Présentation du module

Module RShiny

- 3 x 2H TD
- 3 x 2H TP

L'objectif de cette ressource est de savoir automatiser des traitements et produire des résultats directement intégrables dans un rapport.

Barème du projet

- Application Shiny: /5
- Rapport Markdown:/5
- Documentation et Github:/5
- Connexion à une base de données : /5

Objectifs









- Application Shiny: Construire et publier une application Shiny respectant un cahier des charges d'un client et répondant à une problématique
- Rapport Markdown: Rédiger une note de synthèse statistique sur la qualité des données et des explorations statistiques pour répondre à la problématique du projet

- Documentation et Github : Versionner ou sauvegarder les livrables sur github.

 Documenter l'application et les différentes livrable pour qu'un utilisateur puisse installer le nécessaire pour consulter les livrables
- Connexion à une base de données : Stocker les différentes sources de données ou fichiers plats dans une base de données avec un SGBD gratuit (ex: Big Query) pour que les travaux R consultent directement la base de données

Modalité

- Trouver un jeu de données avec plusieurs tables (spotify, netflix, etc)
- Trouver une problématique
- Des groupes de 2 personnes minimum ou 3 personnes maximum du même groupe de TP
- 6 semaines pour réaliser le projet. La deadline est le dernier TP
- Livrer les rendus sur git à l'issue de cette date



Vision

Rappels R Introduction au RMarkdown Suite du projet L'objectif de la séance est de se remettre doucement Dans cette séance, nous développons un rapport Continuer le projet en autonomie et l'intervenant est sur le langage. automatisé. Valider sa team et les jeux de données. là en support TP3 TD2 TP1 TD1 TP2 TD3 Rappels RShiny Début du projet Fin du projet Dans cette séance, nous allons revoir un peu le L'idée est de commencer le projet avec ses données Finir le projet en

et son équipe. Réfléchir à une problématique

autonomie et l'intervenant est là en

support

RShiny. Réfléchir à un jeu de données

Organisation de la séance

Voici les objectifs du TD1 et TD2

- 01 | Exercices R express
- 02 | Shiny tutorial
- 03 | Construction des teams
- 04 | Recherche jeux de données avec plusieurs tables
- 05 | Réfléchir à une problématique



RMarkdown

Ressource: https://juba.github.io/tidyverse/13-rmarkdown.html

L'idée est de réaliser une étude statistique basée sur une ou plusieurs problématiques.

Cahier des charges :

- Présentation du jeu de données
- Exploration de données (stat univariées et bivariées)
- Tableau
- Visualisation (graphiques et interprétation) avec des commentaires
- 5 à 8 pages (sans le code R)

Livrable: sur GitHub (script.Rmd et l'export en HTML)



RShiny

Ressource: https://shiny.rstudio.com/tutorial/

L'idée est de réaliser une application pour un client avec différentes fonctionnalités et analyser permettant de traiter la ou les problématiques..

Cahier des charges:

- Lire, filtrer et visualiser les données (tableaux et graphiques)
- Pouvoir exporter ses graphiques en png
- Avoir un script CSS (existant ou non) sur l'application Shiny
- Déployer l'application en ligne sur https://www.shinyapps.io/

Livrable: sur GitHub (scripts de application et documentation technique et fonctionnelle pour un utilisateur)



GitHub

Ressource: https://docs.github.com/en/get-started/quickstart/hello-world et https://product.hubspot.com/blog/git-and-github-tutorial-for-beginner

L'idée est de travailler en collaboration et livrer l'ensemble du projet dans un repository sur GitHub.

Cahier des charges:

- Un repository pour le projet
- Une branche par membre et une branche master
- Utiliser les commandes standards (commit, pull request, merge)
- Documentation du repos dans un README

Livrable: sur GitHub (partage du repository avec le projet)



Database

Ressource:

L'idée est de travailler dans un premier temps avec des fichiers plats pour concevoir le rapport et l'application puis dans un second temps migrer vers une base de données avec 3 tables minimum..

Cahier des charges :

- Un MCD avec 3 tables minimums au format CSV
- Intégrer ses CSV dans un SGBD avec votre modèle relationnel que le rapport et l'application utilisera au lieu des CSV de départ

Livrable: sur GitHub (dataset brut et dataset utilisé par l'application)





Soutenance

Lors du dernier TP, chaque groupe présentera son travail en 10min + 5min de question (avec tout le groupe TP).

Cahier des charges:

- Présenter le contexte et les données
- Présenter le MCD
- Présenter l'application Shiny avec DEMO
- Réponse rapide de la problématique

Livrable: Aucun

Ressources

- GIT: https://docs.github.com/en/get-started/quickstart/hello-world et https://product.hubspot.com/blog/git-and-github-tutorial-for-beginner
- SHINY: https://shiny.rstudio.com/tutorial/
- R: https://juba.github.io/tidyverse/ / https://juba.github.io/tidyverse/13-rmarkdown.html