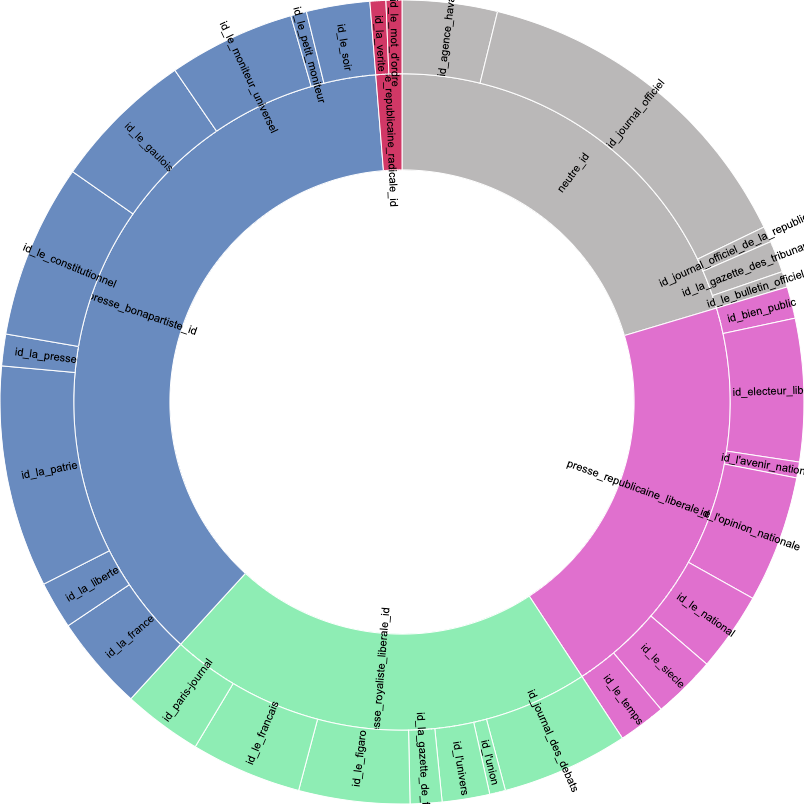
Humanités numériques

# Livrables de visualisation

L’étude qui procède est réalisé dans le cadre d’un examen de l’UE Humanités numériques. Cette dernière est élaborée autour de mon sujet de mémoire : *Le traitement journalistique d’un conflit socio-politique du Siège de Paris 1870 à la Commune de Paris 1871 par la presse provinciale : le* Journal de Chartres.

Dans le cadre de ce travail, trois outils sont exploités, à savoir : *Excel, Xmind* et *Rawgraph* dont voici un exemple ci-dessous :



*Exemple de graphique exécuté via* Rawgraphdont l’objectif de visualisation est de relever la proportion de mention des différents orientation politique dans le *Journal de Chartres.*

L’utilisation de ces trois outils se justifie par l’exploitation de mes sources écrites, à savoir les titres de presse du *Journal de Chartres*, et se fait en quatre étapes.

Néanmoins, avant de faire le catalogue des étapes, je m’attarderai sur mon sentiment envers les Humanités numériques et leur exploitation.

En effet, pour être franc, au début de l’année je ne voyais pas le rapport entre les Humanités numériques et la recherche en Histoire, surtout du point de vue de mes recherches. Bien que je comprenne l’utilité de cette science pour certaines autres. Par ignorance ou par manque de recul sur ma recherche, je ne me senti pas concerné par cette science, même si la curiosité de l’historien et l’attirance pour certains outils numériques plus ludiques ne sont pas à exclure. De plus, je dois rajouter que je ne suis pas très familier avec tout ce qui concerne les outils informatiques. Je ne dirai pas non plus que je me refuse à l’informatique mais disons que je suis plus proche de l’historien avec sa veste en tweed et son livre à la bibliothèque plutôt que de l’historien qui exécute du balisage XML de ses sources. En assistant aux cours, cela m’a permis de découvrir certains outils qui m’étaient alors inconnus. Je me suis surtout fait la remarque que je travaillais essentiellement sur des sources de presse numérisées et qu’a *fortiori* j’utilisais déjà un outil numérique. Je commençais donc à faire un parallèle avec mon étude. La question étant, quel outil utiliser ? Quel est le plus adapté à l’exploitation de mes sources et surtout quelle orientation je veux en faire ? Il est bien joli de faire des graphiques sur une carte du Monde mais encore faut-il que ce soit utile à la recherche. Fallait donc que je trouve le bon angle et la bonne question. Je pouvais faire une base de données étant donné que je travaille sur un gros corpus de sources avec le nombre d’articles que j’ai à disposition. Mais encore une fois, dans quel but et à quelle question je réponds ? Il faut se choisir le bon axe. Puis je me suis rappelé que *Le Journal de Chartres,* en tant que journal de province détenait ses informations d’autres journaux. Ces derniers variaient en fonction de leur provenance mais également de leur orientations politiques. Il est écrit dans les sources préfectorales que le journal chartrain est un journal conservateur. Une source étant ce qu’elle est, pour vérifier sa véracité quoi de mieux que le contenu du journal lui-même et ses diverses mentions. C’est ainsi que je me suis dit que je pouvais relever toutes les mentions à d’autres titres de presse et ainsi en faire une analyse. Celle-ci me permettra de conclure en fonction du nombre et de la fréquence des mentions si le journal est bien à tendance conservateur ou *a contrario* républicain progressiste. De là s’offre à moi les différents outils numériques. Je me retrouve donc devant beaucoup de données à exploiter qu’il faut organiser. En fonction des données récoltées, il est plus pertinent d’avoir recours à des outils de visualisation. Le but étant d’obtenir une vue globale des éléments et ensuite procéder à une analyse plus détaillée en fonction de l’orientation politique et de la région de publication des journaux.

Nous voilà de nouveau aux quatre étapes. J’avais une vision précise de ce que je voulais faire, par contre la manière d’y arriver était tout autre. J’ai erré entre l’utopie et la réalité. Devant l’optique de ne pas y arriver, j’ai opté pour l’option étape par étape, histoire d’être sûr de présenter quelque chose. Cependant même en dehors de l’examen, je ne comptais pas abandonner cette idée. Dans un premier temps, j’ai effectué un relevé plutôt conséquent, pour ne pas dire chronophage, de toutes les mentions de titres de journaux dans le journal sur un an. Ce relevé, je l’ai réalisé à la main sur un cahier. Ceci était déjà une première erreur. Pour optimiser mon temps, il aurait mieux fallu que ce relevé je l’exécute directement sur un tableau *Excel.* Ceci n’étant pas le cas, j’ai inscrit chaque titre de presse mentionné sur un cahier selon trois catégories : la presse nationale, la presse internationale et la presse provinciale. Après avoir réalisé cela, j’ai fait des recherches notamment sur la presse nationale sur leur orientation politique (la presse nationale étant pour la majorité politisée), sur la presse internationale pour leur provenance nationale et sur la presse internationale pour leur provenance régionale (j’ai dans ce cas divisé la France en six : Nord, Est, Sud-est, Sud-ouest, Nord-ouest et la région centre).

Dans un deuxième temps, après avoir récolté ce catalogue, j’ai choisi de travailler sur l’outil de visualisation *Xmind*. Ce dernier est un outil très intuitif et assez facile à prendre en main. Il dispose d’une version d’essai gratuite, son installation est rapide et quand on dispose d’un Mac un petit icône se glisse dans la barre de tâches du haut et en clic on est sur *Xmind.* Ce dernier est aussi exploitable en application smartphone. Hors l’aspect pratique, j’ai fait ce choix dans le but d’organiser et d’avoir une vue plus épurée sur mon catalogue de presse. *Xmind* est un outil de carte mentale très utile quand on dispose d’une base de données épistolaire. Il aide l’utilisateur à trouver des points d’intérêts de données ou même des manquements. Pour autant, ce n’est pas dans un but de faire des correspondances entre chaque donnée que j’ai opté pour celui-ci. Je pourrais trouver des correspondances entre par exemple un journal nationale royaliste de Paris avec un journal royaliste de Londres mais cela n’aurait guère d’utilité pour mon objectif. La carte mentale me sert surtout dans son utilité première de visualisation et de vue d’ensemble. Il est peut-être erroné de parler de visualisation pour *Xmind* qui est en réalité un outil de modèle conceptuel. Pour être honnête je n’ai pas vraiment compris la distinction entre les deux, étant donnée la finalité de *Xmind* qui est d’obtenir une vue plus claire d’une base de données. Avec *Xmind*, j’ai créé trois cartes mentales assez similaires. Les trois sont organisées selon le même ordre, du sujet central du *Journal de* Chartres auquel on ajoute trois branches : presse nationale, presse provinciale et presse internationale. Ces trois dernières branches sont divisées en sous branches selon les orientations politiques pour la presse nationale, selon les régions de France pour la presse provinciale et selon les pays pour la presse internationale. En ce qui ce qui concerne les deuxième et troisième cartes mentales, j’au seulement rajouté pour chaque titre le nombre d’apparition dans le journal eurélien que j’ai esthétisé différemment entre les deux pour une meilleure lecture. Je me suis rappelé pendant le cours qu’il était important de mettre un *id* sur chaque table ou données ce sur quoi on travaille. Cependant, je pense n’avoir pas tout bien assimilé. Mon utilisation de *id* est, je pense, mauvaise ou du moins pas très efficace. Je me suis souvenu que l’*id\_* se met devant le nom et si un nom peut se retrouver dans deux colonnes différentes ou en lien avec deux tables différentes, il fallait après son nom mettre *id\_.* Il me semblait aussi que l’identification servait surtout pour l’outil *MySql Workbench.* Néanmoins, pour limiter les risques, j’ai utilisé ce système d’*id* dans l’optique de transférer mon fichier dans un tableur de visualisation graphique tel que *Tableau Public* ou encore *Rawgraph.* A la place de *Xmind*, j’aurai pu choisir d’autres modèles conceptuels comme nommé ci-dessus *MySql Workbench* ou bien *Lucid chart* ou encore *Bigtable* de *Google.* Afin de justifier mes choix : le premier car il est plus difficile d’accès que *Xmind* et selon moi moins optimisé, le second tout simplement car il est payant dès trois essais ce qui limite grandement son utilisation et le dernier car je ne le connais pas du tout (faible argument je le conçois). Avec *Xmind*, *il faut* se méfier de ne pas surcharger d’informations la carte mentale ainsi que les liens entre chaque table pour ne pas que cela devienne illisible. Pour un outil destiné à une visualisation des données ce serait un comble. J’ai également rencontré un problème et de taille. En effectuant mon travail sur cet outil, j’avais pour objectif de le transférer sur une application graphique par la suite. Cependant, en période d’essai il est impossible d’enregistrer un document en format *.xlsx*. Contretemps et difficulté qui m’ont amené à envisager l’idée de travailler sur un autre outil qu’est *Excel.* Avant de développer la deuxième étape, je vous présente ci-dessous, dans l’ordre, les trois cartes mentales effectuées grâce à *Xmind.* Étant limité dans son format, en termes de taille que de lisibilité, je joindrai dans le dossier complet des copies de mes différents travaux.







Dans un troisième temps et parce que j’ai été contraint, je me suis donc rabattu sur le tableau *Excel* qui bien qu’il soit plus courant reste un outil numérique et jouit d’une grande polyvalence. Malheureusement, je ne maîtrise pas assez *Excel* pour l’exploiter dans son entièreté et qui je pense aurait pu m’exempter la quatrième étape. Il est vrai que le tableur de *Microsoft* permet également de produire des tableaux graphiques. Dès le départ, j’avais dans l’idée d’établir un tableau de stockage orienté colonnes. En effet, je disposais de cinq tables que sont : titres de presse, région, lieu, orientation politique et fréquence. Toujours avec cette notion d’*id* je voulais réaliser plusieurs tableaux qui seront reliés entre eux par l’*id*. Néanmoins, la complexité et la charge de travail en plus m’ont assez vite dissuadé (surtout que j’avais déjà du retard) de continuer dans cette voix. J’ai maintenu l’idée de l’*id* mais cette fois-ci dans un tableau de stockage orienté ligne en gardant les cinq colonnes nommées ci-dessus. Surtout que je me suis rendu compte que sans l’*id* mais grâce aux colonnes il était possible d’utiliser et de réaliser des tableaux graphiques via certaines applications. J’ai réalisé deux tableaux, un premier avec tous les éléments donc presse nationale, internationale et provinciale et le second centré seulement sur la presse nationale et les orientations politiques. Je précise cela car c’est à partir de ces deux différents tableaux que je vais réaliser la visualisation finale de mon travail. J’ai également instauré un code couleur pour le deuxième tableau afin de bien différencier les orientations politiques des journaux. Code couleur qui me servira de base et que j’essaierai de respecter dans les graphiques qui suivront (même si pour des soucis de lisibilité la couleur blanche donnée à la presse royaliste a été changé en verte).

Dans un quatrième et dernier temps et grâce aux précédents travaux, j’ai atteint enfin mon objectif premier qui était de réaliser des graphiques. Dorénavant en possession d’un tableau *Excel*, je pouvais accéder aux outils de visualisation. Parmi ceux-ci, j’ai opté pour *Rawgraph.* Ce dernier est un outil gratuit, accessible sur le Web donc qui ne demande aucun téléchargement à l’inverse de *Tableau Public*. Tandis que le premier est plutôt intuitif, le second est plus complexe à exploiter. J’aurais pu penser également à *Voyant Tools* ou bien *Palladio. Voyant Tools* est plus destiné à travailler avec des documents textes plutôt que des données statistiques donc pas le plus optimisé pour mon catalogue. *Palladio* travaille avec le format *.csv* et *.tab*, je dois l’avouer, je travaille très peu avec ces deux formats et me sont trop étrangers. *Rawgraph* me paraissait le plus adapté à mes besoins et surtout à mes compétences. Il permet d’enregistrer ses résultats directement sur le site et en toute facilité, et donne la possibilité de les enregistrer en format *.png* soit au format image. L’outil de visualisation tableau graphique me permet d’exploiter au mieux mes ressources et surtout comprendre mon catalogue de données. En effet, par ses différents graphiques disponibles, l’outil facilite l’étude et l’approche de ma base de données. Je m’explique, par certains graphiques, il donne une vision plus large du catalogue et ainsi distingue mieux le poids des différents éléments. Pour cela, en allant sur le site web de *Rawgraph,* il suffit de copier le document *Excel* dans la rubrique approprié. Ensuite, une fois analysé, il ne reste plus qu’à choisir le graphique le plus adéquate pour l’étude. Un simple coup de slide vers le bas et de nombreux choix s’offrent à nous. Du fait de sa gratuité, l’outil ne dispose pas non plus d’une multitude de choix mais on peut aisément trouver son bonheur. Mon étude l’atteste car dans mon tableau *Excel*, je ne détiens qu’une catégorie chiffrée, celle des « frequence » et il est difficile de déceler un graphique dont une seule donnée chiffrée ne suffise. Après plusieurs essais en croisant les données et essayant plusieurs variantes, j’ai déniché les plus appropriés. Pour chaque graphique, il faut rentrer certaines tables dans des catégories mise à dispositions telles que : *axe, size, step, colors*… Le choix de la table dans la catégorie est important car elle oriente le graphique. Il faut bien choisir et faire des tentatives différentes pour obtenir le meilleur rendu. Dans mon cas les graphiques proposés les plus utiles étaient : *circle packing, sunburst* et *alluvial diagram.* J’ai utilisé mes données en deux temps. En effet, j’avais deux objectifs : à savoir celui d’analyser le nombre de citations dans le *Journal de Chartres* en fonction de l’orientation politique dans la presse nationale en choisissant les tables « titre\_de\_presse », « orientation\_politique » et « frequence ». Et pour second objectif, d’étudier le nombre de citations dans le journal en fonction de la région de publication en prenant les tables « titres\_de\_presse », « region », « lieu » et « frequence ». J’ai ainsi, effectué cinq graphiques, deux pour le premier questionnement et trois pour le second. Encore une fois, pour chaque graphique, le choix de bien situer les tables dans les bonnes catégories est importante mais aussi l’utilisation d’un code couleur afin de faciliter la lisibilité des résultats. Ce n’est pas tout, à l’instar de la couleur, l’échelle aussi est importante pour la lisibilité comme pour la clarté des informations. Il ne faut pas surcharger d’informations le graphique. J’ai eu un léger problème pour ce cas-ci, j’ai un graphique où deux titres de données se superposent et malheureusement on peut changer la couleur de remplissage mais pas de la police. Le problème de couleur ne se limite pas à cela, il est vrai que sur *Rawgraph* il existe un manque réel de libertés. Ce manque de libertés concerne également le manque de certains graphiques essentiels et également l’absence de pourcentage. On ne trouve pas d’outils disponibles pour calculer des proportions. Néanmoins, je suis satisfait des résultats obtenus. Au lieu de me battre avec des chiffres dans un tableau et à faire mes calculs, cet outil m’a aidé à avoir une meilleure visibilité sur mes données et après analyse d’établir des conclusions qui affirme plus ou moins l’aspect conservateur du journal. A travers l’analyse des graphiques on remarque que la presse bonapartiste est la plus citée par le journal suivi de la presse royaliste à quasi égalité avec la presse républicaine modérée. Cela est à relativiser si on prend le nombre de mention de chaque titre à part. Dans une étude plus poussée, je pourrai aller plus loin et chercher les orientations politiques des titres de presse provinciaux et internationaux mais ceci est plus difficile à trouver et parfois pas très clair. L’interprétation des graphiques est importante mais il ne faut pas oublier que cela ne reste que des chiffres. Pour citer un exemple, chaque titre ici est mentionné dans le *Journal de Chartres* cependant on ne connait pas toujours le contexte dans lequel le titre est mentionné par le journal. Les mentions peuvent être des citations pour contredire un article ou bien en faire l’éloge, un tableau ne peut pas nous le dire. Vous trouverez ci-dessous les graphiques réalisés dans le cadre du devoir mais qui me serviront dans le cadre de mes recherches. Je conçois une fois de plus que cela ne rend pas bien visuellement sur le document *Word* pour cela ils seront également envoyés en pièces-jointes.

En guise de conclusion, je voulais dire quelques mots sur mon ressenti après cette étude. Je voulais d’abord revenir sur mes réserves par rapports aux liens entre Histoire et Humanités numériques surtout dans le cadre de la recherche en Histoire. Il est vrai que je ne comprenais pas en quoi les Humanités numériques pourraient m’aider dans ma recherche. Je voyais cela plus comme une corvée qui retarderait mes recherches en cours. Et bien je me trompais, car en plus d’avoir appris des choses, cela m’a fait au contraire avancer sur mes recherches. Cela m’a permis d’envisager d’un autre angle une question auquel je n’aurai pas pu répondre avec autant de clarté sans l’élaboration de graphiques. Graphiques que je n’aurai connus ni su utiliser avant ce cours. Bien sûr, je connaissais l’apport des Humanités numériques pour les sciences et même pour les sciences sociales telles que l’Histoire. Ce sont bien les Humanités numériques qui ont permis le développement de la numérisation et du balisage de certaines sources, outils plus qu’appréciables pour le chercheur. Les outils numériques sont parmi les acteurs majeurs de certaines avancées notamment dans l’archéologie ou encore des projets comme le *projet Venice.* De plus, ayant pour but d’être professeur d’histoire-géographie, j’envisage même de continuer d’utiliser *Xmind* dans le but de créer des cartes mentales pour les cours.

