

Consigne examen R

Paolo Crosetto

10/21/2020

Quoi?

Vous devez faire un **rapport statistique** sur un jeu de données. Le rapport va être fait en Rmarkdown (`file -> New file -> Rmarkdown...`). Je m'attends à ce que vous produisez deux fichiers: le code source (.Rmd) et le rendu (.html). Le deuxième est créé à partir du premier en appuyant sur **Knit**.

Un rapport de statistique est un rapport qui **présente** les données. Il est en large partie **déscriptif** mais il peut, si vous voulez, avoir aussi un volet d'exploration des liens causals, s'il y en a.

IL s'agit de montrer le jeu de données au lecteur, en illustrant ses caractéristiques; il s'agit aussi de répondre à des questions sur le jeu de données. Ces questions peuvent être des questions simples, même banales (combien d'avion volent-ils de NY?), ou des questions complexes (quelle est la relation entre l'espérance de vie et le PIB d'un pays au fil du temps?). C'est à vous de poser des questions, et d'y répondre grâce aux données.

Vous avez deux semaines pour préparer le rapport; vous aurez 7 minutes le 6 novembre pour le présenter à la classe.

Dois-je présenter?

Oui, vous allez devoir présenter votre travail face à la classe le 6 novembre.

Vous aurez **7 minutes** par personne (18 étudiants fois 7 minutes fait environ 2 heures, plus des pauses et des transitions, voilà nos trois heures de cours).

Vous présenterez votre Rmd transformé en html, pas besoin de faire des slides en plus.

Sur quelles données?

- vous êtes libres d'utiliser le jeu de données que vous aimez parmi ceux présents sur le site: <https://github.com/rfordatascience/tidytuesday>

Pour quand?

Vous faites cela quand vous voulez, mais il doit être chargé sur github, dans le dossier **Exam**, dans un fichier nommé après votre nom (**VotreNom.Rmd**) **avant** (et non pas *pendant*) le cours du 6 novembre à 9h. Donc la date limite est le 6/11 à 9h.

Où?

- vous travaillez chacun dans votre espace.
- le moment venu, vous copiez votre travail fini dans le dossier **Exam**
- et nommez le fichier **<votre_nom>.Rmd** où vous insérez votre nom
- vous vous assurez que cela compile et qu'un fichier .html est créé après le **knit**

Comment vais-je être évalué?

Vous allez être évalué sur les critères suivants.

- vous avez su charger le .Rmd et le .html au bon endroit utilisant git
- le .Rmd compile (je vais le compiler moi même) sans erreurs. [essayez de le compiler sur l'ordi d'un copain pour être sûr, et évitez les parcours absolus pour les données]
- la présentation des données est complète et compréhensible
- vous vous êtes posés des bonnes questions sur les données
- vous avez pu répondre à vos questions (tableaux, graphiques)
- je vais vous poser des questions, du type 'et si on voulait faire X?'; 'comment auriez-vous changé le plot pour montrer Z?', etc. Le degré de confiance dans l'outil montré lors des réponses sera évalué.

Où est-ce que je trouve de l'aide?

- pour vous inspirer et trouver du code pour les graphiques et des idées d'analyse vous pouvez regarder l'application tidytuesday.rocks, qui vous montre des créations graphiques pour chaque dataset, et très souvent du code aussi, ici: <https://nsgrantham.shinyapps.io/tidyuesdayrocks/>
- pour vous inspirer davantage, il y a du code disponible pour la plupart des plots que vous pouvez concevoir ici: <https://www.r-graph-gallery.com/>
- et encore d'autres idées ici: <http://r-statistics.co/Top50-Ggplot2-Visualizations-MasterList-R-Code.html>

Bon courage!