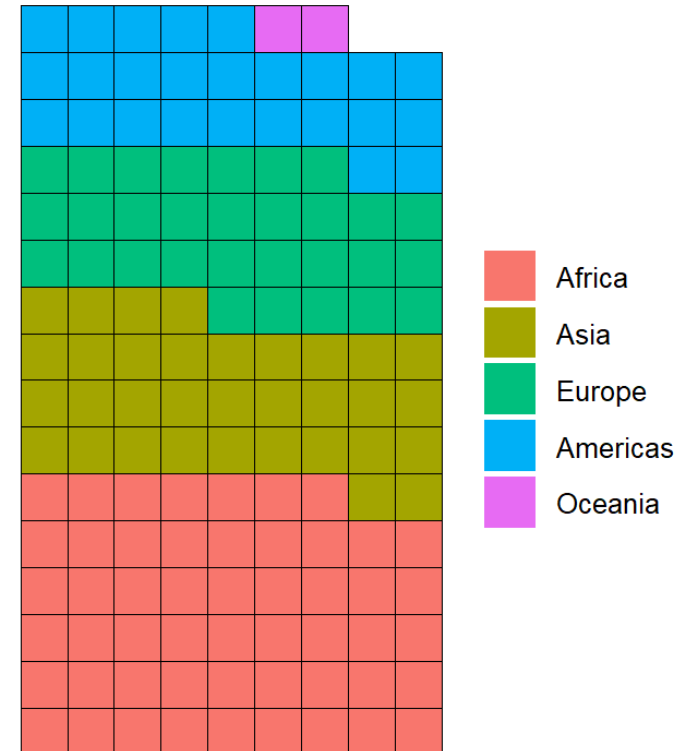


Alternatives : graphiques gaufrés

Afficher les observations individuelles sous forme de carrés

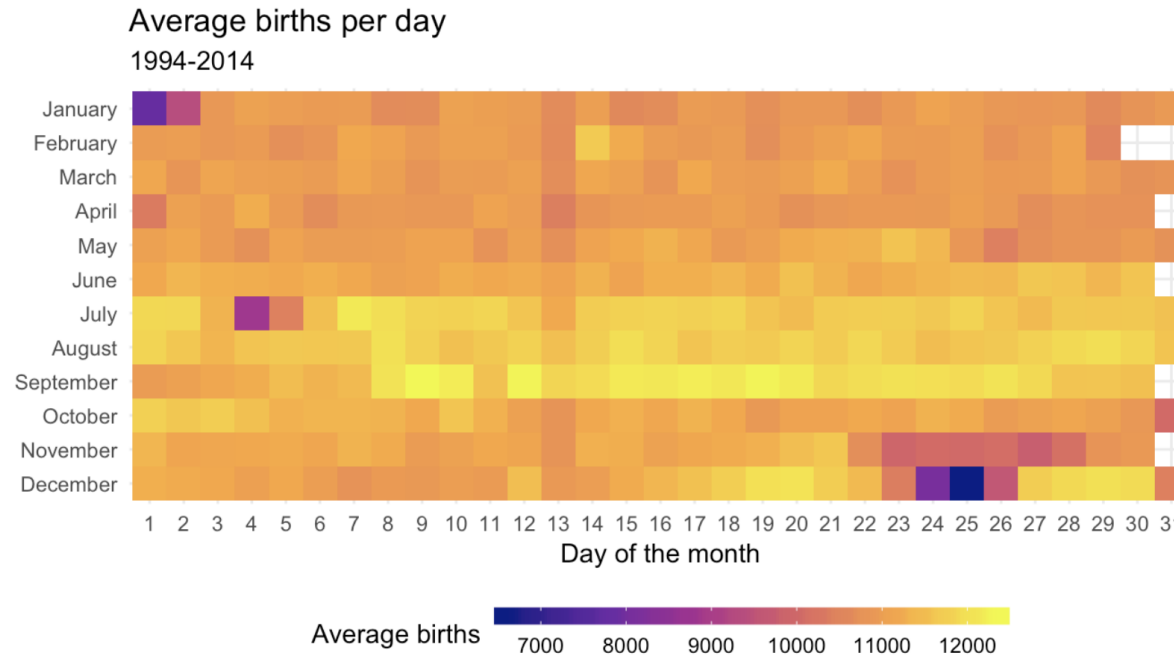
```
# Cela doit être installé d'une manière spéc  
# Exécutez ceci dans votre console  
# devtools::install_github("hrbrmstr/waffle")  
library(waffle)
```

```
ggplot(gapminder_continents,  
       aes(values= n,  
           fill = continent)) +  
  geom_waffle( # geom_waffle() a besoin d'un  
              n_rows = 9, # Il a aussi beau  
              flip = TRUE) +  
  labs(fill = NULL) +  
  coord_equal() + # Rendre tous les carrés  
  theme_void() # Utiliser un thème complet
```



Alternatives : Heatmaps

Si les nombres exacts sont moins importants,
essayez une carte thermique avec `geom_tile()`



Proportions

Pourquoi des proportions ?

Parfois, nous voulons comparer les valeurs de toute une population au lieu de regarder les chiffres bruts

Ne le faites que lorsque cela a un sens analytique !

Comptes de COVID-19 vs proportions

Camemberts

Problèmes de perception avec l'angle et l'espace de remplissage

Seulement correct (ish) s'il y a quelques catégories facilement distinguables

Alternatives

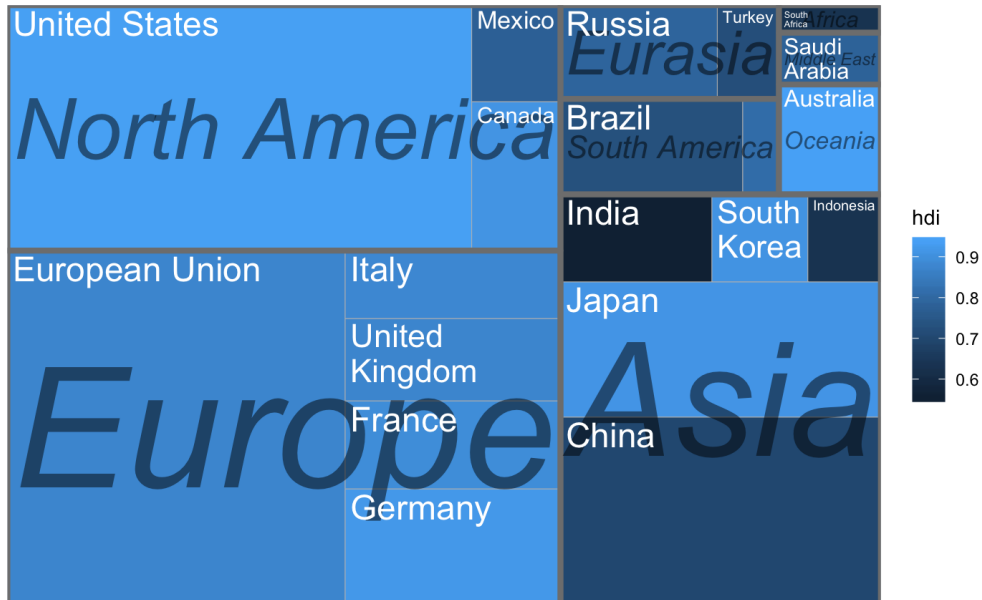
Tracés à barres

L'une des alternatives aux diagrammes à barres

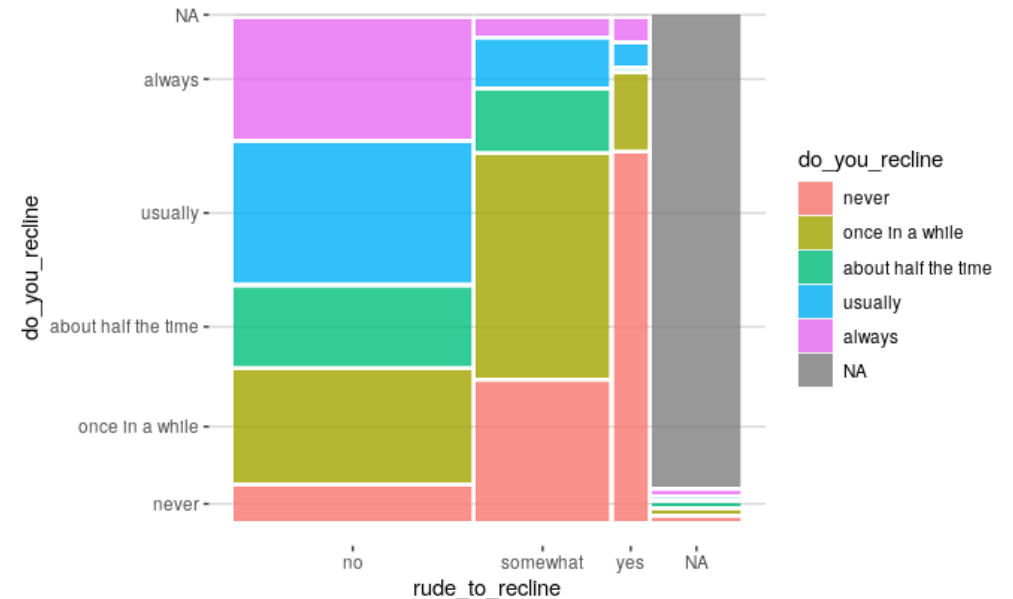
Treemaps et mosaïques
(mais ceux-ci peuvent encore être très difficiles à interpréter)

Treemaps et mosaïques

Treemaps avec **treemapify** package



Mosaïques avec **ggmosaic** package



Alternatives

Tracés à barres

L'une des alternatives aux diagrammes à barres

Treemaps et mosaïques
(mais ceux-ci peuvent encore être très difficiles à interpréter)

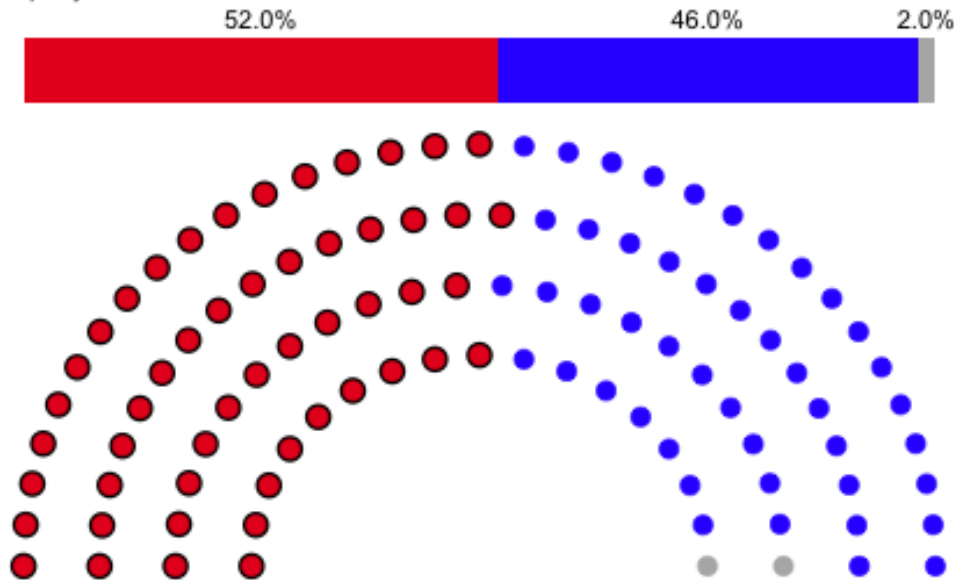
Graphiques spécialisés comme les graphiques parlementaires

Graphique parlementaires

Graphique parlementaire avec **ggparliament** package

United States Senate

The party that has control of the Senate is encircled in black.



UK parliament in 2017

