

Données démographiques relatives à la consommation de bière

par Joel Yan, Statistique Canada, (joel.yan@statcan.ca), John Kezys Mohawk, Collège,
(Activité étudiante actualisée par Shirley Dalrymple, Conseil scolaire du district York, shirleydalrymple@hotmail.com)

« Les plus grands consommateurs de bière nationale sont les jeunes hommes et les jeunes femmes à l'esprit sportif. »
Pendant les années 1980, la cohorte des 20 à 24 ans a présenté la plus forte croissance de consommation de bière.
Foot D., 1996, *Boom Bust & Echo*, Macfarlane Walter & Ross, ISBN 0-921912-97-8.

Tâche à exécuter

Veillez ne pas écrire sur ces feuilles. Inscrivez vos réponses dans un document Word ou à la main. Vous pouvez copier vos graphiques Fathom^{MC} dans le document Word ou les imprimer directement. **Présentez vos réponses et tous les graphiques avant la fin de la période. Vous devez répondre aux questions marquées d'un *.**

Utilisez E-STAT pour vérifier l'hypothèse de M. Foot en créant des tracés de séries chronologiques pour la cohorte des hommes âgés de 20 à 24 ans et de la consommation de bière au Canada entre 1971 et aujourd'hui.

Procédure

Pour accéder aux données sur la consommation de bière

1. Démarrer E-STAT. (<http://estat.statcan.ca>)
2. Extraire les données des bases de données de CANSIM sur E-STAT en cliquant sur « Recherche dans CANSIM » dans la barre latérale gauche. Ce bouton se trouve vers le milieu de la barre latérale gauche.
3. Sous « Méthode de recherche », sélectionner « Mot-clé » et cliquer sur « Continuer ».
4. Dans la boîte 3 **Rechercher :**, taper « consommation de bière », puis cliquer sur « Trouver les tableaux ».
5. Sous **Tableaux actifs**, sélectionner « 002-0011 consommation apparente des aliments par personne au Canada, annuel ».
6. Sous Produits de base, défiler vers le bas et sélectionner « Ale, bière, stout et porter, poids de détail (Litres par année) ».
7. Sélectionner la période de « 1971 » à l'année la plus récente.
8. Cliquer sur l'onglet « Extraire séries chronologiques ».
9. Cliquer sur « Ajouter d'autres séries » dans le bas de la page.

Pour accéder aux données de la cohorte du groupe des 20 à 24 ans

10. Sous « Recherche : », sélectionner « Sujets » et cliquer sur « Continuer ».
11. Cliquer sur la boîte jaune pour afficher « Population et démographie ».
12. Cliquer sur « Caractéristiques de la population ».
13. Sélectionner « Tableau 051-0001 Estimations de la population, selon le groupe d'âge et le sexe, Canada, provinces et territoires, annuel ».
14. Sous **Géographie**, sélectionner « Canada ».
15. Sous **Sexe**, sélectionner « Les deux sexes ».
16. Sous **Groupe d'âge**, sélectionner deux séries chronologiques différentes. Sélectionner d'abord « Tous âges », puis défiler vers le bas et (tout en maintenant la touche Ctrl sous Windows ou la touche Command sous Macintosh) sélectionner également « 20 à 24 ans ».
17. Sélectionner la période de « 1971 » à l'année la plus récente. Vous devez avoir la même période que celle sélectionnée pour la série précédente, consommation de bière.
18. Cliquer sur l'onglet « Extraire séries chronologiques ».

19. Dans la boîte Choix du format de sortie, choisir « WK1 (fichier générique pour tableau), périodes = lignes ».
20. Cliquer sur « Extraire ».
21. Cliquer sur « Open ».

Analyse Fathom

Copier et coller des données

22. Cliquer et parcourir le tableau pour sélectionner toutes les données, sauf la légende du haut de la page ou la ligne de source au bas de la page.
23. Choisir **Copier** dans le menu **Édition**.
24. Passer à **Fathom**. (Si Fathom n'est pas déjà activé, vous devrez le lancer.)
25. Dans un nouveau document, créez un nouveau « cadre » pour un ensemble de données.
26. À partir de la collecte sélectionnée, choisir **Coller les cas** dans le menu **Édition**.
27. Créer un tableau de cas pour la collecte (en choisissant, par exemple, **Tableau de cas** dans le menu **Insérer**).
28. Vous devrez peut-être vérifier les noms d'attribut et supprimer le premier cas.
29. Changer le nom de la collecte pour Consommation de bière.
30. Changer les noms des attributs pour Année, Population totale, Pop20_24, Consommation de bière.
31. Sauvegarder le document Fathom en sélectionnant **Enregistrer** dans le menu **Fichier**.

Présenter les données sous forme de graphique

32. Créer un graphique indiquant la variation de la consommation de bière par personne par rapport aux années.
33. Décrire la tendance qui ressort de ce graphique. Donner une explication possible de la forme de la distribution. *

Manipuler les données – Convertir en pourcentages

Pour étudier l'affirmation de M. Foot, vous devez vérifier la variation dans la cohorte des 20 à 24 ans pour voir si ces personnes sont les plus grands consommateurs de bière, en moyenne. Comme on évalue la consommation de bière en fonction du nombre moyen de litres consommés par année par population totale, nous devons convertir la cohorte des 20 à 24 ans en pourcentage de la population totale.

34. Créer un graphique indiquant la population des gens âgés de 20 à 24 ans par rapport au temps. Comparer le deuxième graphique au graphique sur la consommation de bière. Décrire les liens, le cas échéant. *
35. Pour établir le pourcentage que représente la cohorte des 20 à 24 ans sur la population totale, il faut définir un nouvel attribut : ouvrir le tableau des cas et cliquer sur l'attribut marqué (nouveau) dans le coin supérieur droit du tableau des cas.
36. Donner un nom à ce nouvel attribut; taper Pop20_24commepourcentage.
37. Pour attribuer les valeurs de ce nouvel attribut, sélectionner le menu **Édition**, puis **Formule d'édition**.
38. Dans la boîte de formule, augmenter la liste des attributs (cliquer sur le signe ⊕ à côté de Attributs). Taper la formule suivante : $100 * (\text{Pop20_24} / \text{Pop totale})$.
39. Cliquer sur **Appliquer**. Vous verrez les valeurs calculées dans le tableau des cas.

Établir des liens possibles en utilisant les techniques de création de graphique

40. Créer un graphique montrant la consommation de bière par rapport au nouvel attribut calculé.
41. À l'aide du menu déroulant **Graphique**, superposer une ligne des moindres carrés.
42. Ajouter un graphique de valeurs résiduelles pour faciliter l'évaluation de la pertinence de la ligne des moindres carrés.

*Analyse **

Décrire la corrélation des données en utilisant la ligne la plus pertinente et le graphique de valeurs résiduelles. (Formuler des commentaires sur la force, la direction, les tendances.)

Expliquer la tendance que suit la consommation de bière au fur et à mesure que le pourcentage du groupe des 20 à 24 ans augmente dans la population.

Ce modèle tend-il à appuyer ou à infirmer l'affirmation de M. Foot? Expliquer.

Si vous étiez un spécialiste en commercialisation, de quelle façon utiliseriez-vous ces renseignements pour vendre de la bière?

Préparer un rapport sur un des éléments suivants : *

- ✓ Examiner la relation existant entre la consommation moyenne de bière des 20-24 ans et la consommation moyenne de bière d'une autre cohorte d'âge, comme celle des 15-19 ans ou des 25-29 ans.
- ✓ Examiner la consommation d'une autre boisson (p. ex., vin, lait) en utilisant le même tableau de données de E-STAT et déterminer les cohortes d'âge qui consomment le plus le produit sélectionné, le cas échéant.

Cette tâche sera évaluée sur le plan de la réflexion, de l'interrogation ou de la résolution de problèmes.

	Niveau 1	Niveau 2	Niveau 3	Niveau 4
CONSEILS	- Respecte les directives d'enquête avec beaucoup d'aide - Tire des conclusions et formule des inférences ayant peu rapport aux données	- Respecte les directives d'enquête avec une certaine d'aide - Tire des conclusions et formule des inférences qui se rapportent un peu aux données	- Respecte les directives d'enquête avec très peu d'aide - Tire des conclusions et formule des inférences qui se rapportent aux données	- Respecte les directives d'enquête sans aide - Tire des conclusions et formule des inférences perspicaces qui se rapportent aux données

Notes, graphiques et activités de suivi auxquels l'enseignant doit avoir recours

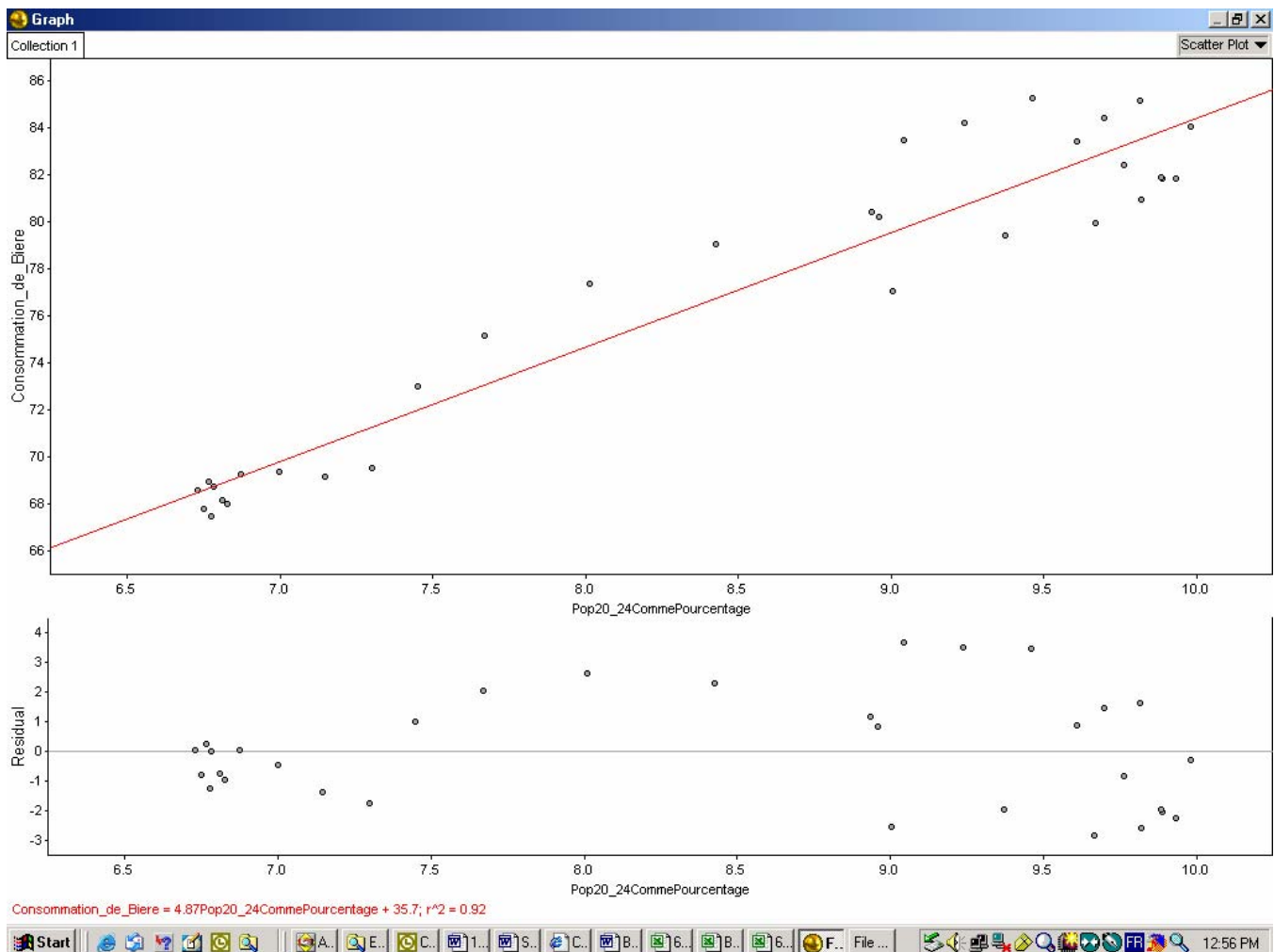
Objectif de cette activité

Cette activité est conçue pour présenter à une classe de mathématique la façon d'utiliser la base de données CANSIM de E-STAT à l'aide de Fathom! Cette activité donne un exemple de la façon de combiner différentes séries chronologiques comportant plus de 1 600 grands tableaux multidimensionnels de CANSIM afin d'analyser les relations possibles entre les ensembles de données. Toutes les séries chronologiques de CANSIM, dont le nombre excède dix millions et qui sont établies à partir d'environ 100 enquêtes différentes, peuvent être regroupées pour l'exploration des relations possibles. Il suffit que les périodes de référence liées aux données se chevauchent.

Attentes liées aux éléments de mathématique du cours de gestion des données

- Résoudre des problèmes complexes à l'aide de diagrammes. (*Gestion des données- Attentes - page 51*)
- Trouver des données pour répondre à une question d'intérêt personnel, en effectuant une recherche dans une base de données bien organisée. (*Organisation des données - page 51*)
- Décrire la relation entre deux variables en utilisant des graphiques en nuage de points et en interprétant le coefficient de corrélation. (*Statistiques- Attentes- page 54*)

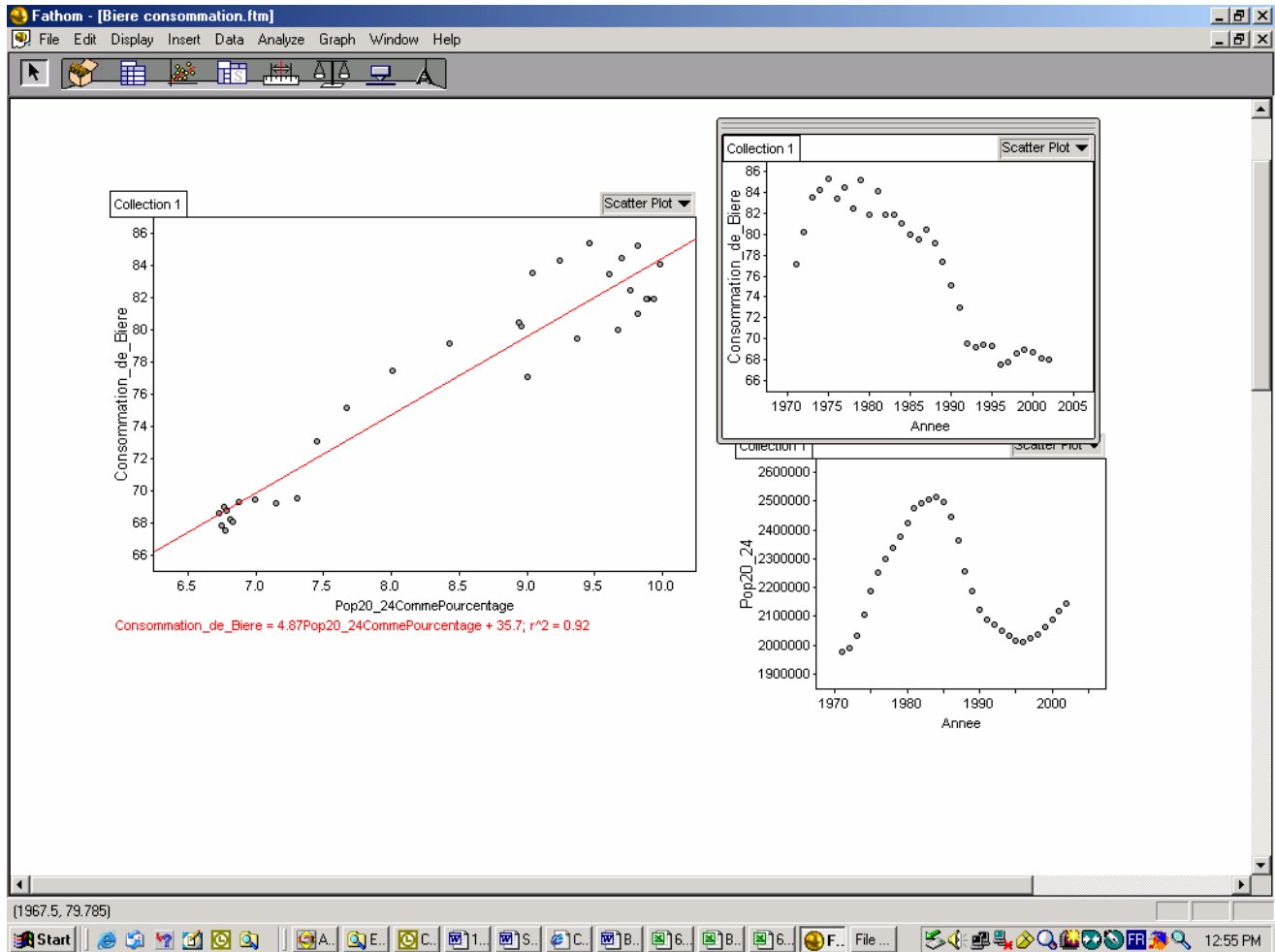
Figure 1 : Graphique en nuage de points et graphique de valeurs résiduelles montrant la relation entre la consommation moyenne de bière des Canadiens en litres par année (sur l'axe des y) et le pourcentage des membres de la population canadienne totale qui sont âgés entre 20 et 24 ans (sur l'axe des x).



Nota : Il existe une très forte corrélation (r^2 de 0,92) entre la consommation moyenne de bière (illustrée sur l'axe des y ci-dessus) et la cohorte des 20 à 24 ans, exprimée en pourcentage de la population totale (illustrée sur l'axe des x ci-dessus).

Question : Expliquer ce que nous révèle le graphique de valeurs résiduelles sous le graphique en nuage de points.

Figure 2 : Graphiques à produire au cours de cette activité



Questions additionnelles à analyser

Comment la population de la cohorte des 20 à 24 ans varie-t-elle au fil du temps?

Quelle fonction semble modeler cette population au fil du temps?

Fathom calcule que le coefficient de corrélation du graphique en nuage de points ci-dessus est de 0,92. Qu'est-ce que cela signifie?

Activité d'enrichissement optionnelle

- ✓ Trouver une autre relation intéressante en utilisant deux tableaux différents à partir des 1 600 tableaux multidimensionnels et plus du CANSIM de E-STAT.