
CONCEPTS AVANCÉS DE LA VISUALISATION DES DONNÉES ET DE LA CRÉATION DE RAPPORTS

APERÇU

1. Visualisation des données avec ggplot2
2. Introduction aux tableaux de bord

OBJECTIFS D'APPRENTISSAGE

Apprendre à créer des visualisations multivariées avec ggplot2.

Se familiariser avec certaines questions et certains concepts liés aux tableaux de bord et à leur conception.

VISUALISATION DES DONNÉES AVEC GGPLOT2

CONCEPTS AVANCÉS DE LA VISUALISATION DES DONNÉES ET DE LA CRÉATION DE RAPPORTS

GRAPHIQUES EN LANGAGE R

En 2018, il existe quatre fonctions graphiques offertes en langage R, soit :

- *base*
- *grid*
- *lattice*
- *ggplot2*

L'accès varie avec la fonction: les fonctions *base*, *grid* et *lattice* sont incluses dans l'installation de base; les fonctions *grid*, *lattice* et *ggplot2* doivent être chargées avant leur utilisation.

INTRODUCTION À GGLOT2

La fonction *ggplot2* est en fait un jeu d'outils permettant de transformer des données en éléments d'affichage visuels. Elle permet à l'utilisateur de commander les détails de l'affichage graphique.

Aspect le plus important, la fonction *ggplot2* peut servir à établir la **structure logique** du graphique.

Un graphique *ggplot2* comporte deux éléments principaux (et des termes optionnels) :

- une fonction esthétique (**aes** – liens entre les données et les éléments graphiques)
- une fonction de géométrie (**geom** – type de graphique)
- *facets, *coordinates, *scales, *labels, *guides, etc.

GRAMMAIRE DE GGLOT2

1. Tidy Data

```
p <- ggplot(data = gapminder, ...
```

gdp	lifexp	pop	continent
340	65	31	Euro
227	51	200	Amer
909	81	80	Euro
126	40	20	Asia

2. Mapping

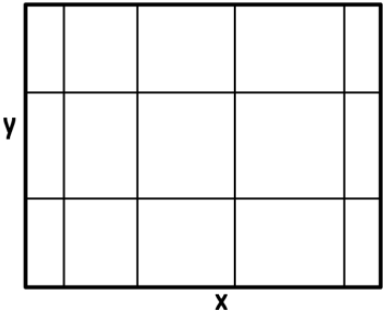
```
p <- ggplot(data = gapminder, mapping =  
  aes(x = gdp, y = lifexp, size = pop,  
      color = continent))
```

3. Geom

```
p + geom_point()
```

4. Co-Ordinates & Scales

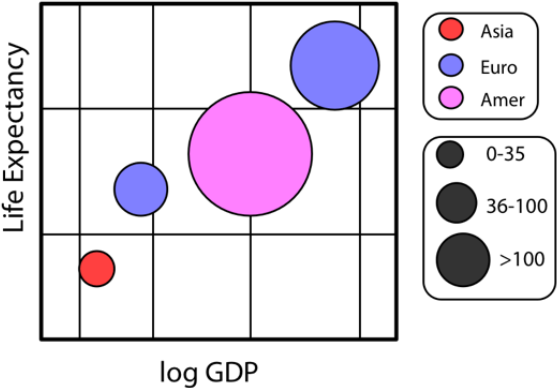
```
p + coord_cartesian() + scale_x_log10()
```



5. Labels & Guides

```
p + labs(x = "log GDP", y = "Life  
Expectancy", title = "A Gapminder Plot")
```

A Gapminder Plot



GRAMMAIRE DE GGLOT2 – GEOM

La source de données et les variables sont précisées au moyen de `ggplot()`.

Les diverses fonctions `geom` précisent la **manière** dont ces variables seront représentées visuellement :

- au moyen de points, de barres, de lignes, de zones ombragées, etc.

Il existe présentement 37 géométries.

GRAMMAIRE DE GGLOT2 – GEOM

Function	Adds	Options
<code>geom_bar()</code>	Bar chart	color, fill, alpha
<code>geom_boxplot()</code>	Box plot	color, fill, alpha, notch, width
<code>geom_density()</code>	Density plot	color, fill, alpha, linetype
<code>geom_histogram()</code>	Histogram	color, fill, alpha, linetype, binwidth
<code>geom_hline()</code>	Horizontal lines	color, alpha, linetype, size
<code>geom_jitter()</code>	Jittered points	color, size, alpha, shape
<code>geom_line()</code>	Line graph	color, alpha, linetype, size
<code>geom_point()</code>	Scatterplot	color, alpha, shape, size
<code>geom_rug()</code>	Rug plot	color, side
<code>geom_smooth()</code>	Fitted line	method, formula, color, fill, linetype, size
<code>geom_text()</code>	Text annotations	Many; see the help for this function
<code>geom_violin()</code>	Violin plot	color, fill, alpha, linetype
<code>geom_vline()</code>	Vertical lines	color, alpha, linetype, size

GRAMMAIRE DE GGLOT2 – GEOM

Option	Specifies
color	colour of points, lines, and borders around filled regions
fill	colour of filled areas such as bars and density regions
alpha	transparency of colors, ranging from 0 (fully transparent) to 1 (opaque)
linetype	pattern for lines (1 = solid, 2 = dashed, 3 = dotted, 4 = dotdash, 5 = longdash, 6 = twodash)
size	point size and line width
shape	point shapes (same as pch, with 0 = open square, 1 = open circle, 2 = open triangle, and so on)
position	position of plotted objects such as bars and points. For bars, “dodge” places grouped bar charts side by side, “stacked” vertically stacks grouped bar charts, and “fill” vertically stacks grouped bar charts and standardizes their heights to be equal; for points, “jitter” reduces point overlap
binwidth	bin width for histograms
notch	indicates whether box plots should be notched (TRUE/FALSE)
sides	placement of rug plots on the graph (“b” = bottom, “l” = left, “t” = top, “r” = right, “bl” = both bottom and left, and so on)
width	width of box plots

GRAMMAIRE DE GGLOT2 – GEOM()

```
library("ggplot2")
data(singer, package="lattice")
# Using data from the 1979 ed. of the
# New York Choral Society

# Histogram of heights
ggplot(singer, aes(x=height)) +
  geom_histogram()

# Boxplot of heights by voice part
ggplot(singer, aes(x=voice.part, y=height)) +
  geom_boxplot()
```

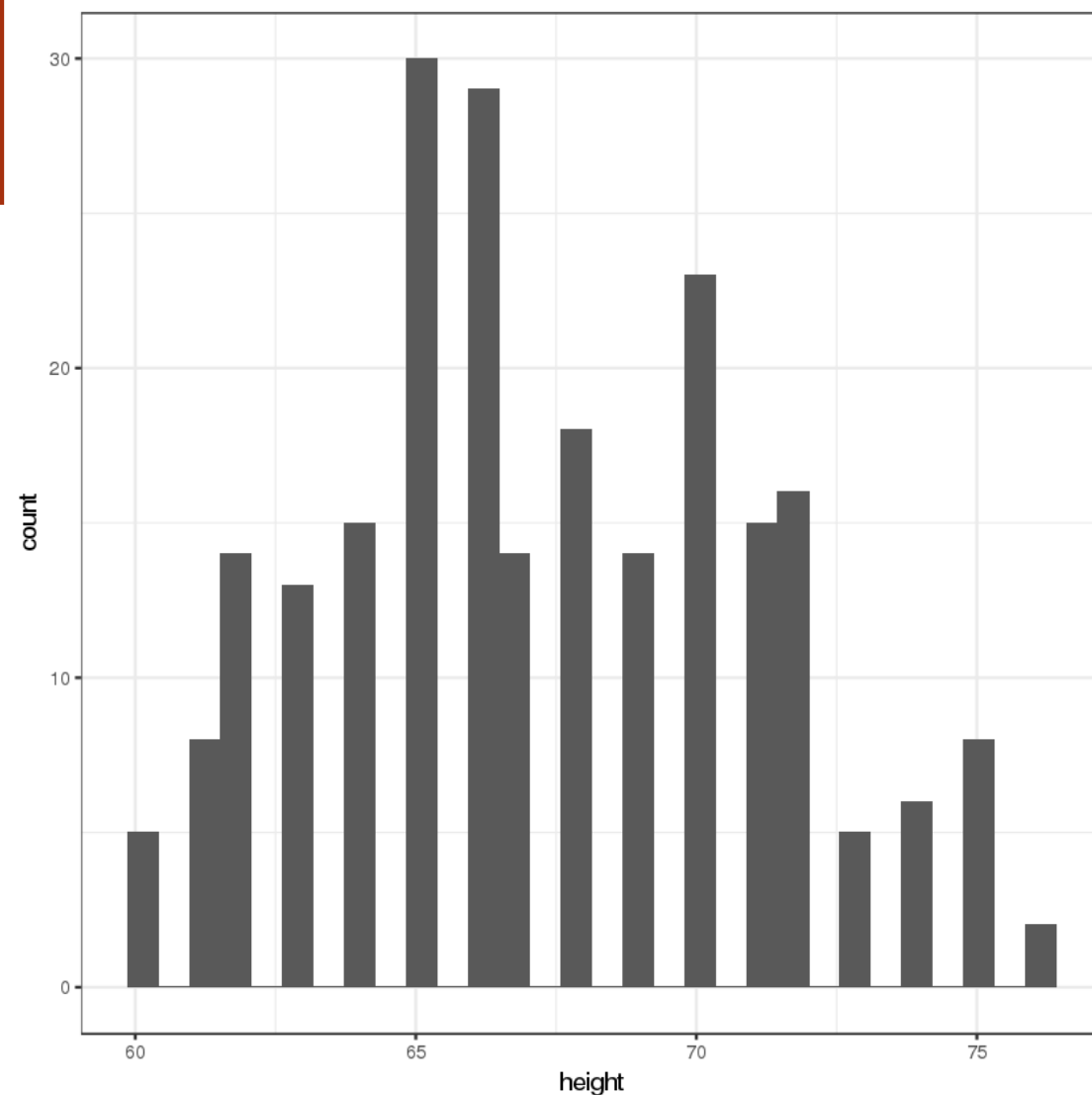
Selon vous, quelle sera la forme du graphique?

GRAMMAIRE DE GGLOT2 – GEOM()

```
library("ggplot2")
data(singer, package="lattice")
# Using data from the 1979 ed. of the
# New York Choral Society

# Histogram of heights
ggplot(singer, aes(x=height)) +
  geom_histogram()

# Boxplot of heights by voice part
ggplot(singer, aes(x=voice.part, y=height)) +
  geom_boxplot()
```

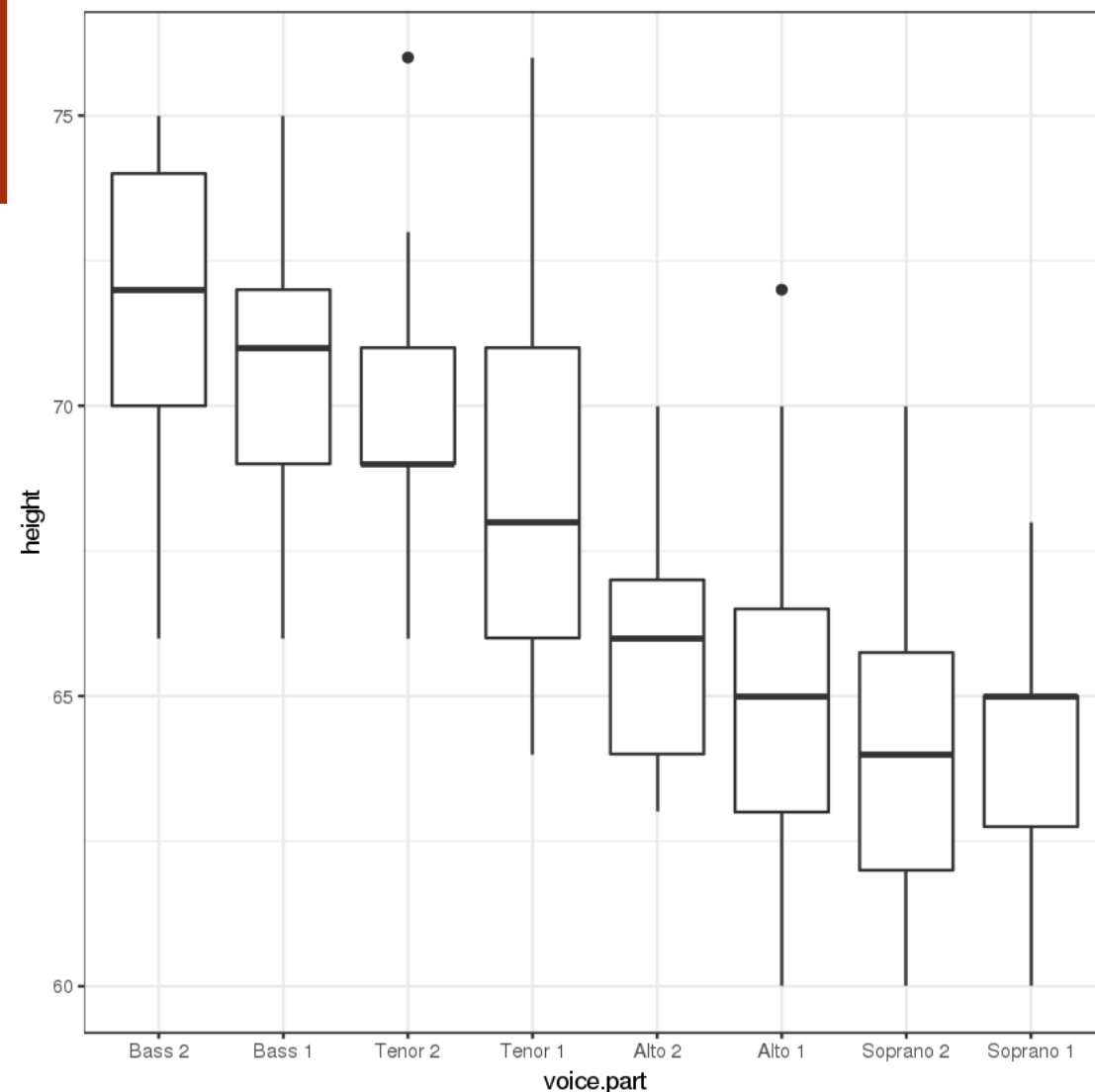


GRAMMAIRE DE GGLOT2 – GEOM()

```
library("ggplot2")
data(singer, package="lattice")
# Using data from the 1979 ed. of the
# New York Choral Society

# Histogram of heights
ggplot(singer, aes(x=height)) +
  geom_histogram()

# Boxplot of heights by voice part
ggplot(singer, aes(x=voice.part, y=height)) +
  geom_boxplot()
```



GRAMMAIRE DE GGLOT2 – GEOM()

```
library(ggplot2)
data(Salaries, package="car")
# Using data on salaries of a sample of
# US university professors (2018-2019)
# var: rank, sex, yrs.since.phd, yrs.service, salary

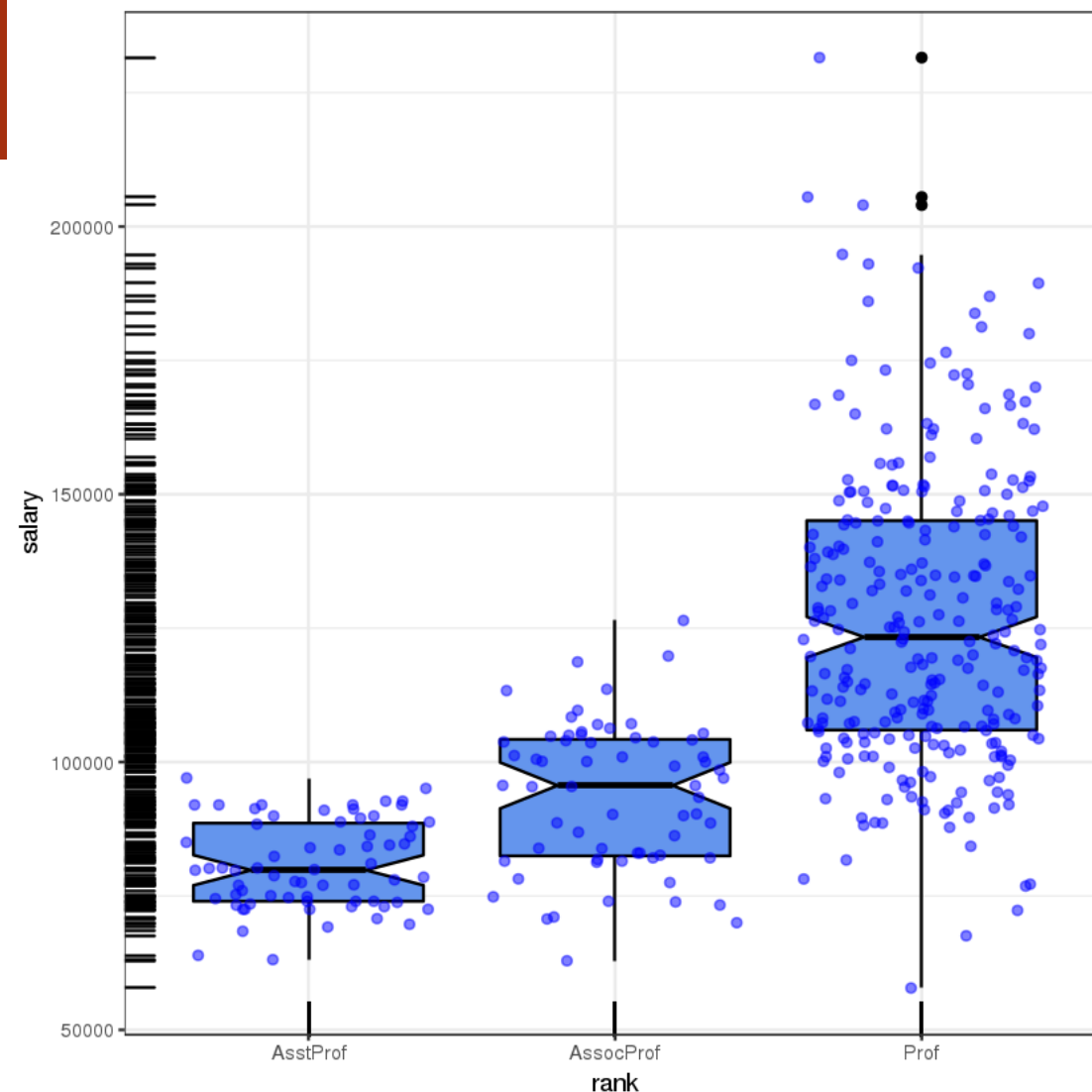
ggplot(Salaries, aes(x=rank, y=salary)) +
  geom_boxplot(fill="cornflowerblue", color="black", notch=TRUE) +
  geom_point(position="jitter", color="blue", alpha=.5) +
  geom_rug(side="l", color="black")
```

Selon vous, quelle sera la forme du graphique?

GRAMMAIRE DE GGLOT2 – GEOM()

```
library(ggplot2)
data(Salaries, package="car")
# Using data on salaries of a sample of
# US university professors (2018-2019)
# var: rank, sex, yrs.since.phd, yrs.service, salary

ggplot(Salaries, aes(x=rank, y=salary)) +
  geom_boxplot(fill="cornflowerblue", color="black", notch=TRUE) +
  geom_point(position="jitter", color="blue", alpha=.5) +
  geom_rug(side="l", color="black")
```



GRAMMAIRE DE GGLOT2 – ESTHÉTIQUE

L'**esthétique** désigne les attributs affichés des données.

Vous devez faire correspondre une donnée à un attribut (comme la taille ou la forme d'un repère) et créer la légende appropriée.

Vous précisez l'esthétique au moyen de la fonction `aes ()`.

Vous pouvez préciser l'esthétique dans la fonction *data* ou *geom*. Si vous précisez l'esthétique dans la fonction *data*, l'esthétique vise alors toutes les fonctions *geom* précisées.

GRAMMAIRE DE GGLOT2 – ESTHÉTIQUE

Les caractéristiques esthétiques offertes avec `geom_point()` (scatterplot), p. ex., sont les suivantes :

- `x, y, alpha, color, fill, shape, size`

Il existe une **différence importante** entre les caractéristiques (comme la couleur et la forme) selon qu'elles sont précisées à l'intérieur et à l'extérieur de la fonction `aes()` :

- à l'intérieur : la couleur ou la forme choisie repose automatiquement sur les données
- à l'extérieur : la caractéristique ne vise pas les données.

GRAMMAIRE DE GGLOT2 – AES()

```
library(ggplot2)
# Using the mpg dataset

# specifying characteristics inside aes()
ggplot(mpg, aes(cty, hwy)) +
  geom_point(aes(colour = class))

# specifying characteristics inside aes()
ggplot(mpg, aes(cty, hwy)) +
  geom_point(colour = "red")
```

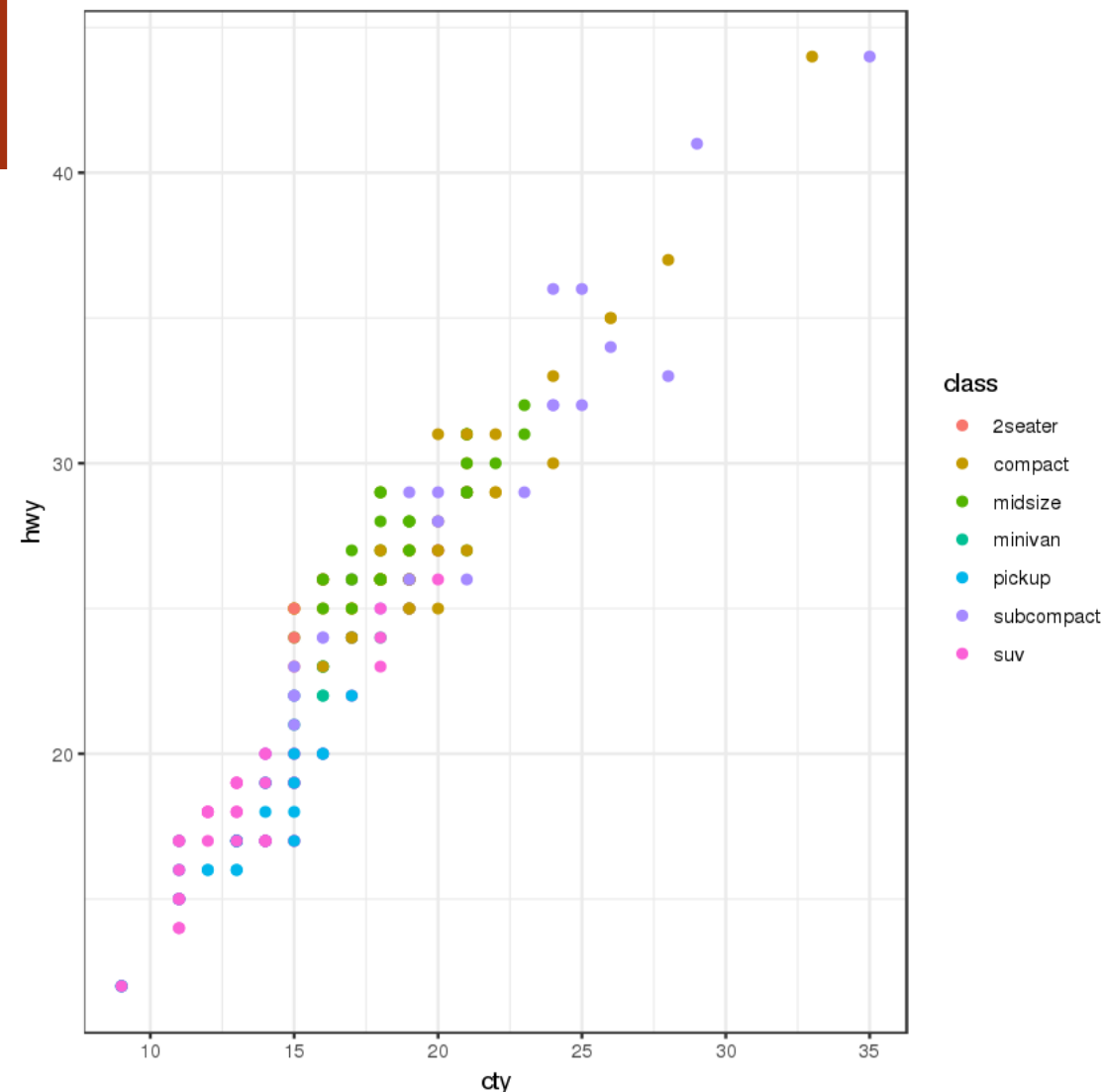
Selon vous, quelle sera la forme du graphique?

GRAMMAIRE DE GGLOT2 – AES()

```
library(ggplot2)
# Using the mpg dataset

# specifying characteristics inside aes()
ggplot(mpg, aes(cty, hwy)) +
  geom_point(aes(colour = class))

# specifying characteristics inside aes()
ggplot(mpg, aes(cty, hwy)) +
  geom_point(colour = "red")
```

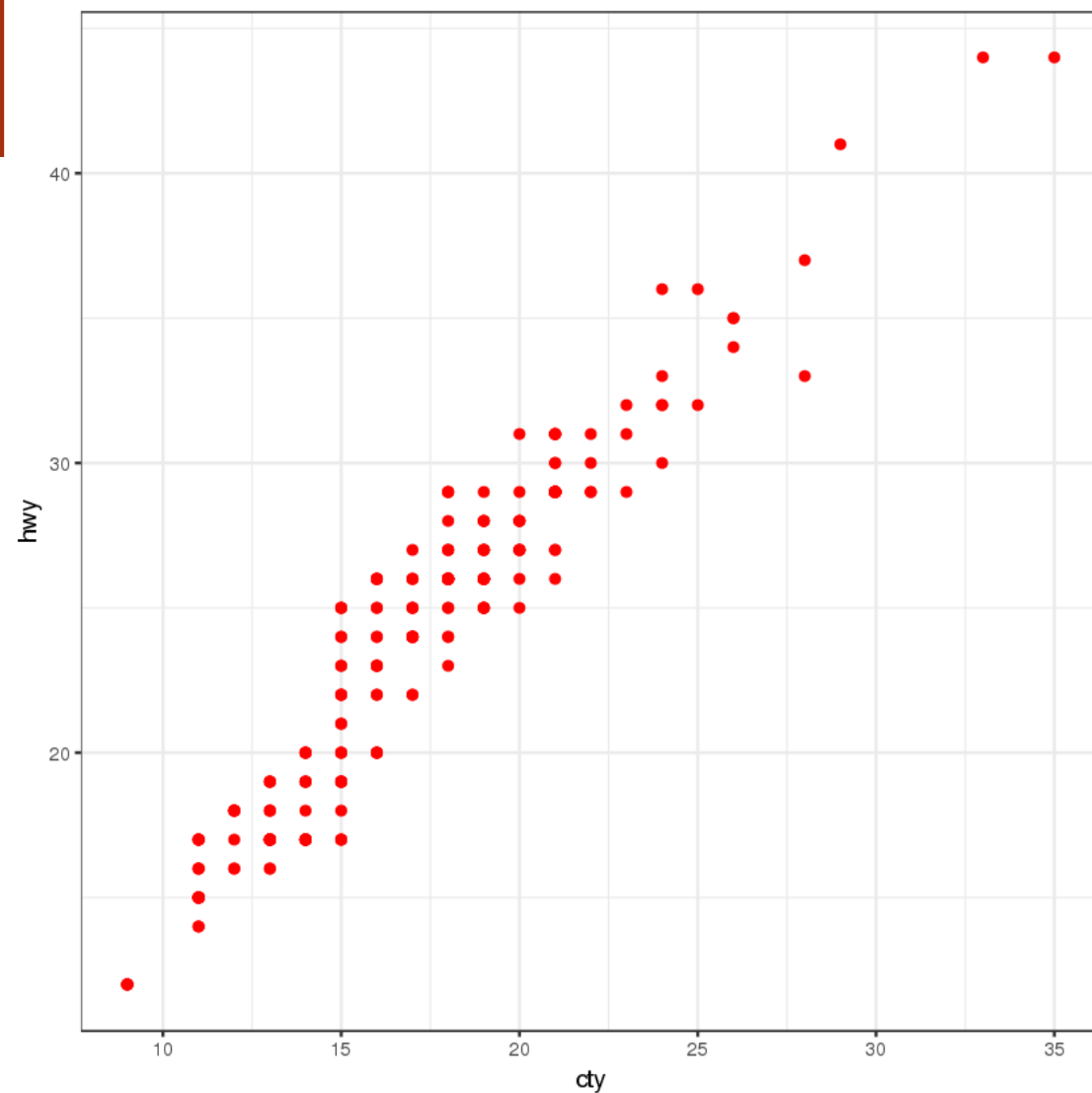


GRAMMAIRE DE GGLOT2 – AES()

```
library(ggplot2)
# Using the mpg dataset

# specifying characteristics inside aes()
ggplot(mpg, aes(cty, hwy)) +
  geom_point(aes(colour = class))

# specifying characteristics inside aes()
ggplot(mpg, aes(cty, hwy)) +
  geom_point(colour = "red")
```



GRAMMAIRE DE GGLOT2 – FACETTES

Avec *ggplot2*, de petits multiples sont désignés des **facettes** :

- `facet_wrap()`, `facet_grid()`

Par défaut, tous les graphiques (un pour chaque facteur) partagent les mêmes axes (*scale-wise*).

La séparation du graphique en une suite de graphiques plus petits et disposés côte à côte facilite les comparaisons.

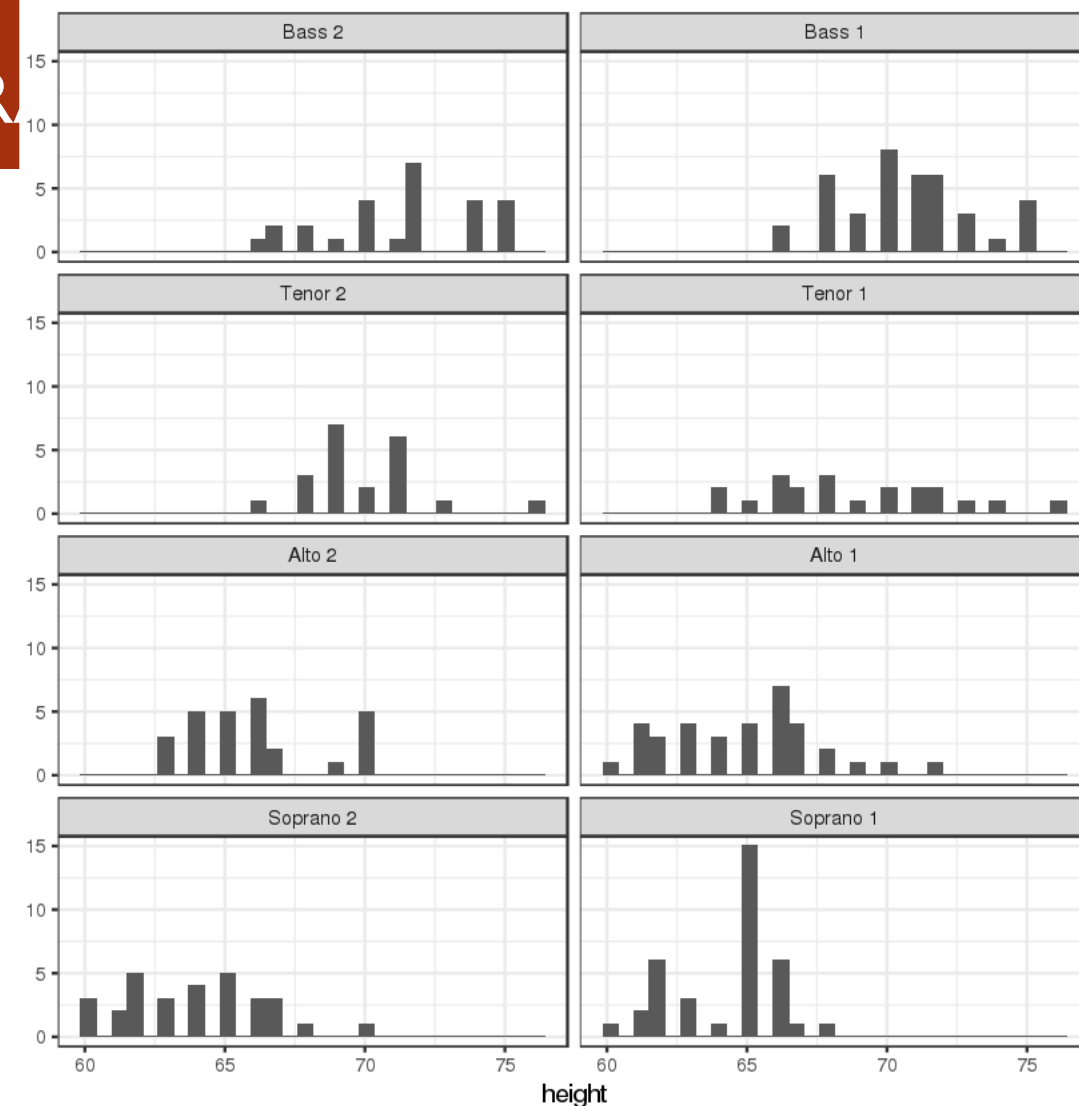
GRAMMAIRE DE GGLOT2 – FACET_WRAP()

```
data(singer, package="lattice")  
library(ggplot2)  
ggplot(data=singer, aes(x=height)) +  
  geom_histogram() +  
  facet_wrap(~voice.part, nrow=4)
```

Selon vous, quelle sera la forme du graphique?

GRAMMAIRE DE GGLOT2 – FACET_WRAP

```
data(singer, package="lattice")  
library(ggplot2)  
ggplot(data=singer, aes(x=height)) +  
  geom_histogram() +  
  facet_wrap(~voice.part, nrow=4)
```



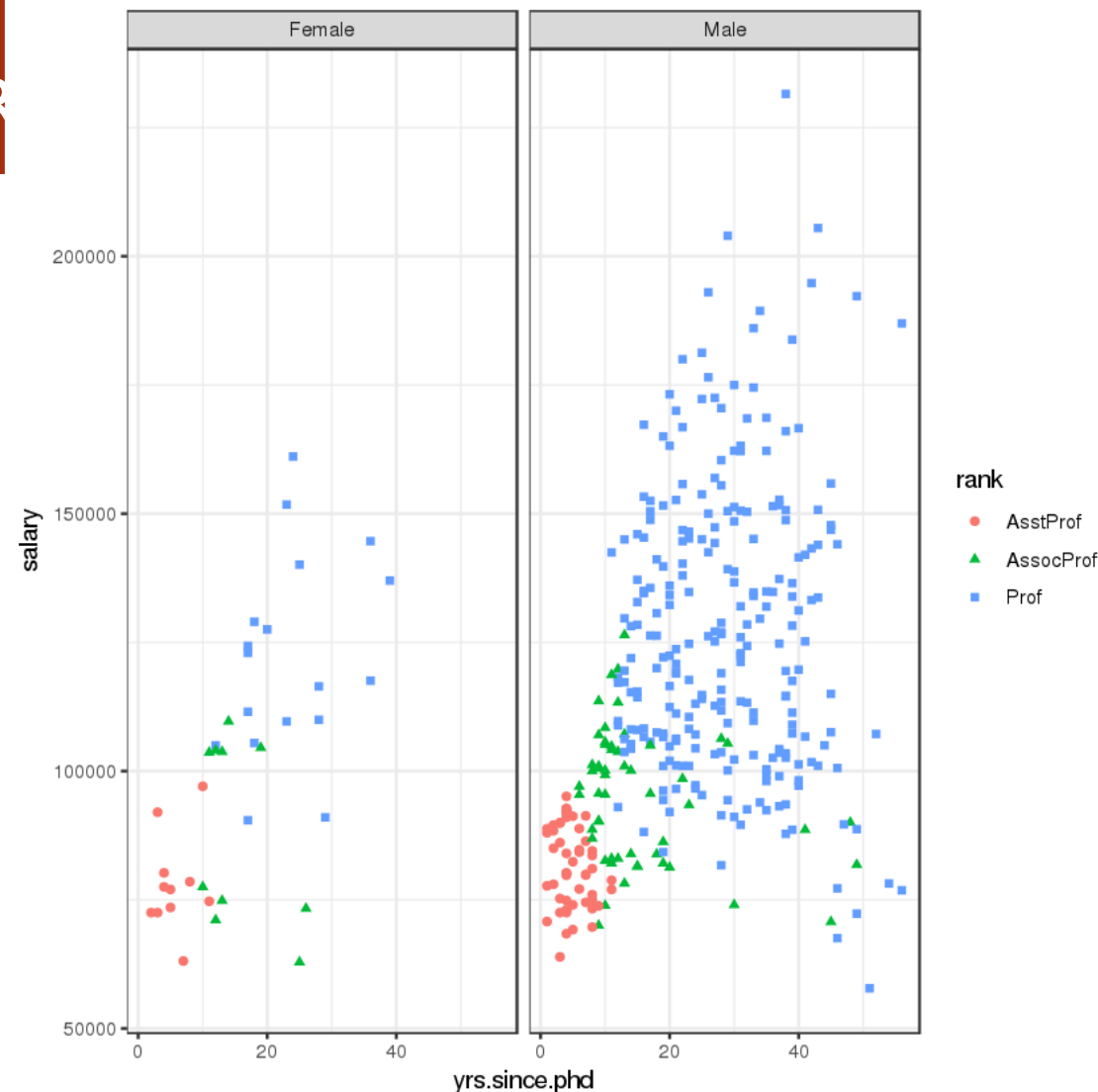
GRAMMAIRE DE GGLOT2 – FACET_GRID()

```
data(Salaries, package="car")  
library(ggplot2)  
ggplot(Salaries, aes(x=yrs.since.phd,  
  y=salary, color=rank, shape=rank)) +  
  geom_point() +  
  facet_grid(~sex)
```

Selon vous, quelle sera la forme du graphique?

GRAMMAIRE DE GGLOT2 – FACET_GRID

```
data(Salaries, package="car")
library(ggplot2)
ggplot(Salaries, aes(x=yrs.since.phd,
  y=salary, color=rank, shape=rank)) +
  geom_point() +
  facet_grid(~sex)
```



EXERCICES

Des explications détaillées et des exemples font l'objet du document *A ggplot2 Primer* accessible à <https://www.data-action-lab.com/2018/11/12/a-ggplot2-primer/>.

Avec un camarade de classe, exécutez quelques exemples fournis dans le cahier R. Il n'est pas nécessaire que vous compreniez tous les détails des visualisations, mais il est utile de connaître les types de résultats possibles.

Créez cinq visualisations au moyen de *ggplot2* à partir des ensembles de données des collisions et de la prolifération d'algues.

INTRODUCTION AUX TABLEAUX DE BORD

CONCEPTS AVANCÉS DE LA VISUALISATION DES DONNÉES ET DE LA CRÉATION DE RAPPORTS

« Si un arbre tombe dans la forêt et si personne ne l'entend tomber, est-ce qu'il fait du bruit? »

(vieille énigme)

RAPPORT ET DÉPLOIEMENT

Une analyse est bonne seulement si on la **transmet** ou si on la **déploie**.

Questions essentielles :

- Qui reçoit le rapport?
- Quels sont les flux de travail menant à sa création?
- Est-ce que les données peuvent donner lieu à des politiques utiles?

La création automatique de rapports devrait **régulièrement** faire l'objet d'une vérification et d'une validation.

RAPPORT ET DÉPLOIEMENT

La **communication** devrait avoir lieu à diverses étapes du projet, et non seulement à son achèvement :

- vous devez maintenir les commanditaires et les clients au fait des principaux points;
- vous pouvez délaissé les détails techniques, mais vous devez tout de même les documenter.

Le **scénario idéal** consiste à utiliser un logiciel d'analyse qui permet aussi de créer des rapports :

- minimise l'erreur humaine liée à la fonction copier-coller;
- supprime le besoin de maintenir la séparation entre l'analyse et la création de rapports;
- facilite le partage du travail avec les autres membres du projet.

Vous pouvez simplifier davantage le processus en procédant à un déploiement directement sur le web.

DISCUSSION

Quels sont vos outils de création de rapport favoris?

Jusqu'à quel point devez-vous tester un produit avant de le déployer?

Quel est le coût du déploiement d'un produit défectueux?

TABLEAUX DE BORD

Un **tableau de bord** est un affichage visuel des données qui sert à surveiller les états et à faciliter la compréhension.

Exemples :

- affichage interactif qui permet à l'utilisateur d'explorer les réclamations d'assurance automobile selon la ville, la province, l'âge du conducteur, etc.;
- fichier PDF qui montre les principaux paramètres de vérification et qui est envoyé par courriel chaque semaine au DG d'un ministère;
- écran monté au mur qui montre en temps réel les statistiques d'un centre d'appel;
- application mobile qui permet aux administrateurs d'un hôpital de voir les délais d'attente chaque heure et chaque jour pour l'année courante et l'année précédente.

QUELQUES QUESTIONS À PRENDRE EN COMPTE

Sur le tableau de bord d'une automobile, l'automobiliste doit comprendre **d'un coup d'œil** un petit nombre d'**indicateurs importants** (vitesse, niveau d'essence, phares, etc.). Un tableau de bord qui ne tient pas compte de ces deux caractéristiques peut donner lieu à des conséquences catastrophiques.

Vous devez répondre aux questions suivantes avant de concevoir un tableau de bord :

- Qui est l'**utilisateur** du tableau de bord?
- Quels renseignements doit **transmettre** le tableau de bord?
- Quelles données (catégories) seront utilisées?
- Qu'est-ce qui **figurera** dans le tableau de bord?
- Comment le tableau de bord va-t-il **aider** l'utilisateur?



LIGNES DIRECTRICES LIÉES À LA CONCEPTION D'UN TABLEAU DE BORD

Nick Smith propose les six règles d'or suivantes :

- **Tenez compte du public** (Qui voulez-vous informer? Est-ce que le DG a réellement besoin de savoir que les serveurs fonctionnent à 88 p. cent de leur capacité?)
- **Sélectionnez le bon type de tableau de bord** (opérationnel, stratégique, analytique)
- **Groupez les données logiquement, utilisez intelligemment l'espace** (séparez les secteurs fonctionnels : produit, ventes et marketing, finances, personnes, etc.)
- **Utilisez des données adaptées au public** (portée et étendue des données, différents tableaux de bord pour différents services, etc.)
- **Évitez d'encombrer le tableau de bord** (présentez seulement les paramètres les plus importants)
- **Actualisez les données à la fréquence appropriée** (en temps réel, chaque jour, chaque semaine, chaque mois, etc.)



✔ Meets or Exceeds Target ⬆ Near Target ✖ Needs Improvement ⬇ Measuring ⬇ Collecting Data

TABLEAU DE BORD DES PARAMÈTRES LIÉS AUX COURS – SCÉNARIO

Vous êtes à la tête d'un département d'études. Vous voulez savoir :

- la note attribuée à un cours d'un professeur comparativement à celle attribuée aux autres cours du département et de l'université en général;
- la charge générale du cours, le nombre d'étudiants et la hausse ou la diminution du nombre d'inscriptions à un cours en particulier;
- le nombre de cours qu'un enseignant a donné au fil du temps;
- les renseignements détaillés sur les notes attribuées au cours le plus récent et les commentaires de l'enseignant.

De quel type de données avez-vous besoin? Comment disposeriez-vous les paramètres dans le tableau de bord pour aider à répondre à ces questions?

TABLEAU DE BORD DES PARAMÈTRES LIÉS AUX COURS – DONNÉES

Year	Semester	Students	Average
'12	S	42	52
	F	16	52
'13	S	71	52
	US	14	52
	F	27	52
'14	S	69	52
	S	55	52
	US	28	52
	F	27	52
	F	61	52
'15	S	46	52
	S	80	52
	US	43	52
	F	61	52
	F	69	52
'16	S	62	52
	S	80	52
	US	50	52
	F	62	52
	F	65	52
	F	69	52

1097

year	enrollments
'12	58
'13	112
'14	240
'15	299
'16	388
	687

Year	# classes
'12	2
'13	3
'14	5
'15	5
'16	6
	21

Year	Semester	Rating
'12	S	6.6
	F	6.5
'13	S	6.7
	US	7.7
	F	6.9
'14	S	6.4
	S	6.7
	US	7.5
	F	7.3
	F	7
'15	S	6.4
	S	7
	US	6.8
	F	7.3
	F	7.7
		7.7

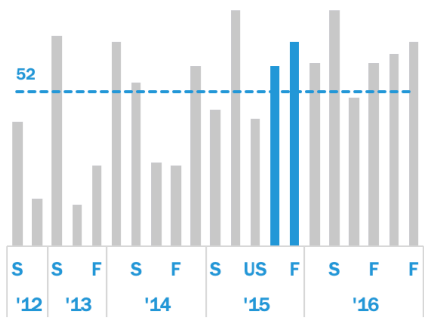
TABLEAU DE BORD DES PARAMÈTRES LIÉS AUX COURS – DONNÉES

Semesters	Questions	Mean Rating	Entity	Shaffer	BANA	College
2015 Fall Semester 002	The instructor was well organized	7.5	Shaffer	7.5	6.8	7
2015 Fall Semester 002	The instructor communicated clearly	7.6	Shaffer	7.6	6.5	6.9
2015 Fall Semester 002	The instructor interacted well with students	7.7	Shaffer	7.7	6.6	7
2015 Fall Semester 002	The Instructor graded fairly	7.6	Shaffer	7.6	6.8	7.1
2015 Fall Semester 002	I developed specific skills and competencies	7.2	Shaffer	7.2	6.3	6.5
2015 Fall Semester 002	Overall, this instructor was excellent	7.7	Shaffer	7.7	6.4	6.8
2015 Fall Semester 002	Overall, this was an excellent course	7.4	Shaffer	7.4	5.9	6.4
2015 Fall Semester 001	The instructor was well organized	7.3	Shaffer	7.3	7	6.9
2015 Fall Semester 001	The instructor communicated clearly	7.4	Shaffer	7.4	6.7	6.7
2015 Fall Semester 001	The instructor interacted well with students	7.3	Shaffer	7.3	6.8	6.8
2015 Fall Semester 001	The Instructor graded fairly	7.5	Shaffer	7.5	7.1	7
2015 Fall Semester 001	I developed specific skills and competencies	6.9	Shaffer	6.9	6.8	6.7
2015 Fall Semester 001	Overall, this instructor was excellent	7.3	Shaffer	7.3	6.7	6.7
2015 Fall Semester 001	Overall, this was an excellent course	7.1	Shaffer	7.1	6.6	6.5

Course Metrics

[<https://bigbookofdashboards.com/dashboards.html>]

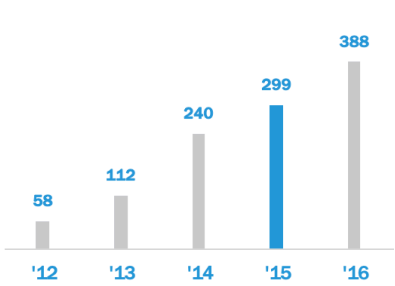
Students



1097

Total Students in five years

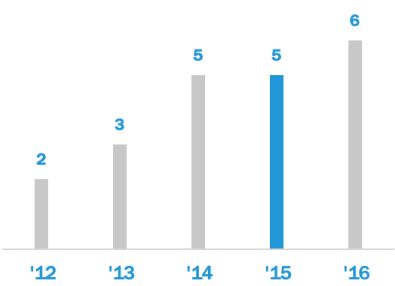
Enrollments



687

Total Students in 2015-2016

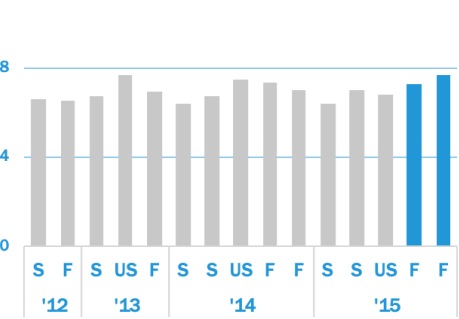
Classes



21

Total Classes in five years

Ratings



7.7 of 8

Most recent instructor rating (out of 8.0)

Semesters

2015 Fall Semester 001

I developed specific skills and competencies
Overall, this was an excellent course
The instructor communicated clearly
The Instructor graded fairly
The instructor was well organized
The instructor interacted well with students
Overall, this instructor was excellent

2015 Fall Semester 002

I developed specific skills and competencies
Overall, this was an excellent course
The instructor communicated clearly
The Instructor graded fairly
The instructor was well organized
The instructor interacted well with students
Overall, this instructor was excellent

Questions

● BANA ■ College ● Shaffer

Ratings

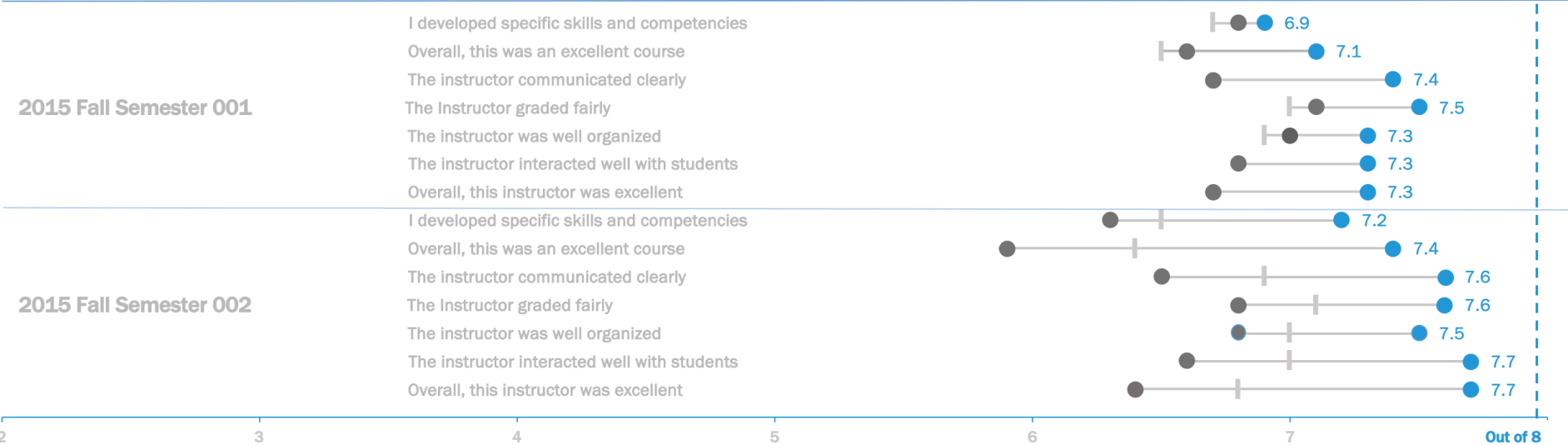


TABLEAU DE BORD – POINTS FORTS

Principaux paramètres faciles à voir

Palette de couleurs simplifiée

Possibilité d'un tableau statique ou interactif

Clarté du sommaire et des détails

DISCUSSION

Aucun tableau de bord n'est parfait – aucun ensemble de graphiques ne conviendra à toutes les personnes qui les utiliseront.

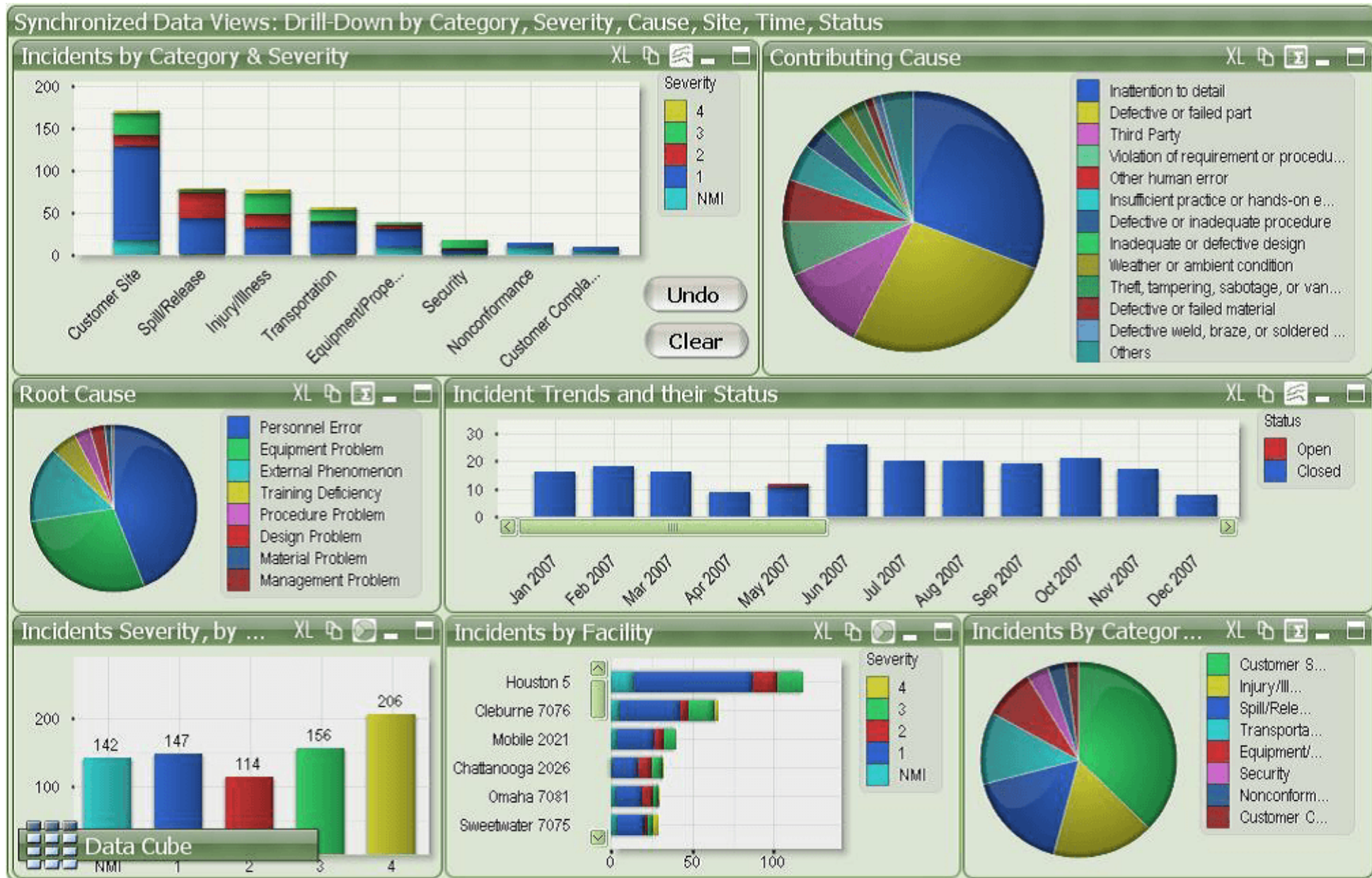
Un tableau de bord doit transmettre la **vérité** et être **fonctionnel**; un tableau de bord **élégant** (agréable, plaisant) convient davantage.

Aucun tableau de bord **n'est complet**. Un bon tableau de bord mène tout de même à un cul-de-sac, mais il devrait permettre à l'utilisateur de se demander « Pourquoi? Quelle est la cause fondamentale du problème? ».

Outils : Excel, Power BI, Tableau, R + Shiny, Geckoboard, Matillion, etc.

EXERCICE

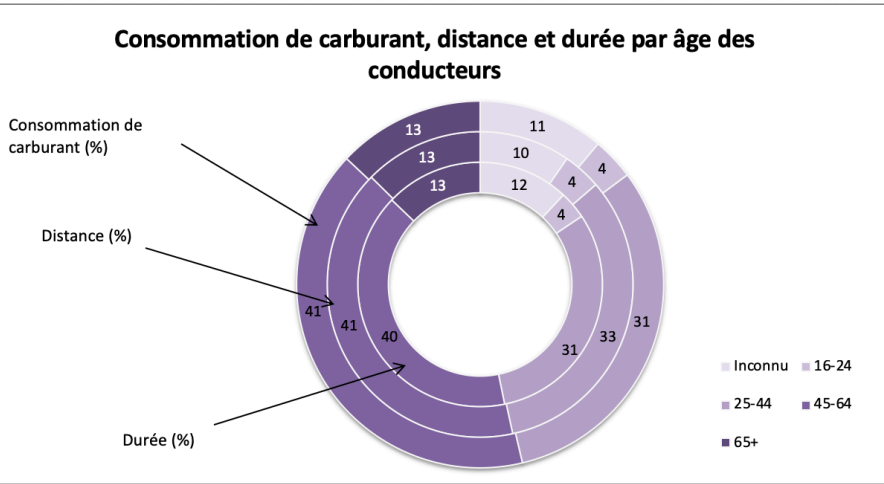
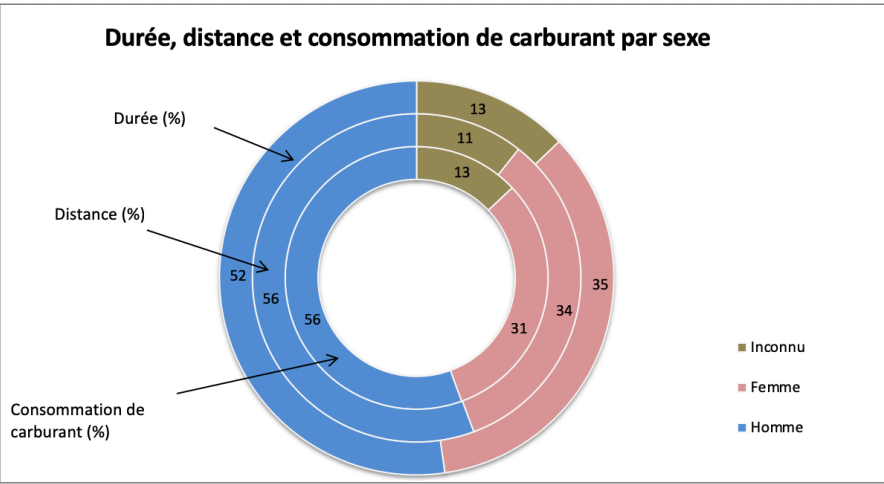
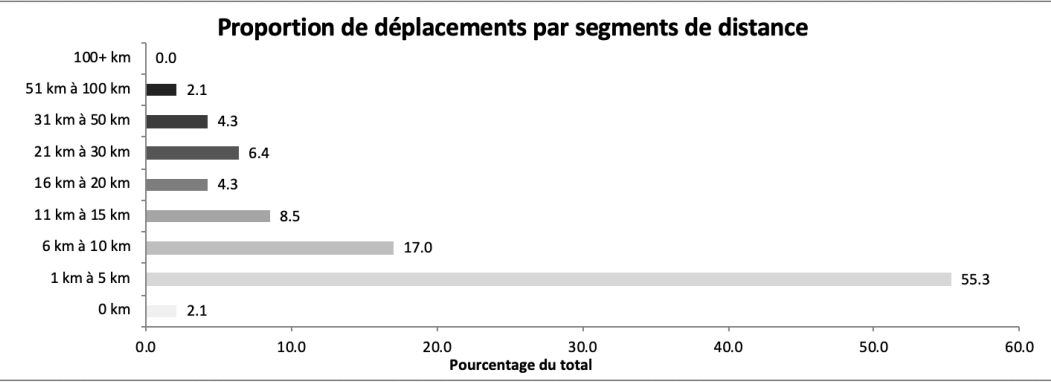
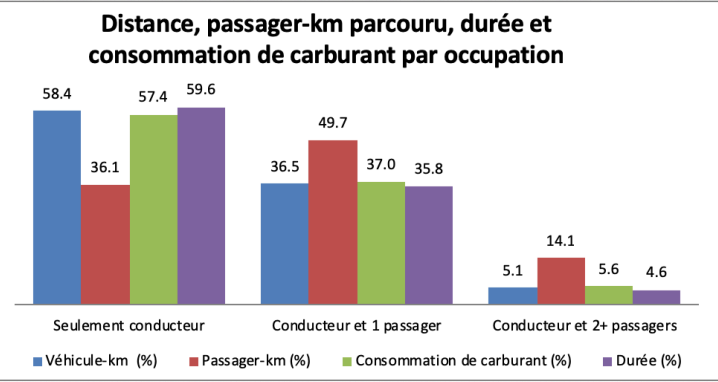
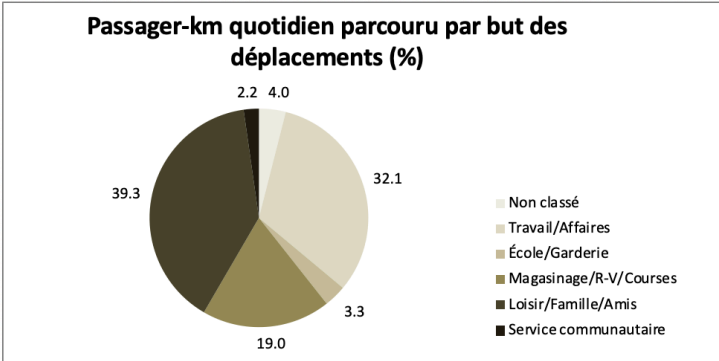
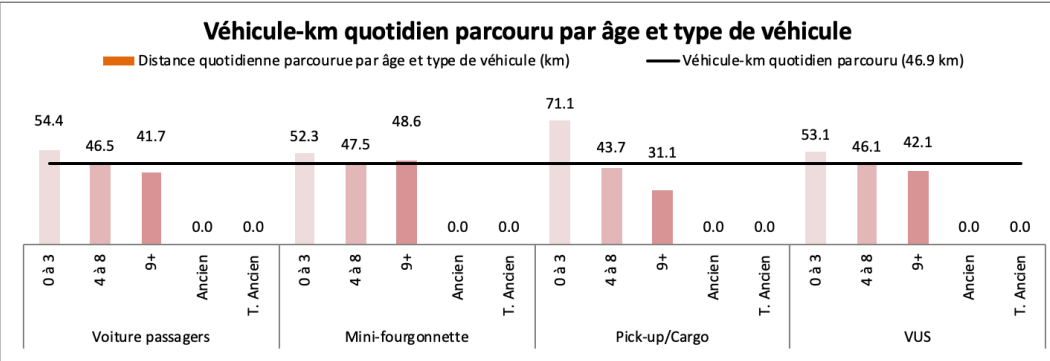
Examinez les prochains tableaux de bord. D'un coup d'œil, pouvez-vous en déterminer le public? Quels sont leurs points forts? Quelles sont leurs limites? Comment pourriez-vous les améliorer?





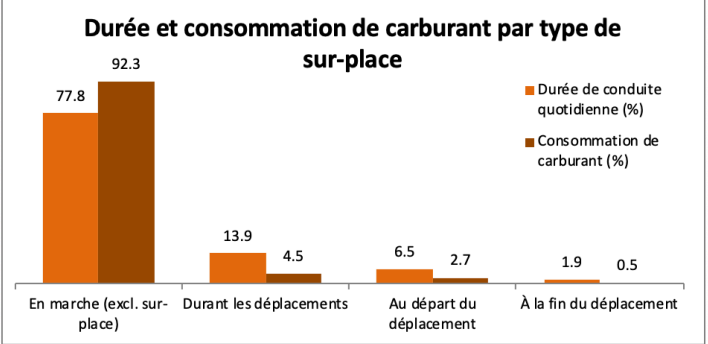
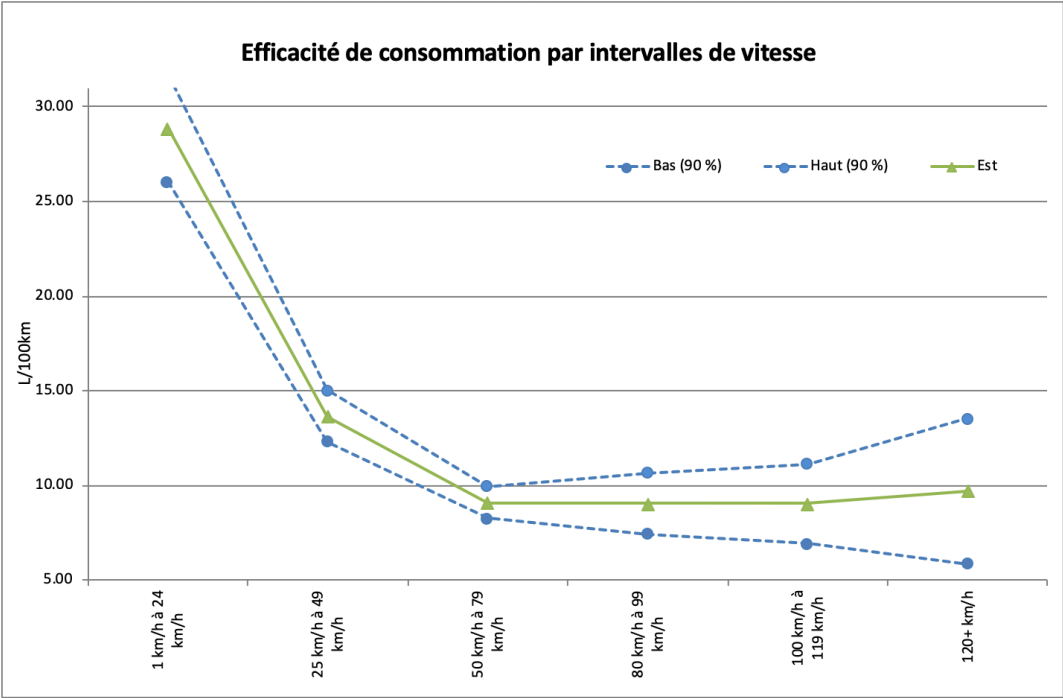
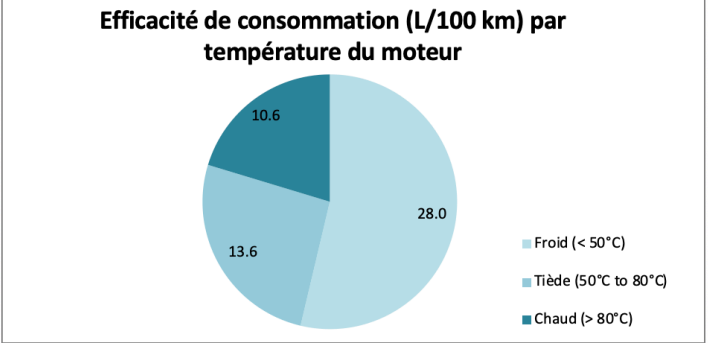
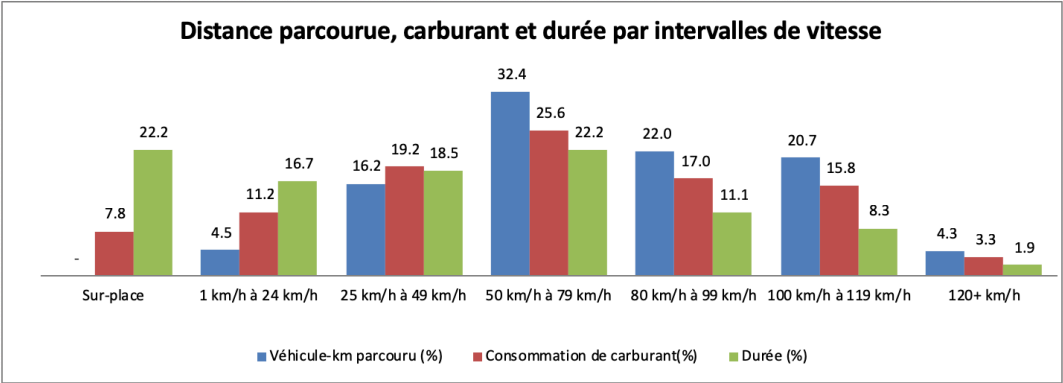
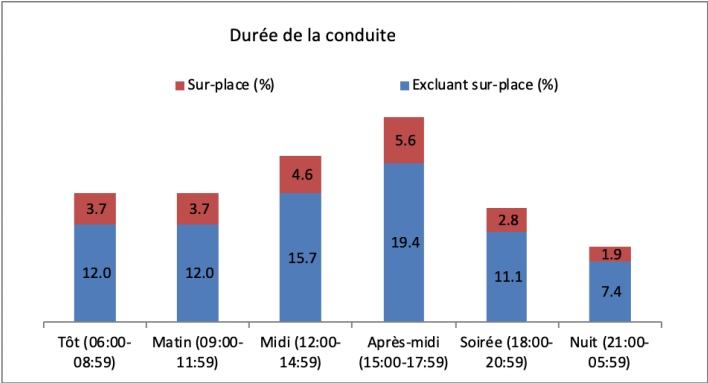
Ontario – 1er trimestre 2012

Caractéristiques des déplacements



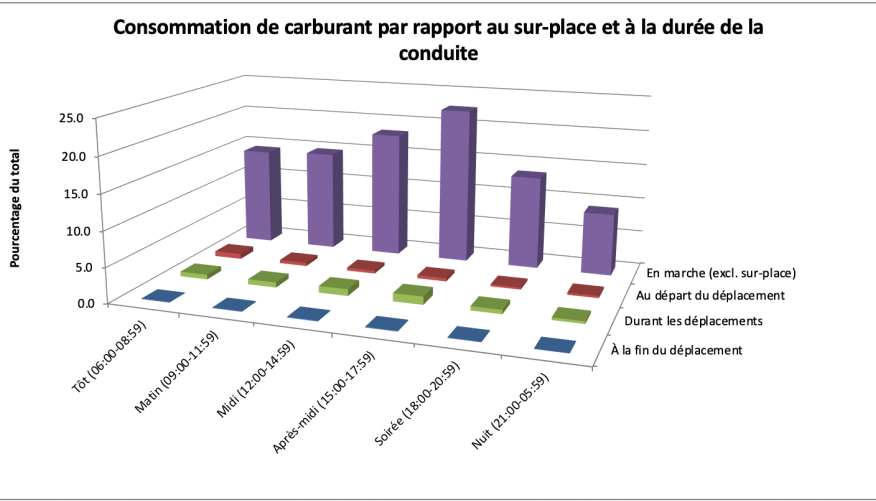
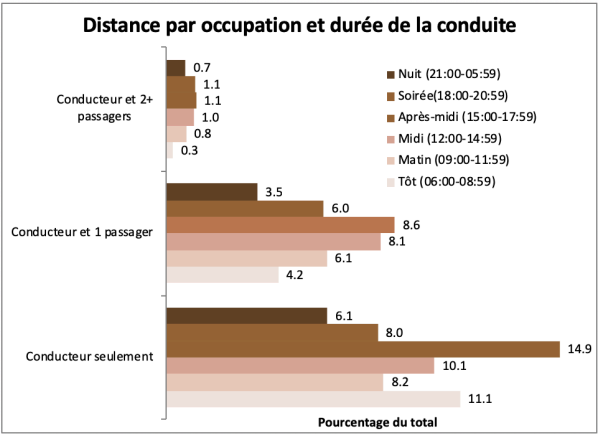
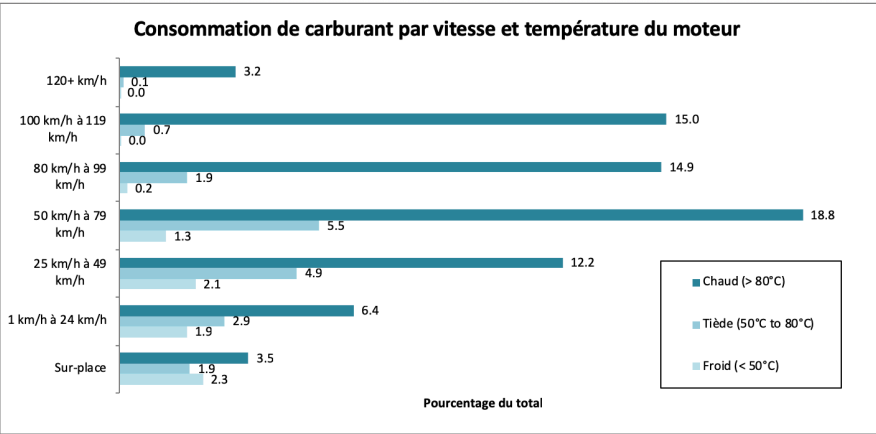
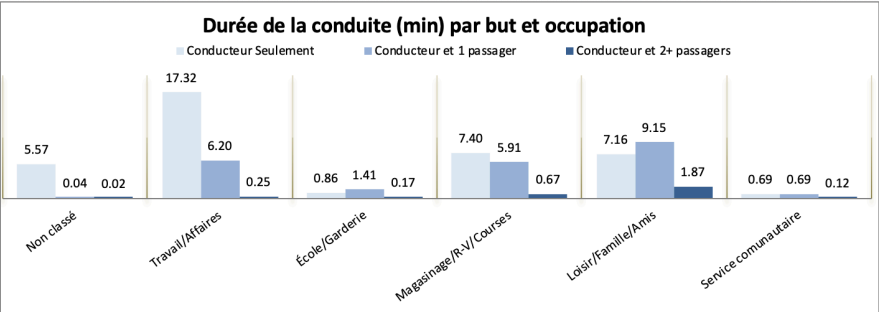
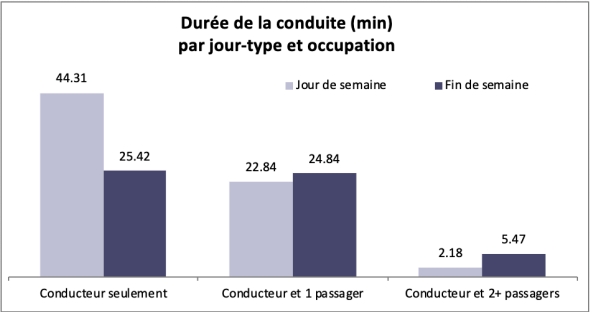
Ontario – 1er trimestre 2012

Sous-caractéristiques des déplacements



Ontario – 1er trimestre 2012

Caractéristiques mixtes sur les déplacements



What-If Analysis: Impact of Minimum Wage

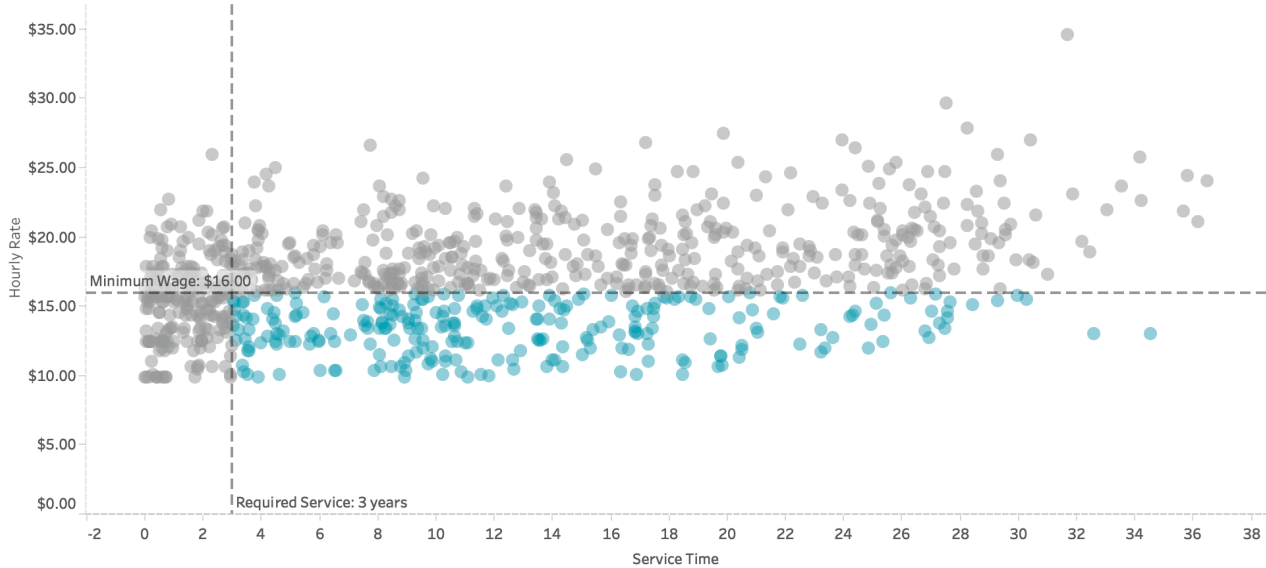
[<https://bigbookofdashboards.com/dashboards.html>]



Proposed Minimum Wage

Required Service

Developed by Matt Chambers
<http://sirvialot.blogspot.com/>



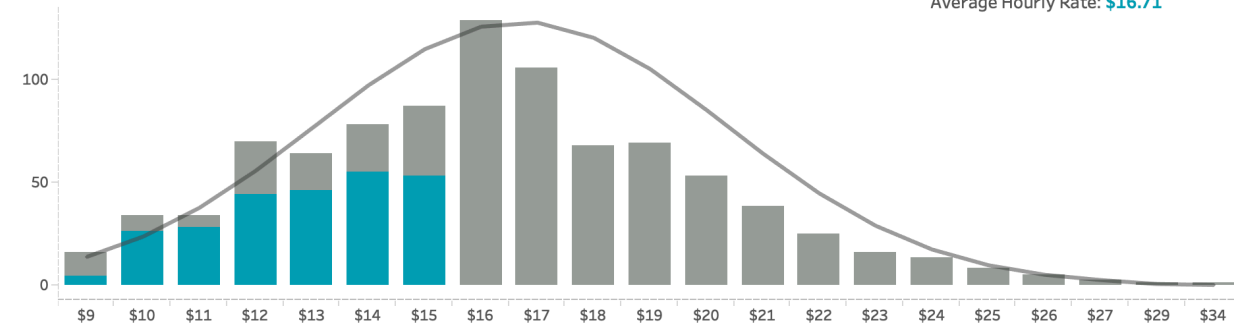
Dollar Impact of Minimum Wage: **\$1,792,206**

Employees Below Minimum Wage: **661**

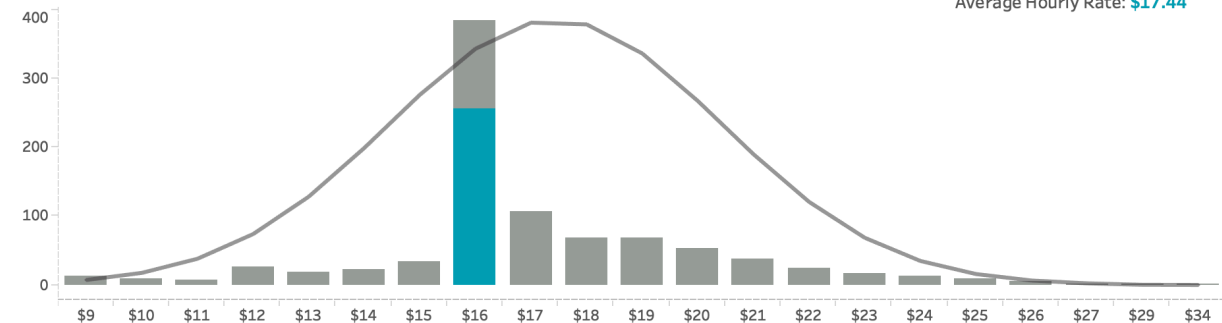
Employees Below Minimum Wage: **256**

Facilities	\$42,440	191
Legal	\$30,108	6
Logistics	\$16,764	38
Engineering	-\$38,645	12
Services	-\$87,052	309
Information Technology	-\$107,696	19
Purchasing	-\$116,048	27
Customer Service	-\$121,224	28
Operations	-\$166,590	35
Marketing	-\$189,834	91
Finance	-\$198,323	15
Research & Development	-\$283,377	39
Human Resources	-\$351,142	32
Supply Chain	-\$528,309	75

Current Distribution



Distribution with Minimum Wage



QUELS EN SONT LES DÉFAUTS?

Tableau de bord 1 : compréhension impossible en un coup d'œil, surutilisation des couleurs, diagrammes à secteurs??

Tableau de bord 2 : visualisations 3D, délimitations et arrière-plan distrayants, absence de données filtrées, nombre insuffisant d'étiquettes et contexte insuffisant.

Tableaux de bord 3 et 4 : ...

EXERCICE

En équipe ou individuellement, préparez un scénario pour lequel un tableau de bord serait utile.

Déterminez les questions particulières auxquelles pourrait répondre le tableau de bord ou les faits qu'il pourrait fournir.

Déterminez les sources de données et les éléments de données qui pourraient servir à votre tableau de bord.

Concevez un affichage (au moyen d'un crayon et d'une feuille) avec des exemples de graphiques.

Quels sont les points forts et les limites de votre tableau de bord? Est-il fonctionnel? élégant?

MATIÈRE SUPPLÉMENTAIRE

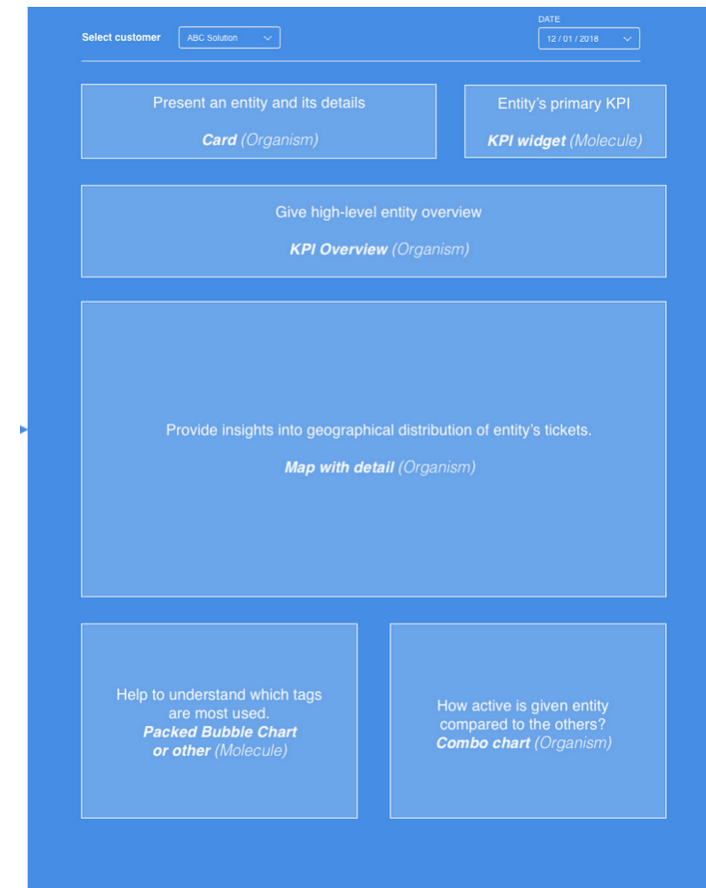
CONCEPTS AVANCÉS DE LA VISUALISATION DES DONNÉES ET DE LA CRÉATION DE RAPPORTS

ÉLÉMENTS DE LA CONCEPTION D'UN TABLEAU DE BORD

Les développeurs de GoodData adoptent une approche intéressante à la conception des tableaux de bord.

Selon eux, un tableau de bord doit se composer :

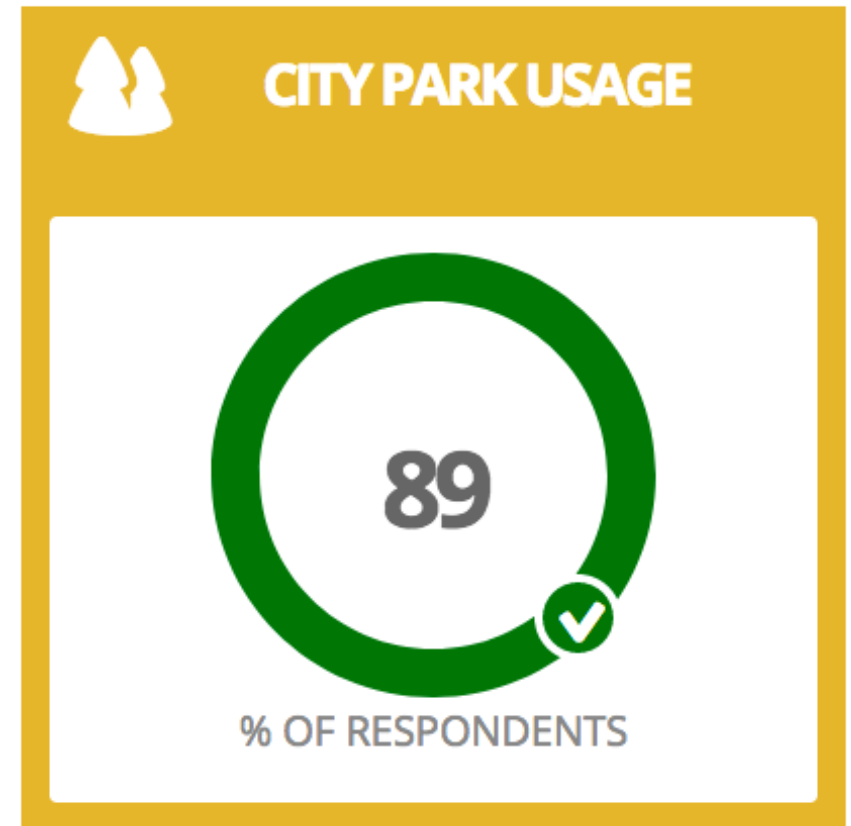
- d'atomes;
- de molécules;
- d'organismes.



COMMENT VISUALISEREZ-VOUS LES DONNÉES?

Pour obtenir des idées de visualisations :

- <https://datavizcatalogue.com>
- <https://datavizproject.com>
- <https://rawgraphs.io>



QUEL TYPE DE DONNÉES AFFICHEREZ-VOUS?

- Quel type de données voulez-vous afficher?
- De quel type de données avez-vous besoin pour décrire votre situation?
- De quels ensembles de données avez-vous besoin?



ÉLÉMENTS D'UNE ANALYSE (I)

Que voulez-vous afficher :

- Qu'est-ce qui s'est passé? – rétrospective
- Qu'est-ce qui se passe? – compréhension de la situation
- Qu'est-ce qui se passera? – prévision



ÉLÉMENTS D'UNE ANALYSE (II)

- Allez-vous montrer à votre public des chiffres « bruts » (nombre d'automobiles sur la route)?
- Allez-vous avoir besoin de mesures agrégées (densité de la circulation)?



QUELQUES QUESTIONS POUR TERMINER

- Qu'avez-vous appris au cours de l'exercice sur les tableaux de bord?
- Est-ce qu'un aspect vous a surpris?
- Est-ce que votre organisation pourrait profiter d'un tableau de bord?
- De quelle équipe ou de quelles compétences aurez-vous besoin pour créer un véritable tableau de bord?

RÉFÉRENCES

CONCEPTS AVANCÉS DE LA VISUALISATION DES DONNÉES ET DE LA CRÉATION DE RAPPORTS

RÉFÉRENCES

Understanding Graphics

Krygier, J. et Wood, D., *Making Maps: A Visual Guide to Map Design for GIS*, Guilford Press, 2016.

Wikipédia, *Interactive Data Visualization*.

NASA, *Is animation an effective tool for data visualization?*.

C.G. Healey, *Perception in Visualization* (très bien!)

Data Physicalizations

Tufte, E., *The Visual Display of Quantitative Information*, Graphics Press, 2001.

Hu, D., *How to Lie With Statistics*, Norton, 1954.

Tufte, E., *Beautiful Evidence*, Graphics Press, 2008.

RÉFÉRENCES

Nussbaumer Knaflic, C., *Storytelling with Data*, Wiley, 2015.

Cairo, A., *The Functional Art*, New Riders, 2013.

Cairo, A., *The Truthful Art*, New Riders, 2016.

Meireilles, I., *Design for Information*, Rockport, 2013.

50 Great Examples of Data Visualization, <http://www.webdesignerdepot.com>.

Wexler, S., Shaffer, J. et Cotgreave, A., *The Big Book of Dashboards*, Wiley, 2017.

Nathan Yau, [*FlowingData*](#).

Wikipédia, [*Data Visualization*](#).

Wikipédia, [*Misleading Graphs*](#).

RÉFÉRENCES

Prabhakaran, S., [*Top 50 ggplot2 Visualizations*](#) (with Master List R Code).

Miller, M., [*The problem with Interactive graphics*](#), Co.Design, 2017.

Wickham, H., *ggplot2: Elegant Graphics for Data Analysis*, 2^e édition, Springer, 2016.

Gorelik, B., [*Data Visualization*](#) (blogue).

Chang, W., *R Graphics Cookbook*, O'Reilly, 2013.

Wickham, H., *A Layered Grammar of Graphics*, *Journal of Computational and Graphical Statistics* 19:3–28, 2009.

Horton, N.J. et Kleinman, K., *Using R and RStudio for Data Management, Statistical Analysis, and Graphics*, 2^e édition, CRC Press, 2016.

Healey, K., *Data Visualization: A Practical Introduction*, 2018.

RÉFÉRENCES

Kabacoff, R.I., *R in Action, Second Edition: Data analysis and graphics with R*, Live, 2011.

Maindonald, J.H., *Using R for Data Analysis and Graphics: Introduction, Code and Commentary*, 2008.

Tyner, S., Briatte, F. et Hofmann, H., *Network Visualization with ggplot2*, The R Journal, vol. 9(1), 2017.

Broman, K., *Data Visualization with ggplot2*, 2016.

Robinson, D., *Visualizing Data Using ggplot2*, à varianceexplained.org.

Manipulating, analyzing and exporting data with tidyverse, à datacarpentry.org.

Wickham, H., *Tidy Data*, Journal of Statistical Software, version 59, numéro 10, 2014.

Gashim, E. et Boily, P., *A ggplot2 Primer*, data-action-lab.com, 2018.