

Controle Continu L2 Pharmacie TD Statistique

Bettega Francois

05 septembre, 2020

Introduction

Le 17 mars paraissait *Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open label non-randomized clinical trial* l'article marseillé d'évaluation de l'efficacité de l'Hydroxychloroquine et azithromycin dans le traitement du COVID-19. Je vous propose dans ce contrôle continu de reproduire une partie des traitements statistiques effectuées dans cet article.

Pour des questions de simplicité je vous fournis des jeux de données nettoyer plus facile d'utilisation. Si vous êtes intéressé par les transformations que j'ai effectuées sur le jeux de données de l'article pour arriver au données fournis avec ce contrôle continu. Vous pourrez retrouver le code R à l'adresse suivante https://github.com/fbettega/TD_L2_pharmacie/ dans [ce fichier](#) si vous êtes curieux.

Vous retrouverez un fichier contenant les données pour chacune des questions les fichiers étant nommer à partir du numéro de la questions.

Vous pouvez me contacter à l'adresse suivante pour toutes question : francois.bettega+L2_pharma@gmail.com

Je n'ai pas réussi a ouvrir le jeux de données ne sera pas une excuse acceptable si vous ne m'avez pas posé de question.

Modalité d'évaluations

Ce devoir sera à rendre pour le 30 septembres à l'adresse mail suivante : francois.bettega+L2_pharma@gmail.com.

Vous serez évaluer sur :

- le choix à justifier des tests statistiques (40%) ainsi que sur les hypothèse qui doivent être vérifié avant de pratiquer le test
- votre rédactions et interprétations des résultats de ces tests (50%)
- Sur les résultats numériques obtenue(10%).

Je vous encourage à coopéré et échangé entre les groupes tout en vous prévenant qu'une attention particulière sera porté à la détections du plagiat y compris dans votre code R.

La forme du compte rendu attendu est constitué de deux fichiers votre compte rendu au format ou vous l'aurez compilé (knit) (HTML,PDF,Word...) ainsi que le fichier Rmarkdown (.Rmd) contenant pour chaque questions : - le test que vous avez effectué pourquoi ? - Le résultat de ce test, ainsi que votre interprétations.

En cas de problème avec la création de votre rapport à partir du Rmarkdown. Ne paniquez pour que je puisse vous noté il me faut juste un fichier PDF¹/Word/Html contenant votre rapport ainsi que votre code au format R ou Rmd par conséquent j'accepte aussi les fichier Rmd qui ne compile pas ainsi que rapport + code.

¹Pour rappel vous pouvez facilement crée un document pdf à partir de word ou google doc en utilisant *enregistrer sous*.

Attention les P-value et test absent du code R mais commenté ou les hypothèse non vérifié dans le code R conduiront à une perte de points -0.5pts par occurrence

Résumés document a rendre

- Un fichier Rmd + rapport knit a partir du fichier (format PDF/word/html) (0.5 pts bonus)
- Un fichier R ou Rmd + Votre rapport un format PDF/Word

TD noté sur 10 points (ramener sur 20) 6 points de questions +1 points si vos code R tourne sans erreur + 1 points sur la qualité générale du rapport + 2points sur l'utilisation de graphique pour visualiser les données (0.5 pts par graphique pertinent) **je m'autorise à retirer des points en cas d'abus par exemple si vous produisez un grand nombres de graphiques par questions sans réfléchir à leur pertinences.**

A cela s'ajoute un points bonus sur les interprétations biologique et pharmaceutique. un points bonus sur une représentation graphique non présenté en TD. 0.5 points Bonus sur le fait de commenter votre code R en utilisant : # suivie de texte pour expliquer ce que vous essayer de faire 0.5 Points pour avoir knit votre rapport depuis votre Rmd

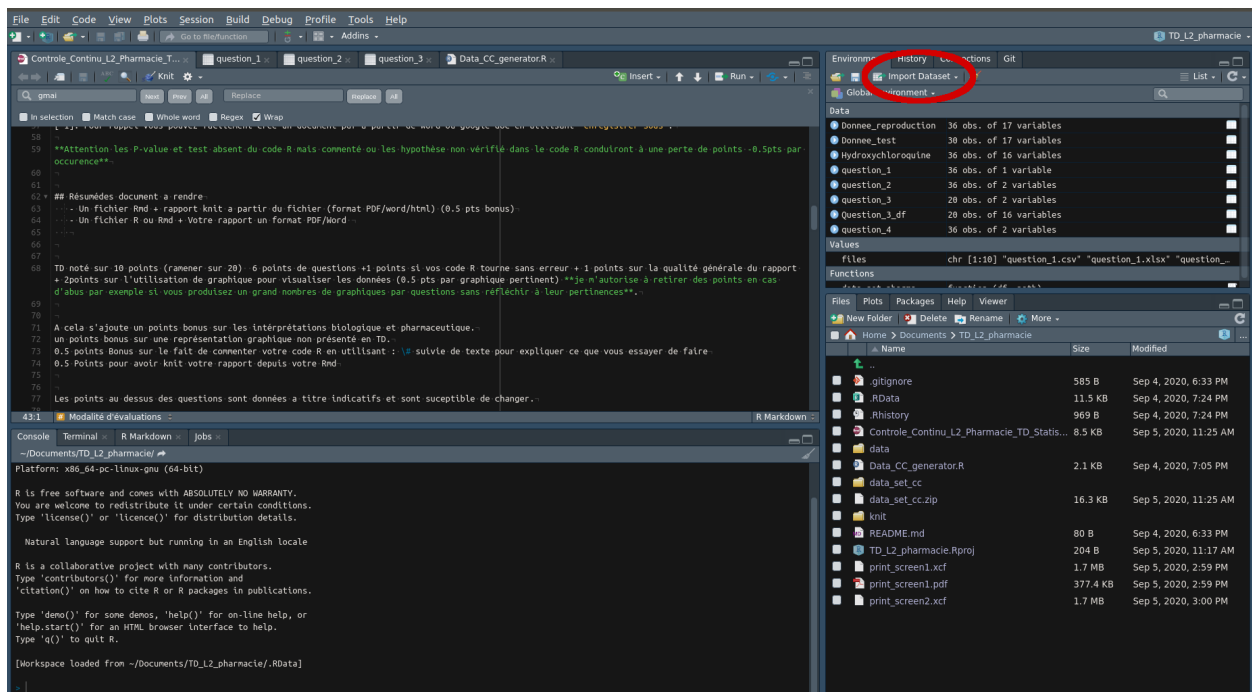
Les points au dessus des questions sont données a titre indicatifs et sont susceptible de changer.

Aide à l'ouverture des fichiers de données

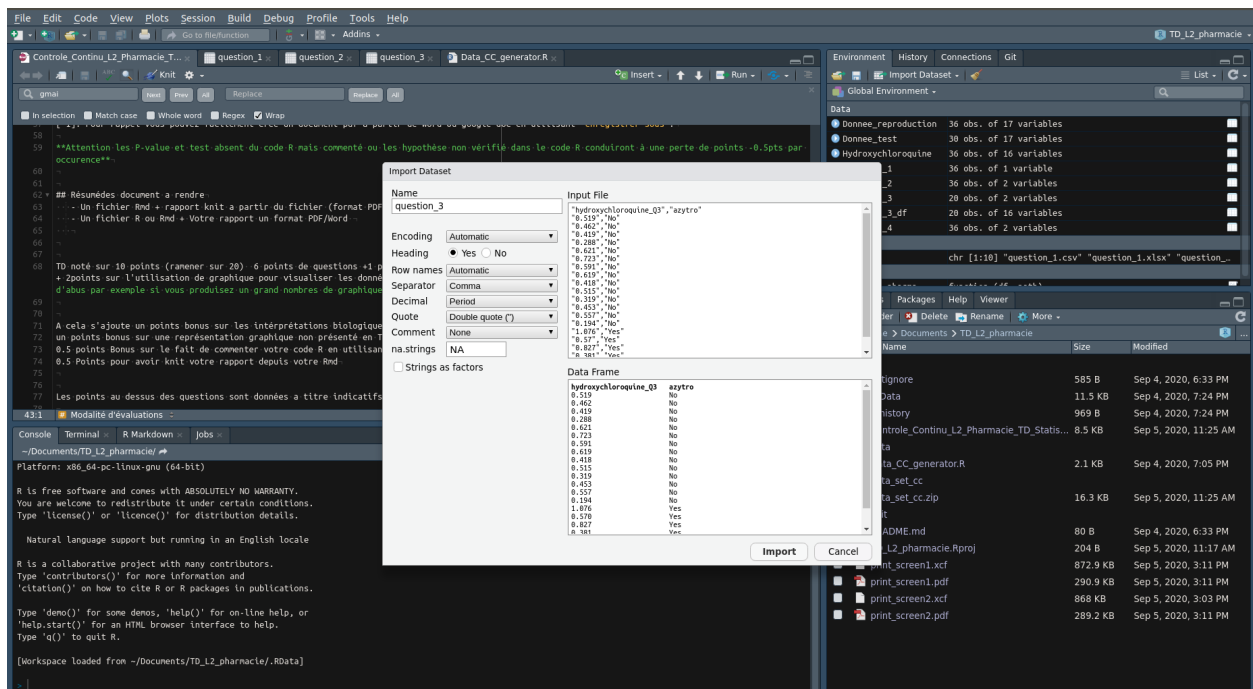
Vous avez un jeux de données par questions il s'appelle **question_x** et je vous les fournis au format xlsx et csv.

Si vous mettez les fichiers dans le même dossiers que votre script Rmd vous pouvez utiliser la ligne de code présenté au début de chaque question pour ouvrir le fichier approprié.

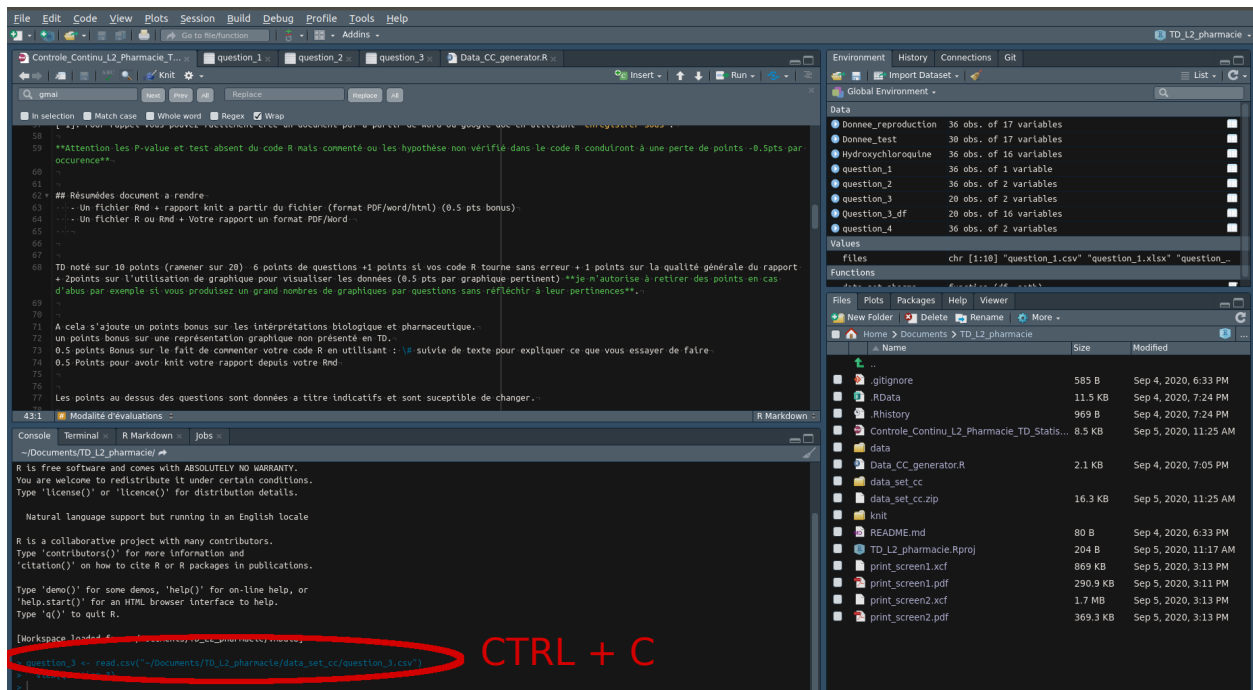
Sinon vous pouvez utiliser le bouton en haut a droite de l'interface de R studio



Selectionner les paramètres approprié



Ensuite faite un copié collé de la seconde ligne de commande dans votre terminal R



Question 1 (1pts)

```
question_1 <- read.csv("data_set_cc/question_1.csv")
```

Ce jeux de données contient l'âge de tous les patients de l'étude mentionner plus haut.

On vous affirme que l'ages des patients n'est pas représentatifs de l'age moyen de la population française qui est de 41.2 ans. Testez cette affirmation.

Question 2 (1pts)

```
question_2 <- read.csv("data_set_cc/question_2.csv")
```

Ce jeux de données contiens : - l'ages des patients - Ainsi que leur statuts traité par Hydroxychloroquine oui ou non

L'age des patients est il différentes dans les 2 groupes traité et non traité par Hydroxychloroquine ?

Pour sectionner les patients par groupes vous pouvez utiliser le code ci-dessous.

```
age_tt <- question_2$age[question_2$hydroxychloroquine == "Yes"]
```

```
age_non_tt <- question_2$age[question_2$hydroxychloroquine == "No"]
```

Question 3 (1pts)

```
question_3 <- read.csv("data_set_cc/question_3.csv")
```

Ce jeux de données contiens : - hydroxychloroquine_Q3 : Concentration d'Hydroxychloroquine dans le serum des patients en µg/ml - azytro: Statuts traité par Azithromycin oui ou non

Pour sectionner les patients par groupes vous pouvez utiliser le code ci-dessous

```
serum_hydroxy_azitro <- question_3$hydroxychloroquine_Q3[question_3$azytro == "Yes"]
```

```
serum_hydroxy_non_azitro <- question_3$hydroxychloroquine_Q3[question_3$azytro == "No"]
```

On vous affirme que la moyenne de la concentration d'Hydroxychloroquine dans le serum des patients traité par Azithromycin est différentes de la concentrations moyennes des patients non traité par Azithromycin. Testez cette affirmation en utilisant **un test non paramétrique**.

Question 4 (4 pts)

Le papier tente d'évaluer si les patients traité par Hydroxychloroquine ont-il des résultat de PCR au jour 6 différents des patients non traité. Je vous propose pour les questions suivante de reproduire cette annalyse de deux manière selon comment on traite les patients qui n'ont pas reçu de PCR au jour 6.

La première (question 4.1) comme dans le papier on considère que les patients traité ont une PCR négative et les patients non traité on une PCR positive

La seconde (question 4.2) les patients n'ayant pas reçu de PCR au jour 6 sont simplement écarté.

Les variables dans les deux jeux de données sont : PCR_1 ou PCR_2 : La PCR au 6eme jour est elle positive au COVID-19

hydroxychloroquine_1 ou hydroxychloroquine_2 : Le patient est il oui ou non traité par Hydroxychloroquine.

Question 4.1 (1.25 pts)

Je vous laisse adapter le code fournis plus haut pour charger les jeux de données et sélectionner les patients par groupes Avec le jeux de données question_4_1 tester l'hypothèse suivante : les patients traité par Hydroxychloroquine ont-il des PCR au jour 6 différents des patients non traité.

Question 4.2 (1.25 pts)

Avec le jeux de données question_4_2 tester l'hypothèse suivante : les patients traité par Hydroxychloroquine ont-il des PCR au jour 6 différents des patients non traité.

Normalement vous pouvez réutiliser exactement le meme code que pour la question 4_1

Commenter et comparer ces résultats (0.5 pts)

Conclusion (1pts bonus)

Discuter ces résultat dans un contexte pharmaceutique ?

Donner votre avis sur ces méthode d'imputations et leurs conséquences potentiel ?

Autres surprenez moi !!