

Manipulation de graphiques

"Lies, damned lies, and statistics" est l'une des premières citations que nous pouvons trouver dès que nous nous intéressons à la littérature autour de la manière de manipuler les opinions avec les statistiques. Son origine a longtemps été attribuée à Léonard Courtney (6 juillet 1832 - 11 mai 1918), un homme politique ayant été Président de la Royal Statistical Society, un organisme professionnel de statisticiens et un organisme de bienfaisance qui fait la promotion des statistiques pour le bien public, créée en 1834 et est encore en activité aujourd'hui.^{1 2} Cependant, son origine n'est à ce jour pas encore déterminée, chaque source indiquant un auteur différent. Cette citation illustre le fait que les statistiques puissent être utilisées pour mentir plus que ne le feraient de « sacrés mensonges », c'est ce que nous allons étudier dans ce dossier.

Un graphique est une représentation des variations d'un phénomène et transcrit des données par le dessin. Les graphiques qui nous sont présentés par des journaux, les journaux télévisés, certaines études peuvent nous manipuler si ces moyens d'accès à l'information ne sont pas objectifs sur celle-ci. La littérature nous renseigne, sur certains facteurs qui peuvent influencer nos pensées et biaiser notre perception de l'information qui se veut rester neutre, tels que l'origine du graphique, l'échelle choisie pour l'axe des abscisses, l'échelle choisie pour l'axe des ordonnées. La perspective est aussi un facteur important. Les graphiques de type secteur en trois dimensions sont à éviter, certains éléments peuvent être placés au premier plan pour leur donner plus de relief et minimiser l'impact d'autres éléments. De plus, la valeur nulle n'est pas représentée dans ce type de graphique.^{3 4} Les graphiques peuvent nous dire quoi penser sur un sujet sans mentir ni modifier les valeurs. Il faut par exemple se méfier des graphiques dont l'échelle ne commence pas à zéro. Nous pouvons nous demander si la façon dont est construit un graphique peut influencer nos opinions face aux idées qu'il véhicule.

¹ (*Royal Statistical Society on JSTOR*, 2022).

² (*Leonard Henry Courtney, Baron Courtney | British Politician*, 2021)

³ (*How to Lie with Statistics*, 1954)

⁴ (*Wheeler*, 1976)

Pour répondre à cette question de recherche, j'ai imaginé un protocole expérimental consistant à présenter à une classe d'élèves de terminale un graphique présentant le chiffre d'affaires fictif réalisé par des employés de commerce, au même poste, dans une même entreprise. Après avoir divisé la classe en deux parties, je présenterai un premier graphique au premier groupe d'élèves. Ce graphique sera produit avec une échelle standard, avec zéro pour origine du graphique, sans coupure d'axe, se voulant objectivement neutre dans les idées qu'il veut véhiculer. Je présenterai un second graphique au second groupe d'élèves. Ce graphique sera produit de manière à observer de grandes différences entre les différents chiffres d'affaires des différents employés, avec une origine du graphique proche du chiffre d'affaires le plus bas.

Trois échelles de Likert seront proposées aux lycéens avec 8 options de réponse aux affirmations suivantes.

- La différence entre les différents chiffres d'affaires est (0 = nulle / 7 = extrême)
- Certains employés ont mieux travaillé que d'autres (0 = pas du tout d'accord / 7 = tout à fait d'accord)
- Certains employés ont plus travaillé que d'autres (0 = pas du tout d'accord / 7 = tout à fait d'accord)

Le but recherché avec cette échelle de Likert est d'éviter les réponses intermédiaires et de couvrir un large spectre d'opinion possible.

Hypothèse générale : La façon dont est construit un graphique peut influencer nos opinions face aux idées qu'il véhicule.

Problématique : La façon dont est construit un graphique peut-elle influencer nos opinions face aux idées qu'il véhicule ?

VI (provoquée) : Type de graphique (neutre vs biaisé).

VD : Résultat à l'affirmation proposée grâce à une échelle de Likert sur 7.

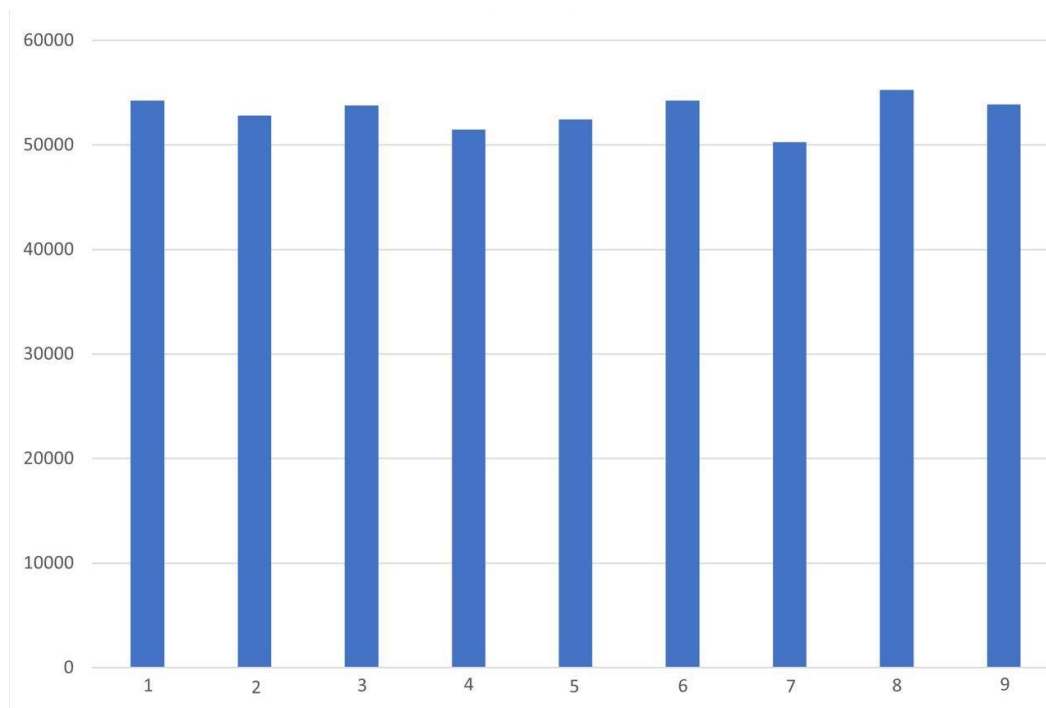
Hypothèses opérationnelles :

- Le premier groupe aura des résultats plus élevés sur les différentes échelles de Likert que le deuxième groupe.
- Le deuxième groupe aura des résultats plus élevés sur les différentes échelles de Likert que le premier groupe.
- Le type de graphique présenté influe sur le résultat à la question posée.

Pour minimiser l'impact des facteurs non-contrôlés, mon protocole expérimental sera réalisé sur des individus de même classe (terminale), l'échantillon sera de même nombre pour les deux groupes. Les spécialités des lycéens seront équitablement réparties entre les groupes car on peut supposer que des lycéens ayant l'enseignement de spécialité de Mathématiques soient avertis face aux techniques pour modifier notre perception d'un graphique car c'est un sujet qui peut avoir été abordé en cours par le professeur.

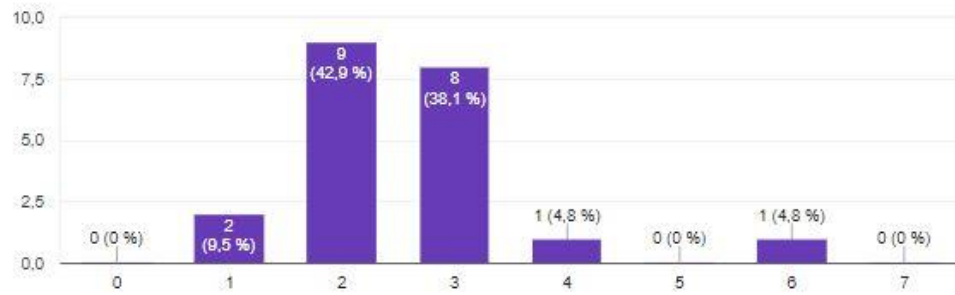
Ce protocole expérimental m'a été refusé au lycée Alphonse Daudet à Nîmes par M. Pascal Lorblanchet, le proviseur. J'ai donc reproduit le même protocole qu'expliqué précédemment, sous la forme de deux questionnaires en ligne. Le premier questionnaire correspondant au premier graphique (neutre). Le deuxième correspondant au second graphique (biaisé). Les réponses ont été fournies, après consentement éclairé des participants, par des étudiants, en Arts, de STAPS, d'Histoire, d'Ergothérapie et de Psychologie. La répartition des étudiants entre les groupes a été effectuée en randomisation simple. Chaque étudiant n'a répondu qu'à un des deux questionnaires. Les questions sont identiques entre les deux questionnaires, seuls les graphiques changent. La seule différence entre les deux graphiques est l'origine. J'ai collecté 21 réponses pour le premier questionnaire et 24 réponses pour le second questionnaire. Cette différence s'explique par la randomisation neutre. Voici les résultats obtenus.

Étude n°1 (graphique objectivement neutre)



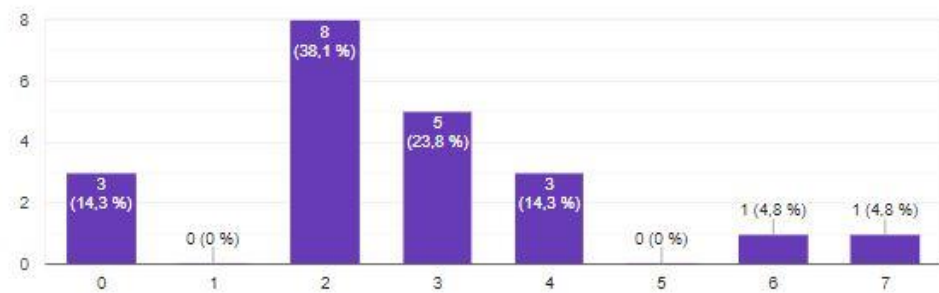
La différence entre les différents chiffres d'affaires est (0 = nulle / 7 = extrême)

21 réponses



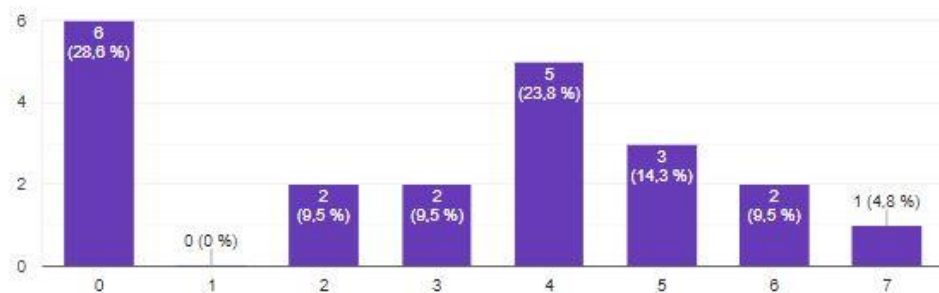
Certains employés ont mieux travaillé que d'autres (0 = pas du tout d'accord / 7 = tout à fait d'accord)

21 réponses

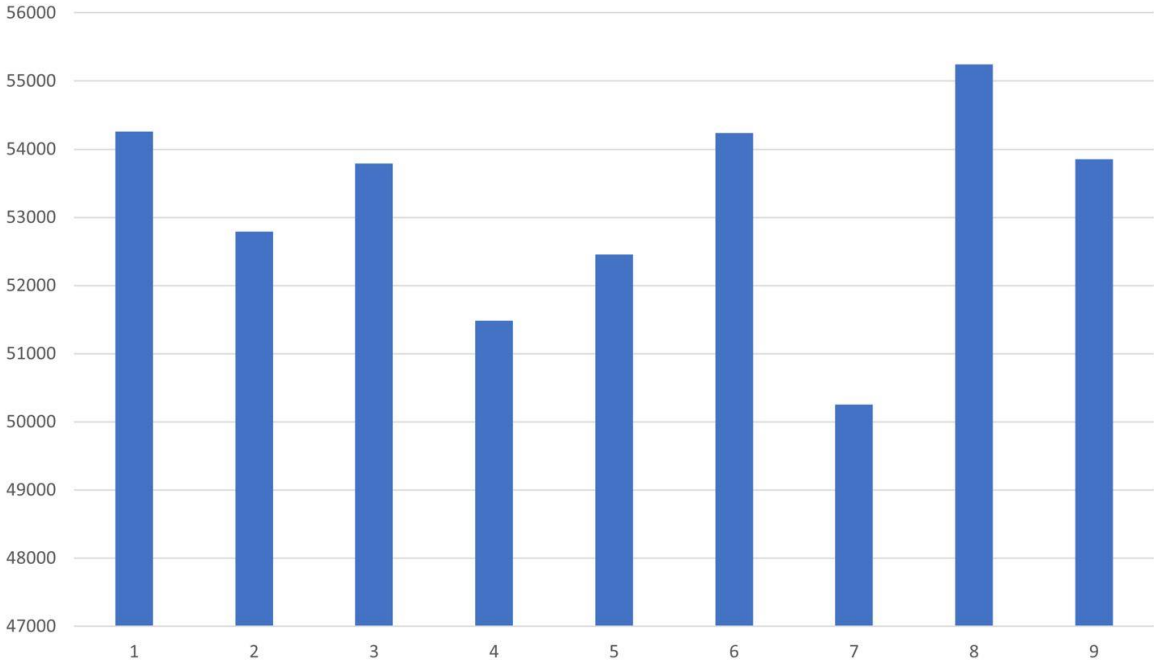


Certains employés ont plus travaillé que d'autres (0 = pas du tout d'accord / 7 = tout à fait d'accord)

21 réponses

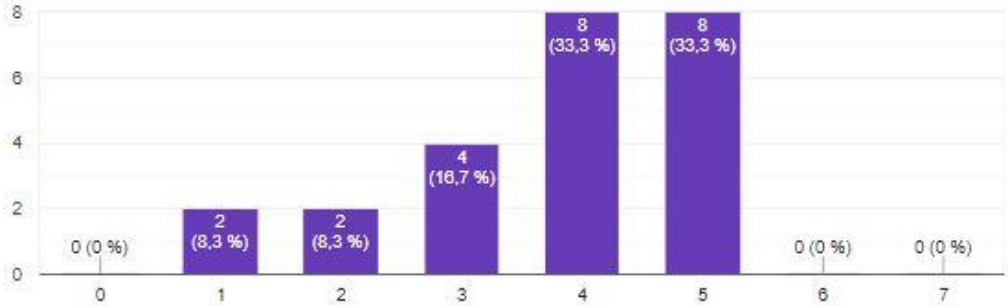


Étude n°2 (graphique biaisé)



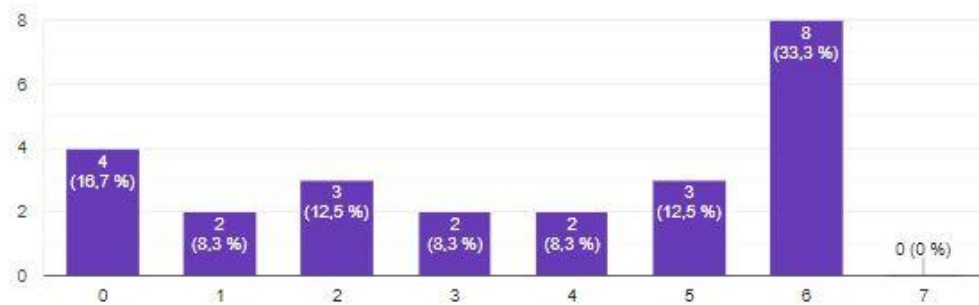
La différence entre les différents chiffres d'affaires est (0 = nulle / 7 = extrême)

24 réponses



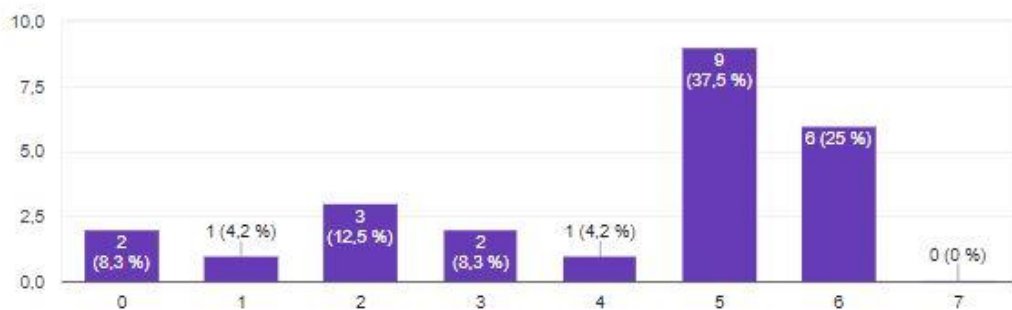
Certains employés ont mieux travaillé que d'autres (0 = pas du tout d'accord / 7 = tout à fait d'accord)

24 réponses



Certains employés ont plus travaillé que d'autres (0 = pas du tout d'accord / 7 = tout à fait d'accord)

24 réponses



Les statistiques inférentielles n'étant pas traitées en première année de Psychologie, je ne suis malheureusement pas en mesure de les appliquer à mes résultats. Je vais donc comparer les pourcentages obtenus des différentes modalités de chaque affirmation entre les différents questionnaires en utilisant certains outils vus en cours de statistiques pour psychologues ce semestre.

Concernant la première affirmation du premier questionnaire, le mode est représenté par la modalité 2, qui concerne 42,9% des réponses des participants. La médiane est aussi de 2. En comparaison, le mode de la première affirmation du deuxième questionnaire est représenté par la modalité 4 (33,3%) et la modalité 5 (33,3%). La médiane est de 4. Nous pouvons observer une différence significative entre les réponses des différents groupes. En moyenne, le premier groupe estime une différence moindre entre les différents chiffres d'affaires des différents employés que le deuxième groupe.

Concernant la deuxième affirmation du premier questionnaire, le mode est représenté par la modalité 2 (38,1% des réponses des participants), tout comme la médiane. En comparaison, le mode de la deuxième affirmation du deuxième questionnaire est représenté par la modalité 6 (33,3%). La médiane est de 4. Nous pouvons observer une différence significative entre les réponses des différents groupes. En moyenne, le premier groupe estime être peu d'accord avec le fait que certains employés aient mieux travaillé que d'autres. En revanche, en moyenne, le deuxième groupe estime être d'accord avec le fait que certains employés aient mieux travaillé que d'autres.

Concernant la troisième affirmation du premier questionnaire, le mode est représenté par la modalité 0 (28,6% des réponses des participants), la médiane est de 4. En comparaison, le mode de la troisième affirmation du deuxième questionnaire est représenté par la modalité 5 (37,5%). La médiane est de 5. La médiane entre les différents groupes est proche, seul le mode tendrait à une différence entre les deux groupes. 28,6% des participants du premier groupe ont répondu 0 contre 8,3% pour les participants du deuxième groupe concernant cette question. Cependant, la médiane est un outil plus pertinent que le mode en statistiques. Nous pouvons donc émettre comme hypothèse que pour cette question concernant la quantité de travail, le manque d'informations ou les idées préconçues des participants ont joué sur leurs réponses.

Pour conclure, mon hypothèse générale, la façon dont est construit un graphique peut influencer nos opinions face aux idées qu'il véhicule, est vraie dans l'étude que j'ai menée. Cependant elle est critiquable sur trois points majeurs. Premièrement, le participant est au courant qu'il passe une étude dès lors qu'il clique sur le lien pour répondre au questionnaire, ce qui fait qu'il est sujet au biais, ses réponses ne sont pas issues d'une situation écologique. C'est d'ailleurs le cas de la majorité des études réalisées. Deuxièmement, la population n'est pas représentative de la population générale, les étudiants interrogés ont 20 ans en moyenne, la population est trop homogène. Troisièmement, l'étude est réalisée en simple aveugle car, ayant créé deux questionnaires, je savais quelles réponses correspondaient à chaque questionnaire puisque deux feuilles de calculs différentes ont été créées. Le double aveugle aurait été possible si un autre éventuel membre du groupe avait analysé les données statistiques, différent de celui qui a créé les questionnaires. Les hypothèses opérationnelles « le deuxième groupe aura des résultats plus élevés sur les différentes échelles de Likert que le premier groupe. » et « le type de graphique présenté influe sur le résultat à la question posée. » sont vraies. En moyenne, le groupe 2 a tendance à affirmer que la différence entre les chiffres d'affaires est importante et

que certains employés ont plus et mieux travaillé que d'autres, là où le groupe 1 émet des avis plus modérés sur ces questions. En outre, à la fin de chaque questionnaire, j'ai laissé une place pour que le participant puisse s'exprimer s'il en avait l'envie. Ce n'est pas une variable qui m'intéressait lors de l'étude. Après réflexion et voyant les différences concernant le nombre de réponses, les participants s'expriment beaucoup plus sur la deuxième étude que sur la première. Les remarques concernant la première étude disent que les chiffres d'affaires sont plus ou moins égaux. Les remarques de la deuxième étude viennent critiquer les questions et certains participants affirment que « On ne peut pas connaître la charge de travail qu'il y a derrière les résultats obtenus. » ou encore que « La différence entre les différents chiffres n'est pas aussi importante que le graphique veut nous faire croire. » J'ai reçu 2 réponses (environ 9,5%) pour la première étude contre 9 réponses (37,5%) pour la deuxième étude. Nous pouvons émettre l'hypothèse qu'apporter des jugements de valeurs aux personnes salariées pose un problème éthique aux participants, c'est pourquoi ils s'expriment plus dans l'étude 2 car c'est dans l'étude 2 que le jugement est le plus inéquitable entre les différents employés. De plus, nous pouvons émettre une autre hypothèse. Au cours de la passation du questionnaire, le participant est actif, réfléchit, il produit lui-même ses réponses. Il sait qu'il passe une étude. Nous pouvons donc penser que c'est ce qui explique la pertinence des remarques effectuées. Nous pouvons nous demander, lorsque le participant est passif, s'il est plus suggestible aux informations voulant être véhiculées par une source d'information quelconque. Cette expérience nous montre l'impact que peuvent avoir certaines stratégies de manipulation de graphiques sur nos pensées. Ici, la conséquence n'est que minime mais ces mêmes manipulations peuvent s'appliquer à des graphiques promouvant un médicament, mettant ainsi en jeu notre santé.

Bibliographie

(1954) *How to Lie with Statistics*. By Darrell Huff.

Wheeler, M. (1976). *Lies, Damn Lies, and Statistics : The Manipulation of Public Opinion in America*. W. W. Norton & Company.

Royal Statistical Society on JSTOR. (2022). Royal Statistical Society.
<https://www.jstor.org/publisher/rss>

Britannica, T. Editors of Encyclopaedia (2021, July 2). *Leonard Henry Courtney, Baron Courtney*. *Encyclopedia Britannica*.

<https://www.britannica.com/biography/Leonard-Henry-Courtney-Baron-Courtney-of-Penwith>