

Cas d'étude : TAMAGO

Scénario : S2 + S3 + S4

Auteurs de l'analyse : Caroline Emin, Valérie Fontanieu, Nadine Mandran

Période de l'analyse : Juillet-août 2015

Dispositif d'apprentissage (Etude de cas de Hubble)

Type de dispositif : Jeu multi-joueurs synchrone en ligne

Finalité de l'apprentissage : Le jeu est destiné à des étudiants passant l'agrégation dans le cadre de la validation du C2i2e.

Utilisation du dispositif et fonctionnalités : Le jeu se joue à distance sur une plage de 2 h avec 3 joueurs par équipe. Les équipes sont créées aléatoirement en fonction du nombre d'étudiants inscrits et connectés à la session de jeu.

finalité du jeu => Il faut nourrir un Tamago qui incarne une classe d'élèves au moyen de ressources pédagogiques en fonction du droit de l'exception pédagogique

Le jeu consiste à sélectionner un item (ressource pédagogique à sélectionner en fonction du type, de la longueur de l'extrait, droits associés...) sur l'étagère, la déplacer vers le frigo et choisir un mode de diffusion, enfin de la donner au Tamagocours. Les joueurs peuvent utiliser le chat pour se donner des conseils sur les ressources licites ou non, à donner ou non.

Le jeu donne un feedback immédiat pour savoir si l'élément donné à manger au personnage était licite ou non.

Production de données => historique

Le jeu a été utilisé une première fois en 2014, toutes les traces recueillies ont donné lieu à une analyse quantitative.

Pour les données de 2015, ce sont les mêmes traces des actions de jeu qui ont été collectées qu'en 2014. **Le jeu a été complexifié pour la session de 2015 en créant une difficulté progressive par niveau (suppression de la majorité des ressources libres de droits des niveaux 4 et 5 et restriction des modes de diffusion disponibles) + quand une action est trop souvent répétée à l'identique (même joueur, même type de ressource, même mode de diffusion), le joueur est bloqué et doit varier l'alimentation du Tamago.**

La problématique posée pour l'analyse :

L'objectif de cette expérimentation est de mettre en pratique la théorie des situations didactiques et de caractériser les rôles d'apprenant que prend le joueur.

Pour se faire, nous cherchons à vérifier :

- quelle est la composition des différentes classes de ludants
- l'existence de certains patterns de jeu significatifs
- l'évolution dans le temps du comportement des joueurs

Les données sont traitées sous la forme de données agrégées par joueur comptant toutes les actions ainsi que deux patterns d'actions identifiés.

Objectifs de l'analyse

- répondre à la problématique posée

- Reproduire les analyses de 2014 pour identifier des classes de joueurs et des pattern de jeu. Chercher de nouvelles variables pour affiner ces classes et voir s'il y a une évolution du jeu dans le temps.

Méthode retenue : Analyse en composante principale (ACP), suivi par une classification ascendante hiérarchique puis extraction des traces des parangons et visualisation de leur traces sous forme de chronogrammes.

Description du stockage des données:

Plateformes/outils utilisés: Undertracks => traces des parangons : le besoin était uniquement d'utiliser l'opérateur de visualisation (chronogramme) (A terme transférer les données brutes de TAMAGO sur Undertracks)

Points forts	Points faibles
partage des données protection des données	pb du codage des données en UTF8 fichier plus lisible sur excel manque le lien avec certains algo de traitement des données (exemple ACP et CAH) ou le lien avec R

Description des pré-traitements:

Production des données:

Extraction des données séquentielles : Utilisation de l'outil BackOffice (<http://tamagocours.ens-lyon.fr/botamagocours/web/app.php/>) pour exporter les traces des actions de jeu d'un ensemble de sessions au format csv

Enrichissement des données:

Codage manuel du chat dans le fichier csv à partir des verbatims du chat, codage à plusieurs mains avec une mise en commun (pourcentage de ressemblance supérieur à 80% pour validation)

Nettoyage 1 :

Suppression des individus ayant été déconnectés

Production des données agrégées :

Ré-Import de l'ensemble traces+codage sur le BackOffice pour le calcul des données agrégées (la table à utiliser dans le back office se nomme Tamagocours2015_log)

- Import N°1 : ajout du codage dans les données agrégées
- Import N°2 : ajout de level et iswon dans les données agrégées
- Fusion des deux fichiers pour garder un seul fichier de données agrégées: on obtient le jeu de données agrégées dataset1

Remarque : 2 soucis ce sont posés :

- une ligne 'Pro' apparaît dans le fichier de traces issu du back Office
- les dates ne gardent pas leur bon format en passant de .xls à csv, il faut passer par LibreOffice et modifier le format de cellule en Date heure = JJ/MM/AAAA HH:MM:SS, ensuite enregistrer en CSV

Bilan sur l'obtention des données brutes

25 944 lignes de log sur (26 146 total)

85 groupes pris en compte sur 87 initiaux

242 joueurs sur 252 initiaux

Séquencer le jeu de données

Pour faire une analyse dans le temps, on coupe le jeu de données agrégées entre les niveaux 1 à 3 et 4 et 5 (pallié de difficulté atteint au niveau 3)

1. Ouverture du fichier de traces nettoyé sous Excel, utilisation de l'option "filtrer" par "level" et créer deux fichiers de traces.
2. Ré-import du fichier sur le back office de Tamagocours avec le codage du chat et exportation des données agrégées en CSV.

Plateformes/outils utilisés: BackOffice Tamago, Excel

Points forts	Points faibles
Accès distant au backoffice qui permet un stockage des données en ligne sécurisé	Succession d'imports et exports de données Codage du chat se fait manuellement Il faut faire attention aux formats de la date

Description des analyses :

Liste des 1ères variables (idem aux analyses de 2014) => dataset 1

helpLink	nombre de fois qu'un joueur clique sur un lien de l'aide juridique, suppose que le joueur cherche activement des bonnes réponses.
P_feedgood	pourcentage de feed good par rapport au nombre total de feed, suppose que le joueur réussit le jeu.
Pattern_A-F	nombre de pattern AddToFridge-FeedTamago, suppose que le joueur gave la tamago sans trop de réflexion, sans même consulter le détail des ressources.
Pattern_S-A-F	nombre de pattern ShowItem CUPBOARD-AddToFridge-FeedTamago, suppose que le joueur a regardé le détail de la ressource avant de la donner au Tamago et par conséquent a réfléchi.

Chat_V	nombre de lignes de chat codées V pour validation (inclue les formulations d'hypothèses, de questions et d'observation), suppose un jeu collaboratif où le joueur est plutôt en attente d'une réponse de ses coéquipiers.
Chat_F	nombre de lignes de chat codées F pour formulation (formulation explicite d'une règle), suppose un jeu collaboratif où le joueur est dans la posture de conseiller ses coéquipiers
Chat_OJ	nombre de lignes de chat codées OJ (observation sur le jeu), suppose que les joueurs ont des interactions à propos du jeu mais sans collaboration
Chat_NC	nombre de lignes de chat codées NC (non comptées), suppose que les joueurs ont des interactions complètement HS.
totalAction	nombre d'actions totales effectuées par un joueur
feedTamagogood	nombre de fois qu'un feed effectué par un joueur est validé
feedTamagobad	nombre de fois qu'un feed effectué par un joueur est invalidé
showItemCupboard	nombre de fois qu'un joueur consulte le détail d'une ressource avant de la sélectionner ou non

Reproduction des analyses de 2014 :

Mode opératoire méthodologique

Analyse en composante principale (ACP), suivi par une classification ascendante hiérarchique.

Mode opératoire technique, logiciels utilisés

Excel + SPAD + Undertracks

A partir du fichier d'export des traces généré par le back office de Tamagocours (données agrégées sur la base totale) :

Ouverture du fichier dans Excel :

Toilettage du fichier :

- suppression de lignes vides (filtres sur 0)
- renommages :

feedTamago -> feedtotal

Liste S-A-F -> pattern_SAF

Liste A-F -> pattern_AF

NC -> chat_Nc

Oj -> chat_Oj

F -> chat_F

H -> chat_H

O -> chat_O

Q -> chat_Q

- création de variables :

chat_V = H + O + Q

p_Feedgood = Feedgood/feedtotal*100 [=SI(G2=0;0;B2/G2*100)]

Sous SPAD

- création d'un projet Spad
- ajout méthode import fichier texte délimité (séparateur = ;)
- ajout méthode ACP (répertoire méthodes factorielles) : sélection des variables actives et illustratives

Variables à conserver (pour besoin de comparaison avec précédentes analyses) :

Actives : Feedgood/feedtotal*100, pattern SAF, pattern AF, helpLink, showItemActionCUPBOARD, chat_Nc, chat_Oj, chat_F, chat_V

Supplémentaires : feedBad, feedGood, total action

- ajout méthode classification CAH: pour paramétrer le nombre de facteurs retenus, se reporter aux résultats de l'ACP pour regarder le nombre d'axes contributif)
- ajout méthode coupure de l'arbre et test suivant plusieurs nombres de facteurs retenus (regarder la pertinence des classes en fonction du vtest et de ce que l'on désire observer)
- exportation des résultats de la classification
- interprétation des résultats et description des classes obtenues
- comparaison avec les résultats de la précédente session

Sous Undertracks : on cherche à visualiser les parangons:

- à partir de SPAD, extraire les individus parangons, puis manuellement sélectionner dans le fichier de traces les log correspondant à chaque individu parangons, créer un nouveau fichier csv UTF8 des traces des parangons
- importer le fichier traces des parangons sur Undertracks
- avec Orange, charger le fichier de traces, ajout d'une méthode "sort" par classe, ajout d'une méthode "timeline" : les résultats s'ouvrant dans une fenêtre sur UT

Scripts produits pour l'analyse des données (*exemple script R*)

Script R enchaînement ACP + CAH (problème d'export des résultats de CAH de description des classes)

```
datatamago <- read.csv2("/Users/Narin/Documents/TAMAGOCOURS/R/Donnés-agrégées-Nettoyage.csv")
```

```
library(questionr)
```

```
library(R2HTML)
```

```
library(FactoMineR)
```

```
library(GDAtools)
```

```
acp <- subset (datatamago, select=c(feedGood, feedBad, p_Feedgood, helpLink, showItemCUPBOARD, chat_F, chat_V, chat_Nc, chat_Oj,totalAction, pattern_SAF, pattern_AF))
```

```
acpact <-subset(datatamago, select=c(p_Feedgood, helpLink, showItemCUPBOARD, chat_F, chat_V, chat_Nc, chat_Oj, pattern_SAF, pattern_AF))
```

```
?PCA
acp1 <- PCA(acp, quali.sup=1,2,10)
summary(acp1)
barplot(acp1$eig[1:10,2], main="Histogramme ACP1", names.arg=1:10)
?barplot
write.infile(acp1, file="/Users/Narin/Documents/TAMAGOCOURS/R/ACP1.xls", sep="\t")
resultacp <- acp1$eig[1:10,]
```

```
dist <- dist(acp, method="euclidean",diag=FALSE, upper=FALSE)
hier <-hclust(dist)
plot(hier,hang=1)
hier2 <-hclust(dist, method="ward") # là avec la méthode de Ward
plot(hier,hang=1)
?hclust
classe5 <-cutree(hier2,5)
```

```
#Avec la méthode des Kmeans.
```

```
k=5
cl <- kmeans(acpact,k,20)
p<- princomp(acpact)
u=p$loadings
x<-(t(u)%*%t(acpact))[1:2,]
x<-t(x)
plot(x,col=cl$cluster,pch=3,lwd=3)
c<-(t(u)%*%t(cl$center))[1:2,]
c<-t(c)
points(c,col=1:k,pch=7,lwd=3)
for(i in 1:k){
  print(paste("Cluster",i))
  for(j in(1:length(acpact[, 1]))[cl$cluster==i])
    {segments(x[j, 1], x[j,2], c[i, 1],col=i)}}
text(x[, 1],x[,2],attr(x,"dimnames")[[1]], col=cl$cluster, cex=0,7)
write.infile(cl, file="/Users/Narin/Documents/TAMAGOCOURS/R/ClusterR.xls", sep="\t")
```

```
## Essai de classification ascendante hierarchique
```

```
# La CAH va operer a partir des coordonnees des individus sur les axes
```

```
?HCPC
```

```
res.cah <- HCPC(acp1, nb.clust=5)
```

```
res.cah
```

```
# Visualisons le resultat sous forme de jeu de donnees
```

```
cah <- as.data.frame(res.cah$data.clust)
```

```
View(cah)
```

```
str(cah)
```

```
table(cah$clust)
```

```

cat(res.cah)
write.infile(cah, file="/Users/Narin/Documents/TAMAGOCOURS/R/CAH.xls", sep="\t")
write.infile(res.cah, file="/Users/Narin/Documents/TAMAGOCOURS/R/CAHTest.xls",
sep="\t")

```

Résultats obtenus:

Classes obtenues avec 3 facteurs retenus
Classe 1=26 GAVEURS BAVARDS 11%
Classe 2=72 GAVEURS 30%
Classe 3=84 EFFICIENTS 35%
Classe 4 =50 PRUDENTS 21%
Classe 5=10 BAVARDS EXPERTS 4%

Points forts des analyses	Points faibles des analyses
les classes la robustesse de la méthode entre 2014 et 2015	pas de ligne de code R pour obtenir la description des classes de la CAH impossible d'exporter les résultats d'UnderTracks Nettoyage incomplet : il y a encore des joueurs qui ont eu des problème de connexion et qui ont peu joué (ils pèsent sur les axes)

Reproduction des analyses de 2014 sur les Niveaux 1 à 3 :

Mode opératoire méthodologique

Analyse en composante principale (ACP), suivi par une classification ascendante hiérarchique.

Mode opératoire technique, logiciels utilisés
Excel + SPAD

1/ A partir du fichier d'export des données agrégées séquencées sur les niveaux 1-3 :
Ouverture du fichier dans Excel :

Toilettage du fichier :

- suppression de lignes vides (filtres sur 0)
- renommages :

feedTamago -> feedtotal

Liste S-A-F -> pattern_SAF

Liste A-F -> pattern_AF

NC -> chat_Nc

Oj -> chat_Oj

F -> chat_F

H -> chat_H

O -> chat_O

Q -> chat_Q

- création de variables :

chat_V = H + O + Q

p_Feedgood = Feedgood/feedtotal*100 [=SI(G2=0;0;B2/G2*100)]

Sous SPAD

- création d'un projet Spad
- ajout méthode import fichier texte délimité (séparateur = ;)
- ajout méthode ACP (répertoire méthodes factorielles) : sélection des variables actives et illustratives

Variables à conserver (pour besoin de comparaison avec précédentes analyses) :

Actives : Feedgood/feedtotal*100, pattern SAF, pattern AF, helpLink,
showItemActionCUPBOARD, chat_Nc, chat_Oj, chat_F, chat_V

Supplémentaires : feedBad, feedGood, total action

- ajout méthode classification CAH: pour paramétrer le nombre de facteurs retenus, se reporter aux résultats de l'ACP pour regarder le nombre d'axes contributif)
- ajout méthode coupure de l'arbre et test suivant plusieurs nombres de facteurs retenus (regarder la pertinence des classes en fonction du vtest et de ce que l'on désire observer)
- exportation des résultats de la classification
- interprétation des résultats et description des classes obtenues
- comparaison avec les résultats de la précédente session

Scripts produits pour l'analyse des données (*exemple script R*)

```
Niv1_3 <- read.csv2("/Users/Narin/Documents/TAMAGOCOURS/R/Donnees-agreées
```

```
Niv1-3.csv")
```

```
library(questionr)
```



```

library(R2HTML)
library(FactoMineR)
library(GDAtools)
acp <- subset(Niv1_3, select=c(feedGood, feedBad, p_feedgood, helpLink,
showItemCUPBOARD, chat_F, chat_V, chat_NC, chat_OJ,totalAction,
Pattern_S.A.F, Patter_.A.F))
acp1 <- PCA(acp, quali.sup=1,2,10)
barplot(acp1$eig[1:10,2], main="Histogramme ACP1", names.arg=1:10)
write.infile(acp1, file="/Users/Narin/Documents/TAMAGOCOURS/R/ACPNiv1_3.xls",
sep="\t")
resultacp <- acp1$eig[1:10,]
hier <- hclust(resultacp, method="ward") # là avec la méthode de Ward
plot(hier,hang=1)

```

Résultats obtenus:

Points forts des analyses	Points faibles des analyses

Reproduction des analyses de 2014 sur les Niveaux 4-5 :

Mode opératoire méthodologique

Analyse en composante principale (ACP), suivi par une classification ascendante hiérarchique.

Mode opératoire technique, logiciels utilisés

Excel + SPAD

1/ A partir du fichier d'export des données agrégées Niveaux 4 et 5 :

Ouverture du fichier dans Excel :

Toilettage du fichier :

- suppression de lignes vides (filtres sur 0)
- renommages :

feedTamago -> feedtotal

Liste S-A-F -> pattern_SAF

Liste A-F -> pattern_AF

NC -> chat_Nc

Oj -> chat_Oj

F -> chat_F

H -> chat_H

O -> chat_O

Q -> chat_Q

- création de variables :

chat_V = H + O + Q

p_Feedgood = Feedgood/feedtotal*100 [=SI(G2=0;0;B2/G2*100)]

Sous SPAD

- création d'un projet Spad
- ajout méthode import fichier texte délimité (séparateur = ;)
- ajout méthode ACP (répertoire méthodes factorielles) : sélection des variables actives et illustratives

Variables à conserver (pour besoin de comparaison avec précédentes analyses) :

Actives : Feedgood/feedtotal*100, pattern SAF, pattern AF, helpLink, showItemActionCUPBOARD, chat_Nc, chat_Oj, chat_F, chat_V

Supplémentaires : feedBad, feedGood, total action

- ajout méthode classification CAH: pour paramétrer le nombre de facteurs retenus, se reporter aux résultats de l'ACP pour regarder le nombre d'axes contributif)
- ajout méthode coupure de l'arbre et test suivant plusieurs nombres de facteurs retenus (regarder la pertinence des classes en fonction du vtest et de ce que l'on désire observer)
- exportation des résultats de la classification
- interprétation des résultats et description des classes obtenues
- comparaison avec les résultats de la précédente session

Vérification de la pertinence des données séquencées :

Mode opératoire méthodologique

Création d'un tableau croisé sur les classes issues de la CAH

Mode opératoire technique, logiciels utilisés

Sous SPAD, à partir des deux analyses séquencées:

- ajout d'une méthode : archivage des classes
- ajout d'une méthode : tableau croisé des deux archives

Résultats obtenus:

Résultat : la tableau croisé montre une dispersion des individus dans les classes après le niveau 3 suffisamment important pour justifier d'une coupure séquentielle à ce moment.

Points forts des analyses	Points faibles des analyses
Il est pertinent de faire une analyse dans le temps des comportements des joueurs On obtient des groupes différents	Les analyses ne peuvent se passer d'une lecture qualitative des traces pour vérifier les hypothèses, au moins sur les individus parangons. L'extraction des parangons se fait manuellement. Après lecture des traces : les variables

	sélectionnées n'indiquent pas suffisamment d'informations
--	---

Description des itérations:

Nouvelles données

Nettoyage des individus.

Grâce aux résultats précédents, on se rend compte qu'il y a besoin de nettoyer mieux le jeu de données car il y a d'autres individu non joueurs (dataset 2) :

Suppression des individus non-joueurs :

-**Ind9_30** (ne peut pas se connecter)

-**Ind11_32** (ne peut pas jouer car rien ne s'affiche sur son écran à part le chat)

-**Ind22_66**(rejoint au niveau 4)

-**Ind42_120** (rejoint en fin de partie)

-**Ind59_171**(rejoint au niveau 4)

-**Ind38_41** (joue seulement un niveau avec 2 min de jeu, et joue ensuite dans le groupe 43).

- **Groupe 18** (constitué d'un seul joueur a abandonné la partie au niveau 2)

-**Groupe 32** a arrêté de jouer faute de joueurs (bloqué pour passer niveau 2)

Pour repérer ces individus non-joueurs :

prendre le premier centile des feed totaux et regarder ceux qui sont en dessous (il est à 6), ensuite, manuellement, aller lire les traces pour vérifier qu'il n'a pas joué autrement.

recherche ceux qui n'ont pas dépassé le niveau 2 et voir pourquoi : deux groupes ont commencé à 2 joueurs et se sont arrêtés rapidement car un des joueurs était déconnecté et l'autre était bloqué car limité dans le nombre d'actions qu'il peut faire sans intervention d'un camarade.

Calcul des nouvelles variables.

Dataset 3

1. SHOW ITEM FRIDGE OTHERS.

On cherche ici à savoir combien de fois un jour clique sur un item déposé dans le frigo par un des membres de son équipe pour vérification.

Depuis le back office de Tamagocours ([http://tamagocours.ens-](http://tamagocours.ens-lyon.fr/botamagocours/web/app.php/analyse/pattern)

[lyon.fr/botamagocours/web/app.php/analyse/pattern](http://tamagocours.ens-lyon.fr/botamagocours/web/app.php/analyse/pattern)) faire une recherche de pattern sur la base Tamgocours2015_logs, en actsimp (ne pas oublier de sauvegarder les paramètres), action 1 : addToFridge, action 2: ShowItem, cocher suivi de l'item oui. Garder le grp-us-id et le nombre d'occurrences du pattern à partir des résultats du back office et les reporter (manuellement) dans le jeu de données agrégées complet sous Excel. on obtient le pattern AS (Add-Show)

Depuis Excel et sur le fichier des données agrégées, soustraire la colonne du pattern Add-Show à la colonne "ShowItemFridge", on obtient la variable "ShowItemFridgeOTHERS".

Refaire la même manipulation pour les données séquencées entre les niveaux 1 à 3 et 4 à 5 (à partir des fichiers de traces correspondants dans le back office : Tamago2015Niv1a3_logs et Tamago2015Niv4à5_logs)

2. Pattern SHOW BEFORE FEED

On cherche ici à étendre la variable SAF (show item cupboard, add to fridge, feed Tamago) aux individus qui ont réalisé un show item avant le feed, qui ont donc réfléchi et tenter d'élaborer une stratégie, mais pas nécessairement à l'étape cupboard. On a besoin de savoir combien ont fait un show item fridge au sein de ce pattern.

Depuis le back office de Tamagocours, recherche de pattern sur la base

Tamgocours2015_logs, en actsimp (ne pas oublier de sauvegarder les paramètres), action 1 : addtoFridge, action 2: showItemFridge, action 3 : feedTamago, cocher suivi de l'item oui. Garder le grpus-id et le nombre d'occurrences du pattern à partir des résultats du back office et les reporter (manuellement) dans le jeu de données agrégées complet sous Excel. On obtient le pattern ASF (Add-Show-Feed)

Puis, depuis le back office de Tamagocours, recherche de pattern sur la base

Tamgocours2015_logs, en actsimp (ne pas oublier de sauvegarder les paramètres), action 1 : showitemcupboard, action 2: showItemFridge, action 3 : feedTamago, cocher suivi de l'item oui. Garder le grpus-id et le nombre d'occurrences du pattern à partir des résultats du back office et les reporter (manuellement) dans le jeu de données agrégées complet sous Excel. On obtient le pattern SASF (Show-Add-Show-Feed)

Sur le fichier de données agrégées sous Excel, additionner la colonne pattern ASF avec la colonne pattern SAF et soustraire la colonne pattern SASF (éviter de compter deux fois le même élément). On obtient le pattern Show Before Feed.

3. Variables sur le TITRES des ressources

On cherche à voir si les ressources portant le même titre sont privilégiées par les joueurs pour mettre en place une stratégie de reproduction.

A partir d'une requête SQL qui comptabilise le nombre d'utilisation d'un même titre par joueur au cours de la partie (voir fichier table_3_R2COUNT

<https://drive.google.com/open?id=0B6J75kmqnSWCU3NCX1paUVgwOGs>), ouvrir sous

Excel. Faire générer un tableau croisé dynamique à Excel avec en ligne l'utilisateur et en

colonne les différents titres. Calcul automatique (Calculs -> résumer les valeurs par...)

de la moyenne d'utilisation d'un même titre, le maximum d'utilisation d'un même titre, le nombre

de titres différents utilisés (variables Title_Moy, Title_Max, Title_Nb). Exécuter pour chaque

ligne la somme d'utilisation de chaque ressource si la valeur est au moins 3 (pour avoir une

répétition liée à une volonté et pas simplement au nombre limité de ressources dans le jeu) :

=SOMME.SI(Intervalle; ">2"). On obtient la variable Title_Somme>2

Faire la même chose sur les fichiers séquencés niveaux 1-3 et niveaux 4-5

<https://drive.google.com/open?id=0B8SoTMZieg2oN1dTWkZ2SGhkWXM>

<https://drive.google.com/open?id=0B6J75kmqnSWCZXVRMkFUcUY0Rms>

4. Variables sur les TYPE de ressources.

On cherche à voir si les joueurs privilégient l'utilisation d'un ou plusieurs type de ressources.

A partir d'une requête SQL qui comptabilise le nombre d'utilisation d'un même type par joueur au cours de la partie (voir fichier table_3_R3COUNT

<https://drive.google.com/open?id=0B6J75kmqnSWCanF2c2haNDRVOHM>), ouvrir sous

Excel. Faire générer un tableau croisé dynamique à Excel avec en ligne l'utilisateur et en

colonne les différents types (journaux, livres etc.). Calcul automatique (Calculs -> résumer les valeurs par...) de la moyenne d'utilisation d'un même type, le maximum d'utilisation d'un même type, le nombre de types différents utilisés (variables Type Moy, Type Max, Type Nb). Exécuter pour chaque ligne la somme d'utilisation de chaque ressource si la valeur est au moins 3 : =SOMME.SI(Intervalle; ">2"). On obtient la variable Type Somme>2
 Faire la même chose sur les fichiers séquencés niveaux 1-3 et niveaux 4-5
<https://drive.google.com/open?id=0B8SoTMZieg2oUnU1bEtNaGF6TUU>
<https://drive.google.com/open?id=0B6J75kmqnSWCMWVoMXI3X1dORUU>

5. Variable sur les MODES D'UTILISATION d'une ressource

On cherche à voir si les joueurs privilégient l'utilisation d'un même mode de diffusion pour leurs ressources.

A partir d'une requête SQL qui comptabilise le nombre d'utilisation d'un même mode de diffusion par joueur au cours de la partie (voir fichier table_3_R4COUNT <https://drive.google.com/open?id=0B6J75kmqnSWCanF2c2haNDRVOHM>), ouvrir sous Excel. Faire générer un tableau croisé dynamique à Excel avec en ligne l'utilisateur et en colonne les différents modes d'utilisation (photocopies, sujet d'examen, projection etc.). Calcul automatique (Calculs -> résumer les valeurs par...) de la moyenne d'utilisation d'un même mode, le maximum d'utilisation d'un même mode, le nombre de modes différents utilisés (variables MoU Moy, MoU Max, MoU Nb). Exécuter pour chaque ligne la somme d'utilisation de chaque ressource si la valeur est au moins 3 : =SOMME.SI(Intervalle; ">2"). On obtient la variable MoU Somme>2

Faire la même chose sur les fichiers séquencés niveaux 1-3 et niveaux 4-5
<https://drive.google.com/open?id=0B8SoTMZieg2oT2FycFFXN3B3N1E>
<https://drive.google.com/open?id=0B6J75kmqnSWCVHZUZG9WOTB2Wnc>

6. Variable relationnelle Modes d'utilisation/type de ressource associé

On cherche à voir si les joueurs privilégient la combinaison d'un mode d'utilisation associé à un type de ressource (ex : journal en photocopies).

A partir d'une requête SQL qui comptabilise le nombre d'utilisation de la même combinaison Mode/type par joueur au cours de la partie (voir fichier table_3_R5COUNT <https://drive.google.com/open?id=0B6J75kmqnSWCXzJic0t1c09VNW8>), ouvrir sous Excel. Faire générer un tableau croisé dynamique à Excel avec en ligne l'utilisateur et en colonne les combinaisons mode/type. Calcul automatique (Calculs -> résumer les valeurs par...) de la moyenne d'utilisation d'une même combinaison, le maximum d'utilisation d'une même combinaison, le nombre de combinaisons différentes utilisées (variables MoU Moy, MoU Max, MoU Nb). Exécuter pour chaque ligne la somme d'utilisation de chaque combinaison si la valeur est au moins 3 : =SOMME.SI(Intervalle; ">2"). On obtient la variable MoU Somme>2

Faire la même chose sur les fichiers séquencés niveaux 1-3 et niveaux 4-5
<https://drive.google.com/open?id=0B8SoTMZieg2oMDJjSEFDLTBSQ1k>
<https://drive.google.com/open?id=0B6J75kmqnSWCZ1RPNI9WeWZIZ2M>

Description des nouvelles variables sélectionnées pour l'analyse :

title_Max	nombre maximum de fois qu'un même titre a été utilisé pour une action addToFridge, suppose une stratégie de répétition
-----------	--

P_Title	pourcentage de répétition de l'utilisation d'un même titre pour l'action addToFridge (à la condition qu'il soit utilisé au moins trois fois pour que la répétition soit volontaire et pas simplement dérivée du jeu) par rapport au nombre total d'addToFridge, suppose une stratégie de répétition.
TypeMoU_Max	nombre maximum de fois qu'une même combinaison type de ressource et mode d'utilisation (ex journal en sujet d'examen) a été utilisée en addToFridge, suppose une stratégie de répétition.
P_TypeMoU	pourcentage de répétition de l'utilisation d'une même combinaison pour l'action addToFridge (à la condition que la combinaison soit utilisée plus de 3 fois) par rapport au nombre total d'addToFridge, suppose une stratégie de répétition.
duree_sec	temps en seconde de la partie.
showItemFridge OTHERS	nombre de fois qu'un joueur regarde le détail d'une ressource déposée par un membre de son équipe dans le frigo, suppose un jeu collaboratif.
Pattern S_Before-Feed	nombre de pattern showitem before feed tamago, suppose que le joueur a regardé le détail de la ressource avant de la donner au Tamago et par conséquent a réfléchi.
tuto	nombre de fois qu'un joueur clique sur le tutorial
feedTamago	nombre de feed total effectués par un joueur

A partir de dataset 3, on reprend l'enchaînement ACP -> CAH -> Parangons -> Chronogramme avec Excel, Spad, Under Tracks et Orange. Ce qui nous donne des nouvelles classes et des nouveaux résultats.

Jeu de données global	Niveaux 1 à 3	Niveaux 4 à 5
Classes obtenues avec 6 facteurs retenus	Classes obtenues avec 8 facteurs retenus	Classes obtenues avec 7 facteurs retenus
Classe 1 =52 21% GAVEURS	Classe 1 =57 23% PERSEVERANTS	Classe 1 = 32 13% GAVEURS
Classe 2=31 13% PERSEVERANTS	Classe 2 = 14 6% PRUDENTS	Classe 2 =23 10% PERSEVERANTS
Classe 3 =16	Classe 3 =15	Classe 3 =4

7% BAVARDS	6% INACTIFS	2% COLLABORATEURS
Classe 4=108 45% INACTIFS	Classe 4 = 109 45% EFFICIENTS	Classe 4 = 15 6% DESABUSES
Classe 5=30 12% PRUDENTS	Classe 5 =8 3% COLLABORATEURS	Classe 5 = 12 5% BAVARDS
Classe 6=6 2%	Classe 6= 39 16% GAVEURS	Classe 6= 60 25%
		Classe 7 = 96 40%

Visualisation des évolutions entre les deux séquences de jeu:

Reproduction de la méthode de tableau croisé sous SPAD.

Création d'un schéma synthétique sous R :

```
df<- data.frame(segment =c("EFFICIENTS", "PERSEVERANTS", "GAVEURS",
"PRUDENTS", "INACTIFS", "COLLABORATEURS"), segpct = c(45, 23, 16, 6, 6, 4),
  INACTIFS=c(42.6, 38.6, 30.8, 78.6, 26.7, 11.1),
  STRATEGES=c(21.3, 35.1, 23.1, 14.3, 40, 0),
  GAVEURS=c(12, 19.3, 20.5, 0, 0, 0),
  PERSEVERANTS=c(11.1, 5.3, 15.4, 7.1, 6.7, 0),
  DESABUSES=c(10.2, 0, 2.6, 0, 13.3, 11.1),
  BAVARDS=c(2.8, 1.8, 7.7, 0, 13.3, 33.3),
  COLLABORATEURS=c(0, 0, 0, 0, 0, 44.4))
```

```
df$xmax <- cumsum(df$segpct)
df$xmin <- df$xmax - df$segpct
df$segpct <- NULL
head(df)
library(ggplot2)
library(plyr)
library(reshape2)
dfm <- melt(df, id = c("segment", "xmin", "xmax"))
head(dfm)
dfm1 <- ddply(dfm, .(segment), transform, ymax = cumsum(value))
dfm1 <- ddply(dfm1, .(segment), transform,
  ymin = ymax - value)
?ddply

dfm1$text <- with(dfm1, xmin + (xmax - xmin)/2)
```

```

dfm1$ytext <- with(dfm1, ymin + (ymax - ymin)/2)
p <- ggplot(dfm1, aes(ymin = ymin, ymax = ymax, xmin = xmin, xmax = xmax, fill = variable))
?ggplot
p1 <- p + geom_rect(colour = l("grey"))
p2 <- p1 + geom_text(aes(x = xtext, y = ytext, label = ifelse(segment == "EFFICIENTS",
paste(variable, " - ", value, "%", sep = ""),
paste(value,"%", sep = ""))), size = 3.5)

p3 <- p2 + geom_text(aes(x = xtext, y = ifelse(segment == "INACTIFS", -2,-5), label =
paste(segment)), size = 4)

last_plot() + scale_fill_brewer(palette = "Set3")

```

Points forts des itérations	Points faibles des itérations
<p>On a pu rajouter des variables et les calculer</p> <p>Les traces permettaient de calculer ces nouvelles variables</p> <p>Les classes obtenues sont plus nuancées</p> <p>On a une évolution dans le temps qui donne des éléments importants sur l'attitude face aux jeux et aux apprentissages</p>	<p>La variable retenue pour la collaboration est insuffisante : il faut un nouveau codage du chat pour plus de précision</p> <p>On se rend compte que l'on aurait pu tracer d'autres éléments dans la collecte des données initiale</p>