

=====
Date de rédaction : 09/11/2015

Nom du rédacteur du document : Tony Doat

Spécialités : ~~Didactique, Sciences de l'éducation,~~ Informatique, ~~Analyse de données~~ (*rayez les mentions inutiles*)
=====

Cas d'étude Hubble : MOOC AZ

Scénario hubble : ?

Personnes impliquées pour la collecte et l'analyse : Mathieu Cisel

Période de la collecte : ?

Periode de l'analyse : ?

Dispositif d'apprentissage (Etude de cas de Hubble)

Type de dispositif :

Finalité de l'apprentissage :

Utilisation du dispositif et fonctionnalités :

Contexte de production de données

Au besoin indiquer les différents moments de la production (savoir si des données ont été produites sur plusieurs années)

La problématique posée pour l'analyse :

Objectifs de l'analyse :

(travail en cours, description ici des premiers retours sur la plateforme UT-Orange)

- créer une grille pour comparer les outils d'un point de vue technique
- Identifier les points forts et les points faibles techniques des outils de collecte
- Identifier les points forts et les points faibles techniques des outils d'analyses
- proposer des pistes d'améliorations

Le tout testé sous Windows.

Description du stockage des données:

Plateformes/outils utilisés: plateforme Undertrack-Orange

Points forts de ces plateformes	Points faibles
<ul style="list-style-type: none">- Dépôt des données en ligne via une interface web (pas d'installation)- Partage possible des données- Restriction possible sur des données- Visualisation en lignes des premières lignes	<ul style="list-style-type: none">- Format d'entrée peu permissif (csv parfaitement formaté et sans la protection des chaînes de caractères)

Production des données avant le traitement :

Décrire le processus de production des données brutes :

Liste des variables initiales : Nom et Description

Plateformes/outils utilisés:

Points forts	Points faibles

Description des pré-traitements:

Objectifs des pré-traitements

Décrire le processus de pré-traitement:

Plateformes/outils utilisés:

Points forts	Points faibles

Description des analyses :

Liste des variables : Nom et Description

Liste des méthodes mise en œuvre :

Mode opératoire technique, logiciels utilisés

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

Résultats obtenus:

Points forts des analyses	Points faibles des analyses
<ul style="list-style-type: none">- Partage du processus d'analyse (succession d'opérations)- Possibilité de définir de nouveaux opérateurs en lignes de code (R, Python)- Installation sur machine locale sans grande difficultés- Présence (non testé) de nombreux opérateurs d'analyses statistiques	<ul style="list-style-type: none">- Quelques bugs dans l'utilisation de l'application Orange qui peuvent aboutir à la non-sauvegarde du travail en cours- Nécessité de télécharger les données, long si données volumineuses- Le processus d'analyse sur des données massive prend du temps

A FAIRE : point non vérifié : tester cas d'une analyse basée sur l'utilisation de plusieurs données d'entrées (local et/ou serveur)

Description des données produites au cours du traitement

Objectif de la création de ces nouvelles données :

Mode de calcul de ces variables :

Description des nouvelles variables : Nom et description

Nom	description

Description des Itérations

Pourquoi le processus d'analyse a été reproduit ?

Points forts des itérations	Points faibles des itérations