



Devoir 3 : Statistiques en sciences politiques et sociales

Année académique 2021/2021

Enseignants : Philippe Bocquier & Lev Lhommeau

**Travail à rendre le :
15 avril 2022**

Université catholique de Louvain
FUCaM Mons
27 mars 2022

Mise en situation

Vous êtes journalistes dans un très grand journal économique. Ce journal réalise un dossier sur l'état des salaires dans le monde en fonction du genre, tout en prenant en compte s'il s'agit d'un pays à revenu faible, intermédiaire, intermédiaire de la tranche supérieure ou supérieur¹. Il a donc demandé à des instituts de sondage d'interroger un échantillon représentatif de la population de chaque pays sur ses revenus. Chacun d'entre vous devra analyser l'une de ces enquêtes. Dans ces sondages, vous trouverez le genre ainsi que les salaires mensuels en dollar états-unis (\$).

Astuce 1: Le fichier de « votre » pays

Chaque étudiant a sa propre base de données qui est nommé *NOMPrenom.RDS*. Si vous vous appelez Jeanne Dubois, le fichier se nommera *duboisjeanne.RDS*. Cette base de données contient deux variables : la variable « Salaire » qui reprend le salaire mensuel en dollar états-unis (\$) et la variable « Genre » reprenant le genre des répondants. Dans cette variable « Genre », 0 correspond à homme et 1 à femme.

Vous trouverez tous les fichiers sur la page suivante : lev.lhommeau.de/article16.html

Afin d'éviter que vous soyez influencé par le nom du pays, le nom des pays a été anonymisé et l'on a attribué votre nom au pays que vous étudiez sous le format suivant : *NOMPrenom* (le fichier se nomme *NOMPrenom.RDS*).

Astuce 2: Charger votre fichier dans R

Vous pouvez télécharger votre fichier sur la page lev.lhommeau.de/article16.html et l'ouvrir ensuite avec RStudio. Pour faire cela, placez le fichier dans votre espace de travail et saisissez la commande suivante dans RStudio :

```
VOTRENOM <- readRDS("VOTRENOMPrenom.rds")
```

Les utilisateurs de Windows peuvent aussi, après avoir téléchargé le fichier, entrer la commande suivante dans RStudio et choisir le fichier dans la fenêtre qui s'ouvre :

```
VOTRENOM <- readRDS(choose.files())
```

Vous pouvez aussi saisir la commande suivante dans RStudio pour charger le fichier directement d'internet :

```
VOTRENOM <- readRDS(url("https://lev.lhommeau.de/IMG/rds/  
VOTRENOMPrenom.rds")) # Pour vous assurer de saisir l'adresse  
correcte, vous pouvez copier l'adresse de votre fichier sur la page internet (  
en fonction du navigateur: clic droit sur le lien et copier l'adresse).
```

Avec ces commandes, vous créez l'objet *VOTRENOM*. Cet objet contient votre base de données à analyser.

Analysez le résultat de l'enquête de « votre » pays avec RStudio en répondant aux six questions suivantes :

1. Vous trouverez davantage d'informations sur cette classification ici sur le site de la Banque mondiale.

Questions

Attention 1: Arrondies

Vos réponses doivent être arrondies à deux décimales après la virgule.

Attention 2: Code R

Fournissez le code R qui vous a permis d'obtenir le résultat pour toutes les questions. Si vous utilisez *RCommander* (Rcmd), copiez le code généré par *RCommander*.

Question 1 (1 point). Quels sont la moyenne et l'écart type des salaires des habitants de « votre » pays ?

Question 2 (1 point). Quel est le salaire moyen des femmes et des hommes dans « votre » pays ?

Question 3 (1 point + 0,5 point bonus). Tracez un histogramme des salaires dans « votre » pays. Bonus : tracez la courbe de distribution normale sur cet histogramme afin de comparer la distribution de la variable à celle-ci.

Question 4 (1 point). À quelle distribution statistique connue les salaires dans « votre » pays correspondent-ils le mieux ? Notez que des déviations aléatoires sont attendues dans l'échantillon de salaires et qu'une distribution doit-être identifiée. Justifiez brièvement votre choix.

Question 5 (1 point). À l'aide de votre réponse à la question 1 et des critères de la Banque mondiale (boîte info 1), déterminez si « votre » pays fait partie des pays à revenu faible, moyen, moyen supérieur ou élevé².

Attention 3: Salaires mensuels

Rappelez-vous que votre base de données contient les salaires *mensuels*.

Boîte info 1: Critères de la Banque mondiale pour classer les pays

La Banque mondiale a défini les bornes suivantes :

\$1045 annuels ou moins (<\$87 mensuels) : pays à revenu faible

Entre \$1046 et \$4095 annuels (>\$87 — \$341 mensuels) : pays à revenu intermédiaire

Entre \$4096 et \$12 695 annuels (>\$341 — \$1057 mensuels) : pays à revenu intermédiaire de la tranche supérieure

\$12 696 annuels ou plus (>\$1057 mensuels) : pays à revenu élevé

Question 6 (1 point / 0,5 point si vous avez mal répondu à la question 1 ou 5, mais que vous calculez correctement la probabilité sur base de vos réponses erronées). Sous l'hypothèse que les salaires suivent une distribution normale, avec la moyenne et l'écart type que vous avez calculés à la question 1 et la classe de pays déterminée à la question 5, quelle est la probabilité qu'une personne y perçoive un salaire supérieur à l'une des valeurs suivantes (voir boîte info 2) ?

2. Au sens strict, cela est calculé sur base du revenu national brut (RNB). Lors de cet exercice, nous faisons cependant comme si les salaires représentaient fidèlement le RNB. C'est donc la moyenne des salaires qui permet de déterminer à quel groupe appartient votre pays. Vous trouverez davantage d'informations sur le RNB ici sur le site de l'OCDE.

Attention 4: Distribution normale

Même si la distribution des salaires n'est pas normale dans « votre » pays, faites comme si elle l'était pour la question 6.

Boîte info 2: Valeurs à partir desquels calculer la probabilité

La valeur à partir de laquelle la probabilité doit être calculée dépend de votre réponse à la question 5. Pour trouver cette valeur, référez-vous à la liste suivante :

Pays à revenu faible : >\$50

Pays à revenu intermédiaire : >\$200

Pays à revenu intermédiaire de la tranche supérieure : >\$900

Pays à revenu élevé : >\$4000

Aide à la lecture : si vous avez, par exemple, estimé à la question 5 que votre pays est un pays à revenu intermédiaire, il faut calculer la probabilité de percevoir un salaire supérieur à \$200.

Astuce 3: Arrondies dans R

Les réponses doivent être arrondies à deux décimales après la virgule. Vous pouvez le faire « à la main », mais R vous propose de le faire à votre place grâce à la fonction « round() ». Si vous voulez connaître l'arrondie à deux décimales du nombre 1,555 tapez la commande suivante dans R :

```
round(1.555, 2)
```

Le premier argument est le nombre que vous souhaitez arrondir (ici : 1,555), le second argument, après la virgule, est le nombre de décimales que vous souhaitez préserver (ici : 2).