

Process Book

Glyphs and Mosaic: Interactive Visualizations of Likert Scale Questionnaires
Raphaël Diana & Alix Gonnot

L'objectif de ce document est de présenter l'avancement du travail effectué dans le cadre du projet de Data Visualization réalisé pendant l'année scolaire 2016-2017.

A la fin de ce document vous pourrez trouver nos *Fives Design Sketches* sur lesquelles nous avons exposé nos idées préliminaires pour ce projet.

12/12/2016 :

<https://public.tableau.com/profile/alix.gonnot#!/vizhome/PremirevisuProjetDV/Feuille1>

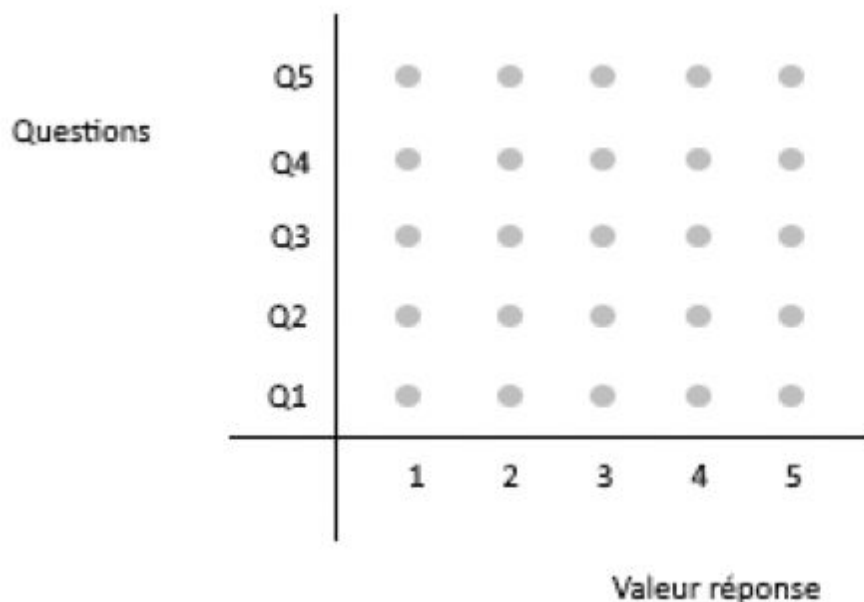
Cette visualisation montre les réponses des utilisateurs en fonction des questions. La valeur de la réponse est représentée par sa couleur.

Problème : on a dû modifier le jeu de données et rajouter un `id_user` pour chaque question.

Sera-t-il nécessaire de réaliser cette étape de prétraitement avec D3.js également ?

13/12/2016 :

Visualisation 2 : Glyphes et sélection de patterns



Le GIF ci-dessus (réalisé à l'aide d'un ensemble d'images) représente la seconde visualisation que nous souhaitons mettre en place.

15/12/2016 :

Retour sur les prototypes de visualisations :

Pour les cercles :

Attention, les échelles de likert ne sont pas toujours équilibrée au niveau des réponses possibles (ex : est ce que la distance entre “très satisfait” et “satisfait” est la même que la distance entre ‘satisfait” et “neutre”

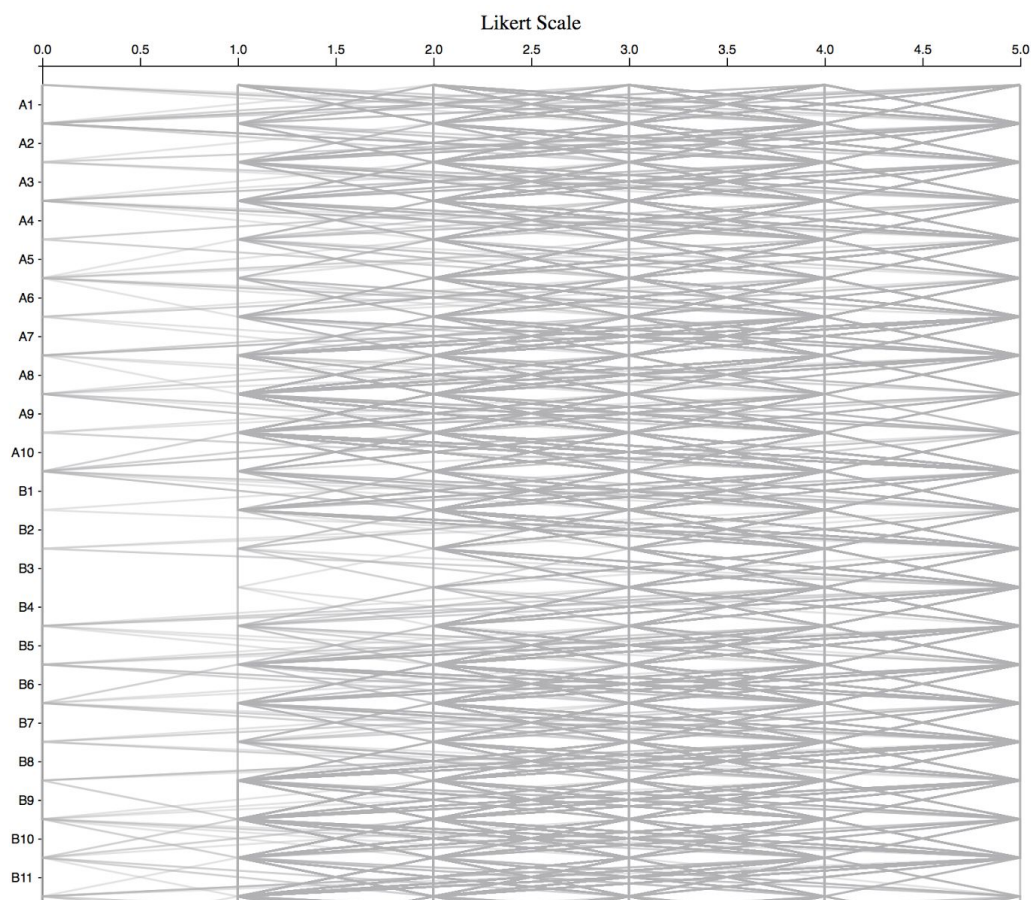
Pour les glyphes :

Il peut être intéressant d'inclure l'affichage sélectif selon les métadonnées et la réorganisation des questions.

29/12/2016 :

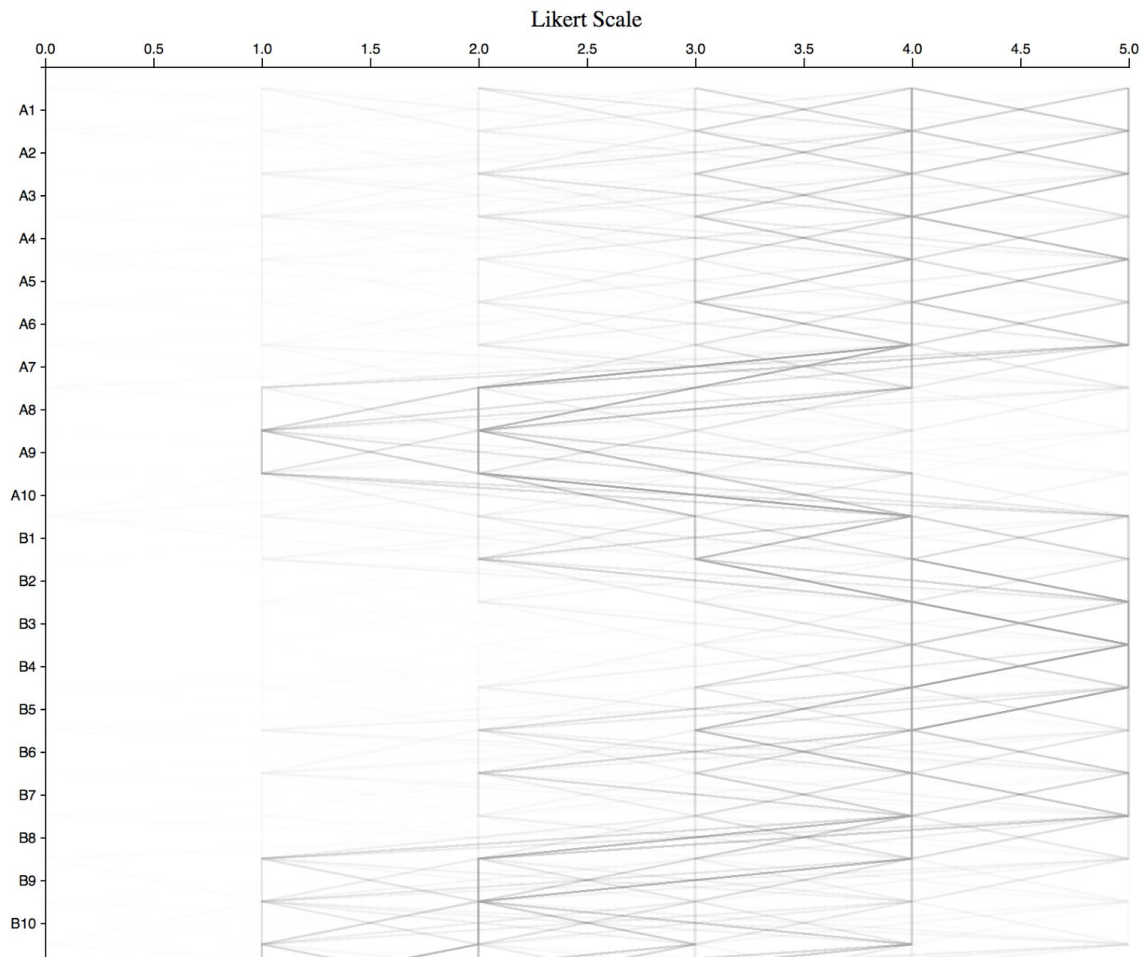
Glyphes :

Première visualisation de questionnaire sous forme de glyphe



Nombre de maximal de questionnaire à afficher pour une meilleure lisibilité

$n = 10 * \text{Likert scale level}$



Mosaïque :

début de la réalisation de la visualisation avec les carrés de couleurs.

On peut s'inspirer du bloc suivant : <http://bl.ocks.org/oyyd/859fafc8122977a3afd6>

Il est très similaire à ce que nous souhaitons réaliser.

1ère étapes, affichage des axes :



On constate rapidement un problème, il sera difficile d'afficher beaucoup d'utilisateur et beaucoup de questions pour des raisons de lisibilité du graphe.
Pour la réalisation de la visualisation on se contente de 20 questions environs et d'une centaine d'individus répondants. On verra pour afficher plus dans la suite.
Il faudra aussi essayer de rendre la création des axes dynamiques pour qu'ils s'adaptent au nombre de questions facilement.

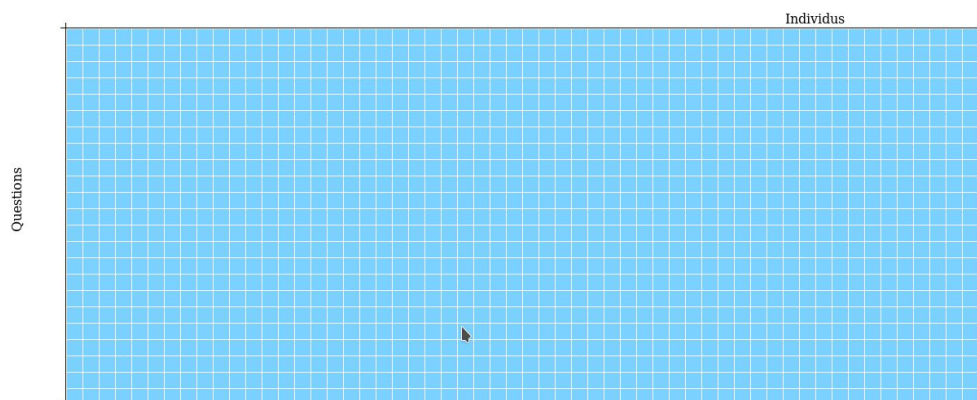
Les données devront avoir un format précis pour pouvoir les traiter -> par exemple, on peut attendre au niveau des entêtes de colonnes une distinction claire entre questions et métadonnées.

01/01/17

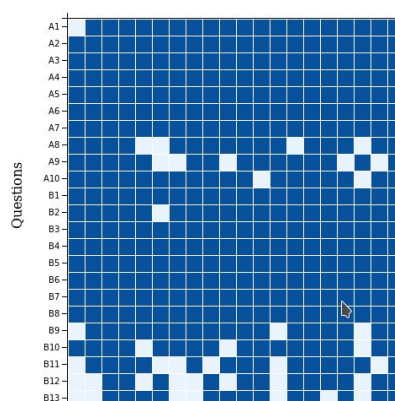
Mosaïque :

Affichage des carrés de couleurs représentant les réponses aux questions par utilisateur.
Il reste à afficher les noms des questions sur l'axes questions pour plus de clarté.

Il faudra ensuite ajouter les couleurs correspondant à la vraie valeur de la réponse puis permettre un affichage selectif en fonction des métadonnées.



Premier affichage avec la couleur dépendant de la réponse.



Pour le moment l'affichage de la couleur est codé en dur mais une échelle de couleur sera définie.

4/01/2017 :

Quand l'échelle de couleur sera ajoutée, l'objectif sera de filtrer les utilisateurs en fonctions des métadonnées présentes dans les données.

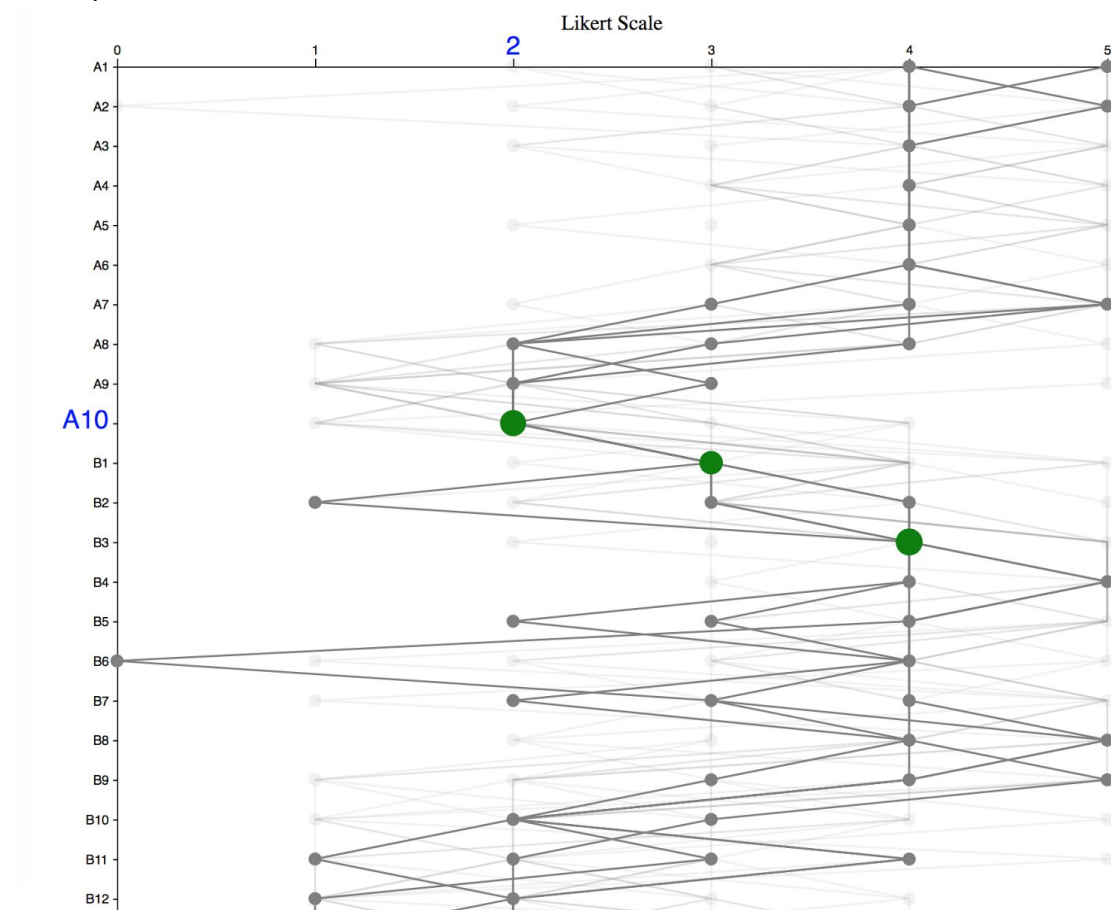
On aimerait mettre en place un endroit dans la page qui contiendrait une section par colonne de métadonnées et qui affichera autant de "checkbox" que de valeurs disponibles pour cette colonne. Un bouton permettra d'appliquer le filtre sur la liste d'utilisateurs à afficher.

03/01/2017 :

Glyphes :

Ajout du filtrage des glyphes lors de la sélection des points :

Les glyphes contenant le points ou les points sont mis en valeur tandis que les autres sont estompés.



6/01/2017 :

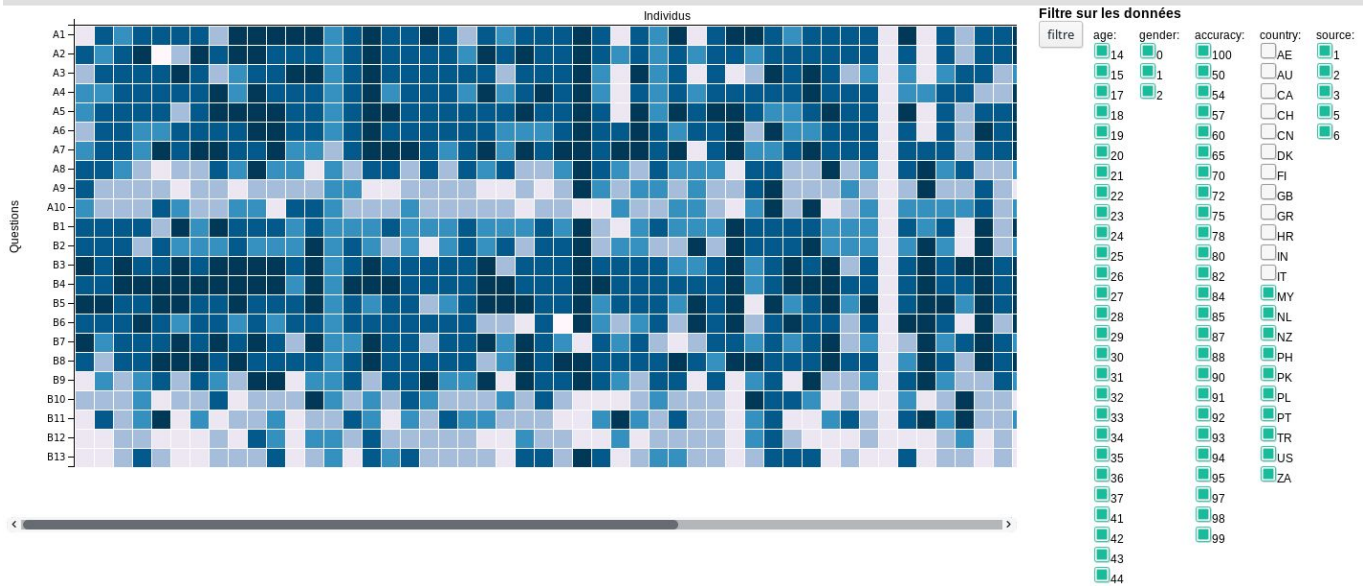
Mosaïque :

Réalisation de la partie filtre en fonction des métadonnées pour la visualisation avec les carrés. Un peu de mise en forme de la page également.

Finalement l'id user ajouté artificiellement dans les données sources ne sert à rien, on peut faire la même chose en considérant uniquement le numéro de la ligne.

Il faudra cependant demander à ce que les colonnes questions et métadonnées soient marquées comme telles pour faciliter le parsing du csv.

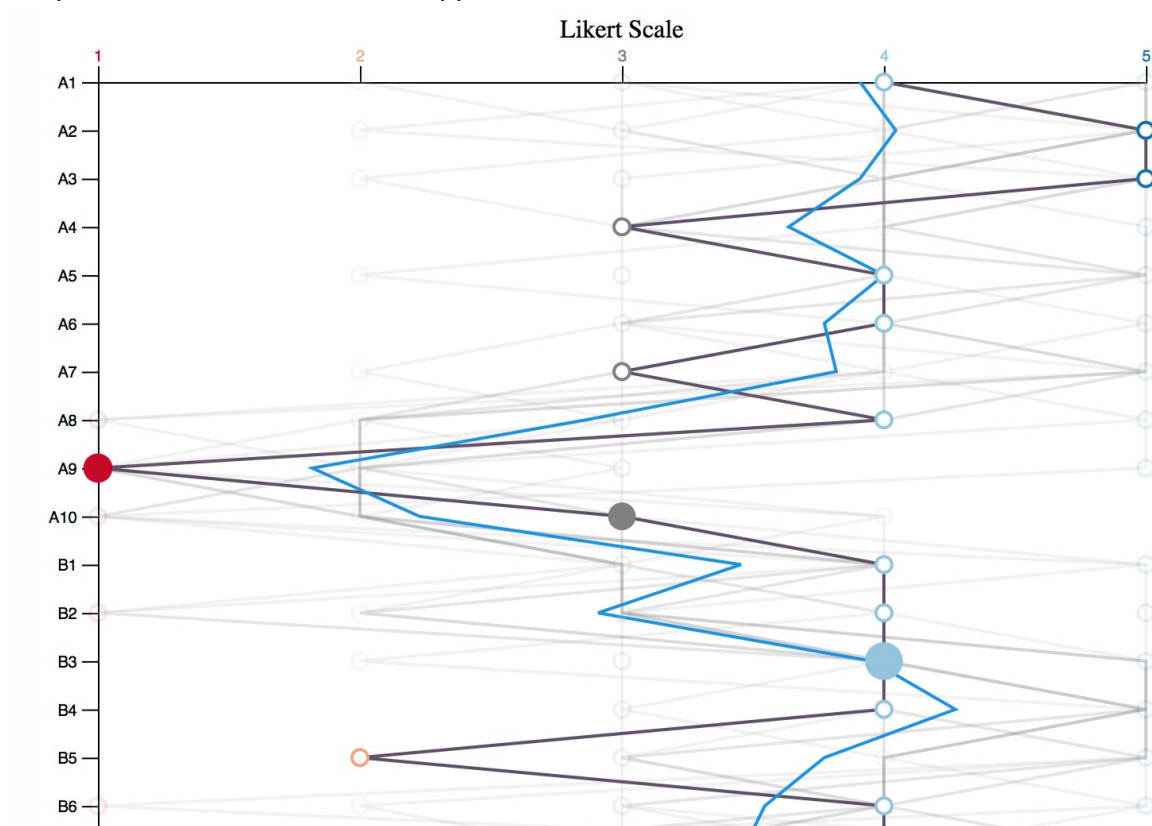
Visualisation 1 - Carrés



07/01/2017

Glyph :

Modification des formes représentant les réponses. Les cercles sont maintenant vides et se remplissent lors de la sélection rappelant des cases à cocher.



08/01/2017 :

Mosaïque :

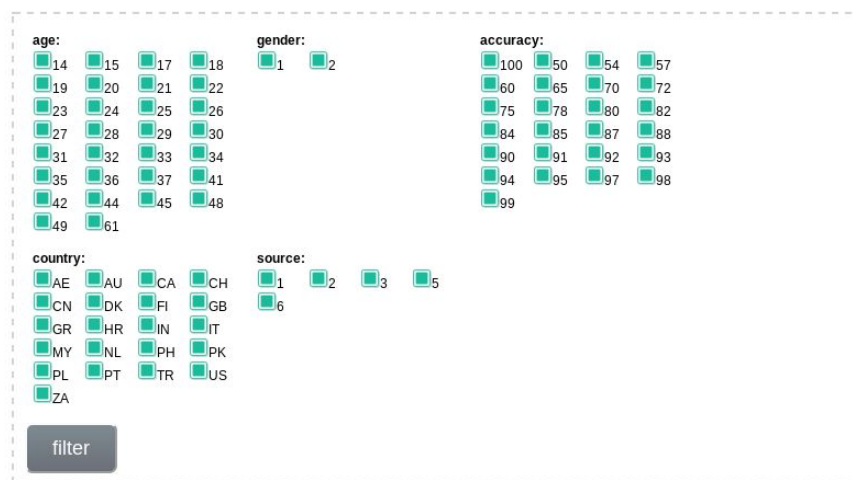
Ajout d'un tooltip permettant de donner l'individu et la question associés à chaque carrés.

Ajout d'une légende pour donner les valeurs associées à chaque couleurs et modification du code pour le rendre plus facilement adaptable à un autre dataset.



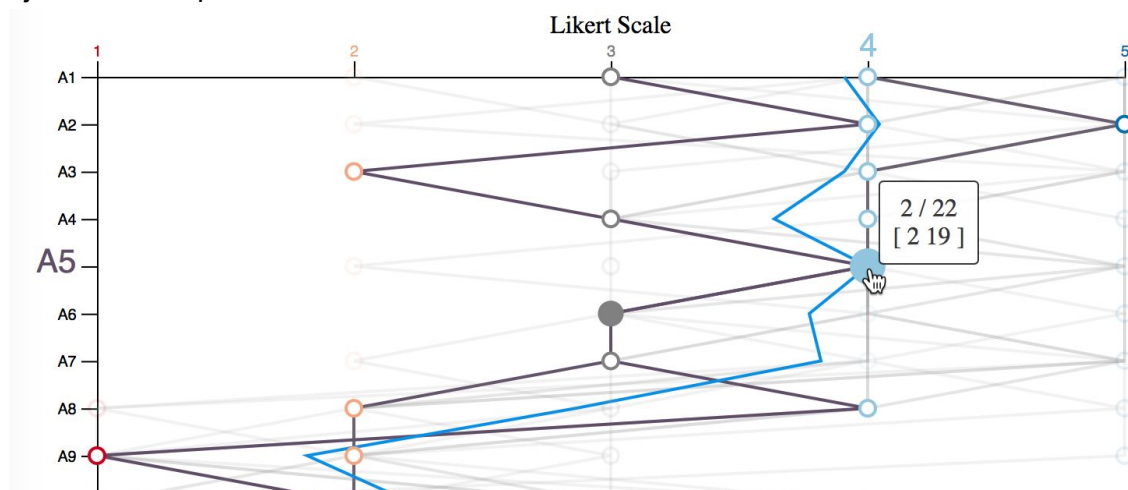
Les colonnes représentant les questions doivent avoir une entête commençant par "q_" et les colonnes représentant les métadonnées doivent avoir une entête commençant par "m_".

Mise en forme de l'affichage des filtres.



Glyph:

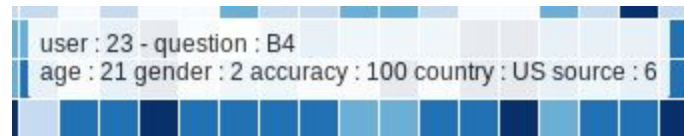
Ajout des tooltips donnant des information sur les noeuds survolé ou sélectionné.



09/01/2017 :

Mosaïque :

Ajout des métadonnées d'un individu dans le tooltip.



Rédaction du rapport et nettoyage du code

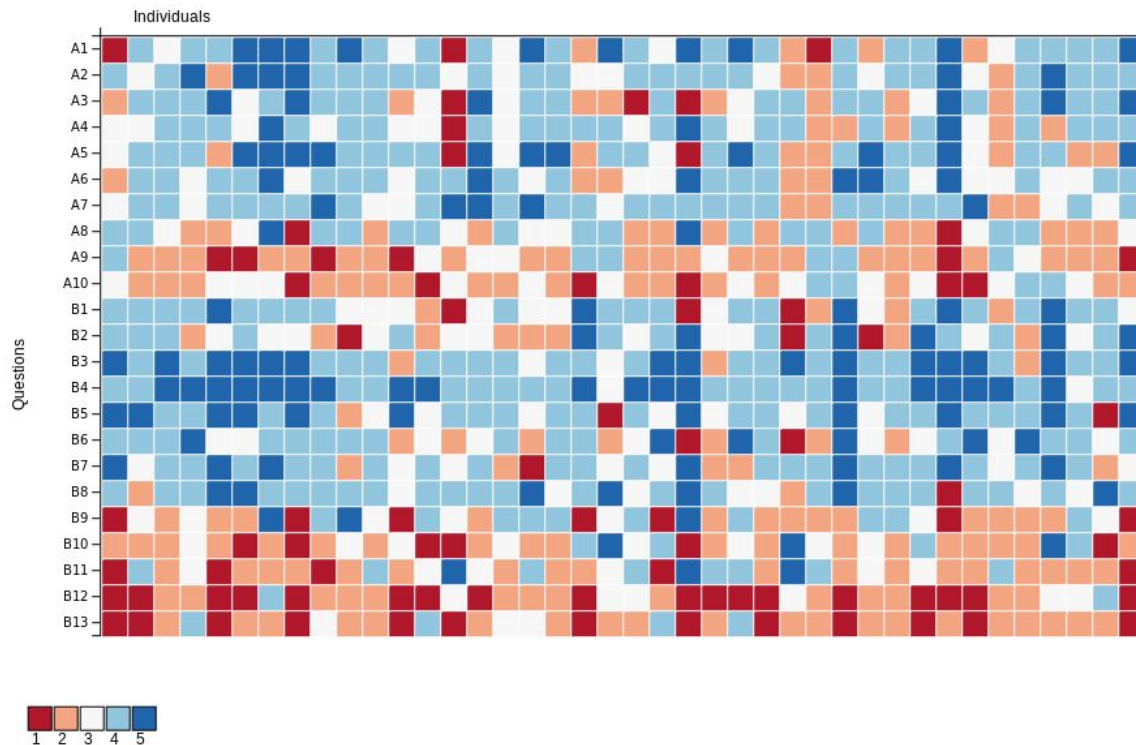
10/01/2017 :

Rédaction du rapport et nettoyage du code

11/01/2017 :

Mosaïque :

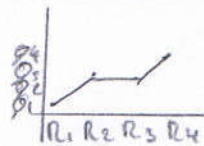
Commentaires ajoutés dans le code et changement de l'échelle de couleurs de nuances de bleus à bleu-rouge pour faciliter la lecture et associer une signification aux couleurs.



Rédaction du rapport et nettoyage du code

1. Ideas

glyphe = profil



aggrégat de glyphes pour visuel simple

→ très adapté au questionnaire de statistique

≠ granularité (groupe, profil, utilisateur)

utiliser les métadonnées dispo pour catégoriser

à déterminer nous même les profils (data mining)

principe de calques pour naviguer dans les glyphes (dans les axes / dans les groupes)



mettre à l'évidence les différences entre les glyphes (couleur) → comparer en utilisant la même catégorie

détecter des questions interdépendantes

→ couleurs ? glyphes à nouveau ?

→ faire un peu de Data Mining pour avoir les NA

→ Calques

glyphe + coloration

possibilité de dépasser des données initiales

Donner des Dats → "dont de personnes a donné cette réponse"

2. Filter

- glyphe pour représenter les réponses à différents niveaux de granularité

- système de calques pour comparer les glyphes (couleur)

- utilisation des métadonnées dispo pour catégoriser

- déterminer des profils visuellement

- données supplémentaires (statistiques → "dont de personnes de telle catégorie ont donné cette réponse")

- détecter visuellement des questions liées

(si l'utilisateur répond 1 à la question 1, il répond 2 à la question 2)

- pouvoir se débarrasser des données non pertinentes

3. Categorize

glyphes → représentation des réponses des utilisateurs / groupes

détecter les réponses non pertinentes

calques → pour comparer les glyphes

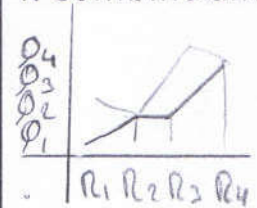
faire apparaître les tendances

déterminer visuellement des profils des assemblées de questions interdépendantes

couleur → souligner les ≠ et les similitudes

utiliser les métadonnées pour faire apparaître des statistiques.

4. Combine and Refine



Aggrégation des différents glyphes

pour dégager des tendances de réponse majoritaire

éventuellement détecter des profils d'utilisateurs

Représenter les réponses non pertinentes

Avoir la possibilité de manipuler les glyphes comme des calques pour les comparer facilement, mettre à l'évidence les ≠ à l'aide de couleurs

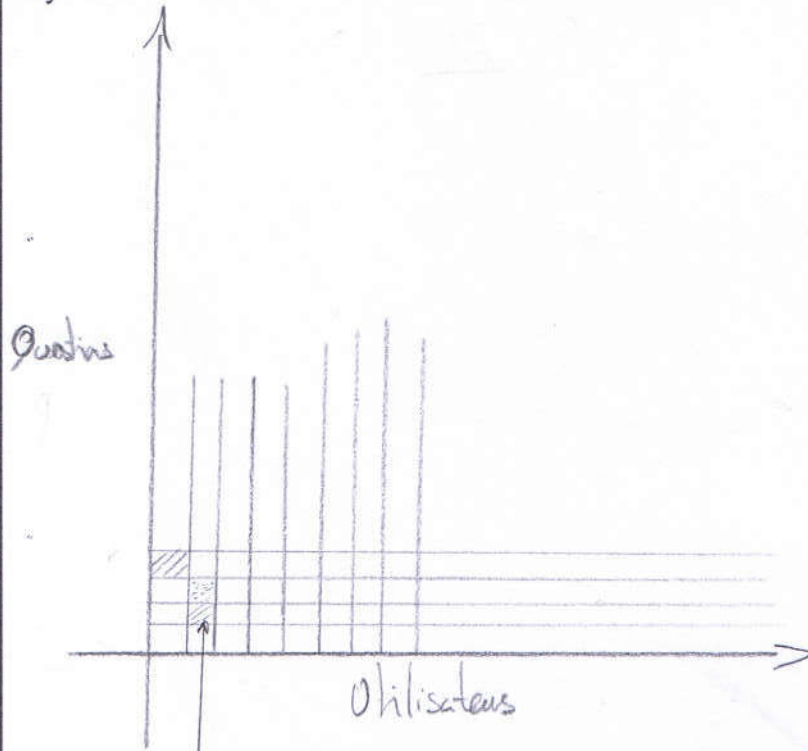
Données supplémentaires des statistiques sur une question ou un utilisateur permet d'apparaître si on passe la souris dessus.

5. Question

Détection des similarités entre glyphes différents

Avoir des questionnaires avec un minimum de méta-données

Layout



Couleur prise dans une échelle pour représenter la réponse

par ex = Orange foncé -2 Orange clair -1 Blanc Neutre 0 Bleu -1 Bleu foncé +2

Title: Couleur des réponses

Author: Alix Gombot / Raphaël Dierckx

Date: 30/11/2016

Sheet: 2

Task:

Operations

donis sur 1 individu
mise en évidence des
réponses de l'individu
(les autres colonnes deviennent
moins opaques)

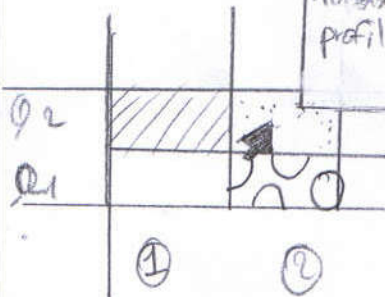
même chose avec les questions
possibilité de masquer un

possibilité de faire varier
la position des utilisateurs sur
l'axe pour trier cela la
méthode est même de les
déplacer à la main

Focus

Quand la souris est positionnée sur une case, apparition
d'un tooltip.

Réponse moyenne à cette question
• écart de la question (si d'usage)
• Réponse à la question
• Réponse moyenne des utilisateurs de
profil similaire
(...)



Discussion

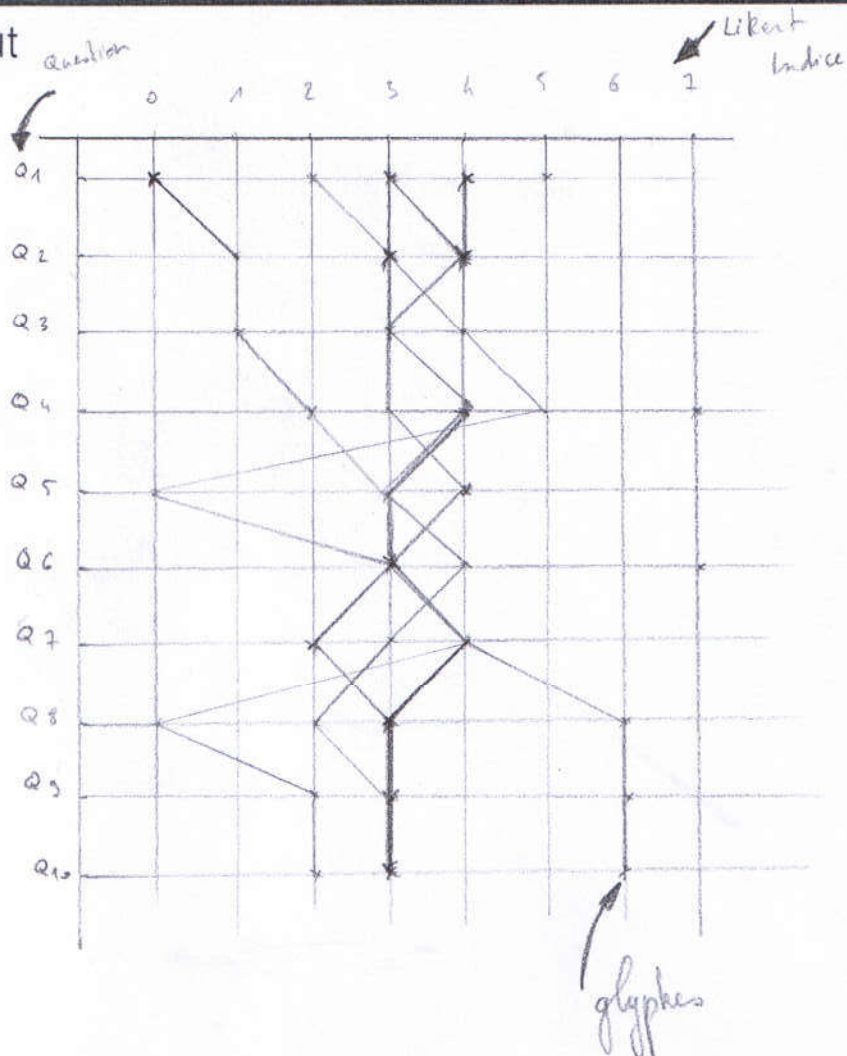
Inconvénients

- dense
- difficile de détecter
des profils sur des
données volumineuses

Avantages

- comparaison individus
facile visuellement quand
ils sont proches sur l'axe
des X.

Layout



Title: Une glyphe

Author:

Date:

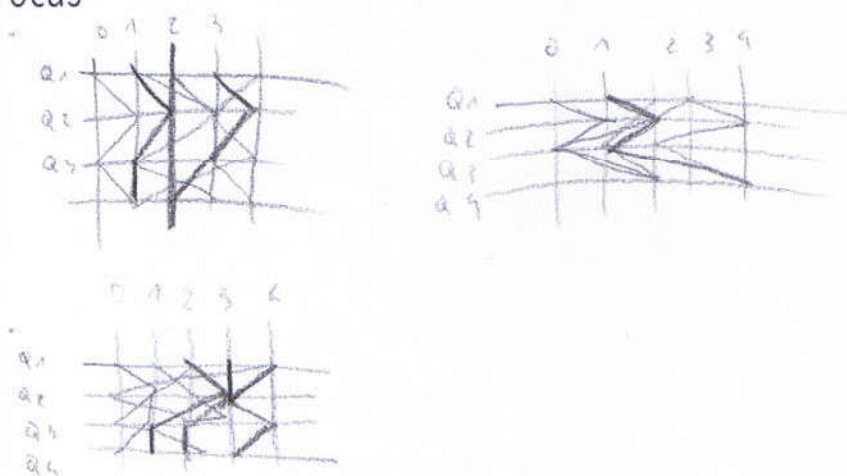
Sheet:

Task:

Operations

- Sélectionner un glyphe
- Sélectionner une réponse
- Sélectionner un point des glyphes

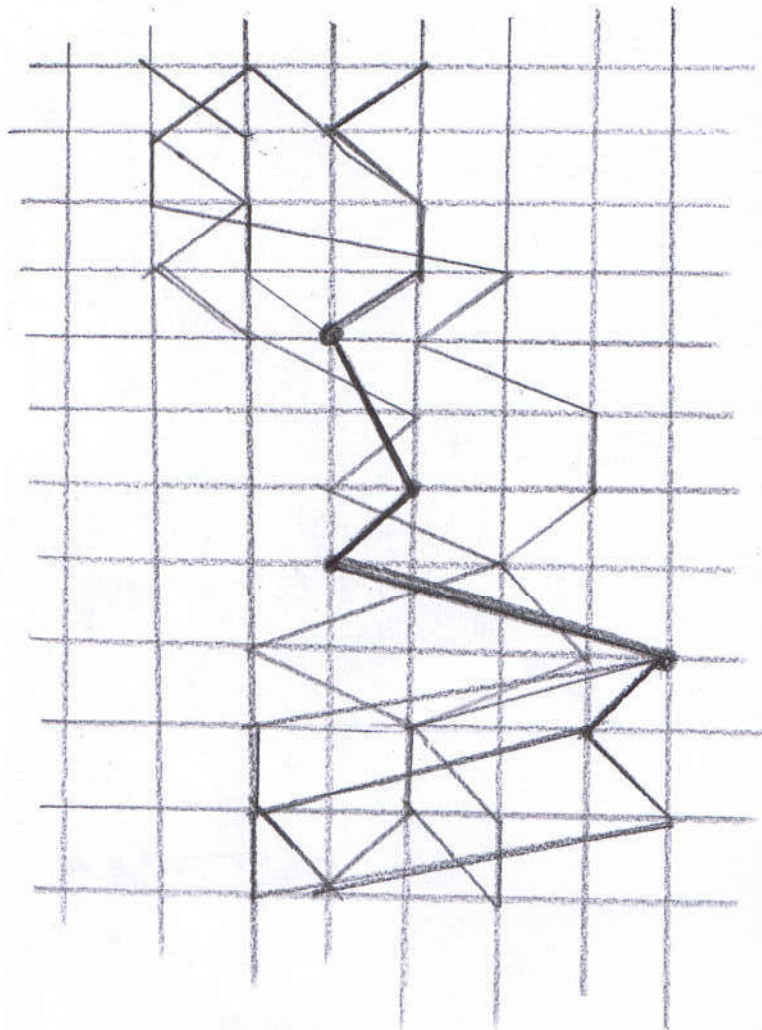
Focus



Discussion

- Problème de densité des données : plein de glyphes superposés.
- ↳ Utilisation de laques d'utilisation.
- Navigation entre les données : passage de la souris pas optimal à cause de la densité de données.

Layout



Title: Matching entre glyphe

Author:

Date:

Sheet:

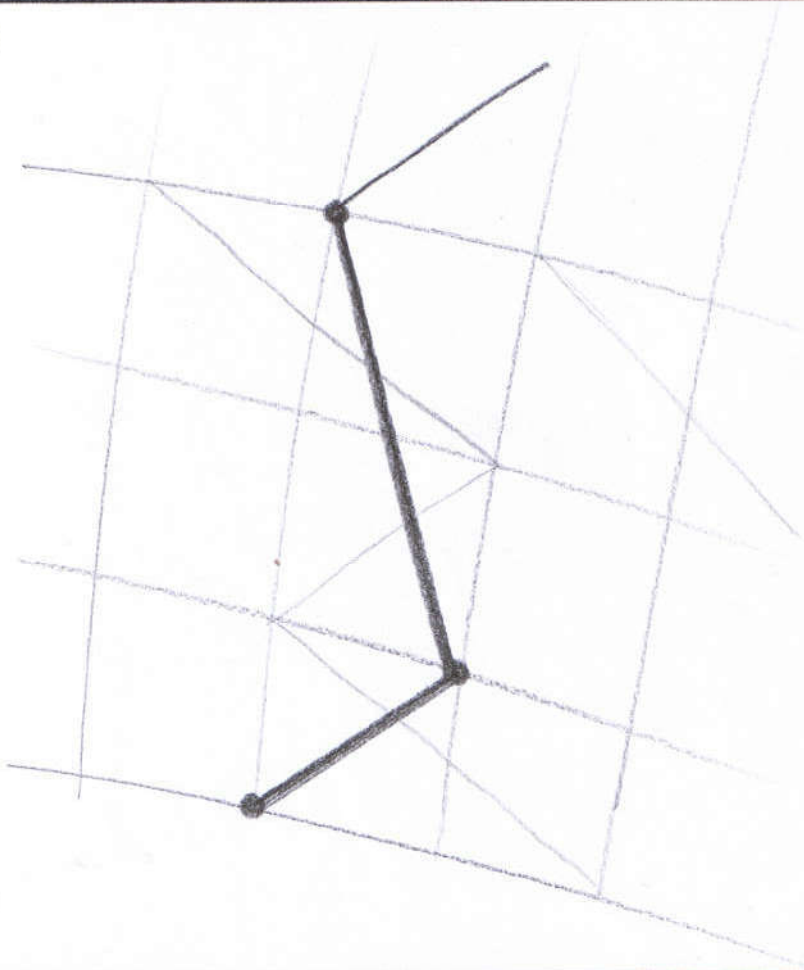
Task:

Operations

Sélectionner un morceau
de glyphe

↳ tous les glyphes
qui contiennent ce
morceau sont mis
en valeur

Focus

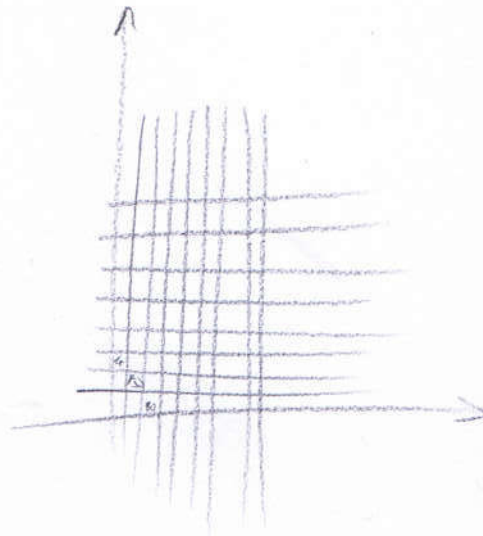
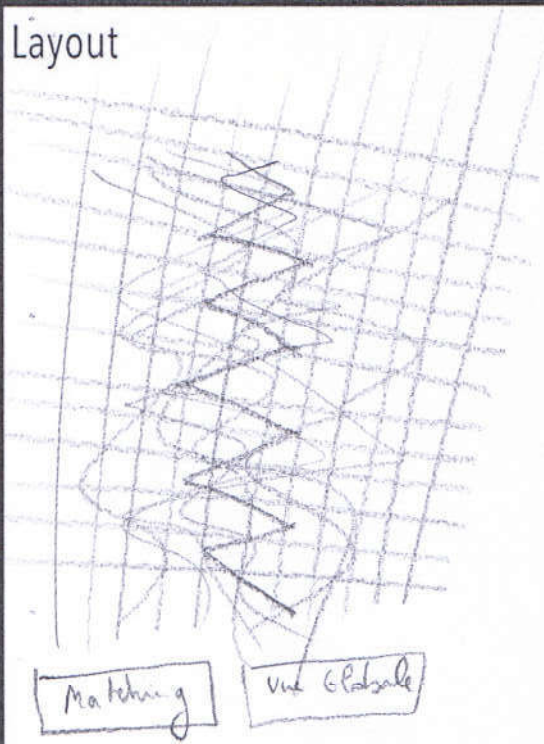


Discussion

- Problème de densité
des données
↳ faire varier la
saturation de la couleur
pour les mises en
valeur

Peut-être à combiner
avec une glyphe

Layout



Title:

Author:

Date:

Sheet:

Task:

Operations

Possibilité de basculer
de la vue matching à
la vue globale à l'aide
d'un composant graphique

Focus

Detail