

Mash-up de données

Les sénateurs français.

I. Délimitation du sujet

Le consigne de l'exercice était “de proposer une analyse prosopographique des sénateurs français et une analyse des groupes politiques d'appartenance via des visualisations de données”.

Dans le cadre de cet exercice, nous avons limité notre sujet aux Sénateurs de la Ve République car les fichiers disponibles sur data.senat.fr ne portaient que sur ces Sénateurs et représentaient déjà un nombre conséquent de sénateurs et sénatrices.

II. Les données

Notre travail prend pour point de départ les données mises à disposition par le Sénat. En effet, nous étions certains de leur exhaustivité et de leur fiabilité. Dans un second temps, nous les avons complétées par les données proposées par wikidata.

A. data.senat.fr

Parmi les différents jeux de données disponibles sur le site data.senat.fr, nous avons choisi de nous concentrer sur General, Histogroupes et Elusen. Nous souhaitions éviter les redondances et nous concentrer sur les thématiques de notre sujet : les sénateurs et les groupes politiques.

Le fichier General regroupe l'ensemble des dernières informations de chaque sénateur (anciens ou actuels).

Le fichier Histogroupes retrace l'historique par sénateur des groupes politiques du Sénat (anciens ou actuels).

Enfin le fichier Elusen retrace l'historique des mandats (date d'élection, de début et de fin de mandat ainsi que le motif) de chaque sénateur.

B. wikidata

Nous avons complété ce premier ensemble de données avec différentes requêtes sur wikidata.

1. Requête n°1 :

```
PREFIX wd: <http://www.wikidata.org/entity/>  
PREFIX wdt: <http://www.wikidata.org/prop/direct/>  
PREFIX wikibase: <http://wikiba.se/ontology#>  
PREFIX p: <http://www.wikidata.org/prop/>  
PREFIX ps: <http://www.wikidata.org/prop/statement/>
```

```
PREFIX pq: <http://www.wikidata.org/prop/qualifier/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX bd: <http://www.bigdata.com/rdf#>
```

```
SELECT ?id_senat ?senateur ?senateurLabel
(group_concat(DISTINCT ?lieu_naissanceLabel;separator=" | ") as ?lieux_naissance)
(group_concat(DISTINCT ?lieu_mortLabel;separator=" | ") as ?lieux_mort)
(group_concat(DISTINCT ?autres_mandatsLabel;separator=" | ") as ?les_autres_mandats)
(group_concat(DISTINCT ?occupationLabel;separator=" | ") as ?occupations)
(group_concat(DISTINCT ?educationLabel;separator=" | ") as ?etudes)
(group_concat(DISTINCT ?autre_organisationLabel;separator=" | ") as ?autres_organisations)
(group_concat(DISTINCT ?distinctionLabel;separator=" | ") as ?distinctions)
```

```
WHERE
```

```
{
  # INFORMATION INDISPENSABLE : sénateur de la 5e République avec un ID dans la
  base ODSSEN.
  # sénateur / occupe la fonction de / sénateur.
  # il y a 3 identifiants pour les sénateurs, on utilise UNION pour tous les
  interroger)
  {{?senateur wdt:P39 wd:Q14828018} UNION {?senateur wdt:P39 wd:Q15686806} UNION
  {?senateur wdt:P39 wd:Q24931583}}
```

```
  #senateur / a un identifiant dans le fichier sénat / ID
  ?senateur wdt:P1808 ?id_senat;
  # et un genre.
  wdt:P21 ?genreLabel.
```

```
  # ENRICHISSEMENT : DONNEES SUPPLEMENTAIRES. On utilise OPTIONAL pour ne pas
  supprimer les entrées qui n'auraient pas d'information sur ce champ.
```

```
  OPTIONAL {?senateur wdt:P19 ?lieu_naissance.}
  OPTIONAL {?senateur wdt:P20 ?lieu_mort.}
  OPTIONAL {?senateur wdt:P39 ?autres_mandats.}
  OPTIONAL {?senateur wdt:P106 ?occupation.}
  OPTIONAL {?senateur wdt:P69 ?education.}
  OPTIONAL {?senateur wdt:P463 ?autre_organisation.}
  OPTIONAL {?senateur wdt:P166 ?distinction.}
```

```
  #Filtre sur le type d'ID de la base sénat que l'on retient pour matcher avec la
  base ODSSEN.
```

```
  FILTER(STRSTARTS(?id_senat, "senateur/"))
```

```
SERVICE wikibase:label {
  bd:serviceParam wikibase:language "fr".
  # Bug avec l'utilisation de group_concat : il faut préciser les champs
  labellisés.
  ?senateur rdfs:label ?senateurLabel.
  ?lieu_naissance rdfs:label ?lieu_naissanceLabel.
  ?lieu_mort rdfs:label ?lieu_mortLabel.
  ?autres_mandats rdfs:label ?autres_mandatsLabel.
  ?occupation rdfs:label ?occupationLabel.
  ?education rdfs:label ?educationLabel.
  ?autre_organisation rdfs:label ?autre_organisationLabel.
  ?distinction rdfs:label ?distinctionLabel.
}
}
```

```
GROUP BY ?id_senat ?senateur ?senateurLabel
```

Le premier élément que nous avons dû prendre en compte dans notre requête est les différentes classes permettant d'identifier un sénateur parmi les classes wikidata. Nous avons ainsi lancé la requête sur trois classes (Q14828018, Q24931583 et Q15686806) afin de s'assurer de disposer d'un maximum d'individus identifiés comme Sénateur.trice.s français.e.s.

Par ailleurs, l'objectif étant d'effectuer une jointure avec les données du Sénat, nous avons récupéré uniquement les individus avec un identifiant dans la base Sénat. Son absence ne nous permettant pas de joindre nos set de données ensuite. A cette étape, nous nous sommes aperçus que plusieurs identifiant Sénat existaient. Dans le cadre de notre projet sur les sénateurs de la Ve République et après vérification de l'identifiant utilisé dans nos premiers jeux de données, nous avons filtré cet identifiant pour garder uniquement ceux que nous pourrions utiliser pour la jointure.

Notre requête nous permet également d'obtenir le genre, bien que nous disposions déjà de cette information dans les jeux de données du Sénat.

Cette première requête nous a permis d'enrichir nos données avec des informations dont nous ne disposions pas par ailleurs. Chacun de ces éléments ont été défini comme optionnel dans notre requête afin de ne pas perdre de la donnée lorsqu'un champ n'était pas complété pour un sénateur. Nous avons ainsi récupéré :

- Le lieu de naissance
- Le lieu de mort
- Les autres mandats occupés par un sénateur.
- Les études
- Les distinctions
- Les "occupations", une catégorie qui s'est avérée très hétéroclite par la suite comprenant les activités professionnelles plus ou moins précises, ainsi que parfois les hobbies ou des fonctions politiques parallèles.
- Les autres organisations auxquelles le Sénateur participe.

Nous avons utilisé le "service" wikibase:label pour obtenir les résultats en français. Malheureusement, cet outil ne fonctionne pas très bien avec la fonction group_concat, nous avons donc dû préciser tous les champs dont nous souhaitions avoir le label.

Enfin, pour ne pas multiplier les lignes et éviter les doublons, nous avons utilisé un pipe (|) pour concaténer les informations lorsque celles-ci étaient multiples dans un même champ.

2. Requête n°2 :

```
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX wikibase: <http://wikiba.se/ontology#>
PREFIX p: <http://www.wikidata.org/prop/>
PREFIX ps: <http://www.wikidata.org/prop/statement/>
PREFIX pq: <http://www.wikidata.org/prop/qualifier/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX bd: <http://www.bigdata.com/rdf#>
```

```
SELECT ?id_senat ?senateur ?senateurLabel ?pere_is_senat ?enfant_is_senat
```

```
WHERE
```

```
{
  {{?senateur wdt:P39 wd:Q14828018} UNION {?senateur wdt:P39 wd:Q15686806} UNION
  {?senateur wdt:P39 wd:Q24931583}}
```

```

?senateur wdt:P1808 ?id_senat.

OPTIONAL {
    ?senateur wdt:P22 ?pere.
        BIND (EXISTS{?pere wdt:P39 wd:Q14828018}||EXISTS{?pere wdt:P39
wd:Q15686806}||EXISTS{?pere wdt:P39 wd:Q24931583} AS ?pere_senat)
        BIND (IF(?pere_senat, "true", "false") AS ?pere_is_senat)
}

OPTIONAL {
    ?senateur wdt:P40 ?enfant.
        BIND (EXISTS{?enfant wdt:P39 wd:Q14828018}||EXISTS{?enfant wdt:P39
wd:Q15686806}||EXISTS{?enfant wdt:P39 wd:Q24931583} AS ?enf_senat)
        BIND (IF(?enf_senat, "true", "false") AS ?enfant_is_senat).
        FILTER (?enfant_is_senat = "true")
}

FILTER (STRSTARTS(?id_senat, "senateur/"))

SERVICE wikibase:label {
    bd:serviceParam wikibase:language "fr".
    ?senateur rdfs:label ?senateurLabel.
}
}

GROUP BY ?id_senat ?senateur ?senateurLabel ?pere_is_senat ?enfant_is_senat
ORDER BY DESC(?pere_is_senat) DESC(?enfant_is_senat)

```

Cette seconde requête reprend un certain nombre d'éléments précédemment exposés. Sa particularité est d'interroger la base wikidata afin d'obtenir si le sénateur a un parent ou un enfant dans la base, et si oui si celui-ci est lui-même sénateur. Nous avons fait du résultat un booléen.

Nous vérifions tout d'abord que le sénateur ait un père renseigné dans Wikidata, s'il en a un on regarde si le père est aussi sénateur et en renvoie *true* s'il l'est, sinon nous rentrons *false*. Puis nous faisons exactement la même requête mais pour l'enfant (en toute logique on devrait retrouver ainsi au moins les sénateurs fils des sénateurs précédents). Une fois cette dernière requête effectuée nous y appliquons un filtre sur les résultats pour ne garder que les valeurs *true* (donc si le sénateur a au moins un enfant sénateur), le but de ce filtre est surtout d'éviter les lignes doublons dans le cas où un sénateur ait au moins un enfant qui soit aussi sénateur et un autre enfant qui ne le soit pas.

III. Traitement des données et jointure des sets

Nous avons préparé et joint nos données en utilisant Dataiku.

A. Fichiers ODSSEN

Nous avons utilisé trois fichiers ODSSEN : General, Histogroupes. et Elusen.

Le fichier General a été préparé en différents points :

- suppression des colonnes inutiles
- renommage des colonnes
- application de la méthode parse aux différentes dates.
- calcul de l'âge au décès
- calcul du dernier âge connu en 2021 tant pour les sénateurs décédés que pour les sénateurs vivants.

- Nettoyage et harmonisation de la colonne "Catégorie Professionnelle"
- Création d'une colonne "Sous-catégorie professionnelle"
- Harmonisation de ces deux colonnes pour réduire le nombre de valeurs (avec un ultime traitement manuel fait dans Excel lors de la manipulation des données dans Tableau).

Le fichier Histogroupes a été préparé en différents points :

- Première étape:
 - suppression des colonnes inutiles
 - renommage des colonnes
 - application de la méthode parse aux différentes dates.
 - calcul de durée
- Seconde étape:
 - les données ont été regroupées par sénateur sur une même ligne en utilisant la fonction de concaténation
- Troisième étape :
 - découpage des données concaténées en autant de colonnes que nécessaire
 - suppression des colonnes inutiles

Le fichier Elusen a été préparé en différents points :

- Première étape:
 - suppression des colonnes inutiles
 - renommage des colonnes
 - application de la méthode parse aux différentes dates.
 - calcul de durée
- Seconde étape:
 - les données ont été regroupées par sénateur sur une même ligne en utilisant la fonction de concaténation
- Troisième étape :
 - découpage des données concaténées en autant de colonnes que nécessaire
 - suppression des colonnes inutiles
 - calcul du nombre de mandats pour chaque sénateur en prenant en compte le fait que certains sénateurs avaient déjà effectué un ou des précédents mandats pour les sénateurs élus pour le mandat débutant en 1955 ou en 1959.

B. Fichiers wikidata

Les set de données issus de wikidata ont été nettoyés chacun de leur côté avant d'être joint ensemble sur l'identifiant Sénat et l'identifiant wikidata.

Le fichier issu de la première requête (*senateur_wikidata_donnees*) a été préparé en différents points :

- Les identifiants sénat et wikidata ont été nettoyés et standardisés.
- Nous avons nettoyé les colonnes comportant les identifiants et nous avons retiré les mentions d'arrondissement pour les lieux de naissance et de mort.
- Ensuite, nous avons créé plusieurs séries de booléens. Nous avons créé des booléens pour les autres mandats : président de conseil régional, conseiller régional, président de conseil départemental ou conseil général, conseiller général ou départemental, maire, conseiller municipal, député, député européen, membre du conseil de l'Europe. Pour autant, dans le souci

de ne pas perdre de données, nous n'avons pas supprimé la colonne d'origine qui comportait des informations supplémentaires pour chacun de ces mandats comme le territoire sur lequel le Sénateur ou la Sénatrice exerçait son mandat.

- Puis, la colonne "occupation" précédemment évoquée a été difficile à structurer à cause de la grande diversité et originalité des entrées. Une seule mention a pu alors être valorisée : "résistant" pour lequel nous avons créé un booléen. Une seconde mention a été extraite "ministre" que nous avons ensuite fusionnée au booléen ministre précédemment créé à partir de la colonne mandat, afin que tous les ministres soient signalés dans une même colonne. Ici encore, nous avons maintenu la colonne originale pour ne pas perdre ces données intéressantes mais très disparates.
- Une opération identique a été effectuée pour la colonne éducation dont nous avons extrait un booléen sur certaines formations qui nous semblait récurrente et intéressante à mesurer : ENS, IEP (Paris et province), ENA, Polytechnique.
- Enfin, nous avons extrait un booléen de la colonne distinction sur les mentions de Légion d'honneur et de médaille de l'Ordre du Mérite.

Le fichier issu de la seconde requête (*filliation_senateur*) a demandé moins de préparation. Les identifiants sénat et wikidata ont été nettoyés et standardisés. Les colonnes redondantes avec le premier fichier issu de wikidata ont été supprimées afin d'obtenir un fichier avec les identifiant nécessaire à la jointure et les deux booléens sur la filiation avec un Sénateur.

Le fichier résultant de la jointure a lui aussi était préparé. En effet, une "recette" python reposant sur la librairie Panda a permis de corriger un problème de doublon : une sénatrice apparaissait quatre fois car elle était en double dans chacun des fichiers d'origine. Nous avons fait cette opération après nous être assuré que les informations étaient identiques sur chacune des lignes et que nous ne perdions pas de données dans l'opération.

C. Fichiers traités de façon isolés sur Dataiku et ré-importés via fichier CSV

Un géocodage des lieux de naissance, mort et de circonscription a été fait via un plugin et l'API MapQuest. Ce traitement pouvant générer des erreurs à chaque traitement et nécessitant une clé API, il a été supprimé du projet Dataiku. Le résultat du traitement a été conservé via un import CSV.

Les groupes politiques ayant été nombreux et fait l'objet de plusieurs renommages ne facilitant pas une compréhension simple et rapide des données, il a été décidé de faire un classement Droite / Gauche / Non inscrit des sénateurs à partir des données Groupe politique. Ce traitement ayant été fait en parallèle d'une nouvelle version du projet Dataiku a nécessité de faire un import des résultats par un fichier CSV.

D. Jonction des fichiers

Les différents fichiers ont été ensuite rassemblés par un LEFT JOIN sur la base de la clé "matricule". Un dernier nettoyage a été effectué (déplacement de colonnes pour faciliter la lecture). Certaines visualisations ont nécessité un ultime nettoyage manuel pour harmoniser certaines données (cf. supra sur les catégories et sous-catégories professionnelles).

IV. Visualisations

A l'aide de Tableau, nous proposons plusieurs visualisations de nos données. Ils sont stockés sur trois Tableau Public :

- <https://public.tableau.com/profile/doriane.hare#!/>
- <https://public.tableau.com/profile/guillaume.cayeux#!/>
- <https://public.tableau.com/profile/julien.fenech#!/>

A. Histoire : Répartition des élu.e.s au Sénat selon leur genre.

Lien vers visualisation :

https://public.tableau.com/shared/2S5JWW39P?:display_count=n&:origin=viz_share_link

Cet ensemble de visualisations propose de visualiser les données en fonction du genre des sénateurs et sénatrices. Plusieurs visualisations proposent de filtrer sur la date d'élection, permettant ainsi de voir les évolutions des données dans le temps. Cette évolution dans le temps est à mettre en perspective avec la représentativité faible des femmes par rapport à celle des hommes. Par ailleurs, la donnée utilisée pour ce filtre sur la date est la date du premier mandat, elle reflète donc le total des données entre les deux dates sélectionnées.

Un premier graphique montre l'entrée des femmes au Sénat dans le temps. Nous voyons alors une progression notable à partir de 2000.

Une seconde visualisation permet de voir la répartition droite, gauche et non inscrit parmi les sénateurs selon le genre. Nous voyons ainsi que la représentativité droite gauche est différente selon le genre, la gauche étant plus représentée parmi les femmes. Il est intéressant de noter qu'il faut attendre 1974 pour avoir une sénatrice de droite.

Un histogramme vertical permet de visualiser la répartition des élus par circonscription selon leur genre. Nous remarquons que les circonscriptions qui élisent le plus de femmes, comme Paris ou le Nord, ne sont pas les circonscriptions où la répartition homme / femme des élus est la plus équitable.

La dernière visualisation propose de montrer la répartition des commissions selon le genre. Si les affaires sociales et la culture sont les plus représentées chez les femmes, c'est aux affaires économiques et à l'aménagement du territoire que leur part est la plus importante. Il est intéressant également de remarquer que les femmes sont très peu représentées à l'économie, à la finance et aux affaires étrangères. Enfin, nous pouvons voir l'évolution de cette répartition dans le temps. Jusqu'en 1977, les femmes sont principalement à la culture et depuis 2000, la parité est presque respectée aux affaires sociales.

B. Entrée des Résistants au Sénat.

Lien vers visualisation :

https://public.tableau.com/views/senateur_resistant/rsistants?:language=fr-FR&:display_count=n&:origin=viz_share_link

Une visualisation propose de montrer l'entrée des résistants au Sénat. Un code couleur permet de connaître le motif de cette entrée. Nous voyons ainsi que la plupart des résistants accèdent au Sénat en 1959 par élection. Une modification du filtre montre qu'un certain nombre était déjà au Sénat avant la première élection sénatoriale de 1959.

C. Répartition par orientation politique dans les commissions.

Lien vers visualisation :

https://public.tableau.com/views/senateurs_commission_DG/gauchedroitecommission?:language=fr-FR&:display_count=n&:origin=viz_share_link

La visualisation de la répartition des orientations politiques dans les commissions montre que les sénateurs de droite sont majoritairement représentés dans les commissions. Ces résultats sont à mettre en perspective avec le nombre plus important de sénateurs de droite dans le total des sénateurs. Avec le pourcentage de la répartition par commission, nous voyons que les couleurs politiques sont représentées de manière assez homogènes dans chacune des commissions entre 30% et 40% pour la gauche et entre 60% à 70% pour la droite. La gauche est moins représentée aux Affaires étrangères, elle l'est davantage aux lois, à la culture et aux affaires sociales.

D. Évolution de la répartition par orientation politique à l'entrée au Sénat.

Lien vers visualisation :

https://public.tableau.com/views/entre_snat_DG/Feuille13?:language=fr-FR&:display_count=n&:origin=viz_share_link

Un graphique permet de visualiser la répartition par orientation politique à l'entrée au Sénat par année. Il permet ainsi de mesurer l'orientation politique du renouvellement au Sénat. Les points les plus intéressants à analyser sont les plus gros, car les plus représentatifs d'une répartition grâce à un nombre d'individu plus important. Nous voyons ainsi que la droite est toujours majoritaire parmi les nouveaux entrants au Sénat à part pour les élections de 2008. Il est intéressant de remarquer que en 1980 et 2011, années précédant l'élection d'un candidat de gauche aux Présidentiels, l'écart entre les entrants de droite et de gauche se réduit. En 2014 et en 2017, la droite connaît son plus important renouvellement de sénateurs, avec respectivement 58 et 72 entrants.

E. Histoire : Evolution de la répartition des sénateurs de la Vème République par CSP selon leur date d'entrée au Sénat

Lien vers la visualisation:

https://public.tableau.com/shared/3KTYD593C?:display_count=n&:origin=viz_share_link

Cet ensemble de visualisations propose de visualiser les données en fonction de la catégorie professionnelle des sénateurs. Il a été choisi de faire varier les données selon le cursus de l'entrée au sénat (i.e. date de début du 1er mandat). Ce curseur permet d'apprécier la dynamique de changement car elle ne prend en compte que les nouveaux visages.

Plusieurs visualisations sont proposées afin d'étudier plusieurs facteurs (genre, bord politique, sous-catégorie ; seuls ou panachés).

On constate une évolution des catégories professionnelles en lien avec la société. Le monde rural, traditionnellement surreprésenté au Sénat, perd progressivement de sa spécificité avec la montée en puissance des fonctionnaires, des salariés et de l'enseignement.

On peut également identifier dans une autre visualisation les catégories d'où les sénateurs de gauche proviennent (enseignement, fonctionnaires et salariés) et constater dans la suivante que la gauche n'a pas plus avantagé les femmes que la droite. Aussi, on ne peut que reconnaître l'efficacité de la loi sur la parité. Enfin, les visualisations par sous-catégorie professionnelle permet d'aller chercher dans le détail la répartition des sénateurs.

On voit que les retraités, les cadres, les hauts fonctionnaires sont surreprésentés toute époque confondue.

F. Les âges des Sénateurs

Lien vers la visualisation:

https://public.tableau.com/app/profile/julien.fenech/viz/gedessnateurs/Age_Senateur

Cette visualisation suggère d'aborder les âges des sénateurs sous différents angles politiques ou socio-économiques. L'âge des sénateurs est une question récurrente dans la société et la vie politique française, et le terme même de "Sénat" renvoie à l'âge de ses membres.

Une première vue permet de visualiser l'âge moyen auquel un sénateur est élu pour son premier mandat et compare si cet âge connaît des variations en fonction de son étiquette politique, de l'année où il a été élu et du motif de son élection (remplacement, renouvellement, réélection, ces deux derniers motifs sont surtout vrai pour des sénateurs élus pendant la IV^e République et réélus pendant la Ve mais qui, de fait connaisse un premier mandat pour la Ve République).

Un premier graphique nous renseigne donc sur la moyenne d'âge des sénateurs en fonction de leur année d'élection et de leur couleur politique. Nous constatons que pendant les grands moments de victoire politique (surtout visible pendant les alternances) les sénateurs élus de la couleur politique victorieuse sont en moyenne bien plus âgés que les sénateurs de l'autre côté du prisme politique (surtout visible pendant les grandes alternances de la période Mitterrand et Chirac), mais que dans l'ensemble l'âge des sénateurs élus n'est pas proportionnel à l'année où ils ont été élus, on ne vote pas plus jeune maintenant qu'auparavant bien que l'âge d'entrée au Sénat a été abaissée. Toutefois ces statistiques sont à pondérer par un par le nombre des non-inscrits qui ne veut pas forcément dire qu'un Sénateur ne veut s'afficher sous aucune couleur politique mais que son inscription n'est pas renseigné dans les données du Sénat au moment de son élection.

Un second graphique nous renseigne sur les motifs du début de ce premier mandat en fonction de l'âge du Sénateur au moment de la prise de ses fonctions et de sa couleur politique. On peut voir qu'en général les remplaçants/suppléants sont en moyenne plus âgés que des nouveaux élus, on constate aussi que les Sénateurs de droite ont été renouvelés en plus grand nombre que les sénateurs de gauche.

Une seconde vue permet d'attaquer les différents âges de la vie politique au Sénat sous l'angle des catégories professionnelles. Cette vue permet de voir l'impact de celles-ci à l'âge de première entrée au Sénat, de l'âge de dernier mandat des sénateurs, et si les catégories professionnelles ont une répercussion sur l'âge de décès des sénateurs. Dans l'ensemble, nous pouvons conclure, qu'en définitive, la catégorie professionnelle semble avoir peu d'impact sur ces trois aspects de la vie des sénateurs. Les gens issus du commerce et de l'industrie sont sensiblement plus âgés au moment de leur premier mandat que les salariés. Les journalistes eux se présentent et sont élus à un âge plus avancé que les autres catégories pour les derniers mandats de sénateurs. Et en moyenne, les sénateurs décèdent aux mêmes âges, les salariés mourant en moyenne 1 an et demi avant leurs collègues sans activité professionnelle déclarée.

Enfin dernière vue compare l'impact du genre sur les moyennes d'âge des sénateurs en fonction de leurs premier mandat, leurs dernier mandat et leurs âge de décès.

On constate que les femmes sont légèrement plus âgées au moment de leur premier mandat que les hommes mais que ce sont les hommes qui sont, cette fois-ci, légèrement plus âgés que les femmes lors de leurs derniers mandats. Sans surprise, la moyenne féminine de l'âge de décès est sensiblement supérieure à celle des hommes. Toutefois les Sénatrices sont beaucoup moins nombreuses que les Sénateurs, ce qui pourrait faire varier ces résultats si une parité totale existait.

G. Relation entre les lieux de naissance, de décès et la circonscription des Sénateurs

Lien vers la visualisation:

<https://public.tableau.com/app/profile/julien.fenech/viz/CartedesSnateurs/CartedesSnateurs>

La visualisation de la carte permet de visualiser les lieux que les Sénateurs ont traversés pendant leur vie : leur ville de naissance, de décès (pour ceux qui sont décédés), et la circonscription où ils ont été élus.

Le but de cette visualisation est de permettre de se rendre compte de la relation d'un Sénateur, qui représente la collectivité territoriale, avec son lieu d'élection. Cette visualisation n'est pas exempte de problème surtout dû à l'abondance des données traitées, ce qui donne ce foisonnement de données.

Les circonscriptions sont représentées sous forme de département complètement remplis à la couleur de leurs sénateurs, les lieux de naissances et de décès sont eux représentés sous forme de point avec la couleur du sénateur là aussi pour pouvoir mieux visualiser toutes ces informations lors du choix du sénateur dans la la zone de la légende. Enfin des lignes reliant les différents objets spatiaux permettent de remonter le fil du sénateur.

Toutefois on s'aperçoit, pour peu de réussir à démêler la pelote affichée à l'écran, que beaucoup de sénateurs sont nés, ont été élus et sont décédés dans une circonscription ou une circonscription limitrophe. Bien sûr on ne peut pas conclure que toutes les personnes nées et éventuellement mortes en dehors de leur circonscription d'élection sont là par parachutage et n'ont aucun lien émotionnel ou personnelle avec leurs circonscriptions. Mais dans l'ensemble la plupart des informations semblent cohérentes.

La grosse difficultés de cette vue tient surtout à la manipulation de données de géolocalisation sur Tableau et de la structuration de ces données. Tableau a des fonctionnalités beaucoup plus poussées qu'utilisées ici quant au traitement de ces données de géolocalisation mais cela suppose un traitement particulier de ces données avec l'ajout de code ou d'identification particulier pour pouvoir suivre les déplacements d'un sénateur. De plus le manque d'expérience ne m'a pas permis de m'approprier pleinement toutes les données pour pouvoir rendre une visualisation graphiquement claire et totalement cohérente, c'est là un des principaux point de blocage de cette visualisation, la difficulté de concevoir intellectuellement la représentation des mouvements ou du moins des liens d'un très grand nombre de personnes avec des régions ou des lieux précis.