# [Réalisez des rapports statistiques clairs et impactants](https://openclassrooms.com/fr/courses/4525336-realisez-des-rapports-statistiques-clairs-et-impactants)

# Table des matières

[Réalisez des rapports statistiques clairs et impactants 1](#_Toc62311041)

[Table des matières 1](#_Toc62311042)

[Comprenez ce qu’est un rapport statistique 5](#_Toc62311043)

[Qu'est-ce qu'un bon rapport statistique ? 5](#_Toc62311044)

[À quoi ressemble un rapport statistique ? 5](#_Toc62311045)

[Le lecteur idéal n'existe pas ! 5](#_Toc62311046)

[En résumé 6](#_Toc62311047)

[**Concevez votre projet de rapport** 7](#_Toc62311048)

[**Votre lecteur** 7](#_Toc62311049)

[**Votre projet d’écriture** 7](#_Toc62311050)

[**Votre message essentiel** 7](#_Toc62311051)

[**Dégagez les idées principales** 8](#_Toc62311052)

[**En résumé** 8](#_Toc62311053)

[**Rédigez en 3 étapes** 9](#_Toc62311054)

[**Première étape : rédigez un plan** 9](#_Toc62311055)

[**Deuxième étape : rédigez rapidement une première version** 9](#_Toc62311056)

[**Troisième étape : reprenez votre texte** 9](#_Toc62311057)

[**En résumé** 9](#_Toc62311058)

[Hiérarchisez vos informations dans un plan 11](#_Toc62311059)

[Commencez votre rapport par le plus intéressant ! 11](#_Toc62311060)

[Les lois de proximité 11](#_Toc62311061)

[Utilisez un angle 12](#_Toc62311062)

[Allez du particulier au général 12](#_Toc62311063)

[En résumé 13](#_Toc62311064)

[Entraînez-vous à réaliser un plan de rapport 14](#_Toc62311065)

[À vous de jouer 14](#_Toc62311066)

[Vérifiez-votre travail 14](#_Toc62311067)

[Clarifiez votre propos 15](#_Toc62311068)

[Le danger du jargon 15](#_Toc62311069)

[Attention à la logique instrumentale 16](#_Toc62311070)

[En résumé 16](#_Toc62311071)

[**Rédigez des textes vivants** 17](#_Toc62311072)

[**Écrivez au présent** 17](#_Toc62311073)

[**Attention aux enchâssements** 17](#_Toc62311074)

[**Utilisez des verbes plutôt que des noms communs** 17](#_Toc62311075)

[**Évitez les expressions « creuses »** 18](#_Toc62311076)

[**En résumé** 18](#_Toc62311077)

[**Choisissez des titres efficaces** 19](#_Toc62311078)

[**Soignez vos titres !** 19](#_Toc62311079)

[**5 erreurs à éviter dans un titre** 19](#_Toc62311080)

[**Découvrez l’intérêt du chapeau** 19](#_Toc62311081)

[**Utilisez des intertitres pour relancer l’intérêt** 20](#_Toc62311082)

[**En résumé** 20](#_Toc62311083)

[Entraînez-vous à rédiger les titres et le chapô d'un rapport 21](#_Toc62311084)

[À vous de jouer 21](#_Toc62311085)

[Vérifiez votre travail 21](#_Toc62311086)

[Comprenez l'intérêt des illustrations dans un rapport statistiques 22](#_Toc62311087)

[Le rôle des illustrations dans votre rapport 22](#_Toc62311088)

[Utilisez les tableaux à bon escient 23](#_Toc62311089)

[L'utilité des tableaux 23](#_Toc62311090)

[Optimisez vos tableaux 23](#_Toc62311091)

[Choisissez des graphiques adaptés à votre message 25](#_Toc62311092)

[L'intérêt des représentations graphiques 25](#_Toc62311093)

[Anatomie d'un graphique : le vocabulaire 26](#_Toc62311094)

[Comment choisir le meilleur graphique ? 27](#_Toc62311095)

[Découvrez les différents types de graphiques 29](#_Toc62311096)

[Diagrammes en bâtons (*Bar chart*) 29](#_Toc62311097)

[Pyramides des âges 29](#_Toc62311098)

[Courbes (*Line chart*) 30](#_Toc62311099)

[Diagramme circulaire (*Pie chart*) 31](#_Toc62311100)

[Nuages de points (*Scatter plot*) 31](#_Toc62311101)

[Boîtes à moustaches (*Box plot*) 32](#_Toc62311102)

[Aires empilées 33](#_Toc62311103)

[Diagramme de Kiviat (*Radar chart)* 34](#_Toc62311104)

[Ressources complémentaires 34](#_Toc62311105)

[Réalisez des graphiques compréhensibles 36](#_Toc62311106)

[Quelques idées reçues sur les graphiques 36](#_Toc62311107)

[Soignez vos graphiques 36](#_Toc62311108)

[« Do » et « Don't » des représentations graphiques 37](#_Toc62311109)

[Entraînez-vous à illustrer un rapport avec des graphiques pertinents 40](#_Toc62311110)

[À vous de jouer 40](#_Toc62311111)

[Vérifiez votre travail 40](#_Toc62311112)

Vous avez passé des heures, ​ voire des jours à faire tourner vos algorithmes​ pour étudier un phénomène​ ​ou une population​.

Malheureusement​, **au moment d’écrire votre rapport,** vous ne savez pas comment ​faire.  
Pire, vous avez **le syndrome de la page blanche**...

Pas​ de panique !

Ce cours est là​ pour vous aider​ à acquérir des compétences de communication indispensables ​qui vont faire la différence.

Vous comprendrez l’**importance​ du travail préparatoire**. Vous connaîtrez **les trois étapes indispensables ​de l’écriture** et vous saurez​ **choisir un plan pertinent** ​pour conserver ​votre lecteur.  
  
À la fin de ce cours, **vos textes ​seront plus vivants** et vos rapports statistiques feront mouche !

**Objectifs pédagogiques :**

* **Concevoir** un plan de rapport statistique impactant
* Améliorer un rapport statistique avec les **techniques rédactionnelles**
* **Représenter graphiquement** des données statistiques

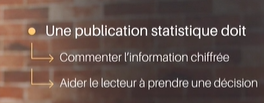
**Prérequis :**

Ce cours est ouvert à tous ceux qui ont des **bases en statistiques et en analyse de données**.

Ce cours fait partie de la formation en ligne "[***Data Analyst***](https://openclassrooms.com/paths/data-analyst)" réalisée en partenariat avec l'ENSAE.



## Comprenez ce qu’est un rapport statistique

Les textes publiés par les data analyst se basent sur des chiffres issus de statistiques. Ce travail statistique a parfois pris des jours, des semaines, des mois. Et l’envie est forte de vouloir présenter aux lecteurs tous les résultats obtenus. Conclusion, le lecteur est noyé sous un amas de chiffres. Pire, le texte ne fait que répéter le contenu des tableaux.

Un rapport ou une publication statistique, ce n’est pas cela !

### Qu'est-ce qu'un bon rapport statistique ?

Une publication statistique doit **commenter l’information chiffrée**, **apporter de l’information au lecteur**, **l’aider à prendre une décision**.

Le texte doit expliquer de quoi il s’agit, qui est concerné, quand cela s’est passé, où cela se passe, quel est l’objectif de l’enquête ou de l’étude, avec quels moyens ou comment s’est déroulé l’action.

Enfin, l’objectif principal d’une publication statistique est que votre lecteur comprenne et retienne le message que vous aurez voulu lui faire passer. **Si vos lecteurs n’ont rien retenu à la fin de leur lecture, ils auront l’impression d’avoir perdu leur temps**. Pour cette raison, votre mode d’écriture doit favoriser la mémorisation. Les techniques rédactionnelles et ce cours vont vous y aider.

### À quoi ressemble un rapport statistique ?

Selon l'entreprise ou l'institution pour laquelle vous travaillez, la longueur et la forme peut changer à la marge. Mais la structure est toujours la même, et les conseils de ce cours vont seront utile, quel que soit le rapport que vous devrez rédiger.

**Quelques exemples de rapport**

* À l'INSEE, un rapport habituel fait [**2 à 4 pages**](https://www.epsilon.insee.fr/jspui/bitstream/1/2789/1/etudes39.pdf).
* Sur OpenClassrooms, pour réaliser les projets des parcours Data Scientist et Data Analyst, nous demandons des rapport de [**ce type**](https://s3-eu-west-1.amazonaws.com/static.oc-static.com/prod/courses/files/R%C3%A9alisez+des+rapports+statistiques+clairs+et+impactants+/rapport+FD.pdf).
* En entreprise, les rapports sont semblables, mais peuvent être de longueur variable. Leur forme dépend, bien sûr, de l'industrie dans laquelle vous travaillez.
* Le format notebook est de plus en plus populaire pour travailler sur vos données et les communiquer. En voici un exemple en ligne, sur **[Kaggle](https://www.kaggle.com/fabiendaniel/film-recommendation-engine" \t "_blank)**.

### Le lecteur idéal n'existe pas !

Quand vous rédigez un texte, n’imaginez pas que tous les destinataires le liront attentivement. La réalité est bien loin de ce scénario idéal. Vos lecteurs, tout comme vous, reçoivent chaque jour une montagne d’informations et de notes à lire.

Ils se contentent dans le meilleur des cas de commencer à les parcourir, et s’arrêtent rapidement dès que la lecture ne les intéresse pas ou leur semble confuse.

N’oubliez jamais : votre lecteur n’est pas obligé de vous lire !

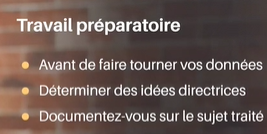
Vous avez de la chance, les publications statistiques visent souvent un public désireux d’avoir une information à la source, plus complète que celle qu’ils auraient dans des publications moins spécialisées. Ils sont prêts à faire un effort de lecture, à condition toutefois qu’il ne soit pas excessif, et qu’il soit récompensé par un contenu intéressant et compréhensible. Ils ont aussi une connaissance plus ou moins approfondie des notions que vous utilisez, et il vous faut en tenir compte.

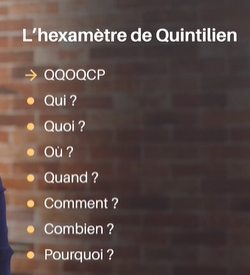
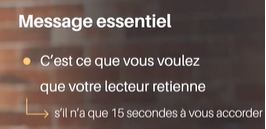
### En résumé

* Un rapport statistique n'est pas une collection de chiffres, de tableaux et de graphiques.
* Au contraire, c'est un ensemble organisé autour d'idées directrices.
* Un bon rapport statistique donne de l'information utile au lecteur et l'aide à prendre des décisions.
* Votre lecteur n'est pas obligé de vous lire, alors trouvez les moyens de conserver son attention.

Vous savez ce que doit contenir un rapport statistique, voyons maintenant comment éviter le syndrome de la page blanche !

**Concevez votre projet de rapport**



Désormais, vous savez ce que votre lecteur attend d’une publication statistique mais vous ne savez pas par quoi ni comment commencer. Vous avez le syndrome de la page blanche. Pas de panique. Nous allons voir comment vous débloque en concevant votre projet de rapport.

**Votre lecteur**

En premier lieu, vous devez vous demander qui est votre lecteur. Est-ce que votre rapport s’adresse à vos collègues, spécialistes, ou au grand public ? Lors de la conception de votre rapport et la rédaction, vous devrez toujours gardez votre lecteur en tête !

**Votre projet d’écriture**

La qualité d’un article est largement déterminée avant que la première ligne ait été écrite.

Autrement dit, la phase préparatoire à l’écriture a une importance capitale !

Tout d’abord, votre texte devra répondre aux questions de « l’hexamètre de Quintilien ». Vous savez, le fameux **QQOQCP** : Qui ? Quand ? Où ? Quoi ? Comment ? Pourquoi ? Vous devez mobiliser vos connaissances pour répondre à ces questions.

Sachez qu’avant tout, **un bon texte se construit autour d'idées**. Vous devrez aller plus loin que le simple commentaire de vos données. Il est indispensable de vous documenter sur le sujet que vous traitez, en lisant ce qui a déjà été écrit dessus.

**Votre message essentiel**

Il vous faut donc, avant d’écrire, classer tous les éléments que vous avez rassemblés et en faire une synthèse. Vous obtiendrez ainsi un petit nombre d’idées qui constitueront votre « message essentiel ».

**Qu’est-ce que le message essentiel ?** C’est ce que vous voulez que votre lecteur retienne s’il n’a que 15 secondes à vous accorder.

Le message essentiel et les autres idées, hiérarchisées de la plus à la moins importante, vont composer le corps de votre article. Ces idées vont structurer votre texte, lui servir de colonne vertébrale.

Avant de commencer à écrire, assurez-vous qu’elles sont tout à fait claires dans votre esprit. Le mieux pour cela est de les écrire sous forme de phrases complètes.

**Dégagez les idées principales**

Rassembler de l’information, en faire une synthèse sous la forme de quelques idées essentielles : tout ce travail préparatoire n’est pas facile et prend du temps. Mais, dégagez les idées principales permet de réfléchir à l’ordre dans lequel vous allez exprimer vos idées. Cela permet aussi de vérifier qu’elles sont suffisamment étayées par les données.

Tout ce travail en amont est essentiel, et vous permettra d’écrire votre texte plus facilement lorsque vous passerez en phase de rédaction.

Pour bien faire comprendre vos idées, n’oubliez pas de les illustrer par des exemples, des chiffres. En outre, votre texte sera bien meilleur : les idées ressortiront mieux, de manière quasi automatique, et vos lecteurs les retiendront plus facilement.

**En résumé**

* Vous écrivez pour un lecteur spécifique. Gardez-le en tête !
* Demandez-vous si votre rapport répond bien aux questions essentielles QQOQCP.
* Définissez votre message essentiel, et insistez dessus dans votre rapport.
* Identifiez aussi les idées principales du rapport, et chassez le superflu.

*Vous savez quel est le message essentiel que vous voulez que vos lecteurs retiennent, Vous avez suffisamment d’éléments pour étayer votre message, l’écriture peut désormais commencer. Voyons maintenant les trois étapes essentielles de la rédaction.*

**Rédigez en 3 étapes**



La rédaction se déroule en trois étapes itératives que je vous conseille vivement de suivre sous peine de perdre beaucoup de temps !

**Première étape : rédigez un plan**

Vous avez dégagé les idées principales, la phase suivante consiste à faire un plan. Trop souvent les data analysts se lancent dans la rédaction sans avoir élaboré de plan. Faire un plan, c'est choisir l'ordre dans lequel vous allez exprimer vos idées essentielles.  
**Comment choisir cet ordre ?** Le critère à adopter s'exprime simplement : **le plus intéressant d'abord**. Le plus intéressant du point de vue du lecteur bien sûr.

Nous reviendrons sur le sujet du plan au chapitre suivant.

**Deuxième étape : rédigez rapidement une première version**

Une fois votre plan terminé, il est temps de vous lancer dans l’écriture proprement dite. Votre travail consiste alors à développer les idées essentielles, à les expliquer, à les concrétiser par des exemples. Mais attention, vous devez rédiger ce premier brouillon le plus vite possible.

Bloquez votre agenda et consacrez plusieurs heures d’affilée à cette rédaction. Il est essentiel de vous consacrer pleinement à cette écriture, sans craindre les problèmes de style. Le but est d'avoir une première version du rapport complet pour fixer le fond de votre propos.

**Troisième étape : reprenez votre texte**

Bien entendu, votre première version ne sera pas d’une excellente qualité. Voilà pourquoi vous reprendrez votre texte lors de cette troisième étape pour mettre au point la version finale du rapport. C’est au cours cette phase que vous supprimez les défauts les plus évidents, comme les répétitions ou les tournures syntaxiques approximatives. Mais surtout à ce moment que vous mettrez en application**les techniques rédactionnelles** que nous allons voir bientôt.

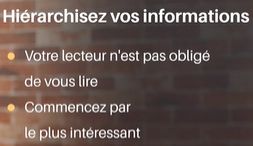
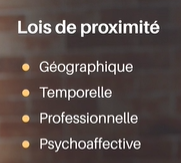
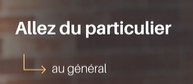
N’hésitez à revoir votre texte plusieurs fois, jusqu’à arriver à une version qui vous convienne. Et surtout n’ayez aucun complexe : tous ceux dont le métier est d'écrire procèdent ainsi.

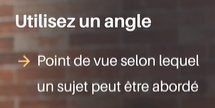
**En résumé**

* Pour rédiger efficacement, adoptez une méthode en 3 étapes.
* Commencez par articuler vos idées dans un plan. Rappelez-vous : le plus intéressant d'abord !
* Rédigez tout le rapport une première fois, le plus vite possible.
* Ensuite, reprenez-le au propre en appliquant les techniques rédactionnelles.

*Vous connaissez désormais les 3 étapes essentielles de la rédaction. Revenons sur un point essentiel de ce travail : comment attirer le lecteur ? Grâce à un plan de rapport judicieux !*

## Hiérarchisez vos informations dans un plan



Revenons à la première étape de la méthode, qui nécessite que l’on s’y attarde.

Faut-il le rappeler, **votre lecteur n'est pas obligé d'aller jusqu'au bout de votre texte**. Même s'il a été attiré par le sujet que vous traitez, il peut s'interrompre rapidement car il sera de moins en moins réceptif au fil de sa lecture.

### Commencez votre rapport par le plus intéressant !

Votre lecteur continuera la lecture de votre rapport si vous lui présentez ce qui l'intéresse le plus au début. Si jamais il s'arrête en cours de route, au moins aura-t-il lu ce qui avait le plus de chances de l'intéresser. Alors que si vous commencez par des idées plus secondaires, il risque de vous quitter avant d'être arrivé au plus important.

#### Prenons un exemple

Regardez ce plan, et demandez-vous ce qui peut intéresser un lecteur potentiel :

I. Contexte et objectifs  
II. Préparation et exploration des données  
III. Modélisation  
IV. Synthèse des résultats et recommandations

Ce qui intéresse votre lecteur, ce sont les **résultats** et vos **recommandations**. Il faut commencer par là. La préparation et l’exploration des données ne l’intéresse pas forcément !

En effet, les data analyst consacrent une grande partie de leur temps à travailler sur les données, à faire tourner des algorithmes. Vous pourriez avoir tendance à parler plus aisément de la méthode de travail. Mais vos lecteurs ne sont pas forcément intéressés par vos problèmes méthodologiques. Traitez donc la méthodologie à la fin du texte, ou mieux, dans un encadré.

De même, attention de ne pas suivre l’ordre d’une nomenclature ou l’ordre des questions d’une enquête. C’est souvent soporifique !

### Les lois de proximité

Pour déterminer ce qui peut intéresser le plus vos lecteurs, utilisez les lois de proximité, qui permettent d'accrocher l’intérêt des lecteurs.

Ces lois sont utilisées par les journalistes. Il en existe 4, mais les data analysts sont plutôt concernés par les trois premières :

1. Loi de proximité **géographique** : votre lecteur est plus intéressé par ce qui se passe dans sa ville ou son pays que dans un pays à l'autre bout du monde.
2. Loi de proximité **temporelle** : nous sommes plus intéressés par le futur immédiat que par le futur lointain, et même le présent !
3. Loi de proximité **professionnelle** : les lecteurs seront plus concernés par des informations qui les touchent directement comme l’emploi, les impôts...
4. Loi de proximité **psycho-affective** : pour les sujets qui touchent le lecteur à un niveau plus personnel.

Retenez donc que vous devez partir de votre lecteur et vous demander quelles informations sont les plus proches de lui !

### Utilisez un angle

Bien souvent vous pouvez vous sentir noyé sous de nombreux de tableaux, graphiques... Bref une tonne d’information !  
Devant cette masse d’information, **la notion d’« angle » est précieuse**. Qu’est-ce qu’un angle ? Il s’agit des divers points de vue, souvent nombreux, selon lesquels un sujet peut être abordé.

Ainsi, si vous voulez parler des délocalisations vous pourrez aborder de nombreux aspects : les conséquences sur l’emploi, les motivations des entreprises qui délocalisent, les actions éventuelles des pouvoirs publics... Dans un article ou un rapport, vous ne devez surtout pas traiter de tous ces aspects à la fois. Vous examinerez vos données et favoriserez le meilleur angle pour traiter vos données.

Quand vous ne pouvez pas appliquer cette méthode, et que vous devez écrire un seul texte divisé en plusieurs parties, la notion d’angle vous sera utile : assurez-vous que chaque partie est consacrée à un angle, et surtout vous ne mélangez pas les angles dans la même partie.

### Allez du particulier au général

Vous avez souvent entendu l'inverse : la méthode de l'entonnoir. C'est à dire aller du général au particulier.

Mais quand on commente des données statistiques, aller du niveau le plus regroupé vers les niveaux les plus fins est souvent plus intéressant.

Donnez par exemple un résultat pour une commune, puis le même résultat pour un ensemble de localités. En effet, le niveau le plus fin est le plus concret, donc celui qui intéresse le plus votre lecteur.

#### Ne copiez pas la version de l’année précédente

De nombreuses sources statistiques sont périodiques, généralement annuelles. Elles donnent lieu à des publications récurrentes.

La tentation est forte de reprendre d’une année à l’autre le même plan. Cela prend moins de temps, et c'est rassurant !

Mais il n’est pas sûr que, d’une année à l’autre, les aspects les plus intéressants, ceux dont il faut parler en premier, soient les plus pertinents. Il faut donc réfléchir chaque année aux idées à mettre en avant, et au plan à adopter.

**Quelques exemples de plan de rapport**

* Plan d'un [**rapport de l'INSEE**](https://www.epsilon.insee.fr/jspui/bitstream/1/2789/1/etudes39.pdf).
* Plan d'un Notebook sur **[Kaggle](https://www.kaggle.com/fabiendaniel/film-recommendation-engine" \t "_blank)**.

### En résumé

* Commencez votre plan de rapport par le plus intéressant pour votre lecteur. Pour cela, réfléchissez à ces attentes et aidez-vous des lois de proximité.
* Adoptez toujours un angle, afin de trouver une cohérence à vos données statistiques et de sélectionner les éléments les plus importants.
* Enfin, n'hésitez pas à aller du niveau le plus fin au niveau global. Le particulier, le concret, intéresse votre lecteur.

Vous avez défini votre projet d’écriture et vous savez concevoir un plan qui intéressera votre lecteur. Dans la prochaine partie, vous verrez comment appliquer des techniques de rédaction efficaces.

## Entraînez-vous à réaliser un plan de rapport

### À vous de jouer

Pour vous entraîner, réalisez cet exercice étape par étape. Une fois terminé, vous pouvez comparer votre travail avec les pistes que je vous propose.

À partir de ce [rapport](https://s3-eu-west-1.amazonaws.com/static.oc-static.com/prod/courses/files/R%C3%A9alisez+des+rapports+statistiques+clairs+et+impactants+/Insee+Flash+Champagne-Ardenne+Exercice+Partie+1.pdf) provenant de l'INSEE, dégagez les éléments suivants :

1. Quel est le message essentiel de ce rapport ?
   * La péiurbanisation du SCoT de Reims depuis 1975 et ces impacts sur l’environement et sur le social

***La périurbanisation au sein du SCoT de la région rémoise***

* + Un impact sur la mixité sociale et l’environnement

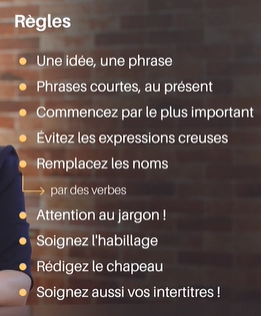
1. Proposez un plan allant du plus important au moins important. Vous classerez les paragraphes de cette manière : "1 / 2 / 3 / ... / 10".

🡺3/ 1/ 2/ 8/ 6/ 5/ 4/ 7/ 9/ 10/

### Vérifiez-votre travail

Voici un [**exemple**](https://static.oc-static.com/activities/2464/evaluation_resources/proposez-un-plan-de-rapport_exemple-2018-05-31T135124.zip) pour vous permettre de vérifier votre travail.

## Clarifiez votre propos



Nous avons vu les différentes phases de la rédaction : le travail préparatoire, le plan, la rédaction rapide d’un premier jet et la phase de réécriture. Voyons concrètement comment vous devriez réécrire votre texte à l’aide des techniques rédactionnelles.

Les techniques rédactionnelles que nous allons examiner sont destinées à vous aider dans deux domaines : rendre vos textes plus vivants, et les rendre plus clairs. Ainsi, vous conserverez vos lecteurs plus longtemps. En plus, ils retiendront mieux ce qu’ils auront lu.

### Le danger du jargon

N’oubliez pas que vos lecteurs ne maîtrisent pas forcément les concepts et les méthodes que vous utilisez. Or malheureusement, s’efforcer de se faire comprendre par des lecteurs non spécialistes n’est pas si facile. D’ailleurs, on retrouve parfois des mauvaises interprétations d’études dans la presse.

#### Identifiez le jargon que vous utilisez

**Tout d’abord, évitez les mots, les expressions et les sigles que vous seul connaissez**. Dans tous les domaines professionnels existe un jargon.

Par exemple, à l’INSEE, on parle de « migrations alternantes » au lieu de dire « déplacements domicile-travail » tout simplement. Ou encore de « risque vieillesse » pour dire « retraite ».

Pour appréhender certains phénomènes, l’emploi du jargon est parfois inévitable. Pour que les non-spécialistes puissent interpréter correctement vos messages, il faut alors expliquer à vos lecteurs les concepts que vous utilisez et leur faire comprendre pourquoi vous les avez introduits.

Ainsi : dire que « La heatmap de la matrice de corrélation permet de mettre en exergue les fortes corrélations, ce qui peut être utile pour le feature engineering. » ne guide pas votre lecteur non-spécialiste.

Les explications du jargon utilisé doivent être données simplement, dans le corps du texte, les explications détaillées pouvant être mise dans un encadré.

**Vous pouvez aussi vouloir apporter trop de précisions qui en fait vont perturber vos lecteurs.**

Par exemple, beaucoup de textes parlent des actifs de 15 ans et plus. Le « 15 ans et plus » est inutile, et mieux vaut enlever cette précision.

Enlever de vos textes tout le jargon statistique et remplacez-le par des mots et des expressions connus de tous.

**Les sigles constituent une sorte de jargon**. Seuls les plus connus d’entre eux peuvent être employés sans explication : SNCF, SMIC…  
Mais les lecteurs non data analyst ne savent pas ce qu’est une « ACP », « RLM », ou encore la « méthode KNN ».

Le recours au vocabulaire spécialisé, et plus encore aux sigles, est très fréquent dans les notes internes et cela n’est pas gênant si le document est pour un cercle restreint. Mais si la diffusion est plus large, une partie de message échappera au lecteur et risque d’être mal perçue.

### Attention à la logique instrumentale

Évitez également ce que l’on appelle la logique instrumentale. Les data analysts calculent bien souvent des taux et étudient leur évolution, ce qui les conduit à écrire des phrases du type : « Le taux de chômage a augmenté ». Je vous conseille plutôt de parler directement du phénomène et d’écrire : « Le chômage augmente. ».

### En résumé

* Pour rendre vos textes plus clairs, évitez l’abus de jargon et de sigles.
* Mais surtout, gardez en tête votre lecteur, pour savoir quel vocabulaire vous pouvez employer.

Votre rapport est maintenant plus clair et adapté à votre lectorat. Voyons à présent comment le rendre plus vivant.

**Rédigez des textes vivants**

Certaines de nos habitudes d’écriture rendent les textes abstraits et compliquent la lecture.

L’application systématique des principes que nous allons voir va vous permettre d’améliorer vos textes.  
Pour rendre vos textes plus impactants, rappelez-vous de cette règle : **une idée = une phrase.**

Au lieu de...

« Puisque l'ACP a permis une réduction du nombre de dimensions, et qu'elle a mis en évidence 3 groupes d'enregistrements, on peut utiliser une méthode de clustering afin de séparer au mieux ces groupes. »

Écrivez plutôt...

« L'ACP permet de réduire les dimensions et met en évidence 3 groupes d’enregistrements . La méthode de clustering permet de les séparer de manière optimale. »

**Écrivez au présent**

Pour un style plus léger et vivant, abusez du présent de l’indicatif. Rien ne vous empêche de commenter les résultats d’une enquête ou d’une opération statistique qui porte sur une période remontant à un an ou plus. C’est le « présent de narration ».

Attention de ne pas mélanger les temps. Si vous écrivez un texte sur des données relatives à 2006, mettez tout ce qui concerne cette année au présent. Si, vous faites des comparaisons avec des années antérieures, employez alors le passé pour ces années.

« En 2014, 25 % des jeunes recherchait un emploi » devient « en 2014, un jeune sur quatre recherche un emploi ».

**Attention aux enchâssements**

Parfois, la phrase va s’allonger par « enchâssement ». Si cet enchâssement a plus de 10 mots, la phrase devient vite incompréhensible.

Comprenez-vous cette phrase ?  
« Puisque le score d'un article est construit en fonction des valeurs de sept variables le décrivant, que ces variables sont quantitatives et correspondent à des cumuls ramenés à une masse de référence (100 g de produit de l'article), on peut faire exactement de même en additionnant les valeurs des différents articles de la recette, pour ces sept variables, que l'on ramène à des valeurs références pour 100g de recette. »

Quand le verbe est trop éloigné du sujet, le lecteur oublie quel est le sujet quand il arrive au verbe. Or, **le début de la phrase est la partie la mieux mémorisée**. Je vous suggère donc de commencer vos phrases par le plus important.

**Utilisez des verbes plutôt que des noms communs**

Remplacer un nom par un verbe permet de rendre une phrase plus impactante.

Plutôt que « L'augmentation des salaires est faible en 2017 », écrivez « en 2017, les salaires augmentent faiblement. »

Plutôt que « La progression du chiffre d’affaire marque le pas », écrivez « le chiffre d’affaire stagne ».

**Évitez les expressions « creuses »**

Évitez absolument les « on observe, on constate, on note, on voit... » ou encore « Comme le montrent les graphiques suivants », « Il peut être ainsi intéressant de voir », « Pour illustrer le problème »  
**Ces expressions sont parfaitement inutiles**, et ne font qu'alourdir le texte !

Les **adverbes** sont aussi bien souvent inutiles, et alourdissent le texte :

« On retrouve globalement la même analyse pour les autres composantes principales. »

Le « globalement » ne doit pas faire illusion. Si on reformule quelque peu, la phrase devient : « L’analyse est similaire aux autres composantes principales. »

**En résumé**

* Prenez conscience de vos habitudes d'écriture et tentez l'alléger vos textes.
* De nombreuses expressions ou termes sont creux : supprimez-les sans vergogne !

*Voici donc quelques éléments pour alléger vos textes et les rendre plus impactants. Dans la suite, nous allons voir comment rédiger des titres efficaces.*

**Choisissez des titres efficaces**

Votre rapport est rédigé : il est intéressant, bien construit, bien écrit.

Mais vos efforts n’auront servi à rien si les lecteurs potentiels n’ont pas envie d’entrer dans l’article. Pour les y inciter, il faut « **habiller** » cet article : lui trouver **un titre** et lui mettre **un chapeau**.

**Soignez vos titres !**

Chaque jour, des millions de nouveaux contenus sont publiés dans les journaux ou sur le web. Pour une majorité de ces contenus rédactionnels, une seule destinée : l’indifférence !

Donc plutôt que « Présentation conjoncturelle des entreprises en 2017 » écrivez « En 2017, 1 entreprise sur 4 a fait faillite » si vous voulez attirer l’attention de vos lecteurs.

**5 erreurs à éviter dans un titre**

**1) Écrire des titres trop longs**  
Vos titres doivent être courts ! D’abord, pour des questions de visibilité. Ensuite, pour un souci de référencement dans les moteurs de recherche : il ne faut pas dépasser 70 caractères espaces compris.

Opter pour un titre long signifie qu’il n’apparaîtra pas complètement dans les résultats de recherche (pensez au mobile). D'autre part, un titre long risque de décourager vos lecteurs en diffusant une image rébarbative de votre rapport.

**2) Écrire des titres trop techniques**  
Vous écrivez pour un public qui n’est pas forcément spécialiste de votre domaine. Essayez de rendre votre titre doit être compréhensible par tous et sans ambiguïté. C’est une première approche de votre contenu. Il doit être simple et, d’un coup d’œil, donner une idée du contenu que vous avez publié.

**3) Écrire des titres que vous êtes le seul à comprendre**  
Vous avez un vrai sens du jeu de mot ? Malheureusement, un bon titre n’est pas nécessairement le titre le plus original. C’est avant tout un titre qui parvient, en six à huit mots, à indiquer sa thématique, et à donner envie d’aller plus loin dans la lecture.

**4) Ecrire toujours les mêmes titres**  
C’est vrai, il y a des titres qui fonctionnent mieux que d’autres. Mais si vous lassez votre lecteur, vous risquez de le perdre.

**5) Oublier de mettre un titre**  
Que dire de plus… N’oubliez pas l’essentiel !

**Découvrez l’intérêt du chapeau**

Lorsqu’un lecteur a été attiré par le titre d’un article, il n'a pas encore vraiment décidé de lire le texte. Le premier paragraphe, appelé « **chapeau** », va confirmer ou infirmer cette décision : c’est sa première fonction.

Le chapeau est un texte assez court, placé au sommet de certains articles comme un chapeau est placé au-dessus de la tête.

Sa deuxième fonction est de donner les informations essentielles contenues dans l’article. Il arrive en effet que le lecteur, en lisant le chapeau, s'aperçoive qu'il n'est pas suffisamment intéressé par le sujet pour entrer dans le texte lui-même. En lisant le chapeau, il aura au moins pris connaissance des informations les plus importantes que l’auteur souhaite donner.

**Le chapeau ne doit pas être trop long** : **pas plus de 10 %** de la taille du texte principal, et surtout pas plus de 30 secondes de lecture.

Alors que les articles écrits par les statisticiens souffrent généralement d’un excès de chiffres, les chapeaux n’en contiennent souvent pas assez, voire aucun. Le chapeau apparaît alors comme une suite de généralités, ce qui ne donne pas envie d’entrer dans le texte.

**Utilisez des intertitres pour relancer l’intérêt**

Le titre et le chapeau ont attiré un certain nombre de lecteurs qui ont commencé à lire le texte. Mais leur intérêt va avoir tendance à s’émousser au fur et à mesure de la lecture. Il est bon de le relancer de temps en temps par des intertitres bien choisis.

Toutefois, la situation est un peu différente dans les textes longs. Ils sont divisés en sous-parties, chacune développant une des idées essentielles qui auront été dégagées lors de la préparation du texte. Chaque sous-partie est alors annoncée par un intertitre, qui est en fait un titre de section, et qui reprend l’idée qu’elle développe. Le choix de l’intertitre répond donc aux règles vues plus haut pour les titres.

Enfin, les encadrés, quand il y en a, méritent eux aussi de vrais titres. Des pseudo-titres comme : « encadré 1, méthodologie, législation » n’encouragent pas à les lire...

**En résumé**

* Vous avez passé du temps à préparer votre rapport et à le rédiger : donnez-lui de bons titres.
* Le titre est la dernière chose que vous allez rédiger, mais la première chose que verra votre lecteur.
* Soignez aussi les intertitres et le chapeau de votre texte

*Vous avez vu plusieurs techniques rédactionnelles qui, je l’espère, vous aideront à rédiger de meilleurs rapports. Dans l’évaluation qui suit, vous allez réécrire un extrait de rapport qui comporte quelques suprises..!*

## Entraînez-vous à rédiger les titres et le chapô d'un rapport

### À vous de jouer

Pour vous entraîner, réalisez cet exercice étape par étape. Une fois terminé, vous pouvez comparer votre travail avec les pistes que je vous propose.

À partir de cet [extrait](https://s3-eu-west-1.amazonaws.com/static.oc-static.com/prod/courses/files/R%C3%A9alisez+des+rapports+statistiques+clairs+et+impactants+/Insee+Flash+Champagne-Ardenne+Exercice+Partie+2.pdf) de rapport de l'INSEE, réalisez les éléments suivants :

1. Proposez un titre pour le rapport.
2. Proposez des intertitres pour les quatre chapitres du rapport.
3. Ecrivez le chapô du rapport.

### Vérifiez votre travail

Voici un [**exemple**](https://static.oc-static.com/activities/2465/evaluation_resources/redigez-les-titres-et-le-chapo-dun-rapport_exemple-2018-05-31T135347.zip) pour vous permettre de vérifier votre travail.

## Comprenez l'intérêt des illustrations dans un rapport statistiques

Vous connaissez désormais les grands principes de la rédaction d’un rapport statistique. Mais les rapports statistiques, ce sont aussi **des chiffres**, représentés sous forme de **tableaux**, de **graphiques** ou encore de **cartes**. Pour bien choisir et concevoir ces illustrations, vous devez comprendre leur rôle dans votre publication.

### Le rôle des illustrations dans votre rapport

1️⃣ Le rôle premier d’une illustration est de **donner de l’information chiffrée** plus ou moins précise, de **montrer une tendance**.

2️⃣ C’est aussi une bonne porte d’entrée dans un article. Nous l'avons vu, le titre et le chapeau ont déjà cette fonction, mais ils ne suffisent pas toujours. **Une bonne illustration peut susciter l’intérêt de vos lecteurs**, leur donner envie d’en savoir plus, ce qui les poussera à lire votre publication.

3️⃣ Une illustration sert à **récupérer les lecteurs susceptibles de quitter une publication**. En effet, nous sommes nombreux à être « visuels » et nous avons tendance à regarder les illustrations en priorité. Si un graphique retient notre attention, nous allons continuer notre lecture.

4️⃣ Les illustrations permettent d’**aérer la publication** que ce soit en ligne ou sur papier.  
Un graphique ou une carte permet de faire ressortir un message bien mieux que de nombreuses lignes de texte.

5️⃣ Enfin, environ 70% des récepteurs sensoriels du corps humain sont dans l’œil. **Les illustrations favorisent ainsi la mémorisation visuelle !**

Comme pour le texte, vos illustrations doivent présenter un **message essentiel**. Tableaux et graphiques sont des compléments ou des zooms sur des points du texte. Mais ils doivent être compréhensibles sans le texte.

« Avec trois nombres vous faites une phrase, avec quatre à vingt nombres vous faites un tableau, au-delà de vingt nombres vous faites un graphique » **E.Tufte, Professeur de statistiques à l’université de Yale**

Nous venons de voir l’intérêt des illustrations dans un rapport statistique, intéressons-nous maintenant spécifiquement aux tableaux et aux graphiques.

## Utilisez les tableaux à bon escient

Les data analyst travaillent sur des tableaux de données. C’est **leur matière première**. Dans une publication statistique, il est nécessaire de restituer l’information traitée mais il faut aussi faire des choix, comme pour la rédaction.

### L'utilité des tableaux

Un tableau dans une publication statistique sert à donner, sur une surface limitée, **une grande quantité d’informations utiles** (le plus souvent quantitative) sur un thème donné.

Si cela permet un gain de place, c’est parce que les informations sont « rangées » en ligne et/ou en colonne.

### Optimisez vos tableaux

Pour être efficace, un tableau doit être référencé dans le texte de manière à ce que le lecteur sache à quel moment précis il doit interrompre la lecture du texte pour regarder le tableau.

Attention, ne laissez pas de tableau « flottant », déconnecté du texte.

#### L'habillage des tableaux

Mais un tableau doit aussi être autonome. Il doit pouvoir être compris sans avoir besoin de se référer au texte. Pour cela, il doit au moins comporter les informations suivantes :

* un titre,
* les intitulés explicites des lignes et des colonnes,
* les unités employées,
* la source des données,
* le champ (= la population étudiée),
* la date de référence.

#### Rendez vos tableaux lisibles

Un tableau bien présenté peut aussi jouer le rôle d'entrée dans un texte. Le tableau est précis. Mais attention, la grande majorité de vos lecteurs ne réclame pas forcément une grande précision. Si vous mettez un tableau, prenez garde qu'il ne soit pas trop indigeste. Évitez les tableaux contenant**trop de chiffres**, même si cela vous oblige à sélectionner l'information.

#### Allégez vos tableaux

Les tableaux figurant dans les articles sont destinés à les **illustrer**. Il ne faut pas les confondre avec des tableaux que les utilisateurs consultent pour y trouver une information précise, en référence. Les tableaux de référence doivent comporter le plus de chiffres possibles, pour augmenter les chances que les utilisateurs trouvent ce qu’ils recherchent.

En revanche, les tableaux illustratifs, ceux inclus dans une publication, n’ont pas à remplir cette fonction.  
Vous devez trier les informations, choisir celles qui sont les plus susceptibles d’intéresser votre lecteur, qui illustreront au mieux le texte.

**Le texte ne doit, à aucun moment, être le double d’un tableau.** Ceci ne veut pas dire que les textes ne doivent comporter aucun chiffre : ils doivent en contenir, mais en nombre restreint. L’auteur doit sélectionner quelques chiffres sur lesquels il souhaite insister, parce que ce sont les plus importants, ou les plus inattendus.

Votre rôle d’auteur est de les commenter, c’est-à-dire de faire ressortir ce qu’ils apportent.

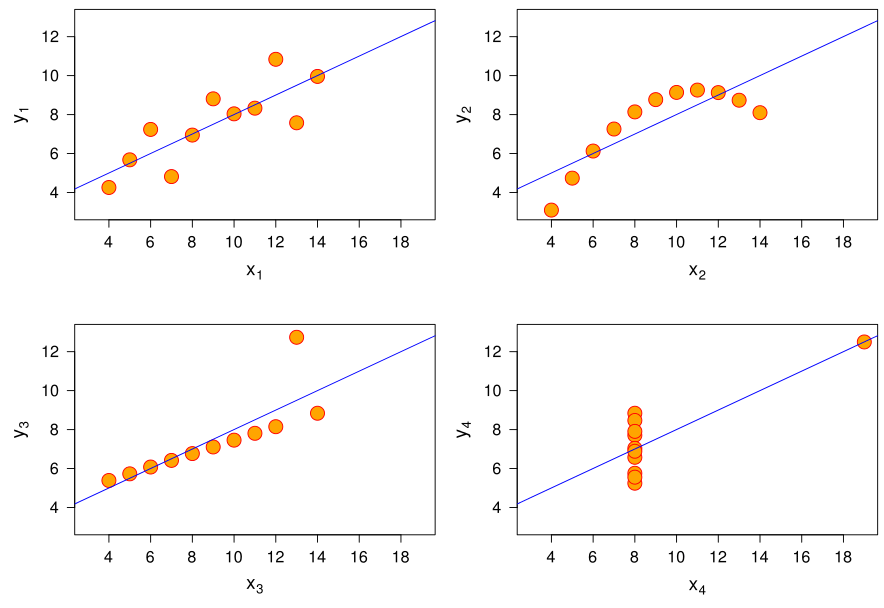
Voici quelques bonnes pratiques au sujet des tableaux. Découvrez maintenant le sel des publications satistiques : les graphiques !

## Choisissez des graphiques adaptés à votre message

Vous savez désormais quand utiliser les tableaux dans vos publications, voyons maintenant **l’intérêt des graphiques.** Vous vous doutez qu’ils sont essentiels à la communication des données… Alors, rentrons dans le vif du sujet !

### L'intérêt des représentations graphiques

Regardez l'exemple suivant :

Exemple du « Quartet d'Anscombe »

Les quatre distributions suivantes présentent des **caractéristiques statistiques communes :**

* Elles ont la même moyenne,
* le même modèle de régression linéaire,
* la même somme des carrés des résidus,
* le même coefficient de corrélation, le même R².

Besoin de vous rafraîchir la mémoire à propos de la [**régression linéaire**](https://openclassrooms.com/fr/courses/4525266-decrivez-et-nettoyez-votre-jeu-de-donnees/4774671-analysez-deux-variables-quantitatives-par-regression-lineaire) et du [**coefficient R²**](https://openclassrooms.com/fr/courses/4525266-decrivez-et-nettoyez-votre-jeu-de-donnees/4725622-analysez-la-correlation-entre-deux-variables-quantitatives) ? Suivez les liens ;)

Toutefois, la répartition des données est très différente ! C’est **la représentation graphique** qui nous permet de le comprendre en un coup d’œil.

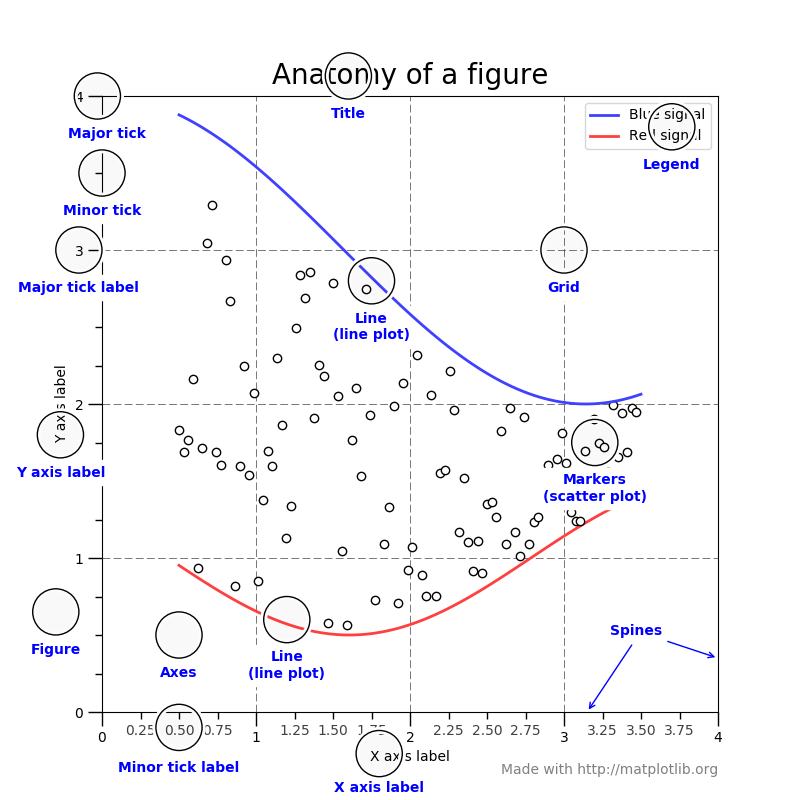
Un graphique **efficace** permet d’apporter de l’information en un seul coup d’œil. Il permet une lecture globale, immédiate et comporte un message simple.

Un graphique doit apporter une réponse correcte et complète à une question donnée, avec un temps d’observation court.

Pour tester l'efficacité de votre graphique, je vous conseille de demander à un lecteur combien de temps il met pour comprendre le graphique et en dégager un message. Au-delà de 30 secondes, il est conseillé de trouver un autre graphique.

### Anatomie d'un graphique : le vocabulaire

Prenez quelques minutes pour étudier les différents éléments d'un graphique. Le graphique ci-dessous, en anglais, a été réalisé avec la librairie Python Matplotlib (retrouvez le code à [cette adresse](https://matplotlib.org/examples/showcase/anatomy.html) !)

Anatomie d'un graphique

### Comment choisir le meilleur graphique ?

Votre représentation graphique, tel un tableau, un graphique, une carte ou encore un schéma dépendra du message que vous souhaitez délivrer.

En vue de choisir le meilleur graphique, vous devez adopter le point de vue de votre lecteur et vous demander :

• Le graphique est-il adapté au niveau de compréhension de mes lecteurs ?

• Quel est l’objectif de ce graphique ? Présenter des données, illustrer une relation... ?

• Quel type de comparaisons veut-on faire ? Une comparaison spatiale, temporelle...

• Quels éléments du graphique vont retenir l’attention du lecteur ? Une couleur, une forme particulière…

• Ces éléments servent-ils la compréhension du message ?

#### Tableau synoptique de choix d'un graphique

|  |  |  |
| --- | --- | --- |
| **Ce que je veux montrer** | **Graphique adapté** | **Pour quelles variables ?** |
| Comparer des valeurs de différentes catégories | Diagramme à barres ou batons | Toutes variables sauf les données quantitatives continues |
| Comparer des valeurs de différentes catégories | Les diagrammes à secteurs, ou « camemberts » | Uniquement pour le binaire, et le qualitatif non ordonné n’ayant pas plus de 4 ou 5 modalités |
| Comparer la valeur d’une variable par rapport à d’autres | Radar | Variable quantitative. Représentation de quantile. Entre 4 et 6 variables. |
| Suivre la valeur au fil du temps (séries chronologiques) | Graphique en courbes | Quantitatives continues |
| Comparer deux séries | Pyramide | Deux variables qualitatives |
| Une distribution | Histogramme (rectangles collés entre eux, de largeurs proportionnelles à l’amplitude des intervalles) | Quantitatives continues |
| Une distribution | Boite à moustaches |  |
| Une distribution entre deux variables | Nuage de points |  |
| Afficher l’interaction entre deux valeurs | Nuage de points | Pour les variables quantitatives (discrètes ou continues) et pour les comparaisons (régionales, internationales, sectorielles…) |
| Données relatives à la géographie | Carte | Pour les comparaisons (régionales, internationales, sectorielles…) |

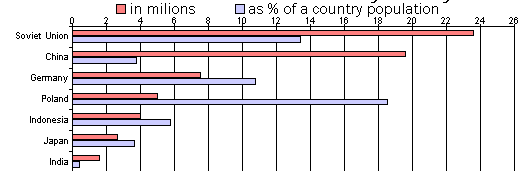
Vous avez découvert l'intérêt des graphiques et savez choisir le graphique adapté à vos données et à votre public. Voyons à présent en détail les graphiques les plus courants dans les publications statistiques.

## Découvrez les différents types de graphiques

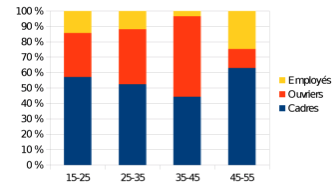
Dans ce chapitre, nous allons voir quelques graphiques courants dans les publications statistiques. Parcourez-le par curiosité, ou consultez-le plus tard pour référence.

### Diagrammes en bâtons (Bar chart)

Les diagrammes en bâtons sont efficaces pour comparer des données entre elles. Ils sont par ailleurs faciles à lire. Il existe différents types de diagrammes en bâtons : en bâtons simples, en bâtons groupés ou en bâtons empilés.



Exemple de diagramme en bâtons groupés

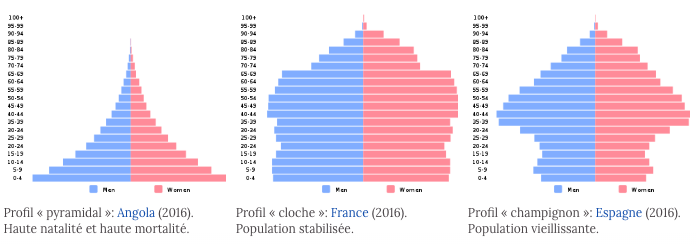


Exemple de diagramme en bâtons empilés

Attention, l’utilisation de diagramme en bâtons empilés est délicate car ils ne sont pas toujours lisibles. Le cas ci-dessus est un exemple réussi de diagramme en bâtons empilés. Il permet de comparer facilement la répartition des salariés selon la classe d’âge.

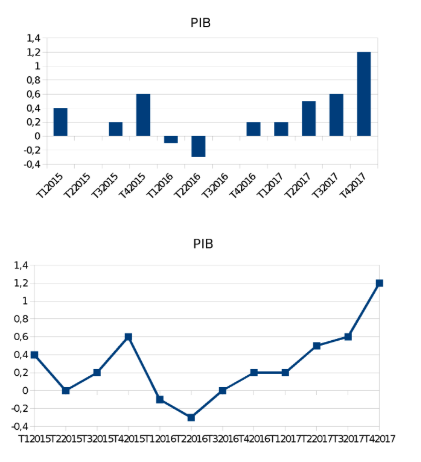
### Pyramides des âges

La pyramide est un exemple particulier mais assez classique de diagramme en bâtons. Il est possible de superposer des courbes, ce qui permet de comparer une pyramide particulière avec une pyramide de référence.

Trois profils de pyramides des âges

### Courbes (Line chart)

Les courbes sont faciles à lire et permettent la représentation de nombreuses modalités. Elles sont bien adaptées aux données temporelles.  
Dans l’exemple suivant, les mêmes données sont représentées en bâtons et en courbes. Les bâtons sont efficaces pour comparer des données entre elles. Mais les courbes montrent plus clairement les changements d’évolution.



Courbe de PIB

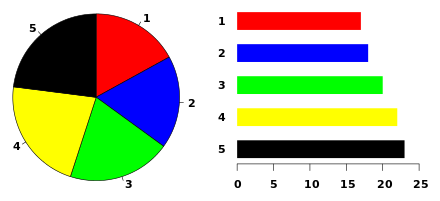
Quelques conseils :

* Bien choisir la graduation du graphique. Commencer à 0 n’est pas obligatoire.
* Penser à utiliser les indices lorsque les variables n’ont pas le même ordre de grandeur.

### Diagramme circulaire (Pie chart)

Attention, pour le diagramme circulaire (ou "camembert" comme on aime l'appeler en français) les aires ne se lisent pas toujours facilement. En effet, c’est une représentation très parlante pour comparer deux “parts” (pourcentage ou quantité). Il devient trop imprécis dès que l’on multiplie le nombre d’entrées, car nous avons du mal à comparer des angles en un coup d’oeil. Il en va de même pour son dérivé le “Donut”.

Veillez également à bien vous assurer que la totalité des secteurs soit égale à 100 !

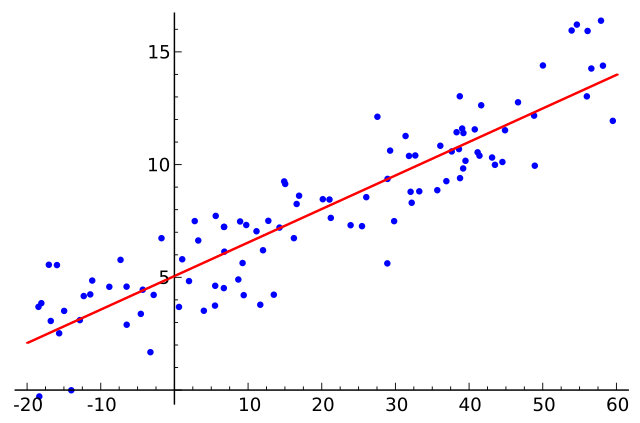


Les mêmes données sous forme de bâtons ou de diagramme circulaire

Sur l'exemple ci-dessus, on comprend que l'œil distingue mieux les différentes proportions sur les bâtons que sur le diagramme circulaire.

### Nuages de points (Scatter plot)

Les nuages de points sont des graphiques permettant de montrer des liens et des corrélations. Ces graphiques, couramment utilisés à l’heure actuelle, délivrent rapidement un message, à la condition qu’il soit simple.

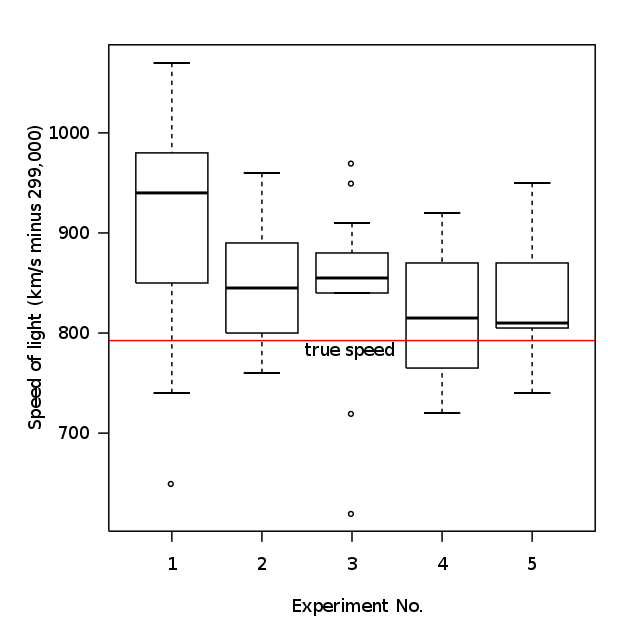


Nuage de point modélisé par une droite de régression linéaire

Pour en savoir plus sur la droite de régression linéaire, rendez-vous [**par ici**](https://openclassrooms.com/fr/courses/4525266-decrivez-et-nettoyez-votre-jeu-de-donnees/4774671-analysez-deux-variables-quantitatives-par-regression-lineaire) !

### Boîtes à moustaches (Box plot)

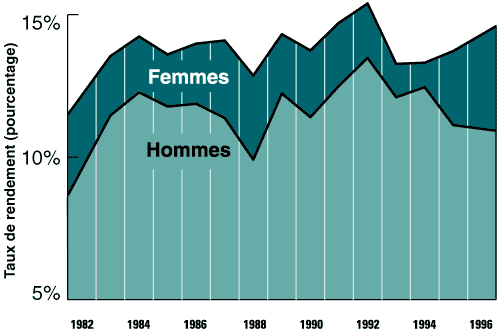
Les graphiques en "boîtes à moustaches" sont utiles à l’analyse mais constituent davantage des graphiques d’études que des graphiques de communication. En effet, ils nécessitent quelques bases en statistiques pour être interprétés.



Résultat d'une expérience de physique

### Aires empilées

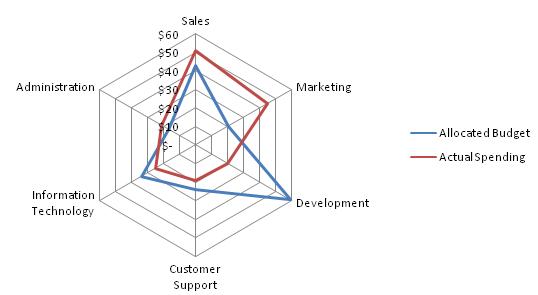
Le graphique en aires empilées ressemble au graphique en courbe, en permettant de bien visualiser les différences d'amplitudes entre les phénomènes étudiés. Veillez à bien régler le contraste entre les aires.



Taux de rendement de femmes et d'hommes (© ebsi.umontreal.ca)

### Diagramme de Kiviat (Radar chart)

Dans un graphique en radar, aussi appelé en "toile d'araignée" il y a autant d'axes que de catégories, tous les axes partent du point central.  
Un graphique radar supporte plusieurs séries, chaque série étant identifiée par une couleur différente.



Budget et dépenses d'une entreprise

### Ressources complémentaires

* Retrouvez encore plus de types de graphiques sur le site [ebsi.umontreal.ca](http://www.ebsi.umontreal.ca/jetrouve/illustre/galerie.htm)
* Apprenez à représenter une distribution avec R grâce à [ce tutoriel](https://flowingdata.com/2012/05/15/how-to-visualize-and-compare-distributions/).
* Réalisez de beaux graphiques en Python grâce à [ce tutoriel](https://www.dataquest.io/blog/making-538-plots/)
* Retrouvez [un cours OpenClassrooms](https://openclassrooms.com/courses/decouvrez-les-librairies-python-pour-la-data-science) sur les librairies Python MatPlotLib et Seaborn pour créer des graphiques.

 Vous avez compris l'intérêt des graphiques et comment choisir la bonne représentation pour vos données. Voyons à présent comment éviter les pièges classiques et obtenir des graphiques clairs et efficaces.

## Réalisez des graphiques compréhensibles

En tant que data scientist ou data analyst, votre défi quotidien est de comprendre des données complexes. Un outil important est la représentation graphique des données, à la fois pour vous-mêmes pour vous construire une intuition sur vos données, mais surtout, dans le cadre de ce cours, pour communiquer vos résultats. Et votre client ou votre supérieur aura toujours besoin de « le voir pour le croire »

### Quelques idées reçues sur les graphiques

Pas la peine de faire de jolis graphiques, avec des légendes etc. Je pourrai les refaire plus tard au propre, non ?

**FAUX !** ❌

* Vous risquez de perdre du temps à lire votre figure (et faire perdre du temps !)
* Vous risquez de faire des erreurs en lisant votre figure, et donc de mal interpréter vos données,  
  ou de manquer un signal intéressant.
* Prendre du temps pour soigner votre graphique n'est pas du temps perdu : vous pourrez le réutiliser dans le futur !

Mais si j'explique à l'oral en même temps, mon interlocuteur n'aura pas de problème pour comprendre, n'est-ce pas ?

**FAUX !** ❌

* C’est dommage, votre client a raté le moment où vous explicitiez le code couleur. Donc vous avez parlé pour rien et il s'est perdu !
* Oups vous avez inversé votre code couleur dans votre explication.
* Votre client a bien compris sur le coup, mais impossible de réexpliquer à son supérieur.

### Soignez vos graphiques

Un graphique doit pouvoir **se comprendre de façon autonome**, sans référence au texte. Il doit permettre une**lecture globale immédiate** et **transmettre un message simple**.

Pour être efficace, un graphique doit donc comporter toutes les informations nécessaires à sa compréhension :

#### Niveau 1 : un graphique compréhensible doit comporter...

* Un titre (inclus dans la figure, ou sur la slide, ou dans le texte)
* Une légende
* Le nom des axes (nom de la variable, éventuellement unité, échelle - données  
  normalisées, pourcentage...)
* Des caractéristiques **lisibles.**Si vous pensez que votre graphique est lisible sur votre ordinateur, n'oubliez pas que sur un vidéoprojecteur, la taille et les couleurs peuvent changer.

#### Niveau 2 : réalisez un graphique lisible...

* Attention à la taille des textes (titres, titres des axes, tick labels...)
* Attention à la taille des marqueurs, à l'épaisseur des lignes.
* Attention aux légendes coupées.
* Attention à la résolution en sauvegardant vos graphiques ! Utilisez la fonction  savefig  ,et ne pas copier-coller vos images depuis un notebook jupyter.
* Si vous faites des graphiques très complexes, avec beaucoup d’éléments, pensez à ajuster la taille du graphique.
* Prenez l’habitude de sauvegarder vos graphiques dans un format vectorisé (pdf par exemple).
* Attention au choix des couleurs. Deux mots d'ordre : **lisibilité** et **pertinence**.

#### Niveau 3 : optimisez vos graphiques...

Il peut être pénible de changer à chaque fois la taille des textes, le fond, savoir si on met une grille arrière ou pas ! Pour cela, plusieurs solutions existent (par exemple ici pour python) :

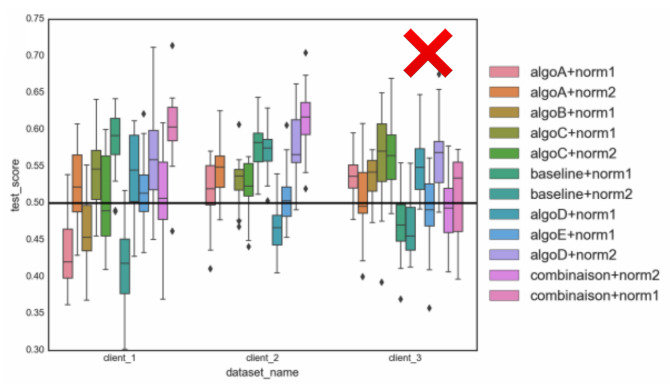
* Utiliser des styles prédéfinis. Vous pouvez ainsi redéfinir le comportement “par défaut” de  
  [matplotlib](https://matplotlib.org/devdocs/gallery/style_sheets/style_sheets_reference.html).
* C'est aussi possible avec [seaborn](https://seaborn.pydata.org/tutorial/aesthetics.html" \t "_blank), qui distingue des styles (couleurs par défaut, fonds) et des contextes.
* Définir des fonctions pour les opérations que vous faites souvent. Par exemple, changer automatiquement l’affichage des dates si vous travaillez avec des données temporelles.

### « Do » et « Don't » des représentations graphiques

Voyons ensembles les différents pièges courants dans lesquels ne pas tomber. Vous comprendrez comment faire des graphiques lisibles, qui serviront avec pertinence votre message.

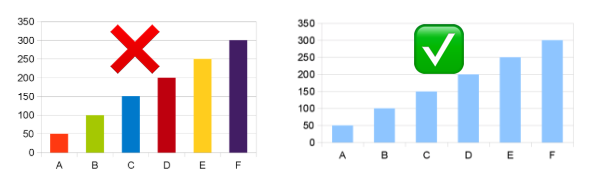
#### Représenter trop d'informations

Un lecteur ne peut pas mémoriser plus de dix objets différents. On ne peut pas tout dire avec un graphique. Il vaut donc mieux faire plusieurs graphiques simples qu’un seul graphique complexe.

Attention à la surcharge d'informations

#### N'abusez pas des couleurs !

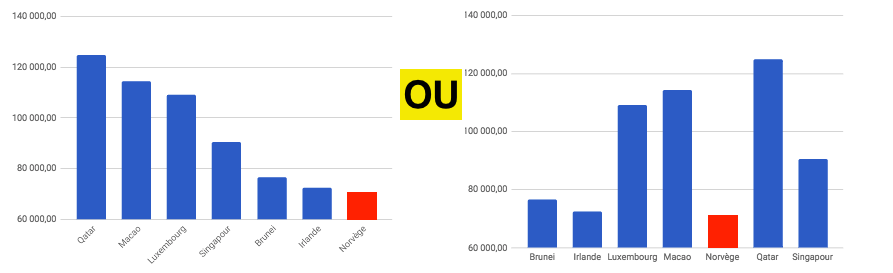
Vous ne devez pas multiplier les informations, même les couleurs sur un même graphique.  
Utilisez des couleurs mais sans en abuser. Ainsi, regardez les deux graphiques ci-dessous : Ils transmettent la même information, mais le message est parasité par les couleurs dans le premier. Le choix des couleurs qui composent une visualisation doit faire intrinsèquement parti du message.

Attention aux couleurs inutiles !

De la même façon certaines couleurs sont communément associées à des groupes, à des catégories, à une zone géographique. En France par exemple, une carte de résultats électoraux où les résultats seraient en rouge pour la droite et bleu pour la gauche, quelle qu’en soit la légende, perdrait assurément le lecteur.

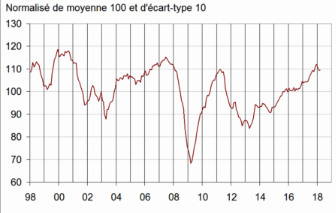
#### Triez vos données !

Pour les diagrammes en bâtons, il est important de trier les données.  
Dans l’exemple suivant, il est plus utile que les données soient triées par valeurs plutôt que par ordre alphabétique. Par ailleurs, il est conseillé de distinguer le territoire d’analyse par une couleur différente (ici, la Norvège).

Dans cet exemple, les données sont triées par PIB ou par ordre alphabétique

#### Représentez le temps

Pour la représentation de séries temporelles, il est important de se demander si la période choisie ne fausse pas le jugement. En effet, on peut facilement induire en erreur en sélectionnant une période particulière. Les données en soi ne sont pas fausses mais c’est leur sélection, plus ou moins volontaire, qui biaise l’information. Par ailleurs, il est recommandé de représenter le temps sur l’axe horizontal des x.

Voici un exemple qui montre une tendance à la hausse entre 2013 et 2018, qui est toute relative dès lors que l’on regarde depuis 1998.

#### Attention au bon découpage en classes

Pour représenter des variables quantitatives continues, on est parfois amené à regrouper des valeurs en classes, en tranches. Par exemple, on dira que 18% d'une population a un âge compris entre 0 et 10ans, 29% entre 10 et 15ans, etc.

❌ Dans le cas de variables quantitatives continues (par exemple l’âge), ce découpage n’est pas neutre.

✅ Pour pouvoir comparer les effectifs par tranche, il faut veiller à ce que l’étendue de ces tranches soit identique !

Bien, vous avez découvert les pièges à éviter lors de la création de graphique, ainsi que des conseils pour rendre vos graphiques clairs et efficaces. Prêt à vous entrainer ? Ça se passe dans le prochain chapitre.

## Entraînez-vous à illustrer un rapport avec des graphiques pertinents

### À vous de jouer

Pour vous entraîner, réalisez cet exercice étape par étape. Une fois terminé, vous pouvez comparer votre travail avec les pistes que je vous propose.

Pour réaliser cet exercice, vous devrez être capable de créer des représentations graphiques à partir de données (en Python ou en R, par exemple).

À partir de cet extrait de [rapport](https://s3-eu-west-1.amazonaws.com/static.oc-static.com/prod/courses/files/R%C3%A9alisez+des+rapports+statistiques+clairs+et+impactants+/Texte+Exercice+Partie+3.pdf) et ce [jeu de données](https://s3-eu-west-1.amazonaws.com/static.oc-static.com/prod/courses/files/R%C3%A9alisez+des+rapports+statistiques+clairs+et+impactants+/Jeu+de+donnees+Partie+3.xlsx), réalisez une illustration adaptée pour chacun des trois paragraphes.

### Vérifiez votre travail

Voici un [exemple](https://static.oc-static.com/activities/2466/evaluation_resources/illustrez-un-rapport-avec-des-graphiques-pertinents_exemple-2018-05-31T135546.zip) pour vous permettre de vérifier votre travail.