RCSO
réseau de compétences HES-SO
Information Services Network

Hes·so
Haute Ecole Spécialisée
de Suisse occidentale
Fachhochschule Westschweiz
University of Applied Sciences and Arts
Western Switzerland

Projet VIDAS :
Data Stories et proposition de modèles

Gaia Bongi

Laetizia Sabatini Choquard

Tables des matières

1. RAPPEL DES DÉFINITIONS 4

1.1. STORYTELLING WITH DATA 4
1.2. DATA STORIES 4

2. ÉLÉMENTS CLÉS D'UNE DATA STORY 4

2.1. DONNÉES 5
2.1.1. INSIGHT 7
2.2. NARRATION 8
2.2.1. RÉCIT 10
2.3. VISUEL 11

3. TYPOLOGIE DE DATA STORIES 14

3.1. TROIS PROTOTYPES DE DATA STORIES 17
3.1.1. ANNOTATED CHART 18
3.1.2. COMICS 19
3.1.3. DIAPORAMA – PRÉSENTATION ORALE 20

BIBLIOGRAPHIE 21

ICÔNES 22

Table des illustrations

Figure 1: Why telling story with data is essential to success (LaptrinhX 2016) 5
Figure 2: 4D Audience Framework (Dykes 2020) 5
Figure 3: Deux phases pour le développement d'une Data Story (Dykes 2020) 6
Figure 4: Contrefaçon de Data Story : Data Cut (Dykes 2020) 6
Figure 5: Contrefaçon de Data Story : Data Cameo (Dykes 2020) 7
Figure 6: Contrefaçon de Data Story: Data Decoration (Dykes 2020) 7
Figure 7: Data Storytelling Continuum (Dykes 2020) 9
Figure 8: The Three Acts of a DataStory (Duarte 2019) 10
Figure 9: La pyramide de Freitag (Writers 2020) 11
Figure 10: Prototype Annotated Chart (HEG 2022) 18
Figure 11: Prototype Data Comics (HEG 2022) 19
Figure 12: Prototype Diaporama (HEG 2022) 20

Tableaux

Tableau 1: Attributs clé pour envisager le storytelling 9
Tableau 2: Visuels utilisés pour les Data Stories 11
Tableau 3: 4 principes universel de design 13
Tableau 4: Eléments utiles à la valeur informative 14
Tableau 5: Genres de Data Stories 15
Tableau 6: Typologies de Data Stories 16
Tableau 7: Structures de Data Stories 17

1. Rappel des définitions

Avant de se plonger dans les éléments clés et la typologie des Data Stories, il semble utile de se remémorer les définitions que nous avons sélectionnées et utilisées dans l'élaboration de la classification et de la typologie des visualisations, étant donné que les Data Stories en font partie.

1.1. Storytelling with data

Une visualisation de données, quand elle est complétée par d'autres éléments que les graphiques, est capable de guider le lecteur dans un parcours de compréhension. Le Storytelling with data, c'est par conséquent raconter les données et guider le lecteur dans leur compréhension en utilisant du texte et des éléments visuels pour mettre en avant certaines informations plutôt que d'autres et pour suggérer un ordre narratif. Raconter des données n'implique pas forcement l'objectif d'un appel à l'action.

> « Data storytelling takes the viewer on a journey through the data. Data storytelling marries data visualization with a guided narrative. It pairs the data and the graphics with words, not only describing what can be seen in the image, but telling a story to lead you through the analysis process » (Kent 2017).

1.2. Data Stories

> « Data Stories combine compelling data, valuable visualizations, and a strong narrative in a way that encourages organizational action » (Aanderud 2018).

Nous utilisons le terme *Data Story* pour faire référence à des visualisations qui combinent des éléments visuels et de texte pour communiquer des informations et amener le lecteur à une prise de conscience et à l'action. Ce qui permet de faire une distinction claire entre une Infographie et une Data Story est l'objectif attendu : une Data Story est réalisée pour communiquer une information afin de produire un changement, alors qu'une Infographie est vouée à informer.

2. Éléments clés d'une Data Story

Toute communication de données qui implique la visualisation de données d'une manière graphique ou picturale ne devient pas une Data Story. Selon Dykes (2020), qui rejoint la définition d'Aanderud (2018), la simple présence d'une visualisation de données ne signifie pas qu'une histoire de données est racontée ou qu'une histoire a été prévue.

Une Data Story est en effet constituée de trois éléments importants : la donnée, le visuel et le récit. L'équilibre entre ces trois aspects est primordial et il est indispensable de toujours partir de la donnée et de l'insight (Dykes 2020), à savoir l'information qui émerge des données.

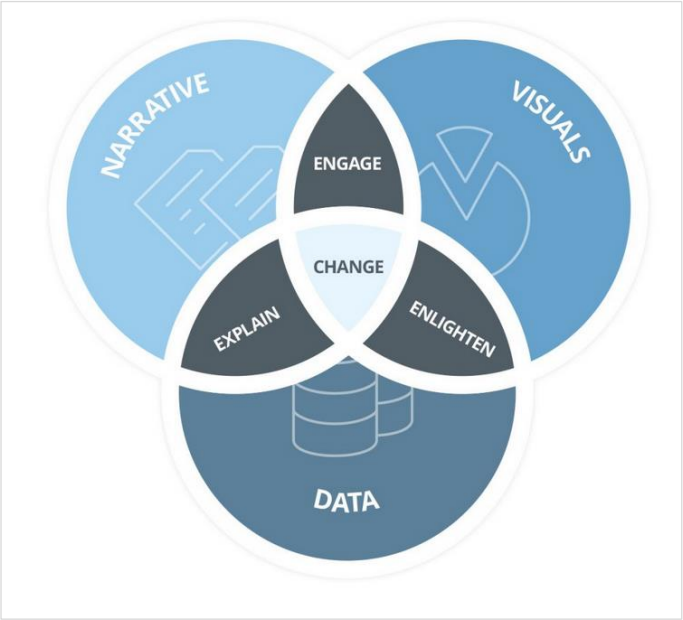


Figure 1: Why telling story with data is essential to success (LaptrinhX 2016)

2.1. Données

Une Data Story se base sur des données pertinentes et fiables, se construit autour d'un insight et il implique obligatoirement de prendre en compte l'audience dans le contexte de son élaboration.

Parce qu'il crée le sens, Nussbaumer Knaflic (2015) pointe en effet l'importance du contexte dans la création d'une visualisation et d'un contenu, dans le fait de raconter une histoire avec les données. Il est ainsi primordial de se poser trois questions:

- Qui : qui est le destinataire ?
- Quoi : que veut-on que le destinataire sache ou fasse ?
- Comment : quelles données sont disponibles ?

Il est ainsi essentiel de se poser les bonnes questions sur les données, car cela permet de trouver la direction et l'objectif adéquats. Chaque question que l'on se pose doit être envisagée en pensant au destinataire de la Data Story (Dykes 2020).

Lorsque l'on connaît suffisamment bien le public auquel on s'adresse, il est possible d'appliquer le modèle 4D qui permet de bien naviguer dans les données (Dykes 2020). Il s'agit de quatre éléments interconnectés permettant de s'orienter dans les données et de mieux cibler l'analyse.

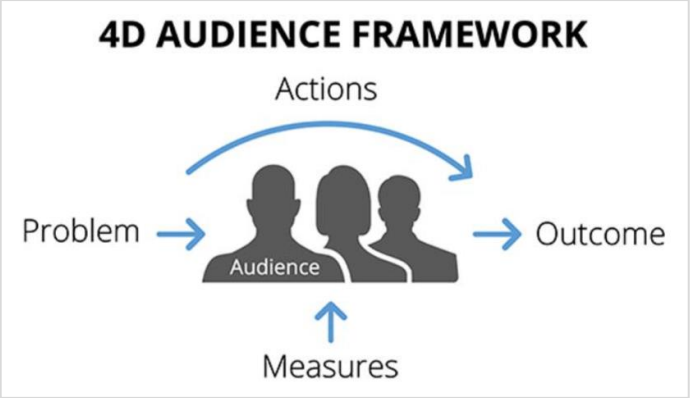


Figure 2: 4D Audience Framework (Dykes 2020)

Le *problème* est défini comme un défi ou une question clé que le public auquel on s'adresse aimerait aborder ; le *résultat* (*outcome*) est l'objectif stratégique à atteindre ou le résultat souhaité par l'audience; les *actions* sont les activités clés et les initiatives stratégiques mises en place pour résoudre un problème ou obtenir un résultat souhaité ; les *mesures* sont les mesures clés et autres données utilisées pour mettre en lumière les précédents points (Dykes 2020).

En augmentant sa connaissance pour couvrir ces 4 dimensions pour une audience spécifique, on obtient une perspective plus focalisée qui permet de poser les bonnes questions aux données (Dykes 2020).

Une fois que l'insight, qui permet la connaissance approfondie de quelque chose, a été trouvé, il s'agit de déterminer s'il peut être la base d'une Data Story (Dykes 2020). Pour ce faire, il est recommandé de le passer sous la loupe de trois questions permettant de comprendre si l'insight est bien significatif et exploitable :

- Pourquoi l'audience cible devrait s'y intéresser, s'en soucier ?
- Qu'est-ce que l'audience devrait faire de cet insight ?
- Quel est l'impact possible ?

S'il n'est pas possible de répondre à ces questions, alors l'insight ne doit pas être partagé et n'a pas besoin que son histoire soit racontée à l'aide d'une Data Story.

L'analyse des données passe par deux étapes et pour aboutir à une Data Story il faut passer de la découverte des données, à savoir lorsque des insights sont trouvés, à leur communication, c'est-à-dire quand ils sont expliqués à l'audience choisie (Dykes 2020).

Figure 3: Deux phases pour le développement d'une Data Story (Dykes 2020)

Si ces deux phases ne sont pas respectées, on peut aboutir à des résultats qui s'apparentent à des Data Stories, mais qui n'en sont pas. En effet, si la simplicité, la clarté et la cohésion sont des éléments essentiels dans la phase de data communication, il existe cependant des contrefaçons de Data Story issues de trois erreurs (Dykes 2020) :

1. Data Cut : dans la phase exploratoire, un résultat semble suffisant et l'on pense que les données et les visualisations explorées permettent de comprendre directement notre insight. Par conséquent, on les propose ainsi à notre audience. De cette manière, il n'est pas possible de prévoir la façon dont notre audience va les interpréter. On l'informe, mais on ne communique pas.

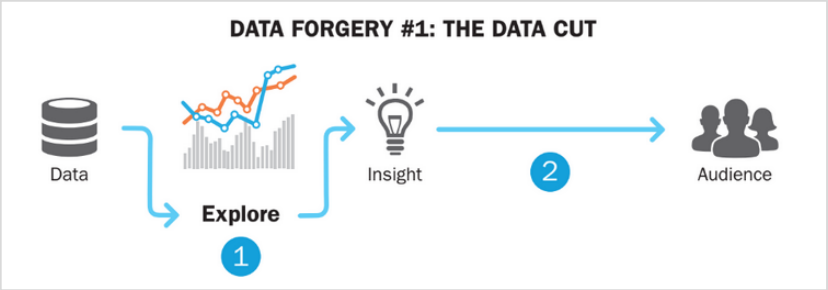


Figure 4: Contrefaçon de Data Story : Data Cut (Dykes 2020)

2. Data Cameo : on a une histoire avant d'avoir sélectionné les données. Par conséquent, on cherche les données en fonction de l'histoire pour la soutenir et bien évidemment on cherche ainsi uniquement les données qui confirment l'histoire.

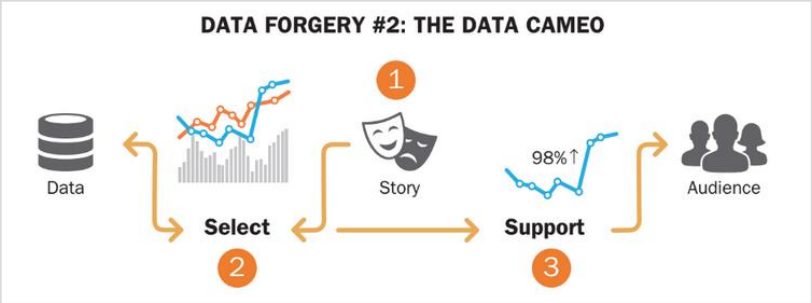


Figure 5: Contrefaçon de Data Story : Data Cameo (Dykes 2020)

3. Data Decoration : il n'y a pas d'objectif clair pour les visualisations que nous sommes en train de créer puisque l'attention est plus focalisée sur l'outils et ces performances plutôt que sur les données. Par conséquent, on visualise des données afin que les autres personnes avec plus d'expertise puissent leur donner du sens.

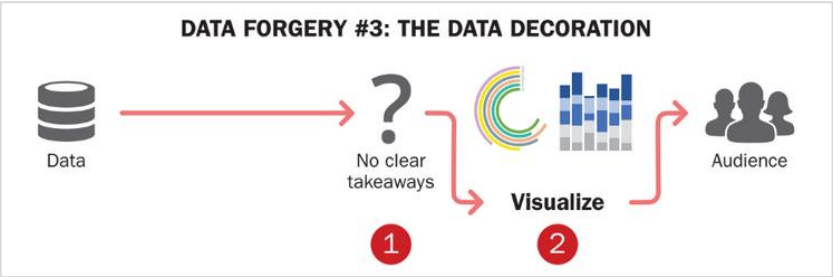


Figure 6: Contrefaçon de Data Story: Data Decoration (Dykes 2020)

2.1.1. Insight

L'insight, ou plus précisément le fait de fournir un insight, est primordial dans le cadre des Data Stories :

> « The ability to take data – to be able to understand it, to extract value from it, to visualize it, to communicate it – that's going to be a hugely important skill in the next decades » (Hal Varian 2009, économiste en chef chez Google, in Dykes 2020, p.7).

L'insight est perçu comme un processus en deux parties: « first, the insight experience of gaining new understanding or knowledge; and secondly, the consequence or product gained from the exploration of a visualization » (Rodriguez et al., 2015, citant Dove and Jones, p.8)

Toujours selon Rodriguez et al., (2015) fournir un insight se décline en quatre étapes :

> 1. Provide Overview: it allows a person to grasp the big picture concept of a dataset, allowing to understand what they know and don't know. 2. Adjust: the exploration of a dataset by adjusting the level of abstraction, enabling the person to change perspective on the dataset making sense of different aspects. 3. Detect Pattern: the person usually finds what he was looking for and gains new knowledge by finding trends, frequencies, and structures in the dataset. 4. Match Mental Model: metaphors enable an effective mapping of data by allowing people to match the visual representation of the dataset with a mental model (Rodriguez et al., 2015, p.8)

Selon Dykes (2020), la compréhension est un changement inattendu dans nos connaissances et elle se produit lorsque nous analysons des données. Un seul insight peut déboucher sur de nouvelles possibilités ayant un

impact sur une grande variété d'activités. Les insights, idéalement, ne font pas que changer notre façon de penser, mais ils nous inspirent à faire les choses différemment. Ils convertissent les données en une direction nous menant alors vers des lieux nouveaux et imprévus : « These unexpected shifts in our knowledge can occur as we analyze and examine data. For example, we may uncover a new relationship, pattern, trend, or anomaly in the data that reshapes how we view things » (Dykes 2020, p.8). Ils sont également susceptibles d'affecter les gens de différentes manières, en ce qui concerne leurs convictions, leur façon de travailler ou leurs priorités.

Dykes met également en avant que bien communiquer les insights permet d'amorcer une action et un changement. « [Key insights] can be true game changers but only if we know how to share them effectively with the people who will decide their fate and help make them a reality » (Dykes 2020, p. 10). Il rend par ailleurs attentif quant à la qualité de leur communication, qui elle a un impact sur l'action. « [Unless] we can improve the communication of these insights we will also see a poorer insight-to-value conversion rate. If an insight isn't understood and isn't compelling, no one will act on it and *no change will occur* » (Dykes 2016).

Quatre objectifs sont avancés pour communiquer l'insight :

1. L'insight doit capter et retenir l'attention de l'audience. « If your insight fails to capture an audience's attention, in its limited window of opportunity, other competing information and distractions will draw their focus, and your insight will go nowhere » (Dykes 2020, p.35).

2. Le public doit comprendre l'insight, car dans le cas contraire, cela peut mener à une mauvaise décision ou un résultat erroné. « When your idea is concrete and clear in the minds of your audience members, you have a shared understanding to build on - whether or not they agree with your findings » (Dykes 2020, p36).

3. L'audience doit se souvenir de l'insight et pour cela il doit résonner suffisamment longtemps. « Narrative will be critical to ensuring your insights take root in the minds of your audience and are not easily forgotten and lost » (Dykes 2020, p.37).

4. Le public doit agir suite à la communication de l'insight. « While it is nice to have an insight understood and remembered, it won't drive any value if it never leads to any form of action. Your data communication must be persuasive if it's going to inspire or motivate people to act » (Dykes 2020, p.37).

Si l'on est déterminé à faire comprendre ses idées et à y donner suite, il faut ainsi passer de la simple information à la communication. L'objectif de l'information est de s'assurer que l'information est reçue, tandis que celui la communication est de s'assurer que le public comprenne la signification des données (Dykes 2020).

La data story est une étape d'un long processus d'analyse qui permet de trouver les bons insights. Il faut d'abord analyser les données afin d'identifier les informations à communiquer. Il s'agit d'un processus qui prévoit en premier lieu le *storyframing*, qui est la sélection des informations, puis il y a le *storytelling* où sont mises en avant seulement ½ informations à présenter en profondeur (Dykes 2020).

2.2. Narration

Afin de mettre en place une communication des insights efficace, il faut ainsi faire appel à la narration et au récit :

> « Storytelling can further amplify the power of your numbers, providing an engaging narrative that connects the dots for your audience and compels them to act. When you craft your data insight into a data story, you have a powerful vehicle for conveying meaning, engaging your audience and driving change » (Dykes 2020, pp. 15-16).

Dykes (2020) identifie cinq attributs clé placés sur un continuum qui aident à comprendre si et quand le storytelling a du sens.

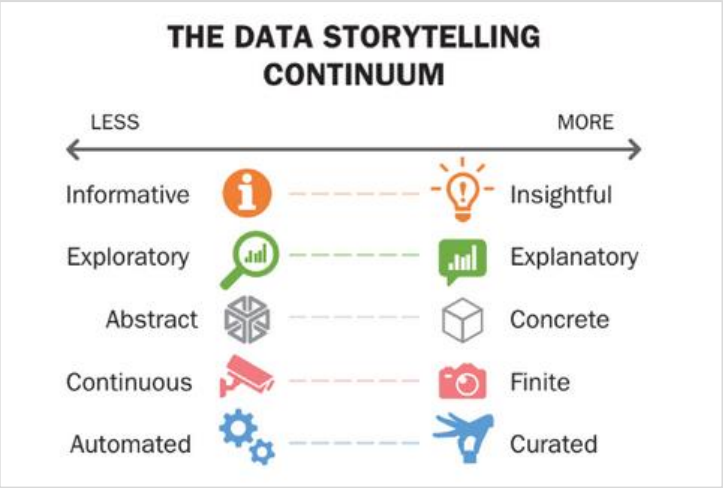


Figure 7: Data Storytelling Continuum (Dykes 2020)

Plus on tend vers la partie droite du continuum et plus le storytelling fait sens. Les attributs clé sont définis ainsi (Dykes 2020) :

Tableau 1: Attributs clé pour envisager le storytelling

1	Informative	L'informatif fournit une information intéressante et utile.
	Insightful	Insightful va au-delà de l'informatif et démontre une perception approfondie.
2	Exploratory	En proposant une approche exploratoire (par exemple des data visualisations interactives), cela permet à l'audience de se faire ses propres insight en explorant les jeux de données choisis pour elle.
	Explanatory	« When you have particular insight in mind it's much easier to explain – in a form of a data story – what your insight is and why it's important » (p.84).
3	Abstract	« When you keep the data more abstract, it frees up the data to be interpreted in a variety of ways [but] you forego the ability to tell a specific data story » (p.84).
	Concrete	« [When]s the insights […] are more concrete and specific, it is much easier to build a data story because you're highlighting a particular view of the data » (p.84).
4	Continuous	Les tableaux de bord automatisés s'actualisent constamment, par conséquent les résultats intéressants fluctuent (vont et viennent), ce qui rend la création d'histoire plus difficile. Il faut ainsi capturer un instantané des données pour pouvoir examiner un insight.
	Finite	
5	Automated	« [The] automated data compositions attempt to display information in meaningful ways, [but] they can often miss or not fully comprehend the significance of certain insights » (p.85).

	Curated	« In order to bring out a story and tailor it for a specific audience, the data often needs to be curated by human hands […]. While technology can easily handle the automated detection of anomalies in the data as well as the automated dissemination of information, it can struggle with the identification and communication of truly meaningful insights. Skilled human intervention is often required to assemble and weave together meaningful visual narrative for important insights » (p.85).

2.2.1. Récit

Dans sa « Poétique », le philosophe Aristote explique qu'un bon récit doit avoir un début, un milieu et une fin. Il s'agit ici d'une structure narrative en trois actes, ou trois temps. Le premier acte met en place l'histoire et présente le protagoniste qui rencontre une difficulté. Le deuxième acte plonge le protagoniste encore plus dans ses mésaventures. Enfin, le troisième acte donne à voir la libération du protagoniste grâce à la solution trouvée.

Selon Nussbaumer Knaflic (2015), les éléments du storytelling, les trois temps, peuvent aider un public cible à comprendre le point de vue qui est énoncé et prendre une décision. Il est possible ainsi de raconter une histoire basée sur les données.

Le début de l'histoire, l'acte 1, présente clairement la situation la situation actuelle, à savoir un problème à résoudre ou une opportunité à poursuivre. Le milieu, l'acte 2, révèle le conflit central et des symptômes mesurables sont à changer. Enfin, la résolution, l'acte 3, propose de résoudre les éléments révélés dans l'acte 2 et de créer un résultat meilleur pour l'avenir (Nussbaumer Knaflic 2015).

Duarte (2019) met elle aussi en avant ces trois actes (situation, complication, résolution) appliqués à un résumé exécutif pour que ce dernier soit engageant et mémorable :

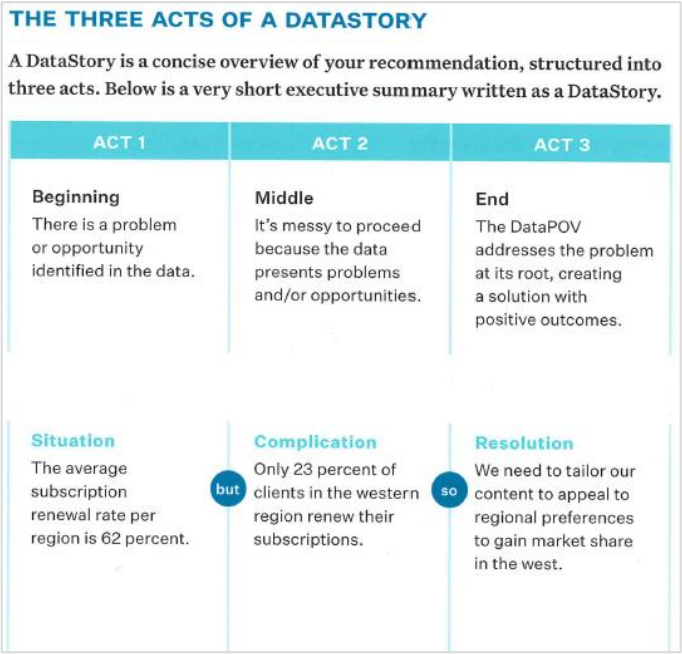


Figure 8: The Three Acts of a DataStory (Duarte 2019)

Cependant, Dykes (2020) affirme que cette structure pensée par Aristote ne suffit pas, car elle ne présente pas suffisamment d'étapes pour construire une Data Story, et s'inspire de la pyramide de Freytag pour proposer son modèle de *Data Story Arc* constitué de quatre étapes. Amini et al. (2015) empruntent également à Freytag cette structure. Elle se distingue du modèle d'Aristote par l'introduction d'un nouvel et quatrième

élément, le *Inciting*, qui est « a major point in the plot where it transitions from the introduction to the beginning of the action or conflict » (Dykes 2020, p. 161).

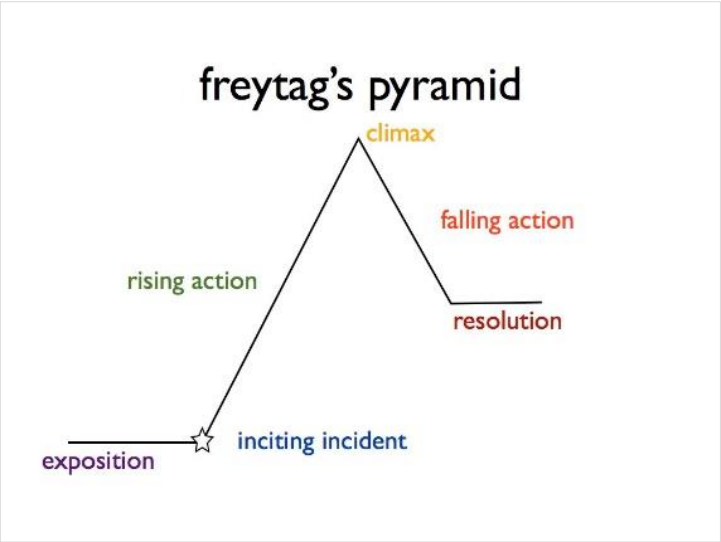


Figure 9: La pyramide de Freitag (Writers 2020)

Toujours d'après Dykes (2020), si dans un roman l'écrivain combine différents « plot points » (éléments de l'intrigue) pour faire avancer l'histoire, dans une Data story il faut assembler des data points. Ces derniers, ces « story points », sont une série de données clés qui viennent de résultats clés ayant attiré l'attention du storyteller au moment de la phase exploratoire.

Le Hero's Journey, ou le concept du monomythe, est une autre structure qui émerge des écrits. Il est développé par Joseph Campbell dans ses livres et ses conférences à partir de la fin des années 1940. Il avance l'idée que tous les mythes du monde, de l'Antiquité à nos jours, racontent essentiellement la même histoire, dont ils ne seraient que des variations, à savoir qu'un héros se retrouve projeté dans une aventure, un périple jalonné de défis qu'ils surpassent, et de laquelle il revient vainqueur. Selon Dykes (2020), ce modèle circulaire, composé de 17 différentes étapes réparties en trois grandes catégories (départ, initiation et retour), est cependant trop compliqué pour assembler des Data Stories.

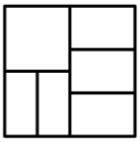
2.3. Visuel

Choisir le visuel approprié dépend du besoin et des données et il doit être facilement et rapidement lisible par le public auquel il s'adresse. Il existe des graphiques plus adaptés pour l'exploration, mais ils obscurcissent le point principal, tandis que d'autres sont plus adaptés pour l'explication et tout le monde les comprend (Nussbaumer Knaflic 2015).

Nussbaumer Knaflic (2015) met en avant 12 types de visuels communément utilisés et qui répondent à la majorité des besoins :

Tableau 2: Visuels utilisés pour les Data Stories

Visuels	
Tt	Quand il n'y a que un nombre ou deux à partager, le faire simplement avec du texte peut être le meilleur moyen de le communiquer.

	Un tableau est utile pour communiquer plusieurs unités de mesures différentes.
	Une carte thermique permet de visualiser les données sous forme de tableau mais en colorant des cellules, cela transmet l'importance des nombres.
	Le nuage de point permet de montrer la relation entre deux éléments, car les données sont encodées simultanément sur l'axe x et sur l'axe y.
	La ligne et le slopegraph sont utiles pour représenter des données continues.
	Les barres permettent facilement de voir rapidement quelle catégorie est la plus grande, par exemple, ou la plus petite, et également la différence incrémentale entre les catégories. Il est important que la base de référence soit à 0 (que l'axe x et l'axe y se croisent à la valeur 0). Les barres verticales ou horizontales empilées comparent les totaux entre les catégories et leurs sous-éléments. Les waterfall charts sont utilisés pour séparer les éléments et se concentrer sur un à la fois, ou pour afficher un point de départ, des augmentations et des diminutions, ainsi que le point final résultant.
	Le square area est utile pour visualiser des nombres de grandeurs très différentes et cela de manière plus compacte. Il est cependant compliqué pour l'œil humain d'attribuer une valeur quantitative à un espace à deux dimensions, ce qui peut le rendre moins lisible.

Certains visuels seraient par ailleurs à éviter :

- Secteurs (camemberts): la littérature émet des avis divers quant aux secteurs. Pour Nussbaumer Knaflic (2015), l'œil humain n'est pas doué pour attribuer une valeur quantitative à un espace bidimensionnel. Si les parts ont une dimension proche, il est peu aisé de savoir laquelle est plus petite ou plus grande et de combien. Cependant, pour Duarte (2019), les secteurs devraient être utilisés car tout le monde les traite facilement et les comprend.
- Anneaux (beignets) : il est difficile d'attribuer une valeur quantitative à une longueur d'arc et de la comparer aux autres (Nussbaumer Knaflic 2015).
- 3D : la perspective affichée trompe l'œil et fausse les chiffres, elle les rend difficiles ou impossibles à interpréter ou à comparer (Nussbaumer Knaflic 2015).

- Axe vertical (y) secondaire : il faut du temps et de la lecture pour comprendre quelles données doivent être lues par rapport à quel axe. Cela peut impliquer une relation qui peut ou non exister (Nussbaumer Knaflic 2015).

De plus, lorsqu'il s'agit de choisir les visuels efficaces à utiliser, et plus précisément en termes de design, il est essentiel de réfléchir à ce qu'ils représentent/signifient et au fait qu'ils peuvent mener à des interprétations différentes (Feigenbaum, Almalhodaei, 2020). Il existe cependant également des principes universels de design pouvant servir de base à la création de visuels. Feigenbaum et Almalhodaei (2020) identifient ainsi 4 piliers dans les principes universels de design qui répondent à la majorité de leurs besoins. Ils sont recensés ci-dessous :

Tableau 3: 4 principes universel de design

4 principes universels de design	
Symboles	« Symbols should be resonant. When something is resonant it relates on an embodied level. One responds to it instantly […]. Ideally, not only would they comprehend your symbol, they would also connect with it, opening up space for meanings to become stickier » (p.128).
	« Symbols should be relevant. For symbols to resonate they need to be relevant. This means that the symbols you create or choose to use should be the right match for your audience […]. [Asking] contextual questions about the connotation of symbols can help you consider their cultural and social relevance to your audience » (p.128).
Couleur	Les couleurs doivent elles aussi résonner auprès de l'audience et être pertinentes pour être efficaces. L'interprétation de la signification des couleurs est tributaire de la culture. Il est ainsi primordial de comprendre les coutumes, les besoins et les attitudes du public cible ; de connaître les perceptions culturelles de la couleur.
	« Creating Colour Contrast. When selecting colours, keep in mind how contrast can be used to highlight significance or show difference. Contrast can be created through your choices of colour hue (type of colour), colour value (lightness or darkness of a colour), and colour saturation (purity of a colour) » (p.130).
	« Colour metaphors and gradients […]. Here colour is used to represent scales or increases and decreases in values. Colour metaphors refer to the use of different colours that carry particular meanings or connotations in a shared cultural context (i.e. green for go, red for stop). A colour gradient refers to a range of colour transitions that is often gradual and blends together as it progresses (i.e. from blue for cold to red for hot) » (p.132).
Légende (Caption)	« As the same visual image can mean different things to different people, we need to guide viewers to the intended message of our data story […]. This means that we need to limit the possible meanings a visual image can have. Captions work to guide audiences toward your Think, Feel, and Do. By crafting carefully worded text that anchors the images in your data story, you can stop viewer's interpretations drifting away from you intended messages » (p.135).
Editorial	« Editorial layout refers to how the different visual elements of your data story are strategically arranged onto a page, screen, poster, or whatever platform you are using in order to effectively and efficiently reach your target audience. The

	layout you choose should work well to get across your intended message […]. How many elements you choose to include, your colour scheme, the size and detail of your symbols, and other illustrated elements all depend on the amount of space you have to work with » (p.136).

Il est essentiel de porter son attention sur la valeur informative des éléments que l'on souhaiterait utiliser ou ajouter. En effet, « identifying anything that isn't adding informative value, or isn't adding *enough* informative value to make up for its presence, and remove those things » (Nussbaumer Knaflic 2015, p. 71).

Pour ce faire, il faut prendre en compte les éléments suivants :

Tableau 4: Eléments utiles à la valeur informative

ELEMENTS UTILES A LA VALEUR INFORMATIVE	
Charge cognitive	Il s'agit de l'effort mental nécessaire pour apprendre de nouvelles informations et il faut éviter de la solliciter si elle n'aide pas l'audience à comprendre l'information. « What matters most when it comes to our visual communications is the perceived cognitive load on the part of our audience: how hard they believe they are going to have to work to get the information out of you communication » (Nussbaumer Knaflic 2015, p.73).
Clutter (encombrement)	Il s'agit d'éléments qui prennent de la place sans accroître la compréhension et donnent l'impression que le visuel est plus compliqué. Le risque étant ainsi que l'audience s'en détourne (Nussbaumer Knaflic 2015).
Gestalt	Cette théorie est basée sur l'idée que « the human brain will attempt to simplify and organize complex images or designs that consist of many elements, by subconsciously arranging the parts into an organized system that creates a whole, rather than just a series of disparate elements » (Chapman 2021). Elle présente divers principes qui sont utiles lors de l'élaboration d'un visuel.
Ordre visuel	Il s'agit de la conception/structure générale et l'emplacement et/ou agencement des éléments, en particulier l'alignement des éléments et l'utilisation d'espaces blancs, à savoir des espaces vides, qui peuvent être utilisés pour attirer l'attention sur les autres éléments (Nussbaumer Knaflic 2015). « [Blank] spaces can be deliberately introduced into a data visualisation in order to tell a more compelling story about the data. Blank spaces can be used to heighten tension, call attention to an issue, or to show gaps between two or more amounts » (Feigenbaum, Almalhodaei, 2020, p.159)
Contraste	Il s'agit de la différence/différenciation dans les éléments présentés. Le contraste aide l'audience à savoir où elle doit porter son attention. Plus les éléments sont différenciés et moins ils se distinguent les uns des autres. Il faut donc rendre différent du reste ce que l'on souhaite mettre en avant (Nussbaumer Knaflic 2015).

3. Typologie de Data Stories

La littérature nous a permis de mettre en avant 5 genres de Data Stories, qui combinent donc données, visualisation et une narration pour communiquer un message clair et susciter une prise de position auprès de l'audience cible.

Tableau 5: Genres de Data Stories

Genres[1]	
	Magazine / article
	Poster
	Bande dessinée
	Diaporama
	Film, Vidéo, Animation

Il s'agit ici de genres qui s'ancrent plutôt dans la forme dans laquelle la Data Story est présentée et dans le support ou médium utilisé. On a ainsi des Data Story sous forme d'article de journal, poster, bande dessinée (ou Comics), diaporama et enfin images en mouvement, comme un film, une vidéo ou encore une animation.

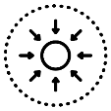
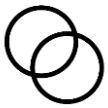
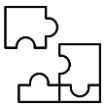
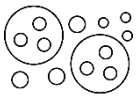
Selon Segel et Heer (2010) certains genres fonctionnent mieux selon le type d'histoire. Le choix adéquat dépend donc de divers facteurs « including the complexity of data, the complexity of the story, the intended audience, and the intended medium » (Segel, Heer 2010, pp. 1145-1146). Pour Feigenbaum et Almalhodaei, ces genres peuvent devenir hybrides et aucun n'est voué qu'à une seule audience:

> « [These] genres can me mixed, and while it may be more traditional to see an annotated chart in a boardroom and a comic strip in a classroom, there are no rules on what genre is necessarily best for your audience. What matters is how each of these different styles of layout supports some narrative structures, while limiting others » (Feigenbaum, Almalhodaei, 2020, p.138).

Ben Jones (2015) évoque lui sept types de Data Stoies, alors que Dykes (2020) mentionne différents types de « story points », qui sont similaires à ceux de Ben Jones (2015), mais il en ajoute deux autres. Ils sont tous regroupés dans le tableau suivant :

[1] Les icônes proviennent du site Web The Noun Project, voir les références dans la section Icônes.

Tableau 6: Typologies de Data Stories

Typologie (Dictus)[2]		
	Change over time	Révèle les changements au fil du temps.
	Drill down	Passe du général au spécifique.
	Zoom out	Passe du spécifique au général.
	Contrast	Compare les différences.
	Intersection	Révèle le(s) point(s) de croisement.
	Factors	Assemble divers facteurs, additionne différents facteurs.
	Outlier	Révèle les cas particuliers, les exceptions.
	Project forward	Montre ce qui est prévu dans le futur.
	Cluster	Révèle une concentration ou une distribution des résultats au sein d'un ensemble de données.

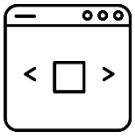
Il s'agit ici plutôt du style d'histoire ou de Dicti de la Data Story, ce que l'on cherche à montrer, l'histoire que l'on veut raconter. Ces types sont particulièrement ancrés dans les données et l'insight qui en découle. On a ainsi neuf typologies ou Dicti qui répondent chacune à un objectif précis, comme démontrer un changement à travers le temps, faire la comparaison des différences ou encore révéler un cas particulier ou une exception.

Il existerait par ailleurs, selon Segel et Heer (2010), deux formes délimitant un spectre sur lequel viennent se placer les Data Stories, ou *Narrative Visualisation*, pour employer leur terme. La forme « author-driven » qui

[2] Les icônes proviennent du site Web The Noun Project, voir les références dans la section Icônes.

se veut une séquence linéaire, avec un message fort et aucune interactivité. A son opposé, pour la « reader-driven », il n'y a aucune séquence élaborée, aucun message et elle offre de l'interactivité. Ces structures encadrent et dessinent ainsi le chemin qu'emprunte le destinataire de la Data Story.

Tableau 7: Structures de Data Stories

Structure[3]		
	Martini Glass	Elle débute avec une approche « Author-Driven », qui peut prendre la forme de questions, d'observations, ou encore d'un article. Puis, une fois que l'utilisateur a suivi le chemin pavé pour lui, une fois que le récit voulu par l'auteur est complété (la tige du verre de Martini), intervient l'approche « Reader –Driven ». L'utilisateur est ici libre d'explorer de façon interactive les données à travers les divers chemins offerts par l'interactivité (le réceptacle du verre).
	Interactive Slidewhow ou Listicle	Elle ressemble à ce que l'on connaît du diaporama, qui est lui orienté « Author-Driven », à la différence qu'il est ici construit avec une interactivité en milieu de récit. L'utilisateur peut ainsi, dans une approche « Reader-Driven » placée donc au milieu du parcours, explorer plus amplement des points particuliers avant de poursuivre à l'étape suivante de l'histoire racontée.
	Drill Down Story	Elle est plus axée « Reader-Driven ». Elle présente une thématique générale puis permet à l'utilisateur de sélectionner des éléments en particulier qui lui donneront davantage de détails et de précisions. L'utilisateur décide ainsi ce qui est raconté et quand. Segel et Heer (2010) précise cependant que cette structure implique tout de même une grande part d'élaboration, étant donné les diverses histoires à envisager et préparer que les interactions de l'utilisateur impliquent.

3.1. Trois prototypes de Data Stories

Pour l'élaboration des différents modèles ou prototypes de Data Stories, il a fallu prendre en compte divers éléments, tel le destinataire de la Data Story ou le support/medium utilisé. Ce qui implique notamment de devoir modifier à la fois les aspects visuels et la manière dont la Data Story est racontée.

Dans le cadre du projet MIDAS, dans lequel le présent projet s'inscrit, des Data Stories pour des publics différents devront être créés, à savoir pour les décideurs/politiques, les enfants et le grand public. Cet objectif a ainsi délimité le cadre de l'élaboration de prototypes.

[3] Les icônes proviennent du site Web The Noun Project, voir les références dans la section Icônes.

3.1.1. Annotated Chart

L'*annotated chart* est un visuel qui combine du texte avec des graphiques.

L'*annotated chart* est classé, selon notre classification et typologie de visualisation, sous la catégorie *storytelling with data*, mais parce que ses caractéristiques peuvent aussi se rapprocher de la Data Story, à savoir qu'il peut aussi impliquer une prise de position de la part du lecteur, il a semblé intéressant de le sélectionner également pour représenter un prototype.

Le public cible se veut ici le grand public, mais certains retours préliminaires nous laissent penser qu'il serait vraisemblablement plutôt adapté à un public familier des données et de ce genre de visualisation. L'objectif sera par conséquent soit de l'adapter pour qu'il convienne à un public plus large, soit d'envisager un autre genre, notamment le poster, peut-être plus adapté pour ce public.

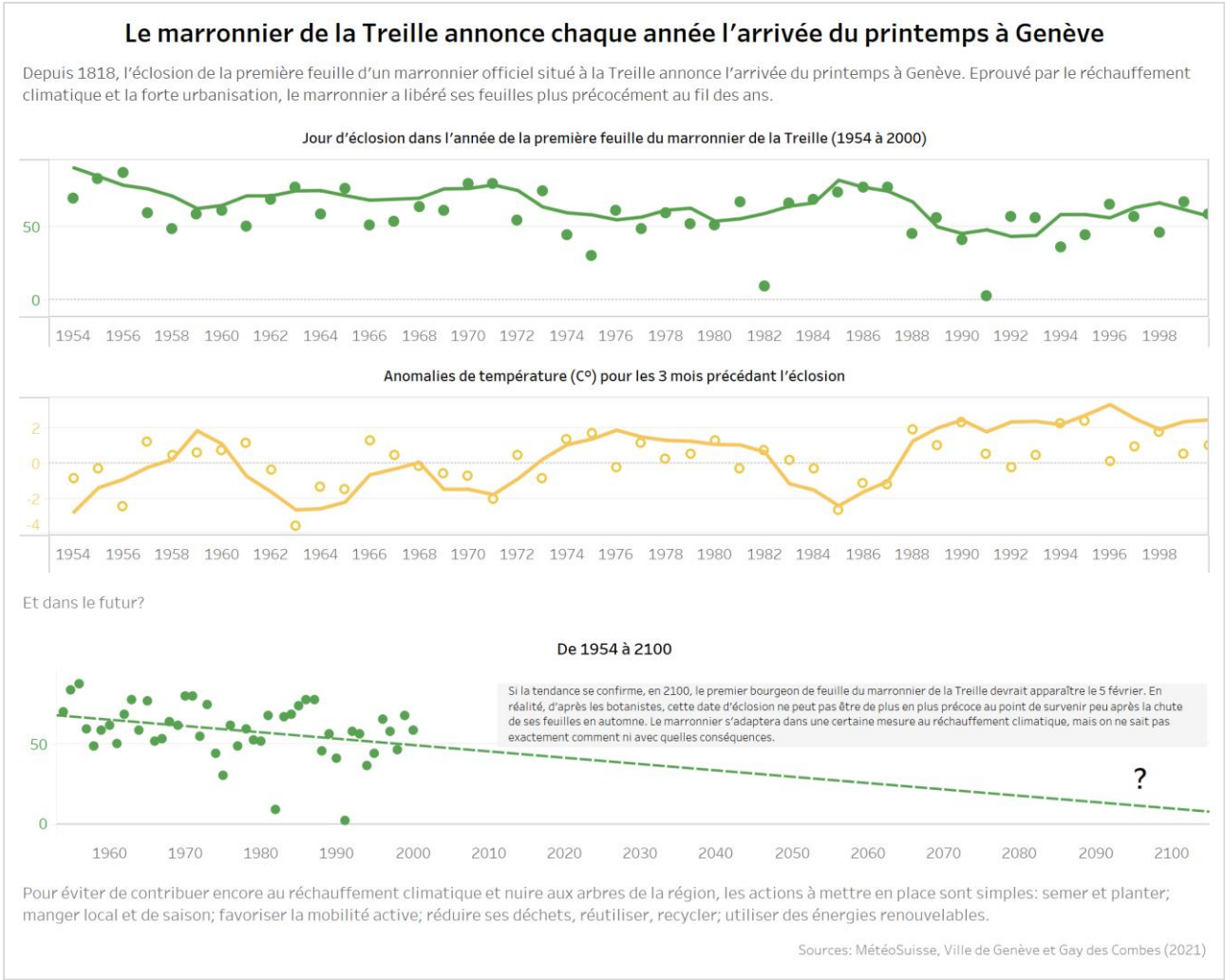


Figure 10: Prototype Annotated Chart (HEG 2022)

3.1.2. Comics

Les Data Comics, ou Comics, sont les bandes dessinées, une forme qui combine du texte (réduit) et des dessins.

Le public cible est ici le jeune public, à savoir les enfants ou adolescents. Le discours doit ainsi lui être adapté, par le vocabulaire, la formulation et le visuel, ainsi que le « call to action », c'est-à-dire la proposition de solutions qui soient applicables par ce public.

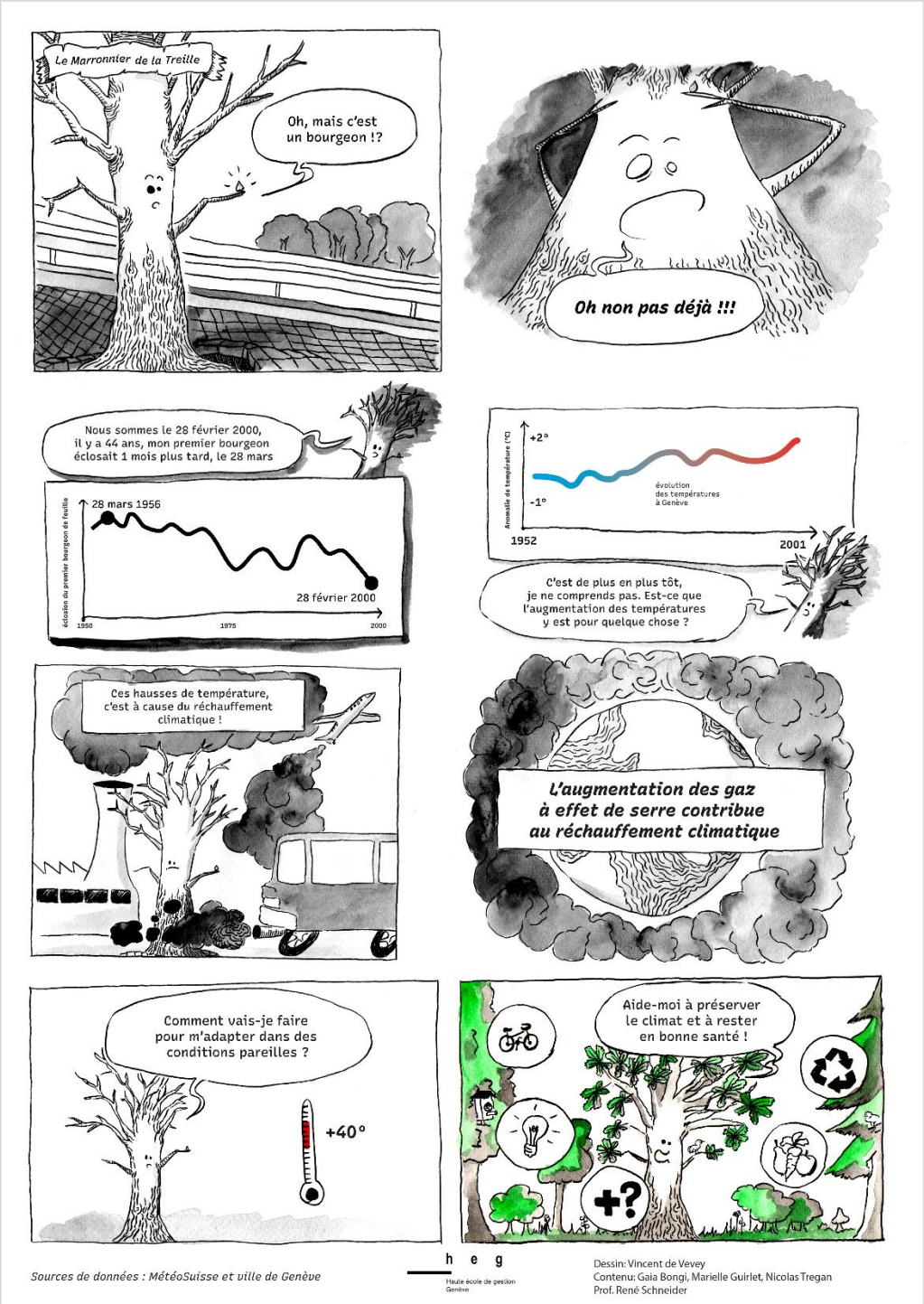


Figure 11: Prototype Data Comics (HEG 2022)

Les adultes étant aussi des lecteurs de bandes dessinées, ce prototype pourrait également être destiné aux adultes qui ne seraient pas sensibilisés à ce thème. Pour ceux qui y seraient déjà sensibilisés, il s'agira alors de modifier la bande dessinée pour qu'elle leur soit adaptée, notamment les éléments de langage et les visuels.

3.1.3. Diaporama – Présentation orale

La présentation orale, avec comme support le diaporama, consiste à parler devant une audience avec en arrière-plan un support pointant les éléments qui sont expliqués verbalement.

Le public cible est représenté ici par les décideurs ou les politiciens. La présentation orale requiert de prendre en compte divers éléments : le ton de la voix, le rythme de la présentation, ou encore la structure de la présentation. En effet, Nussbaumer Knaflic (2015) et Duarte (2019) mentionnent, lorsque l'on effectue une présentation, d'exprimer ce qui est ressenti sur les données, par le vocabulaire, par des sons, interjections ou encore des pauses.

La dimension temporelle est également essentielle et il s'agit de l'exploiter (Nussbaumer Knalfic 2015, Duarte 2019), par exemple ne pas montrer toutes les données d'un coup, ou alors de commencer par la fin, mais user des effets de surprise quand sont révélées les données cachées ainsi que de suspens quand l'histoire est racontée.

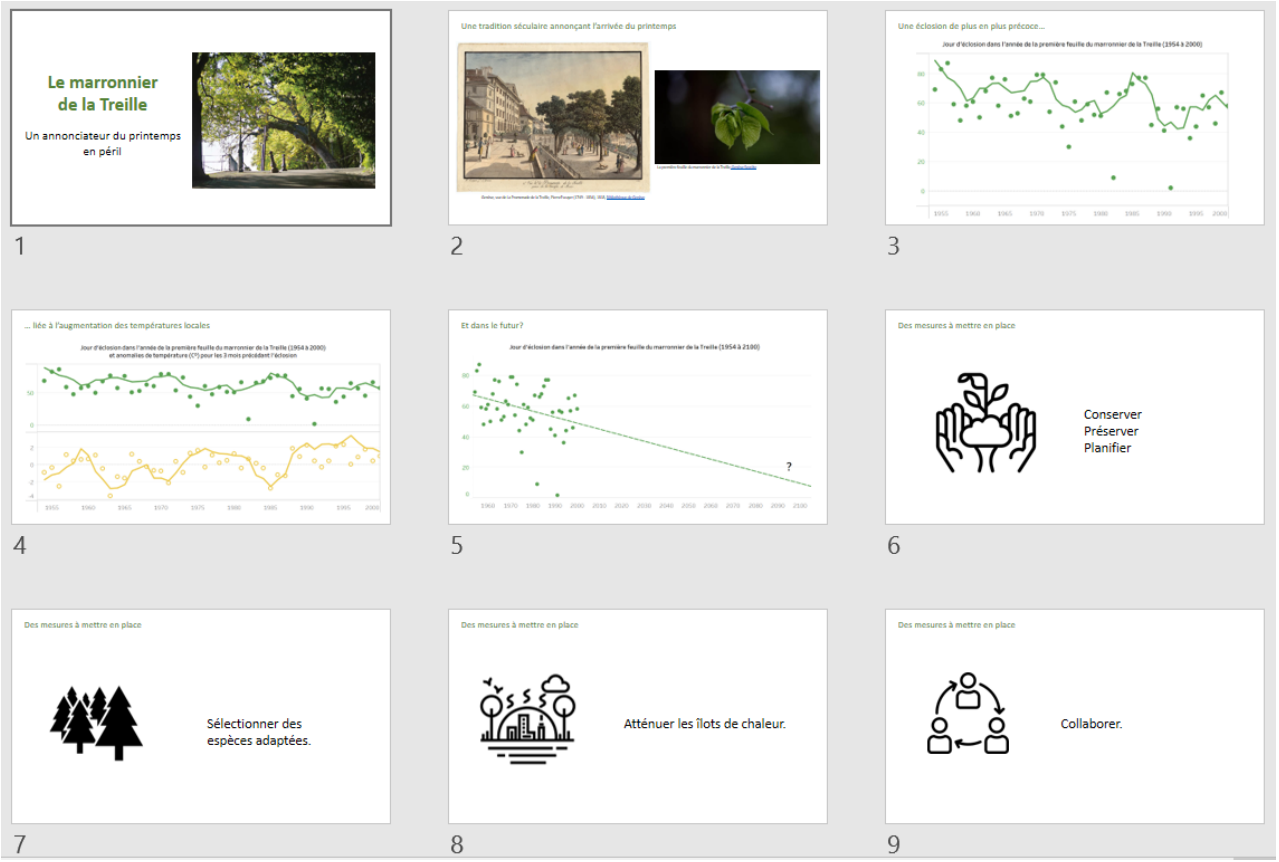


Figure 12: Prototype Diaporama[4] (HEG 2022)

[4] Les icônes proviennent des sites Web The Noun Project et Flaticon, voir les références dans la section Icônes.

Bibliographie

AANDERUD, Tricia, 2018. The Most Effective Data Storytellers Persuade, Not Inform. *Toward Sata Science* [en ligne]. 4 février 2018. [Consulté le 9 décembre 2021]. Disponible à l'adresse: https://towardsdatascience.com/should-data-stories-inform-or-persuade-2753330f3f37

AMINI, Fereshteh, et al., 2015. Understanding Data Videos : Looking at Narrative Visualization through the Cinematography Lens. *CHI* [en ligne]. 18-23 avril 2015. CHI 2015. [Consulté le 12.02.2021]. Disponible à l'adresse : https://dl.acm.org/doi/pdf/10.1145/2702123.2702431 [accès par abonnement]

CHAPMAN, Cameron, 2021. Exploring the Gestalt Principles of Design. *Toptal* [en ligne]. 2021. [Consulté le 21 décembre 2021]. Disponible à l'adresse : https://www.toptal.com/designers/ui/gestalt-principles-of-design

DUARTE, Nancy, 2019. Data Story. Explain Data and Inspire Action through Story. Ideapress Publishing. ISBN 978-1-940858-98-2

DYKES, Brent, 2020. Effective Data Story. How to drive change with data, narrative, and visuals. World Scientific. New Jersey : World Scientific. ISBN 978-1 11961-571-2

DYKES, Brent, 2016. Data Storytelling: The Essential Data Science Skill Everyone Needs. *Forbes* [en ligne]. 31 mai 2016. [Consulté le 26.11.2021]. Disponible à l'adresse: https://www.forbes.com/sites/brentdykes/2016/03/31/data-storytelling-the-essential-data-science-skill-everyone-needs/?sh=39bba92752ad

FEIGENBAUM, Anna, ALAMALHODAEI, Aria, 2020. *The Data Storytelling Workbook.* Oxon et New York: Routledge. ISBN 978-1-138-05211-6

JONES, Ben, 2015. Seven Data Story Types [enregistrement video]. *YouTube* [en ligne]. [Consulté le 10.12.2020]. Disponible à l'adresse: https://www.youtube.com/watch?v=sEZj-eUfbNo

KENT, LeAnna, 2017. Data Storytelling: Bringing Life to Your Data. *Elder Research* [en ligne]. 5 mai 2017. [Consulté le 10 décembre 2021]. Disponible à l'adresse : https://www.elderresearch.com/blog/data-story-telling-bringing-life-to-your-data/

NUSSBAUMER KNAFLIC, Cole, 2015. Storytelling with data. A data visualization guide for business professionals. Wiley. Hoboken : John Wiley & Sons. ISBN 978-1-119-00225-3

RODRIGUEZ, Maria Teresa, et al., 2015. Telling Stories with Data Visualization. *NHT '15: Proceedings of the 2015 Workshop on Narrative & Hypertext* [en ligne]. 1er septembre 2015, pp. 7-11. [Consulté le 12 février 2021. Disponible à l'adresse : https://doi.org/10.1145/2804565.2804567 [accès par abonnement]

SEGEL, Edward, HEER, Jeffrey, 2010. Narrative Visualization: Telling Stories with Data. *IEEE* [en ligne]. 24 octobre 2010. [Consulté le 12 mars 2021]. Disponible à l'adresse : https://cpb-us-e1.wpmucdn.com/sites.northwestern.edu/dist/3/3481/files/2015/02/Narrative_Visualization.pdf

Icônes

p.12

« texte » de Gregror Cresnar : https://thenounproject.com/icon/text-3548104/

« dashboard » de Anconer Design : https://thenounproject.com/icon/dashboard-analytics-4616494/

« heatmap » de ProSymbols : https://thenounproject.com/icon/heat-map-2063413/

«scatterplot» de Med Marki: https://thenounproject.com/icon/scatterplot-1686085/

«line charts» de Minh Do: https://thenounproject.com/icon/line-chart-5141901/

«bar charts» de ProSymbols : https://thenounproject.com/icon/waterfall-chart-2030706/

p.13 :

«square map » de Bernd Lakenbrink : https://thenounproject.com/icon/treemap-chart-4163357/

p.15

« magazine » de Esa Adhiatma: https://thenounproject.com/icon/magazine-3837212/

« Poster » de Creative Mania: https://thenounproject.com/icon/poster-1896870/

« Comics » de Becris: https://thenounproject.com/icon/comics-3249600/

« slide show » de Muhammad Atiq: https://thenounproject.com/icon/slide-show-4516385/

« Video » de shuai tawf: https://thenounproject.com/icon/video-2824122/

p.16

« clock » de ainul muttaqin: https://thenounproject.com/icon/clock-4536240/

« specific » de WEBTECHOPS LLP: https://thenounproject.com/icon/specific-3327403/

« Zoom » de Gendis Studio: https://thenounproject.com/icon/zoom-4876531/

« difference » de Chintuza: https://thenounproject.com/icon/difference-4659424/

« intersection » de Mohamed Mb: https://thenounproject.com/icon/intersection-3681107/

« Puzzle Piece » de Amiryshakiel: https://thenounproject.com/icon/puzzle-piece-1118151/

« Outliers » de sachin modgekar: https://thenounproject.com/icon/outliers-1503810/

« forecast » de Kamin Ginkaew: https://thenounproject.com/icon/forecast-4328082/

« Clusters » de Xela Ub: https://thenounproject.com/icon/clusters-486221/

p.17

« Martini » de OB: https://thenounproject.com/icon/martini-1026256/

« Slideshow » de Linseed Studio: https://thenounproject.com/icon/slideshow-838541/

« Zoom » de Timofei Rostilov: https://thenounproject.com/icon/zoom-3535769/

p.20

« Conservation » de Chattapat: https://thenounproject.com/icon/conservation-2585921/

« forest » de Pablo Rozenberg: https://thenounproject.com/icon/forest-867264/

« Urban Heat Island » de Creative Mania: https://thenounproject.com/icon/urban-heat-island-1030202/

« Team icons » de mynamepong: https://www.flaticon.com/free-icons/team