

**POLYTECHNIQUE
MONTREAL**

**LE GÉNIE
EN PREMIÈRE CLASSE**



INF8808E – Visualisation de données

Été 2020

PROJET – MAQUETTE DU SITE WEB

OUTIL D'ANALYSE PAR LA VISUALISATION

PARCOURS DES USAGERS D'UN LOGICIEL VERS LA VERSION PAYANTE

Guillaume Breton Vincent 1844381

Guillaume Gfeller 1252934

Élise St-Aubin Fournier 2038735

Table des matières

Table des matières.....	1
Mise en contexte, objectifs du projet et audience.....	2
Ensemble de données (the what?).....	3
Problématiques adressées par les visualisations (the why?).....	4
Maquette (mockup) de l’outil d’analyse par la visualisation.....	5
Visualisation principale.....	6
Onglet 1 : Taux de conversion vers la version payante (<i>Conversion rates</i>)	7
Onglet 2 : Analyse de la version d’essai (<i>Trials Analysis</i>)	9
Onglet 3 : Parcours du client (<i>Customer Journey</i>)	10
Éléments de design	11

Mise en contexte, objectifs du projet et audience

Officevibe est un outil développé au Québec et conçu pour accompagner les gestionnaires à devenir de meilleurs leaders pour leur équipe¹. Il s'agit d'une plateforme qui offre une version gratuite (le mode *freemium*, avec des fonctionnalités limitées), un essai gratuit (le mode *trial*, utilisation de la version payante gratuitement pour une courte période de temps) et une version payante (le mode *premium*, incluant toutes les fonctionnalités).

L'acquisition de nouveaux clients payants (*premium*) est clé pour assurer la continuité de l'entreprise, mais demeure un défi de taille. *Officevibe* obtient plus de 1000 nouvelles inscriptions par mois (nouveaux *signups*) qui génèrent une centaine de nouveaux clients payants (*premium*).

Les équipes responsables de l'acquisition chez *Officevibe* (Marketing & Ventes) sont intéressés à en apprendre davantage sur la conversion (de *signup* à *premium*). Ces équipes seront l'audience principale de notre analyse.

L'objectif du projet est d'analyser la conversion des nouvelles inscriptions et d'effectuer le portait de ces compagnies qui deviennent des clients payants.

¹ Pour en savoir plus sur *Officevibe*: <https://youtu.be/1fUEG6EGYmU>

Ensemble de données (the what?)

Les données obtenues sont des données réelles de la compagnie. Toutes les variables sont décrites ci-dessous. Les données incluent toutes les nouvelles inscriptions depuis le 1er janvier 2020 (5 692 inscriptions). Les données ont été extraites en date du 25 mai 2020.

Variable	Description	Valeurs
NetworkID	Identifiant unique pour la compagnie	Valeur unique (clé)
NetworkCreationDate	Date à laquelle la compagnie s'est inscrite sur l'application (<i>signup date</i>)	Champ de type date entre janvier 2020 et mai 2020
TrialStartDate	Date à laquelle la compagnie a débuté son <i>trial</i>	Champ de type date entre janvier 2020 et mai 2020 (NULL si la compagnie n'a pas fait de <i>trial</i>)
TrialEndDate	Date à laquelle la compagnie a terminé son <i>trial</i>	Champ de type date entre janvier 2020 et juin 2020 (NULL si la compagnie n'a pas fait de <i>trial</i>)
PremiumStartDate	Date à laquelle la compagnie a débuté son compte <i>premium</i>	Champ de type date entre janvier 2020 et mai 2020 (NULL si la compagnie n'a pas jamais été dans un compte payant)
PremiumEndDate	Date d'expiration du compte <i>premium</i>	Champ de type date entre mars 2020 et mai 2021 (NULL si la compagnie n'a pas jamais été dans un compte payant) Si la date est inférieure à la date d'extraction, il s'agit de comptes qui ont annulés leur abonnement payant. Si la date est supérieure à la date d'extraction, il s'agit de comptes qui sont encore dans leur abonnement payant.
NetworkDeactivationDate	Date à laquelle la compagnie a annulé son abonnement à l'application	Champ de type date entre janvier 2020 et mai 2020 (NULL si non-applicable)
NetworkCurrentLanguage	Langue choisie par l'administrateur de la compagnie	Anglais / Français / Espagnol / Portugais / Allemand / Multiple (plusieurs langues pour la même compagnie)
NetworkSignupChannel	Canal d'acquisition principal de la compagnie	Display / Email Marketing / Office365 / Paid Search / Paid Social / slack / Social Media / Website / Yammer
NetworkSurveyEmailFrequency	Fréquence choisie par l'administrateur de la compagnie pour le sondage dans l'application	Weekly / Bimonthly / Monthly
CompanySizeNumeric	Taille de la compagnie	Champ de type <i>integer</i> ayant des valeurs entre 1 et 500 000 (Unknown dans certains cas)
Industry	Industrie de la compagnie	Champ de type texte ayant 20 valeurs différentes (exemples: ComputerSoftwareITServices, Manufacturing, NaturalResources, etc.)
Country	Pays principal dans laquelle la compagnie évolue	Champ de type texte ayant plus de 100 valeurs différentes (exemples: Canada, France, United States, South Africa, etc.) – pourrait éventuellement être groupé par continent
NetworkActivated	En plus des stades freemium – trial – premium, le network « activé » signifie que la personne ayant inscrit la compagnie a effectué au moins une action dans l'outil post-inscription	Champ de type booléen : 0 ou 1
NetworkFreemiumActivationDate	Date à laquelle la compagnie est considérée comme activée	Champ de type date entre janvier 2020 et mai 2020 (NULL si non-applicable)
CurrentSubscriptionPlanCode	Pour les compagnies en mode premium, le type d'abonnement choisi	admin_procedures (abonnement spécial) / annual-1000-standard / monthly-1000-standard / trial-7-days / trial-14-days / trial-30-days / NULL si non-applicable

Problématiques adressées par les visualisations (the why?)

Les questions soulevées ci-dessous fournissent toutes une information pertinente à la prise de décision dans le processus d'acquisition de nouveaux clients. Elles se divisent en trois catégories : une question sommaire (dans le but de fournir de l'information utile à la décision), qui se divise en deux types d'analyse : une analyse descriptive (en lien avec les caractéristiques des compagnies) et une analyse comportementale (en lien avec certaines actions effectuées par les compagnies).

Questions	Niveau de priorité	Viz princ.	Onglet 1	Onglet 2	Onglet 3
Question sommaire					
S1. Quelles actions les équipes d'acquisition (marketing & ventes) pourraient-elles prendre pour augmenter le nombre de conversions?	* * *	Question générale			
Analyse descriptive					
AD-1. Quels types de compagnies sont les plus susceptibles de convertir à un abonnement payant?	* * *		X		
AD-2. Est-ce que certains types de compagnies ont des taux de conversion très faibles et devraient donc être exclus des compagnies cibles?	* *		X		
AD-3. Est-ce que la distribution des types d'abonnements demeure constant à travers le temps?	* *				
AD-4. Est-ce que le taux de conversion démontre des variations mensuelles ou saisonnières?	*				
AD-5. Est-ce que le canal d'acquisition <i>paid search</i> est efficace et permet d'augmenter la conversion?	* *		X		
Analyse comportementale					
AC-1. Est-ce que la longueur du <i>trial</i> a une influence sur la conversion?	* *				
AC-2. Est-ce que le moment auquel on débute le <i>trial</i> a une influence sur la conversion?	* *			X	
AC-3. Est-ce que les compagnies qui sont <i>activés</i> convertissent davantage?	* *				
AC-4. Est-ce que les compagnies qui font un <i>trial</i> convertissent davantage?	* * *			X	
AC-5. Quels sont les parcours typiques que font les clients après leur inscription dans le produit?	* * *				X

Maquette (mockup) de l'outil d'analyse par la visualisation

L'objectif de l'outil est de permettre une analyse stratégique du parcours des nouveaux clients et des facteurs qui influencent leur conversion vers la version payante de l'application. Le site web présentera donc les données introduites ci-dessus dans des visualisations qui faciliteront l'analyse stratégique.

L'objectif est que l'outil soit utilisé par les équipes de Marketing et de Ventes lors de la prise de décisions stratégiques. La présentation des visualisations se fera via une approche tableau de bord, qui permettra aux utilisateurs d'effectuer l'exploration des données et de répondre à plusieurs questions clés.

Les visualisations sont découpées en quatre sections. La visualisation principale sera présentée dans le haut de la page et elle sera toujours visible. L'objectif est de donner du contexte à haut niveau à l'utilisateur, et que cette information soit toujours présente. Ensuite, l'utilisateur pourra sélectionner trois onglets différents qui adressent chacun un thème spécifique.

Pour chacune des quatre sections de visualisation, cette section du rapport contient une description sommaire, les questions ciblées par la visualisation (sauf pour la page principale), la maquette de la visualisation ainsi que les interactions possibles.

Visualisation principale

Description sommaire

La visualisation principale sera toujours visible durant l'utilisation du tableau de bord. Celle-ci sera présentée au-dessus des onglets et présentera des caractéristiques générales en lien avec tous les utilisateurs de la plateforme Officevibe. Plus précisément, la vue sommaire affichera la moyenne mensuelle de nouvelles inscriptions (signups), du nombre d'essais gratuits effectués (trials) et le nombre de nouveaux clients payants (premiums). De plus, des diagrammes à bande afficheront la distribution des compagnies par pays et par taille.

Cette information demeure toujours visible afin que l'utilisateur puisse toujours se référer au contexte global de l'analyse (population totale).

Maquette de la visualisation

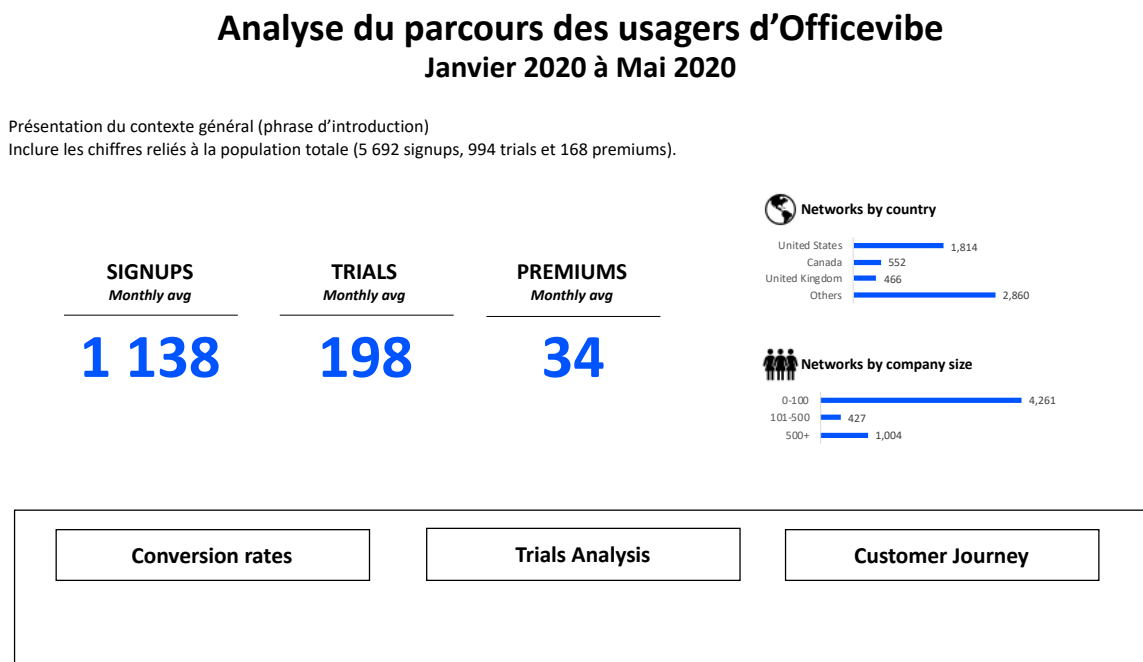


Figure 1: Visualisation principale

Interaction avec la visualisation

La visualisation agit comme menu principal pour les autres visualisations. La seule interactivité dans ce cas se trouve à être les onglets que l'utilisateur peut sélectionner.

Onglet 1 : Taux de conversion vers la version payante (*Conversion rates*)

Description sommaire

Le premier onglet présente une analyse plus détaillée des taux de conversion des clients vers la version payante. Le taux de conversion estimé (\hat{p}) pour une catégorie i est défini comme le nombre de clients convertis (CC) sur le nombre total de clients inscrits (N) qui est la somme des clients inscrits et des clients non-convertis (CN).

$$\hat{p} = \frac{CC_i}{N_i} = \frac{CC_i}{CC_i + CN_i}$$

Typiquement, le taux de conversion sera évalué suite à un certain délai après l'inscription afin de permettre au client qui font l'essai avec la version freemium ou trial d'avoir suffisamment de temps pour convertir. Par exemple, le taux de conversion 90 jours des clients du Canada, sera le nombre de clients canadiens qui se seront convertis à la version payante après 90 jours sur le nombre total de clients canadiens sur cette période.

Dimensions d'analyse

Le taux de conversion sera analysé visuellement selon trois différentes dimensions, soient :

- La taille du client (*Company Size*)
- Le pays d'origine du client (*Country*)
- Le canal d'enregistrement (*Signup Channel*)

Intervalles de confiance

Pour chaque estimation du taux de conversion \hat{p} , il est possible de calculer des intervalles de confiance. Ces intervalles de confiance sont fonction du nombre de clients convertis (CC), du nombre total de clients inscrits (N) et du niveau de confiance choisi (e.g. 80%). Les détails de la méthodologie de calcul seront rendus disponibles à l'utilisateur.

L'avantage d'utiliser des bornes de confiance est de donner à l'utilisateur une idée de l'incertitude sur les taux de conversion estimés. Cette incertitude peut être très élevée pour certaines catégories avec des N petits et l'absence d'incertitude pourrait donc amener à des conclusions ou des décisions qui ne sont pas appropriées.

Questions ciblées

L'analyse des taux de conversion permettra de répondre aux questions suivantes :

AD-1. Quels types de compagnies sont les plus susceptibles de convertir à un abonnement payant?

AD-2. Est-ce que certains types de compagnies ont des taux de conversion très faibles et devraient donc être exclus des compagnies cibles?

AD-5. Est-ce que le canal d'acquisition paid search est efficace et permet d'augmenter la conversion?

Maquette de la visualisation

Il y aura deux sections principales. La première présentera les taux de conversion selon les trois dimensions dans un graphique de type *box plot*, et la deuxième section affichera un graphique de type *heatmap*, incluant les 3 dimensions décrites ci-dessus.

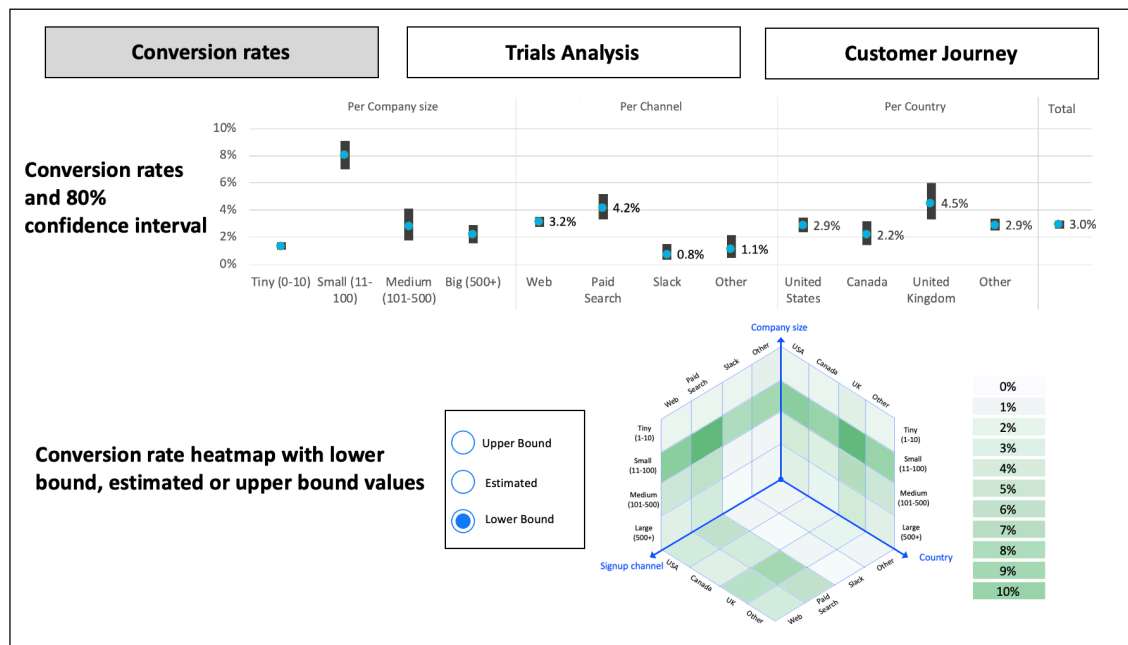


Figure 2: Maquette de l'onglet "Conversion rates"

Interaction avec la visualisation

Pour la première composante qui présente les taux de conversion et leurs intervalles de confiance, quand la souris planera sur un élément, une info-bulle (tooltip) spécifiant les informations suivantes s'affichera : le nombre de clients total (N), le nombre de conversions (CC), le taux de conversion estimé (p), ainsi que les valeurs inférieure et supérieure de l'intervalle de confiance à 80%.

Pour la deuxième portion qui présente le *heatmap*, l'utilisateur aura accès à un bouton radio pour sélectionner laquelle des trois valeurs sera utilisée dans le *heatmap*. En planant sur une case du *heatmap*, il verra aussi une info-bulle qui donnera la valeur représentée dans la cellule.

Onglet 2 : Analyse de la version d'essai (*Trials Analysis*)

Description sommaire

Une idée importante qui guide le modèle d'affaire est qu'offrir un essai de la version premium augmente le taux de conversion. Cependant, il est difficile de déterminer si c'est concrètement le cas considérant les multiples paramètres qui peuvent influencer le client. Le deuxième onglet aura donc pour objectif de visualiser l'impact d'offrir une période d'essai du service payant.

Questions ciblées

L'analyse des taux de conversion permettra de répondre aux questions suivantes :

AC-2. Est-ce que le moment auquel on débute le trial a une influence sur la conversion?

AC-4. Est-ce que les compagnies qui font un trial convertissent davantage?

Maquette de la visualisation

Les pourcentages de conversion avec et sans trial seront présentés directement dans la visualisation. Un second graphique viendra présenter davantage d'information en indiquant le taux de conversion selon le moment (*timing*) de début du trial.

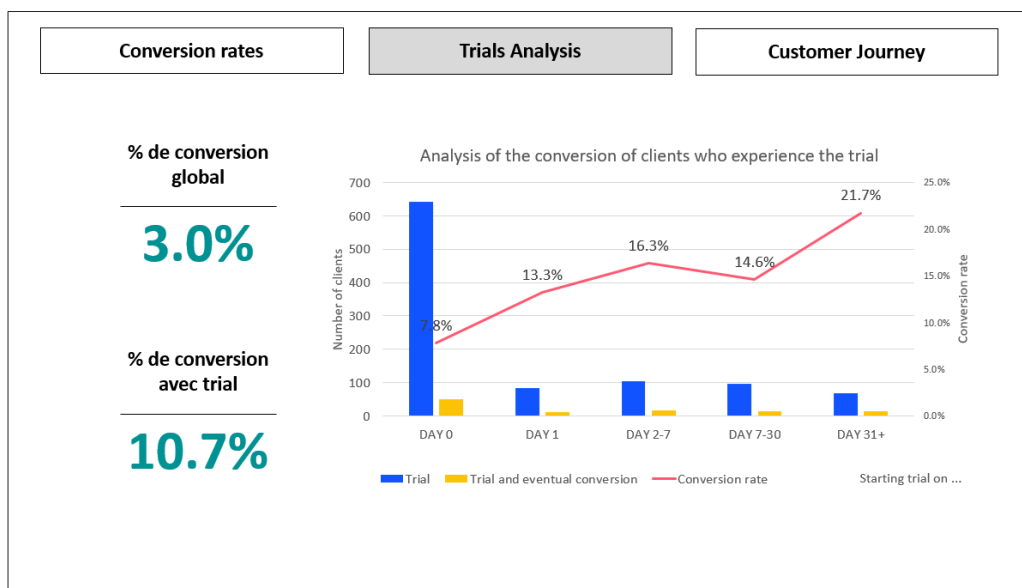


Figure 3: Maquette de l'onglet "Trials Analysis"

Interaction avec la visualisation

Aucune interaction prévue pour cet onglet.

Onglet 3 : Parcours du client (*Customer Journey*)

Description sommaire

Pour le moment, il est difficile d'avoir une vue d'ensemble des parcours des clients avant de s'abonner au service payant. De plus, il est difficile de connaître les temps passés dans chacune des étapes. Le troisième onglet viendra fournir davantage d'information en visualisant les parcours typiquement effectués par les clients.

Questions ciblées

L'analyse des taux de conversion permettra de répondre aux questions suivantes :

AC-5. Quels sont les parcours typiques que font les clients après leur inscription dans le produit?

Maquette de la visualisation

Un premier graphique de type *stacked area* offrira une vue de la distribution des différents statuts de la clientèle à travers le temps. Le moment 0 sur l'axe des X correspond au moment où le client se crée un compte. Le deuxième graphique est un *stacked bar chart* qui présente les 7 parcours possibles des clients, avec le nombre de jours moyens dans chaque étape, ainsi que le pourcentage de la population qui a effectué ce parcours. Cette visualisation présente une information qui est très difficile à comprendre en regardant seulement les données.

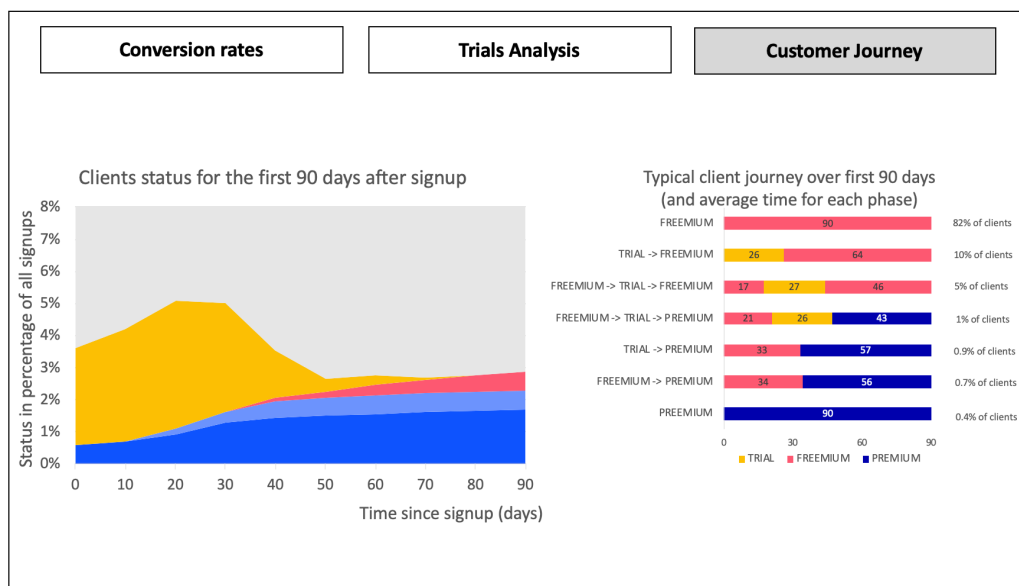


Figure 4: Maquette de l'onglet "Customer Journey"

Interaction avec la visualisation

L'utilisateur pourra passer sa souris sur le graphique de gauche, et ainsi voir une info-bulle avec les pourcentages contenus dans chaque catégorie à un moment précis dans le temps.

Éléments de design

La palette de couleur utilisée sera celle qui correspond au *branding* de *Officevibe*. Celle-ci comporte les couleurs présentées dans la figure ci-dessous.

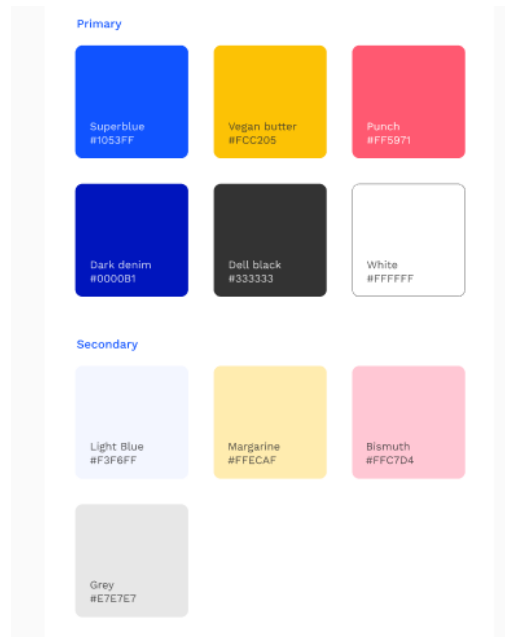


Figure 5: Liste des couleurs primaires et secondaire utilisées

Le tableau de bord sera rédigé en anglais, afin d’accommoder les utilisateurs anglophones.

De plus, puisque l’outil sera majoritairement accédé avec un ordinateur, le tout sera optimisé pour un visionnement sur écran d’ordinateur avec des dimensions 16:9.