

Projet de statistique

On s'intéresse à la stature des Français. Téléchargez le fichier de données `stature` sur le module spiral `statistiques`. Ce fichier contient quatre échantillons de même longueur, chacun des échantillons constituant une colonne. Le premier échantillon est constitué des statures d'hommes habitant en France, le second de femmes habitant en France. Le troisième et le quatrième échantillons sont mesurés respectivement sur des hommes et femmes habitant au Canada.

- Question 1 Réalisez une étude descriptive séparée pour chaque échantillon (par exemple : résumé numérique, histogramme et boxplot, détection de valeurs aberrantes).
- Question 2 Comparez les moyennes des statures des 4 populations, en comparant les moyennes deux à deux, grâce à des tests statistiques. Expliquez le principe du test de moyenne, ainsi que la décision en fonction de la p-valeur. Puis rédigez une conclusion compréhensible du grand public.
- Question 3 Les statistiques de l'INSEE font apparaître que la taille des hommes a pour moyenne 1.741 et pour écart-type 0.071. La taille des femmes a pour moyenne 1.619 et pour écart-type 0.065. Supposons de plus que la loi de ces variables est la loi normale. Testez si ces moyennes sont cohérentes avec les moyennes observées sur nos deux premiers échantillons.
- Question 4 Mettez en œuvre un test du χ^2 pour vérifier que le premier échantillon suit la loi normale $\mathcal{N}(1.741, 0.071)$. Vous rappellerez le principe du test et vous expliquerez les différentes étapes de votre algorithme.
- Question 5 On cherche à estimer la proportion d'hommes qui mesurent plus de 1,85 m pour réévaluer la longueur des lits dans une chaîne d'hôtels. Au vu de notre premier échantillon, donnez un intervalle de confiance pour cette proportion.
- Question 6 Estimez cette proportion pour les hommes français et pour les hommes canadiens à partir de nos observations. La différence entre ces deux proportions est-elle significative ?
- Question 7 À partir de la loi normale avec les paramètres donnés par l'INSEE, calculez la proportion des hommes en France qui mesurent plus de 1,85 m. Même question pour les femmes.

Vous rédigerez un compte-rendu en précisant bien : nom, prénom, formation, année universitaire, nom de l'UE, date... Soignez la rédaction, ainsi que la mise en page. N'oubliez pas les numéros de page et votre nom sur chaque page. Vous pouvez donner des extraits de code dans votre texte pour les expliquer. Donnez le code en entier en annexe (avec commentaires). Vous déposerez sur spiral (dans le module statistiques) votre travail sous forme d'un fichier pdf de 10 pages maximum avant le 8 mai, minuit. En cas de problème sérieux, vous pouvez écrire à `Anne.Perrut@univ-lyon1.fr`.