

# travail de session

Svetlana Zhuk

2024-04-09

## Introduction

En 2015, Justin Trudeau, l'actuel Premier ministre du Parti libéral, a promis de réparer “le système électoral défaillant du pays” en passant du scrutin majoritaire à un tour à une représentation plus proportionnelle afin que les votes soient reflétés plus fidèlement à la Chambre des communes (Carmichael 2017).

François Legault, l'actuel Premier ministre du Québec issu de la CAQ, a fait une promesse similaire : lors de l'élection de 2018, M. Legault a promis de réformer le système électoral en passant du scrutin majoritaire à un tour à un système électoral mixte de la représentation proportionnelle (Sidhartha 2022).

En théorie, un mode de scrutin proportionnel permettrait une meilleure représentation et une plus grande visibilité des petits partis concurrents, comme ceux composés de politiciens de la génération Millennials et de la génération Z, qui se concentrent souvent sur des questions essentielles pour ces générations. Cependant, tant au niveau fédéral que provincial au Québec, ces promesses ne sont toujours pas tenues. Les raisons pour lesquelles ces promesses n'ont pas été tenues sont au mieux ondulantes, comme “le statu quo servirait mieux les Québécois” (dans le cas de Legault) et le fait de souligner que le scrutin proportionnel permettrait à des “voix marginales” d'entrer au Parlement (Rands 2017). Si les Millennials et la Génération Z n'ont pas leur parti politique, peut-être que d'autres partis peuvent aborder les questions concernant ces générations ?

## Problématique de recherche

Abacus Data montre qu'il existe un conflit générationnel concernant les problèmes les plus importants auxquels le Canada est confronté. Pour 57 % des baby-boomers et 42 % des membres de la génération X, les soins de santé sont un enjeu majeur. En revanche, la génération Z se préoccupe davantage de l'accessibilité du logement, 54 % des membres de la génération Z et 51 % des milléniaux considérant qu'il s'agit d'une préoccupation essentielle. En outre, il existe

un clivage permanent sur des questions telles que le changement climatique et l'environnement (Kishchuk 2023).

Une analyse révèle que les gouvernements canadiens dépensent entre 33 321 et 40 152 dollars par personne âgée de 65 ans et plus, entre 13 635 et 14 800 dollars par personne âgée de 45 à 64 ans, et entre 10 406 et 11 614 dollars par personne âgée de moins de 45 ans (Kershaw et Anderson 2016).

L'analyse suggère que les dépenses préférentielles en faveur des personnes âgées peuvent s'expliquer de deux points de vue : les préférences d'un électorat vieillissant et l'influence de l'âge des législateurs sur leurs décisions politiques.

Vallée-Dubois (2023) montre comment l'âge peut façonner les attitudes à l'égard des dépenses publiques. En moyenne, les personnes âgées soutiennent moins les dépenses de l'État en matière d'éducation et d'environnement, alors qu'elles sont nettement plus favorables aux dépenses en matière de transport (p. 469). En outre, la recherche révèle que les adultes dans la quarantaine sont plus favorables aux investissements publics dans les services aux personnes âgées que les autres groupes d'âge (Vallée-Dubois 2023).

Cependant, une étude de McClean (2019), axée sur le Japon, a révélé que l'âge des hommes politiques influence également de manière significative l'élaboration des politiques publiques. Les jeunes maires ont tendance à donner la priorité à différentes politiques de protection sociale par rapport à leurs homologues plus âgés, notamment en augmentant les dépenses liées à la protection de l'enfance par rapport à celles liées à la protection des personnes âgées. Le biais lié à l'âge au sein des institutions politiques peut donc avoir de profondes implications pour la représentation politique et la formulation des politiques publiques (McClean 2019).

Alors, qui s'occupera des questions importantes pour les Millennials et la génération Z, si les assemblées législatives sont majoritairement composées d'autres générations ? Munger, dans son livre "Generation Gap", montre que le Congrès des États-Unis est principalement composé de la génération des baby-boomers, qui détient le pouvoir politique. Les institutions politiques et le système de vote uninominal à un tour sont probablement un obstacle pour un nouveau parti représentant les jeunes générations (Munger 2022). De la même manière que McClean (2009), Munger prévient que les partis, composés principalement de personnes âgées, sont plus susceptibles de répondre aux besoins de leurs électeurs plus âgés (Munger 2022, 47).

Compte tenu du grand nombre de recherches sur la corrélation entre le vieillissement de l'électorat et son influence sur les politiques publiques, notre étude vise à effectuer une analyse descriptive de la composition générationnelle des assemblées législatives provinciales, en mettant l'accent sur le Québec et l'Ontario, les deux plus grandes provinces du Canada. Nous analyserons la période de service dans la législature de 1961 à 2018. L'année de départ est importante car elle marque le moment où les premiers baby-boomers sont devenus éligibles à la participation politique à l'âge de 18 ans. Cette approche nous permettra de retracer le début de leur carrière politique et de déterminer la durée de leur mandat au Parlement.

La question de recherche est la suivante :

## **Quelle est la composition générationnelle des Assemblées législatives de l’Ontario et du Québec, entre 1961 et 2021 ?**

Dans cette étude, nous prévoyons d’examiner les conclusions de McClean et Munger dans le contexte de la composition des assemblées législatives du Québec et de l’Ontario. Nous cherchons à comprendre si la prédominance des générations plus âgées au sein de l’assemblée législative limite la participation des milléniaux, entraînant une représentation insuffisante et un possible mépris pour les politiques qui répondent aux besoins des milléniaux et de la génération Z.

### **Données et méthodes**

Notre étude sera basée sur une analyse quantitative des ensembles de données présentés, en se concentrant sur les statistiques descriptives. Cela nous permettra de généraliser et de décrire la composition des cohortes générationnelles des institutions politiques au niveaux provincial. Nous utiliserons la visualisation pour comparer les données entre les cohortes générationnelles et rendre les résultats accessibles.

Pour répondre à notre question de recherche sur la composition générationnelle des assemblées nationales du Québec et de l’Ontario, nous utiliserons l’ensemble de données ‘Provincial Parliamentary Biographies’ de Rivard et al, 2024.

Cet ensemble de données d’observation fournit des informations électorales sur les législateurs de quatre provinces canadiennes depuis la création de leurs assemblées coloniales au XVIII<sup>e</sup> siècle, soit plus de 7 000 législateurs de l’Ontario, du Québec, du Nouveau-Brunswick et de la Nouvelle-Écosse. Puisque cet ensemble de données couvre une période du XVIII<sup>e</sup> siècle jusqu’au XXI<sup>e</sup> siècle, il convient à notre analyse qui se concentre sur la période de 1961 à 2018, car il contient les informations qui nous intéressent.

Nous nous concentrerons sur le Québec et l’Ontario, les deux plus grandes provinces du Canada. Les variables dont nous aurons besoin sont ‘year\_of\_birth’ pour analyser la cohorte générationnelle, ‘year’ pour connaître l’année d’élection du législateur et ainsi déterminer son âge au début de sa carrière politique, et ‘exit\_year’ pour déterminer combien de temps le député est resté à l’Assemblée nationale.

L’ensemble de données que nous avons enregistré dans notre code sous le nom de “provinces” contient 17 038 observations et 48 variables. Après avoir nettoyé les données et créé les nouvelles variables, nous avons utilisé deux ensembles de données :

- ‘provinces\_clean\_unique\_1961’ : 1,253 observations and 13 variables
- ‘legislative\_age’ (based on ‘provinces\_clean\_unique\_1961’) : 1253 observations and 14 variables

Dans les données “provinces\_clean\_unique”, en plus des variables initiales que nous avons utilisées, telles que ‘*first\_last*’ pour les noms des députés, ‘*province*’ pour les comparaisons entre provinces, ‘*year\_of\_birth*’, ‘*year*’ marquant le début de la carrière politique, et *exit\_year* représentant la fin de la carrière politique, nous avons créé de nouvelles variables

- **gen\_cohort**: Nous classons les députés en fonction de leur *année de naissance* dans des cohortes générationnelles. Pour la définition de ces cohortes, nous nous sommes appuyés sur les données suivantes [the Strauss-Howe Generational Theory](#), qui fournit un cadre pour la répartition des âges. Nous avons donc utilisé les catégories suivantes : Silent, Boomer, Génération X et Millennial.
- **year\_range**: Représente les cohortes générationnelles auxquelles appartiennent les députés.
- **years\_total**: Indique le nombre total d’années pendant lesquelles un député a siégé à l’assemblée législative, calculé comme suit :  $\text{exit\_year} - \text{year} + 1$ . Cette formule inclut l’année de départ dans le décompte total.
- **sessions\_total**: Etant donné que chaque année comprend deux sessions, ce chiffre est calculé comme  $\text{years\_total} * 2$ , représentant le nombre total de sessions durant le mandat du député.
- **start\_year**: Il s’agit de l’année initiale de la carrière du député au sein de l’assemblée, ce qui permet de calculer des années complètes de service sans dupliquer les données. C’est l’année où ils ont commencé leur carrière, marquant le début de leur mandat.

Dans l’ensemble de données ‘legislative\_age’, nous avons ajouté la variable :

- **age\_start**: Cette variable est utilisée pour analyser l’âge auquel le député a commencé sa carrière, calculé comme  $(\text{start\_year} - \text{year\_of\_birth})$ .

Cependant, au cours du processus de nettoyage des données, j’ai rencontré plusieurs problèmes liés à des incohérences dans l’ensemble de données initial. Ces problèmes expliquent le nombre réduit d’observations et ont influencé les résultats.

1. 62 députés, dont les années de début de mandat s’échelonnent de 1990 à 2018, n’ont pas d’année de naissance enregistrée, ce qui m’empêche de déterminer leur cohorte générationnelle et de calculer l’âge auquel ils ont commencé leur mandat.
2. Pour les députés exerçant plusieurs mandats, l’année de sortie enregistrée ne signifie pas la fin de leur mandat en cours, mais plutôt la conclusion de leur dernier mandat avant toute interruption. Cette nuance complique le calcul précis du nombre total d’années de service. Afin de simplifier les calculs pour les députés ayant plusieurs mandats, l’utilisation de l’année de sortie comme début du mandat suivant pourrait permettre d’éviter le surcomptage des années. Cependant, avec 17 000 observations, la vérification individuelle des dossiers de chaque adjoint représente un défi important. Par exemple,

dans le cas de Bob Chiarelli, nous observons des périodes telles que 1987 - 1997, 1990 - 1997, et 1995 - 1997 pour l'année et l'année de sortie, ce qui indique un chevauchement des mandats. Inversement, pour Mike Davison (1975 - 1981, 1981 - 1985) et Evelyn Gigantes (1975 - 1981, 1981 - 1987), l'année de sortie coïncide avec le début du mandat suivant. Ces incohérences posent des problèmes pour déterminer le nombre total d'années passées à l'assemblée législative.

first_last <chr>	province <chr>	gender <chr>	year_of_birth <dbl>	year_of_death <dbl>	year <dbl>	exit_year <dbl>	party <chr>
chiarelli_bob	Ontario	m	1941	NA	1987	1997	Liberal
chiarelli_bob	Ontario	m	1941	NA	1990	1997	Liberal
chiarelli_bob	Ontario	m	1941	NA	1995	1997	Liberal
chiarelli_bob	Ontario	m	1941	NA	2007	2018	Liberal
chiarelli_bob	Ontario	m	1941	NA	2011	2018	Liberal
chiarelli_bob	Ontario	m	1941	NA	2014	2018	Liberal
denis_lazure	Quebec	m	1925	2008	1976	1984	Parti Quebecois
denis_lazure	Quebec	m	1925	2008	1981	1984	Parti Quebecois
denis_lazure	Quebec	m	1925	2008	1989	1996	Parti Quebecois
denis_lazure	Quebec	m	1925	2008	1994	1996	Parti Quebecois

Figure 1: year and exit\_year

3. J'ai 505 députés dont le mandat a commencé entre 1990 et 2018 et qui n'ont pas d'année de sortie enregistrée. Je suppose que cela signifie qu'ils siègent toujours à l'assemblée législative. Dans ces cas, lorsque 'exit\_year' est NA, je l'ai défini comme s'étendant jusqu'à 2024.
4. J'ai 11 observations où l'année\_de\_naissance n'est pas correctement alignée avec l'année et l'année\_de\_sortie, et ces observations doivent être éliminées. Prenons l'exemple de Duncan McFarland, dont l'année\_de\_naissance est 1973, ce qui placerait ce député dans la cohorte de la génération X. Cependant, il est indiqué que le mandat s'est déroulé de 1848 à 1851.

## Résultats

La figure 1 s'inspire de la figure figurant dans le livre de Munger (2022) sur les études américaines. Elle montre que le temps passé à l'Assemblée législative n'est pas réparti uniformément entre les cohortes générationnelles. La ligne pointillée représente une moyenne naïve (une estimation non ajustée à la population) du nombre de mandats que chaque cohorte effectuerait si les mandats étaient répartis de façon égale entre les cohortes. Cependant, nous observons que la majorité des années passées à l'Assemblée est concentrée sur les députés nés entre 1930 et 1960. Cette période englobe les dernières années de la génération silencieuse et couvre la génération des baby-boomers (1943-1960). Les députés nés entre 1930 et 1960 ont les barres

	first_last	province	gender	year_of_birth	year_of_death	year	exit_year	party
1	chiarelli_bob	Ontario	m	1941	NA	1987	1997	Liberal
2	chiarelli_bob	Ontario	m	1941	NA	2007	2018	Liberal
3	davison_mike	Ontario	m	1950	NA	1975	1981	CCF/NDP
4	davison_mike	Ontario	m	1950	NA	1981	1985	CCF/NDP
5	eves_ernie	Ontario	m	1946	NA	1981	2001	Progressive Conservative
6	eves_ernie	Ontario	m	1946	NA	1999	2005	Progressive Conservative
7	gigantes_evelyn	Ontario	f	1942	NA	1975	1981	CCF/NDP
8	gigantes_evelyn	Ontario	f	1942	NA	1981	1987	CCF/NDP
9	gigantes_evelyn	Ontario	f	1942	NA	1990	1995	CCF/NDP
10	havrot_edward	Ontario	m	1927	2017	1971	1975	Progressive Conservative
11	havrot_edward	Ontario	m	1927	2017	1977	1985	Progressive Conservative
12	hayes_pat	Ontario	m	1927	2011	1985	1987	CCF/NDP
13	hayes_pat	Ontario	m	1942	2011	1990	1995	CCF/NDP
14	kells_morley	Ontario	m	1936	NA	1981	1985	Progressive Conservative
15	kells_morley	Ontario	m	1936	NA	1995	2003	Progressive Conservative
16	lessard_wayne	Ontario	m	1956	NA	1990	1995	CCF/NDP
17	lessard_wayne	Ontario	m	1956	NA	1995	1999	CCF/NDP
18	makarchuk_mac	Ontario	m	1931	NA	1967	1971	CCF/NDP
19	makarchuk_mac	Ontario	m	1931	NA	1975	1981	CCF/NDP
20	patten_richard	Ontario	m	1942	NA	1987	1990	Liberal
21	patten_richard	Ontario	m	1942	NA	1995	2007	Liberal
22	rinaldi_lou	Ontario	m	1947	NA	2003	2011	Liberal
23	rinaldi_lou	Ontario	m	1947	NA	2014	2018	Liberal
24	sorbara_greg	Ontario	m	1946	NA	1985	1995	Liberal
25	sorbara_greg	Ontario	m	1946	NA	1999	2012	Liberal

Figure 2: year and exit\_year\_2

first_last <chr>	province <chr>	gender <chr>	year_of_birth <dbl>	year_of_death <dbl>	year <dbl>	exit_year <dbl>	party <chr>	gen_cohort <chr>	year_range <fctr>
Moses Gamble	Ontario	m	1942	NA	1816	1817	NA	Silent	1925–1942
Roger B. Conger	Ontario	m	1950	NA	1844	1848	NA	Babyboom	1943–1960
Reed Burritt	Ontario	m	1946	NA	1848	1851	NA	Babyboom	1943–1960
Duncan McFarland	Ontario	m	1973	NA	1848	1851	NA	GenerationX	1961–1981
George Wright	Ontario	m	1951	NA	1851	1854	Reformer	Babyboom	1943–1960
Daniel McKerlie	Ontario	m	1931	NA	1854	1857	Clear Grits	Silent	1925–1942
Donald Matheson	Ontario	m	1970	NA	1854	1857	Clear Grits	GenerationX	1961–1981
Robert Ferris	Ontario	m	1970	NA	1854	1857	Clear Grits	GenerationX	1961–1981
John Fraser	Ontario	m	1977	NA	1854	1857	Reformer	GenerationX	1961–1981
John R. Clark	Ontario	m	1941	NA	1857	1861	Reformer	Silent	1925–1942

Figure 3: year and exit\_year\_3

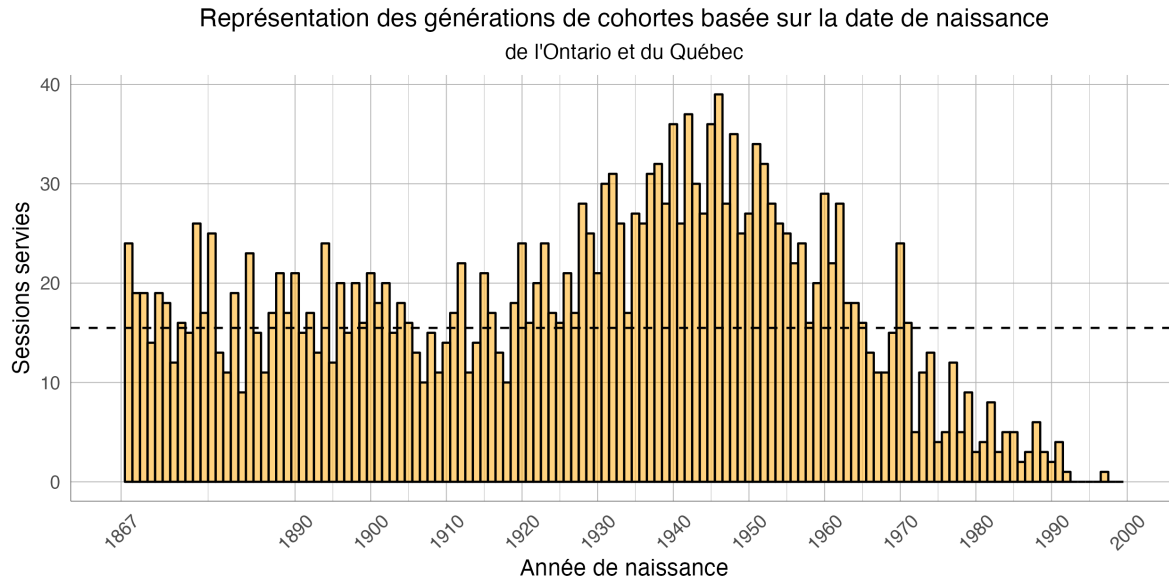
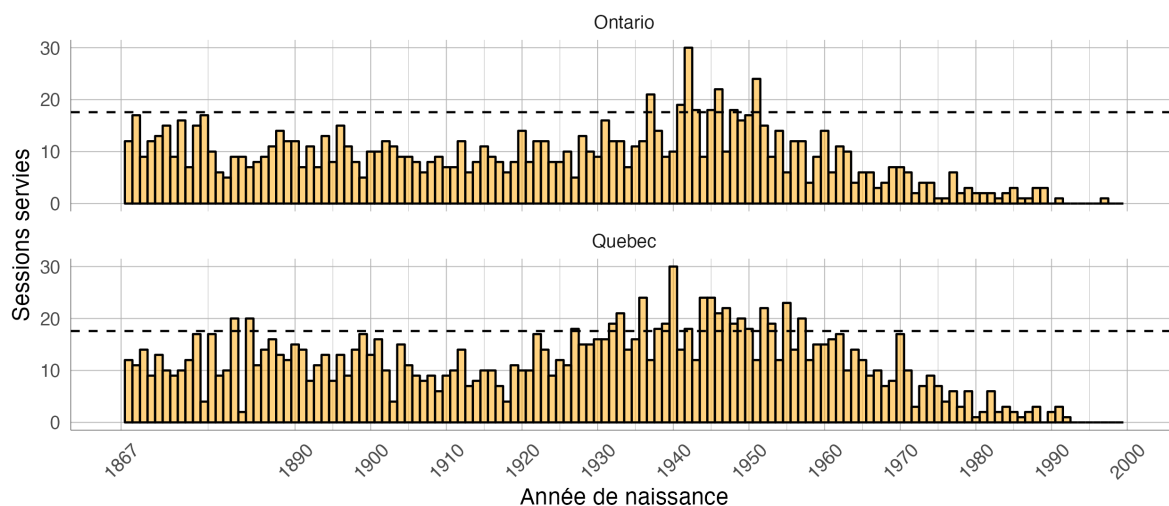


Figure 4: fig.1

les plus élevées, ce qui indique qu'un grand nombre de sessions législatives ont été effectuées par des personnes de ces années de naissance.

Munger qualifie donc ce phénomène de “boomer ballast” et prévient qu'il restera actif jusqu'en 2030 environ. Cela suggère qu'il continuera à influencer les générations suivantes, qui n'ont pas pu entrer dans la vie politique en raison de la domination des baby-boomers à l'Assemblée législative. Le graphique indique que cela pourrait être le cas pour les corps législatifs du Québec et de l'Ontario.

## Représentation des générations de cohortes basée sur la date de naissance



Nous avons décidé d'examiner cette répartition à l'échelle du Québec et de l'Ontario. Nous constatons que la tendance à la domination des membres nés entre 1940 et 1960 reste présente dans les deux provinces. Il y a un pic notable au Québec pour les députés nés en 1940, alors qu'en Ontario, il correspond à 1938, 1944 et 1951. On note également la présence active de députés nés en 1970 à l'Assemblée législative du Québec. Ainsi, la tendance de lestage des boomers peut encore persister dans les deux assemblées législatives, offrant moins d'opportunités de participation aux jeunes générations.

## Âge de début à l'Assemblée par génération de l'Ontario et du Québec

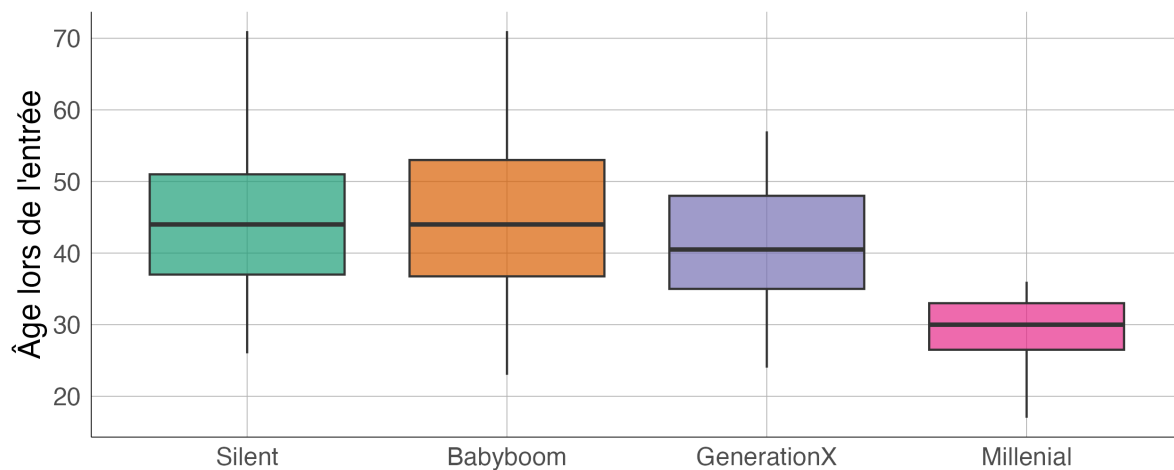
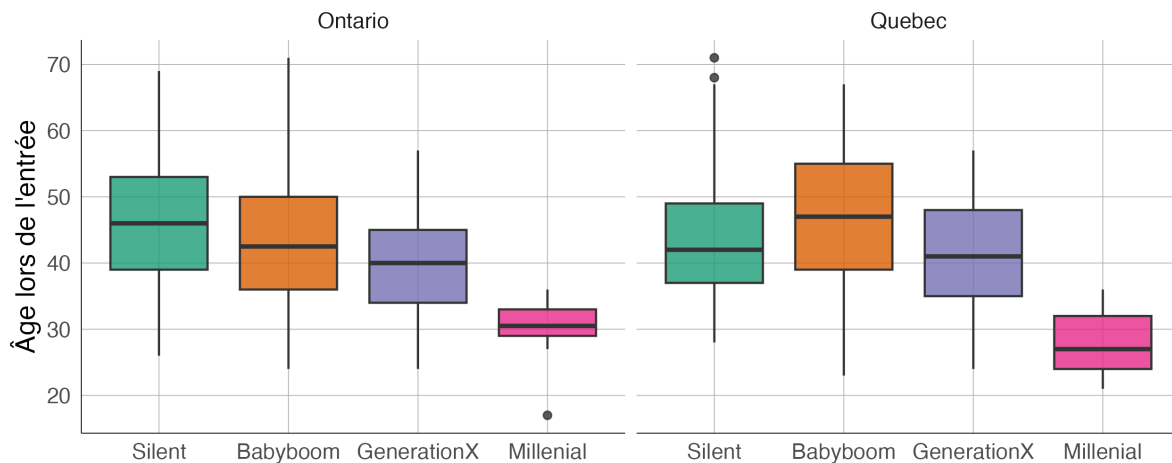


Figure 5: fig.3

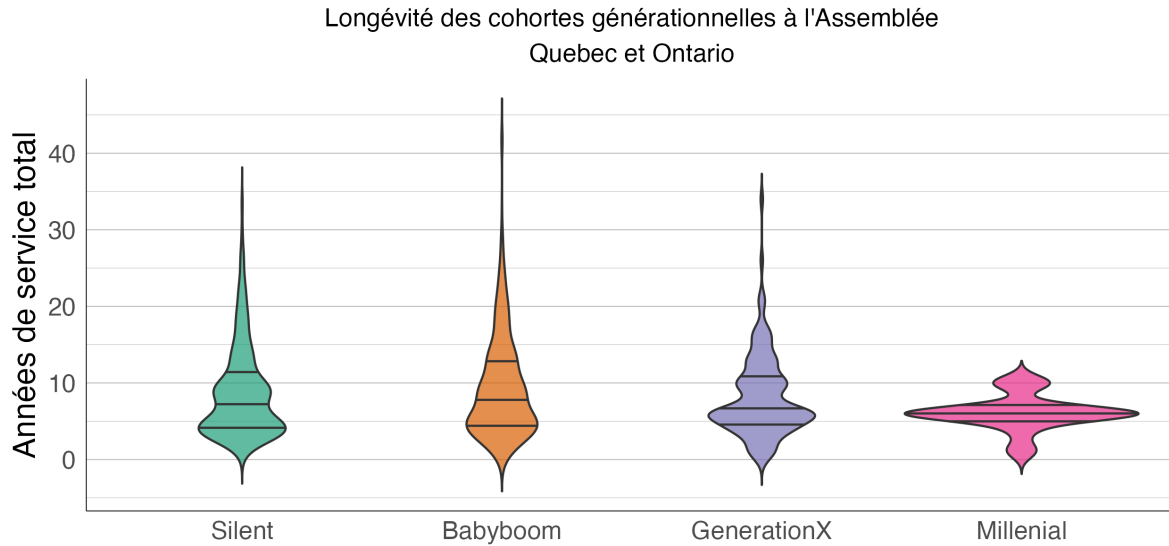


En outre, nous examinons l'âge auquel les députés ont commencé leur mandat à l'Assemblée législative, en nous concentrant sur les générations suivantes : Génération silencieuse, Génération du baby-boom, Génération X et Génération du millénaire. Le graphique indique que l'âge médian pour les générations du baby-boom et du silence est d'environ 45 ans, bien que la distribution des âges soit plus large chez les baby-boomers. Daryl Kramp et Roland Richer ont commencé leur carrière à l'âge de 71 ans. Pour la génération X, l'âge médian est plus bas, à 41 ans. L'âge le plus jeune de début de mandat est observé chez les Millennials, avec une médiane de 30 ans. Malgré le fait que les Millennials n'en sont qu'à leurs débuts en politique, nous émettons l'hypothèse qu'un début de carrière politique plus précoce, comparé aux générations du Silent et du Baby Boom, leur permettra de rester plus longtemps au sein de l'assemblée législative. Toutefois, le boulet des baby-boomers constituera un obstacle, rendant difficile la formation d'un contrepoids significatif à la domination de la génération du baby-boom.

### Âge de début à l'Assemblée par génération



Cependant, lorsque nous analysons les années d'entrée à l'Assemblée législative pour l'Ontario et le Québec, nous pouvons observer certaines différences : L'âge médian des baby-boomers de l'Ontario est de 43 ans, alors que celui des baby-boomers du Québec est de 47 ans. Pour la génération silencieuse, l'âge médian est de 42 ans. Nous notons que l'âge médian d'entrée en Ontario est de 31 ans, ce qui inclut le cas d'une personne commençant à 18 ans - Sam Oosterhoff, né en 1997, en est un exemple. Pour les Millennials au Québec, l'âge médian d'entrée est de 27 ans. Pour la génération silencieuse, les âges médians sont presque les mêmes dans les deux provinces, soit 40 et 41 ans. Étant donné que les milléniaux commencent leur carrière politique plus tôt que leurs collègues de la génération des baby-boomers, ils pourraient avoir la possibilité de mener des carrières politiques plus longues à l'avenir.



Le dernier graphique présente le nombre total d'années de service à l'Assemblée législative par cohorte générationnelle. Il illustre non seulement la durée typique de service pour chaque génération (indiquée par la largeur des violons), mais aussi la diversité au sein de chaque cohorte. Ainsi, nous observons que la cohorte des baby-boomers présente la plus longue variation, avec plus de 40 ans, et que la génération silencieuse affiche plus de 35 ans. Les milléniaux ont la durée de mandat la plus courte, car ils constituent la génération la plus jeune, ce qui explique pourquoi ce groupe passe généralement environ 8 ans à l'Assemblée législative. Il est à noter que la médiane des années de mandat des baby-boomers et des milléniaux, par exemple, est assez similaire et n'est peut-être plus élevée pour les baby-boomers qu'en raison de la présence de quelques députés ayant un mandat plus long que les milléniaux n'ont pas encore.

## Conclusion :

Conformément à notre question de recherche, nous avons étudié la composition générationnelle de l'Assemblée législative au Québec et en Ontario. L'ensemble des données indique que, pour l'essentiel, l'Assemblée législative est composée de la génération silencieuse, des baby-boomers et de la génération X. Les milléniaux ont siégé moins d'années au total et en moyenne, ce qui pourrait expliquer l'apparente négligence de leurs questions. Toutefois, notre analyse (figure 4) suggère également que cet effet pourrait simplement être dû au fait que les Millennials sont encore trop jeunes pour atteindre l'âge moyen typique des députés. Ils commencent déjà, en moyenne, plus tôt que les générations précédentes, et ont donc peut-être simplement besoin de temps pour rattraper leur influence sur la prise de décision. Mais le danger du boulet des baby-boomers persiste. Il est important de prêter attention à la représentation des jeunes générations à l'Assemblée législative, en particulier maintenant que nous vivons dans une ère

de changements techniques et de développement rapide de l'IA. L'Assemblée législative sera chargée d'élaborer des politiques juridiques pour régir l'intelligence artificielle. C'est pourquoi nous avons besoin de la participation de la jeune génération à l'Assemblée. Les milléniaux et la génération Z, qui sont plus avancés sur le plan technique et utilisent la technologie dans la vie de tous les jours, peuvent apporter une contribution précieuse.

Pour améliorer la représentation des jeunes générations au sein de l'assemblée législative, nous pouvons envisager de mettre en place des quotas spéciaux pour les jeunes politiciens afin de représenter les Millennials et la Génération Z. En outre, des recherches supplémentaires sont nécessaires. Par exemple, il est crucial d'examiner la combinaison des générations par année parlementaire et la corrélation entre les votes en faveur des politiques publiques qui intéressent les générations plus âgées et plus jeunes.

## Annexe

```
library(tidyverse)
```

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v dplyr      1.1.4      v readr      2.1.5
v forcats    1.0.0      v stringr    1.5.1
v ggplot2    3.5.0      v tibble     3.2.1
v lubridate  1.9.3      v tidyr      1.3.1
v purrr      1.0.2
-- Conflicts ----- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()     masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

```
library(lubridate)
library(rvest)
```

Attaching package: 'rvest'

The following object is masked from 'package:readr':

```
guess_encoding
```

```
#importation de données
```

```
provinces <- read.csv("~/Dropbox/fas_1001_Zhuk/_travail_session/Data/Provincial Parliament")
```

```
dim(provinces)
```

```
#étudier les données : provinces
```

```
glimpse(provinces)
```

```
dim(provinces)
```

```
colnames(provinces)
```

```
table(provinces$year)
```

### Annexe : variable gen\_cohort, year\_range

```
#nettoyage des données: provinces
```

```
table(provinces$province)
```

New Brunswick	Nova Scotia	Ontario	Quebec
2795	2969	5592	5682

```
table(provinces$year_of_birth)
```

1696	1697	1701	1705	1709	1710	1711	1712	1713	1714	1715	1716	1717	1718	1719	1720
1	1	1	1	4	3	1	1	1	1	10	10	6	4	2	6
1721	1723	1724	1725	1726	1727	1728	1730	1731	1732	1733	1734	1735	1736	1737	1738
1	4	9	3	8	7	7	5	7	2	9	9	16	7	7	14
1739	1740	1741	1742	1743	1744	1745	1746	1747	1748	1749	1750	1751	1752	1753	1754
6	14	7	9	22	28	16	15	12	8	8	44	56	38	46	21
1755	1756	1757	1758	1759	1760	1761	1762	1763	1764	1765	1766	1767	1768	1769	1770
26	29	34	17	17	24	34	40	30	54	40	43	36	35	28	41
1771	1772	1773	1774	1775	1776	1777	1778	1779	1780	1781	1782	1783	1784	1785	1786
26	29	27	22	29	73	50	28	29	28	34	23	36	38	51	72
1787	1788	1789	1790	1791	1792	1793	1794	1795	1796	1797	1798	1799	1800	1801	1802
52	43	64	74	60	55	69	53	50	41	68	77	63	75	40	30

1803	1804	1805	1806	1807	1808	1809	1810	1811	1812	1813	1814	1815	1816	1817	1818
43	62	72	52	47	88	48	56	65	72	52	56	68	57	86	71
1819	1820	1821	1822	1823	1824	1825	1826	1827	1828	1829	1830	1831	1832	1833	1834
54	88	89	67	81	62	88	73	61	81	80	53	36	86	44	70
1835	1836	1837	1838	1839	1840	1841	1842	1843	1844	1845	1846	1847	1848	1849	1850
78	62	78	63	54	112	73	91	93	79	63	77	102	83	77	52
1851	1852	1853	1854	1855	1856	1857	1858	1859	1860	1861	1862	1863	1864	1865	1866
78	78	77	81	85	67	110	68	77	99	77	71	71	72	82	76
1867	1868	1869	1870	1871	1872	1873	1874	1875	1876	1877	1878	1879	1880	1881	1882
95	77	102	106	57	62	72	76	64	83	88	73	76	73	56	80
1883	1884	1885	1886	1887	1888	1889	1890	1891	1892	1893	1894	1895	1896	1897	1898
48	68	62	60	74	94	92	83	67	58	69	83	60	93	77	65
1899	1900	1901	1902	1903	1904	1905	1906	1907	1908	1909	1910	1911	1912	1913	1914
73	70	77	67	39	92	67	58	50	60	55	58	74	89	60	69
1915	1916	1917	1918	1919	1920	1921	1922	1923	1924	1925	1926	1927	1928	1929	1930
86	90	62	53	78	86	81	85	95	78	77	88	72	112	81	71
1931	1932	1933	1934	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946
117	94	94	86	107	102	98	112	77	130	86	190	118	131	128	146
1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962
128	131	132	112	140	102	96	102	123	82	98	68	83	52	67	71
1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978
63	63	48	52	31	37	36	67	39	28	30	35	29	7	25	14
1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1997	
18	6	6	14	9	11	12	7	4	8	3	3	4	2	2	

```
provinces_clean <- provinces |> select(first_last,
                                     province,
                                     gender,
                                     year_of_birth,
                                     year_of_death,
                                     year,
                                     exit_year,
                                     party) |>
  filter(str_detect(province, "Quebec|Ontario"))
```

```
#vérification
provinces_clean |> select(province) |> table()
```

```
province
Ontario  Quebec
5592     5682
```

```
#ajout d'une nouvelle variable de cohorte générationnelle (gen_cohort) et (year_range)
#The Strauss-Howe Generational Theory
```

```
provinces_clean <- provinces_clean |> mutate(gen_cohort = case_when(
  year_of_birth >= 1701 & year_of_birth <= 1723 ~ "Awakening",
  year_of_birth >= 1724 & year_of_birth <= 1741 ~ "Liberty",
  year_of_birth >= 1742 & year_of_birth <= 1766 ~ "Republican",
  year_of_birth >= 1767 & year_of_birth <= 1791 ~ "Compromise",
  year_of_birth >= 1792 & year_of_birth <= 1821 ~ "Transcendental",
  year_of_birth >= 1822 & year_of_birth <= 1842 ~ "Gilded",
  year_of_birth >= 1843 & year_of_birth <= 1859 ~ "Progressive",
  year_of_birth >= 1860 & year_of_birth <= 1882 ~ "Missionary",
  year_of_birth >= 1883 & year_of_birth <= 1900 ~ "Lost",
  year_of_birth >= 1901 & year_of_birth <= 1924 ~ "G.I.",
  year_of_birth >= 1925 & year_of_birth <= 1942 ~ "Silent",
  year_of_birth >= 1943 & year_of_birth <= 1960 ~ "Babyboom",
  year_of_birth >= 1961 & year_of_birth <= 1981 ~ "GenerationX",
  year_of_birth >= 1982 & year_of_birth <= 2004 ~ "Millenial",
  year_of_birth >= 2005 & year_of_birth <= 2025 ~ "Homeland",
  TRUE ~ NA_character_)) |>
mutate(year_range = case_when(
  year_of_birth >= 1701 & year_of_birth <= 1723 ~ "1701-1723",
  year_of_birth >= 1724 & year_of_birth <= 1741 ~ "1724-1741",
  year_of_birth >= 1742 & year_of_birth <= 1766 ~ "1742-1766",
  year_of_birth >= 1767 & year_of_birth <= 1791 ~ "1767-1791",
  year_of_birth >= 1792 & year_of_birth <= 1821 ~ "1792-1821",
  year_of_birth >= 1822 & year_of_birth <= 1842 ~ "1822-1842",
  year_of_birth >= 1843 & year_of_birth <= 1859 ~ "1843-1859",
  year_of_birth >= 1860 & year_of_birth <= 1882 ~ "1860-1882",
  year_of_birth >= 1883 & year_of_birth <= 1900 ~ "1883-1900",
  year_of_birth >= 1901 & year_of_birth <= 1924 ~ "1901-1924",
  year_of_birth >= 1925 & year_of_birth <= 1942 ~ "1925-1942",
  year_of_birth >= 1943 & year_of_birth <= 1960 ~ "1943-1960",
  year_of_birth >= 1961 & year_of_birth <= 1981 ~ "1961-1981",
  year_of_birth >= 1982 & year_of_birth <= 2004 ~ "1982-2004",
  year_of_birth >= 2005 & year_of_birth <= 2025 ~ "2005-2025",
  TRUE ~ NA_character_
))
```

```
provinces_clean |> select(gen_cohort, year_range)|> table()
```

year\_range

gen_cohort	1701-1723	1724-1741	1742-1766	1767-1791	1792-1821	1822-1842
Awakening	2	0	0	0	0	0
Babyboom	0	0	0	0	0	0
Compromise	0	0	0	698	0	0
G.I	0	0	0	0	0	0
GenerationX	0	0	0	0	0	0
Gilded	0	0	0	0	0	1020
Liberty	0	42	0	0	0	0
Lost	0	0	0	0	0	0
Millenial	0	0	0	0	0	0
Missionary	0	0	0	0	0	0
Progressive	0	0	0	0	0	0
Republican	0	0	385	0	0	0
Silent	0	0	0	0	0	0
Transcendental	0	0	0	0	1140	0

	year_range					
gen_cohort	1843-1859	1860-1882	1883-1900	1901-1924	1925-1942	1943-1960
Awakening	0	0	0	0	0	0
Babyboom	0	0	0	0	0	1398
Compromise	0	0	0	0	0	0
G.I	0	0	0	1176	0	0
GenerationX	0	0	0	0	0	0
Gilded	0	0	0	0	0	0
Liberty	0	0	0	0	0	0
Lost	0	0	919	0	0	0
Millenial	0	0	0	0	0	0
Missionary	0	1235	0	0	0	0
Progressive	959	0	0	0	0	0
Republican	0	0	0	0	0	0
Silent	0	0	0	0	1270	0
Transcendental	0	0	0	0	0	0

	year_range	
gen_cohort	1961-1981	1982-2004
Awakening	0	0
Babyboom	0	0
Compromise	0	0
G.I	0	0
GenerationX	563	0
Gilded	0	0
Liberty	0	0
Lost	0	0
Millenial	0	53
Missionary	0	0

Progressive	0	0
Republican	0	0
Silent	0	0
Transcendental	0	0

## NA year\_of\_birth

```
# données : provinces_clean NA

provinces_clean_na <- provinces_clean |>
  select(first_last, year_of_birth, gen_cohort, year) |>
  filter(is.na(year_of_birth)) |>
  filter(year >= 1987) |>
  distinct(first_last, .keep_all = TRUE)

summary(provinces_clean_na$year)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1990	2011	2018	2013	2018	2018

```
dim(provinces_clean_na)
```

```
[1] 62 4
```

```
# limits : J'ai 62 députés dont je ne connais pas l'année de naissance et je ne peux pas v
#png : NA_year_of_birth
```

## Variable : years\_total, sessions\_total

```
#créer une nouvelle colonne combien d'années de service et de sessions

provinces_clean <- provinces_clean |>
  mutate(years_total = exit_year - year + 1,
         sessions_total = years_total * 2)

provinces_clean |>
  filter(str_detect(first_last, "jacques_parizeau|chiarelli_bob|denis_lazure"))
```



	first_last	province	gender	year_of_birth	year_of_death	year	exit_year
1	chiarelli_bob	Ontario	m	1941	NA	1987	1997
2	chiarelli_bob	Ontario	m	1941	NA	1990	1997
3	chiarelli_bob	Ontario	m	1941	NA	1995	1997
4	chiarelli_bob	Ontario	m	1941	NA	2007	2018
5	chiarelli_bob	Ontario	m	1941	NA	2011	2018
6	chiarelli_bob	Ontario	m	1941	NA	2014	2018
7	denis_lazure	Quebec	m	1925	2008	1976	1984
8	denis_lazure	Quebec	m	1925	2008	1981	1984
9	denis_lazure	Quebec	m	1925	2008	1989	1996
10	denis_lazure	Quebec	m	1925	2008	1994	1996
11	jacques_parizeau	Quebec	m	1930	2015	1976	1984
12	jacques_parizeau	Quebec	m	1930	2015	1981	1984
13	jacques_parizeau	Quebec	m	1930	2015	1989	1996
14	jacques_parizeau	Quebec	m	1930	2015	1994	1996

	party	gen_cohort	year_range	years_total	sessions_total
1	Liberal	Silent	1925-1942	11	22
2	Liberal	Silent	1925-1942	8	16
3	Liberal	Silent	1925-1942	3	6
4	Liberal	Silent	1925-1942	12	24
5	Liberal	Silent	1925-1942	8	16
6	Liberal	Silent	1925-1942	5	10
7	Parti Quebecois	Silent	1925-1942	9	18
8	Parti Quebecois	Silent	1925-1942	4	8
9	Parti Quebecois	Silent	1925-1942	8	16
10	Parti Quebecois	Silent	1925-1942	3	6
11	Parti Quebecois	Silent	1925-1942	9	18
12	Parti Quebecois	Silent	1925-1942	4	8
13	Parti Quebecois	Silent	1925-1942	8	16
14	Parti Quebecois	Silent	1925-1942	3	6

```
summary(provinces_clean$year)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1792	1875	1929	1925	1977	2018

```
# Check how many NA in exit_year
```

```
serving_1 <- provinces_clean |>
  select(first_last, year, exit_year) |>
  filter(is.na(exit_year)) |>
```

```

    filter(year <= 1980)

# J'ai 11 députés qui n'ont pas d'année de fin (1871-1921), I should eliminate them

provinces_clean <- provinces_clean |>
  filter(!(is.na(exit_year) & year < 1980))

# verification

serving_1 <- provinces_clean |>
  select(first_last, year, exit_year) |>
  filter(is.na(exit_year)) |>
  filter(year <= 1980)

summary(provinces_clean$year)

```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1792	1875	1929	1925	1977	2018

## NA exit\_year - 2024

```

#NA dans exit_year correspond à ceux qui sont toujours à la Chambre des communes, je prend

# NA qui reste dans "exit_year"

serving_2 <- provinces_clean |>
  select(first_last, year, exit_year) |>
  filter(is.na(exit_year)) |>
  filter(year >= 1961)

dim(serving_2)

```

```
[1] 505    3
```

```
summary(serving_2$year)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1990	2011	2014	2014	2018	2018

```
#NA dans exit_year correspond à ceux qui sont toujours à la Chambre des communes, je prend
#test_Chat_GPT
```

```
current_year <- as.numeric(format(Sys.Date(), "%Y"))
```

```
provinces_clean <- provinces_clean |>
  group_by(first_last) |>
  mutate(start_year = min(year[!is.na(year)])) |>
  ungroup() |>
  mutate(exit_year = if_else(year == start_year & is.na(exit_year) & gen_cohort != "Silent
  group_by(first_last) |>
  mutate(years_total = exit_year - start_year,
         sessions_total = years_total * 2) |>
  ungroup()
```

```
#verification
```

```
provinces_clean |> select(years_total) |> summary()
```

```
years_total
Min.   : -1.00
1st Qu.:  6.00
Median : 10.00
Mean   : 12.33
3rd Qu.: 17.00
Max.   :144.00
NA's   :302
```

```
provinces_clean |> select(first_last, year, exit_year, years_total) |>
  filter(years_total == "-1" | years_total == "75")
```

```
# A tibble: 5 x 4
  first_last    year exit_year years_total
  <chr>        <int>    <dbl>    <dbl>
1 carrere_john  1949     1948      -1
2 cragg_charles 1949     1948      -1
3 John White    1857     1867      75
4 John White    1861     1867      75
5 John White    1863     1867      75
```

```
provinces_clean |> filter(first_last == "John White", .keep_all = TRUE)
```

```
# A tibble: 5 x 13
```

```
  first_last province gender year_of_birth year_of_death  year exit_year party
  <chr>      <chr>   <chr>      <int>          <int> <int>      <dbl> <chr>
1 John White Ontario m          1761          1800  1792        1796 <NA>
2 John White Ontario m          1811          1897  1851        1854 <NA>
3 John White Ontario m          1811          1897  1857        1867 Reform~
4 John White Ontario m          1811          1897  1861        1867 Reform~
5 John White Ontario m          1811          1897  1863        1867 Reform~
# i 5 more variables: gen_cohort <chr>, year_range <chr>, years_total <dbl>,
#   sessions_total <dbl>, start_year <int>
```

```
summary(provinces_clean$year)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1792	1875	1929	1925	1977	2018

```
#Turns out I have two John Whites, I need to rename one
```

```
provinces_clean <- provinces_clean |>
  mutate(first_last = if_else(first_last == "John White" &
    year_of_birth == 1761 &
    year_of_death == 1800, "John_White", first_last))

provinces_clean |> filter(first_last == "John White", .keep_all = TRUE)
```

```
# A tibble: 4 x 13
```

```
  first_last province gender year_of_birth year_of_death  year exit_year party
  <chr>      <chr>   <chr>      <int>          <int> <int>      <dbl> <chr>
1 John White Ontario m          1811          1897  1851        1854 <NA>
2 John White Ontario m          1811          1897  1857        1867 Reform~
3 John White Ontario m          1811          1897  1861        1867 Reform~
4 John White Ontario m          1811          1897  1863        1867 Reform~
# i 5 more variables: gen_cohort <chr>, year_range <chr>, years_total <dbl>,
#   sessions_total <dbl>, start_year <int>
```

```
# copy the code after renaming John White to John_White
```

```
provinces_clean <- provinces_clean |>
  group_by(first_last) |>
  mutate(start_year = min(year[!is.na(year)])) |>
  ungroup() |>
  mutate(exit_year = if_else(year == start_year & is.na(exit_year) & gen_cohort != "Silent",
    start_year, exit_year) |>
  group_by(first_last) |>
  mutate(years_total = exit_year - start_year,
    sessions_total = years_total * 2) |>
  ungroup()
```

```
provinces_clean |> filter(first_last == "John White", .keep_all = TRUE)
```

```
# A tibble: 4 x 13
```

	first_last	province	gender	year_of_birth	year_of_death	year	exit_year	party
	<chr>	<chr>	<chr>	<int>	<int>	<int>	<dbl>	<chr>
1	John White	Ontario	m	1811	1897	1851	1854	<NA>
2	John White	Ontario	m	1811	1897	1857	1867	Reform~
3	John White	Ontario	m	1811	1897	1861	1867	Reform~
4	John White	Ontario	m	1811	1897	1863	1867	Reform~

```
# i 5 more variables: gen_cohort <chr>, year_range <chr>, years_total <dbl>,
# sessions_total <dbl>, start_year <int>
```

```
#nettoyage d'observations -1 en "years_total" et <100 "years_total"
```

```
#andre_peltier something weird is happening with that observation. I don't know, how to so
```

```
provinces_clean <- provinces_clean |>
  filter(!str_detect(years_total, "-1")) |>
  #filter(str_detect(gen_cohort, "Silent|Babyboom|GenerationX|Millenial")) |>
  filter(!str_detect(first_last, "andre_pelletier"))
```

```
provinces_clean |> select(gen_cohort) |> table()
```

```
gen_cohort
```

Awakening	Babyboom	Compromise	G.I	GenerationX
2	1278	698	1176	445
Gilded	Liberty	Lost	Millenial	Missionary

1018	42	917	48	1234
Progressive	Republican	Silent	Transcendental	
959	385	1262	1135	

```
provinces_clean |> select(years_total) |> summary()
```

```
years_total
Min.   : 0.0
1st Qu.: 6.0
Median : 10.0
Mean    : 12.3
3rd Qu.: 17.0
Max.    :144.0
```

```
provinces_clean |> filter(years_total == "42")
```

```
# A tibble: 37 x 13
```

	first_last	province	gender	year_of_birth	year_of_death	year	exit_year	party
	<chr>	<chr>	<chr>	<int>	<int>	<int>	<dbl>	<chr>
1	nixon_harry	Ontario	m	1891	1961	1923	1961	Unit~
2	nixon_harry	Ontario	m	1891	1961	1926	1961	Prog~
3	nixon_harry	Ontario	m	1891	1961	1929	1961	Prog~
4	nixon_harry	Ontario	m	1891	1961	1934	1961	Libe~
5	nixon_harry	Ontario	m	1891	1961	1937	1961	Libe~
6	nixon_harry	Ontario	m	1891	1961	1943	1961	Libe~
7	nixon_harry	Ontario	m	1891	1961	1945	1961	Libe~
8	nixon_harry	Ontario	m	1891	1961	1949	1961	Libe~
9	nixon_harry	Ontario	m	1891	1961	1951	1961	Libe~
10	nixon_harry	Ontario	m	1891	1961	1955	1961	Libe~

```
# i 27 more rows
```

```
# i 5 more variables: gen_cohort <chr>, year_range <chr>, years_total <dbl>,
```

```
# sessions_total <dbl>, start_year <int>
```

```
summary(provinces_clean$year)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1792	1875	1927	1922	1976	2018

```

# I will try to choose the unique values of "years_total" for each deputy with the start y

provinces_clean_unique <- provinces_clean |>
  distinct(first_last, start_year, years_total, sessions_total, .keep_all = TRUE)

# verification

provinces_clean_unique |> select(first_last) |> table()

# I still have someone who is mentioned several times

provinces_clean_unique |>
  filter(str_detect(first_last, "jacques_parizeau|chiarelli_bob|denis_lazure"))

# verification
# It turns out that I have 68 deputies for whom the year of release is written differently
# year and exit_year 2.png

name_counts <- provinces_clean_unique |>
  group_by(first_last) |>
  summarise(count = n(), .groups = 'drop') |>
  filter(count > 1) |> distinct()

repeated_names <- provinces_clean_unique |>
  filter(first_last %in% name_counts$first_last)

# so, I'll eliminate these observations

provinces_clean_unique <- provinces_clean_unique |>
  anti_join(repeated_names, by = "first_last")

# verification

provinces_clean_unique |> select(years_total) |> summary()

```

```

years_total
Min.   : 0.000
1st Qu.: 4.000
Median : 6.000
Mean   : 7.746

```

3rd Qu.:10.000  
Max. :42.000

```
repeated_names <- provinces_clean_unique |>  
  filter(first_last %in% name_counts$first_last)  
  
provinces_clean_unique |> select(gen_cohort) |> table()
```

gen_cohort				
Awakening	Babyboom	Compromise	G.I	GenerationX
2	503	217	394	263
Gilded	Liberty	Lost	Millenial	Missionary
399	28	307	43	412
Progressive	Republican	Silent	Transcendental	
341	147	475	409	

```
# filter by "year" from 1961 when first boomers were 18  
  
provinces_clean_unique_1961 <- provinces_clean_unique |>  
  filter(year >= 1961) |>  
  filter(str_detect(gen_cohort, "Silent|Babyboom|GenerationX|Millenial"))
```

```
#test verification  
  
summary(provinces_clean_unique$year)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1792	1875	1926	1922	1976	2018

## Visualisation\_1

```
# visualisation_1  
# Le graphique montre que le temps passé à la Chambre législative n'est pas réparti également  
  
class(provinces_clean_unique$year_of_birth)  
  
dim(provinces_clean_unique)
```



```

graph_1 <- ggplot(data = provinces_clean_unique, aes(x = year_of_birth)) +
  geom_histogram(binwidth = 1,
                 color = "black",
                 fill = "orange",
                 alpha = 0.5) +
  geom_hline(yintercept = mean(provinces_clean_unique$sessions_total, na.rm = TRUE),
             linetype = "dashed", color = "black") +
  scale_x_continuous(limits = c(1867, 2000),
                     breaks = c(1867, seq(1890, 2000, by = 10))) +
  labs(title = "Représentation des générations de cohortes basée sur la date de naissance",
       subtitle = "de l'Ontario et du Québec",
       x = "Année de naissance",
       y = "Sessions servies") +
  theme_minimal() +
  theme(panel.grid.major.x = element_line(colour = "grey70", size = .2),
        panel.grid.minor.x = element_line(colour = "grey70", size = .1),
        panel.grid.major.y = element_line(colour = "grey70", size = .2),
        panel.grid.minor.y = element_blank(),
        axis.line = element_line(colour = "black",
                                  size = .1),
        axis.text.x = element_text(angle = 45,
                                    vjust = 0.5),
        plot.title = element_text(size = 12,
                                   hjust = 0.5),
        plot.subtitle = element_text(size = 10,
                                      hjust = 0.5),
        text = element_text(face = "plain")
  )

```

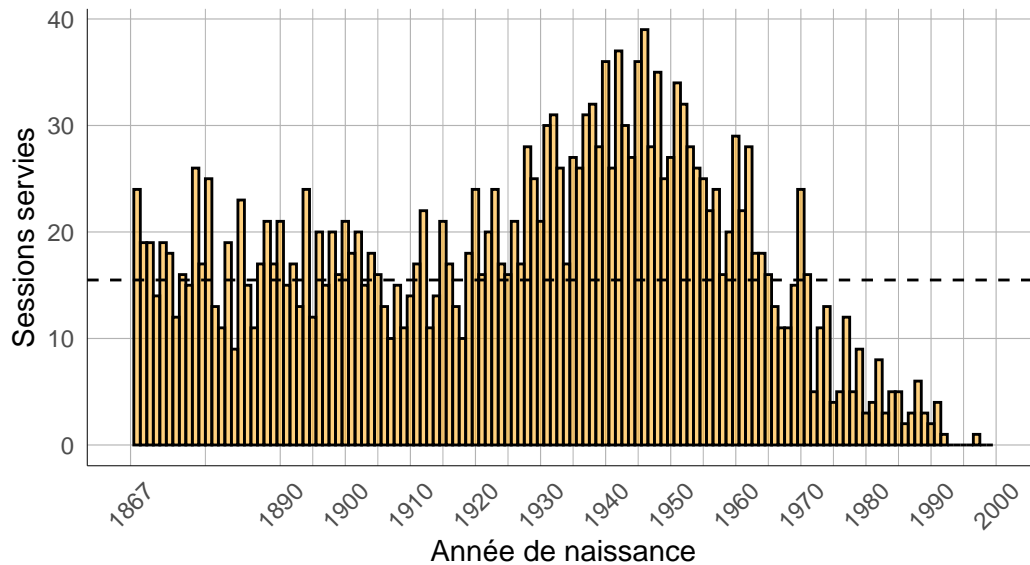
Warning: The `size` argument of `element\_line()` is deprecated as of ggplot2 3.4.0.  
 i Please use the `linewidth` argument instead.

graph\_1

Warning: Removed 1922 rows containing non-finite outside the scale range  
 (`stat\_bin()`).

Warning: Removed 2 rows containing missing values or values outside the scale range  
 (`geom\_bar()`).

Représentation des générations de cohortes basée sur la date de naissance  
de l'Ontario et du Québec



## Visualisation\_2

```
# Divisé entre le Québec et l'Ontario
```

```
graph_2 <- ggplot(data = provinces_clean_unique, aes(x = year_of_birth)) +
  geom_histogram(binwidth = 1,
    color = "black",
    fill = "orange",
    alpha = 0.5) +
  geom_hline(yintercept = mean(provinces_clean_unique$sessions_total, na.rm = TRUE),
    linetype = "dashed", color = "black") +
  facet_wrap(~ province, ncol = 1) +
  scale_x_continuous(limits = c(1867, 2000),
    breaks = c(1867, seq(1890, 2000, by = 10))) +
  labs(title = "Représentation des générations de cohortes basée sur la date de naissance",
    subtitle = "",
    x = "Année de naissance",
    y = "Sessions servies") +
  theme_minimal() +
  theme(panel.grid.major.x = element_line(colour = "grey70", size = .2),
```

```

panel.grid.minor.x = element_line(colour = "grey70", size = .1),
panel.grid.major.y = element_line(colour = "grey70", size = .2),
panel.grid.minor.y = element_blank(),
axis.line = element_line(colour = "black",
                          size = .1),
axis.text.x = element_text(angle = 45,
                            vjust = 0.5),
plot.title = element_text(size = 12,
                           hjust = 0.5),
plot.subtitle = element_text(size = 10,
                              hjust = 0.5),
text = element_text(face = "plain")
)

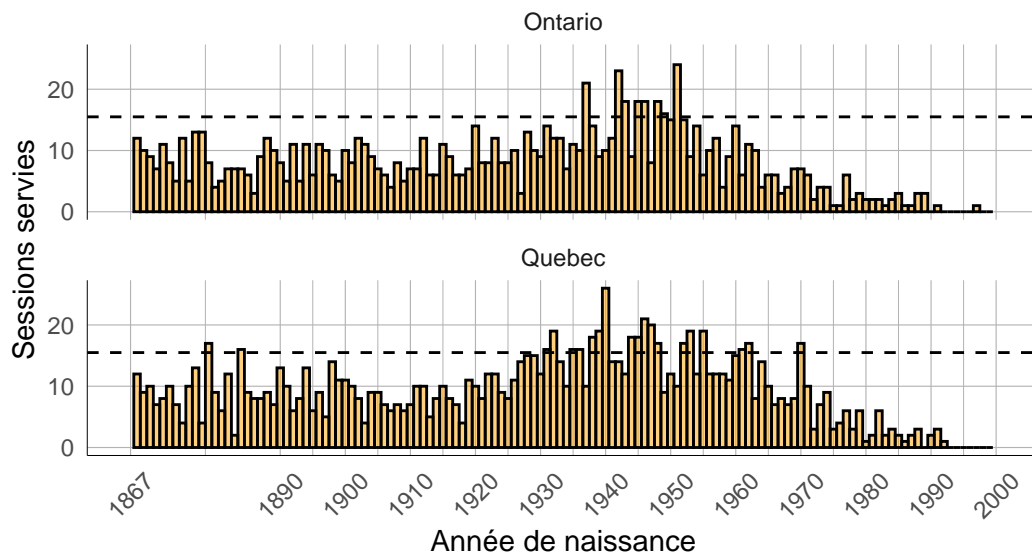
```

graph\_2

Warning: Removed 1922 rows containing non-finite outside the scale range (`stat\_bin()`).

Warning: Removed 4 rows containing missing values or values outside the scale range (`geom\_bar()`).

## Représentation des générations de cohortes basée sur la date de naiss



## Données : legislative\_age, variable age\_start

```
#créer de nouvelles variables, pour comprendre à quel âge et dans quelle cohorte générationnelle on se trouve

legislative_age <- provinces_clean_unique_1961 |>
  mutate(age_start = start_year - year_of_birth) |>
  drop_na(age_start, year_of_birth)

legislative_age |> glimpse()
legislative_age |> select(age_start) |> summary()

legislative_age |> filter(age_start == "-128")

# error : Donald McDonald year_of_birth 1969, year 1841, exit_year 1844, I also have to el

legislative_age <- provinces_clean_unique_1961 |>
  mutate(age_start = start_year - year_of_birth) |>
  drop_na(age_start, year_of_birth) |>
  filter(first_last != "Donald McDonald")

legislative_age |> select(age_start) |> summary()

legislative_age |> filter(age_start >= -126 & age_start <= 0)

#error :I have 11 observations where birth_year does not correctly match start_year and ex

legislative_age <- provinces_clean_unique_1961 |>
  mutate(age_start = start_year - year_of_birth) |>
  drop_na(age_start, year_of_birth) |>
  filter(!(age_start >= -128 & age_start <= 0)) |>
  mutate(gen_cohort = factor(gen_cohort,
                             levels = c("Silent",
                                           "Babyboom",
                                           "GenerationX",
                                           "Millenial")))

legislative_age |> select(age_start) |> summary()

legislative_age |> filter(age_start == "71")
```

```

provinces |> filter(first_last == "kramp_daryl")
# wow, Daryl Kramp was elected to the Ontario Legislative Assembly at the age of 71! never

dim(legislative_age)

legislative_age |> filter(age_start < "20") |> filter(province == "Ontario")

```

### Visualisation\_3

```

# at what age deputies started to work, by generational cohort

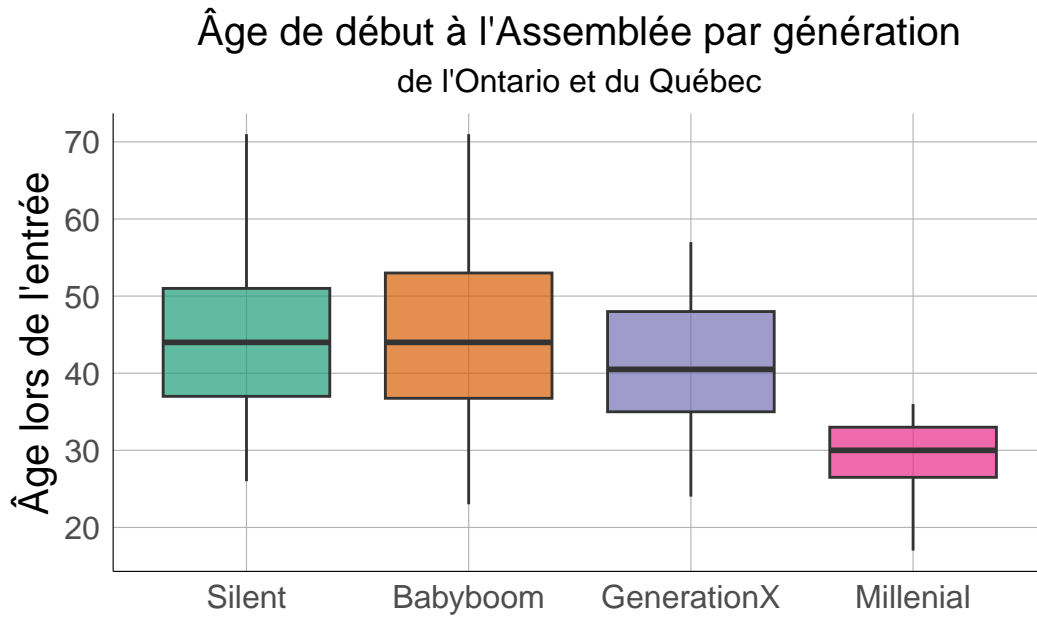
library(RColorBrewer)

graph_3 <- ggplot(data = legislative_age, aes(x = gen_cohort,
                                              y = age_start,
                                              fill = gen_cohort)) +

  geom_boxplot(alpha = .7) +
  scale_fill_brewer(palette = "Dark2") +
  labs(title = "Âge de début à l'Assemblée par génération",
       subtitle = "de l'Ontario et du Québec",
       x = "",
       y = "Âge lors de l'entrée") +
  theme_minimal() +
  theme(legend.position = "none",
        panel.grid.major.x = element_line(colour = "grey70", size = .2),
        panel.grid.minor.x = element_line(colour = "grey70", size = .1),
        panel.grid.major.y = element_line(colour = "grey70", size = .2),
        panel.grid.minor.y = element_blank(),
        axis.line = element_line(colour = "black",
                                  size = .2),
        plot.title = element_text(size = 15,
                                   hjust = 0.5),
        plot.subtitle = element_text(size = 12,
                                       hjust = 0.5),
        text = element_text(face = "plain", size = 15))

graph_3

```



**Visualisation\_3\_2 :**

```
#graph Ontario Quebec

graph_3_2 <- ggplot(data = legislative_age, aes(x = gen_cohort,
                                              y = age_start,
                                              fill = gen_cohort)) +

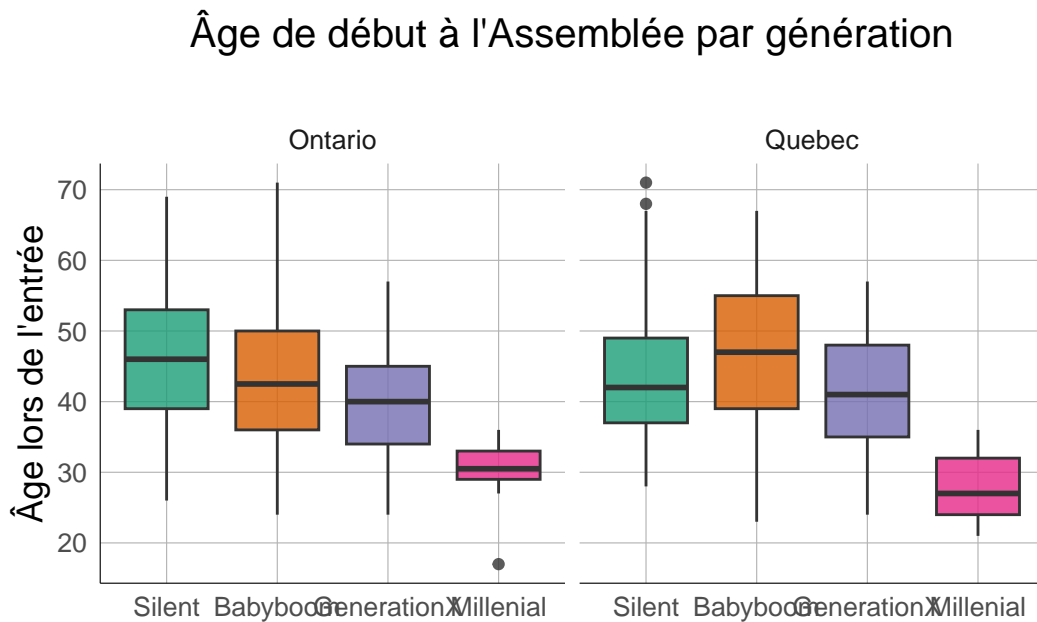
  geom_boxplot(alpha = .8) +
  facet_wrap(~ province) +
  scale_fill_brewer(palette = "Dark2") +
  labs(title = "Âge de début à l'Assemblée par génération",
       subtitle = "",
       x = "",
       y = "Âge lors de l'entrée") +
  theme_minimal() +
  theme(legend.position = "none",
        panel.grid.major.x = element_line(colour = "grey70", size = .2),
        panel.grid.minor.x = element_line(colour = "grey70", size = .1),
        panel.grid.major.y = element_line(colour = "grey70", size = .2),
        panel.grid.minor.y = element_blank(),
        axis.line = element_line(colour = "black",
```

```

        size = .2),
plot.title = element_text(size = 15,
                           hjust = 0.5),
plot.subtitle = element_text(size = 12,
                              hjust = 0.5),
text = element_text(face = "plain", size = 13))

```

graph\_3\_2



#### Visualisation\_4:

```

# graph_total_years

graph_4 <- ggplot(data = legislative_age, aes(x = gen_cohort, y = years_total, fill = gen_cohort)) +
  geom_violin(trim=FALSE, alpha = .7, draw_quantiles = c(0.25, 0.5, 0.75)) +
  #facet_wrap(~ province, ncol = 1) +
  scale_fill_brewer(palette = "Dark2") +
  labs(title = "Longévité des cohortes générationnelles à l'Assemblée",
        subtitle = "Quebec et Ontario",

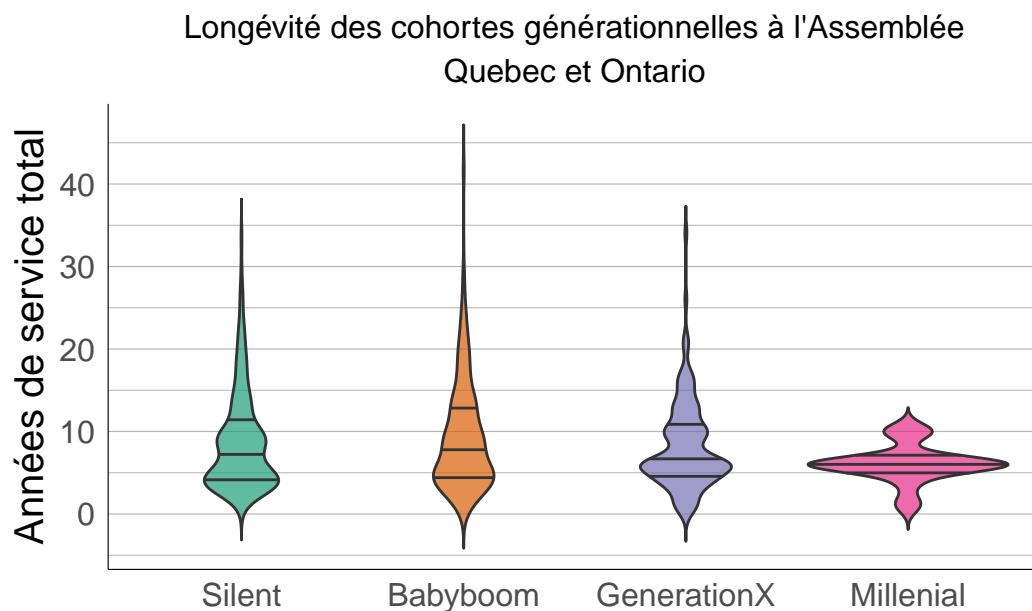
```

```

x = "",
y = "Années de service total") +
theme_minimal() +
theme(legend.position = "none",
      panel.grid.major.x = element_blank(),
      panel.grid.minor.x = element_line(colour = "grey70", size = .1),
      panel.grid.major.y = element_line(colour = "grey70", size = .2),
      panel.grid.minor.y = element_line(colour = "grey70", size = .1),
      axis.line = element_line(colour = "black",
                               size = .2),
      plot.title = element_text(size = 12,
                                hjust = 0.5),
      plot.subtitle = element_text(size = 12,
                                    hjust = 0.5),
      text = element_text(face = "plain", size = 15))

```

graph\_4





## Bibliography notes :

The Strauss-Howe Generational Theory : for description of ages division

- Amanda van Eck Duymaer van Twist and Suzanne Newcombe. “Strauss-Howe Generational Theory.” In James Crossley and Alastair Lockhart (eds.) *Critical Dictionary of Apocalyptic and Millenarian Movements*. 15 January 2021. Retrieved from [www.cdamm.org/articles/strauss-howe](http://www.cdamm.org/articles/strauss-howe). (First published 12 June 2017 [censamm.org/resources/profiles/strauss-howe-generational-theory](http://censamm.org/resources/profiles/strauss-howe-generational-theory).)
- Banerjee, Sidhartha. 2022. «CAQ Leader Rejects Electoral Reform on Campaign Trail, Says He Will Work with Other Quebec Parties». CBC News. 27 septembre 2022. Consulté le 10 mars 2024. <https://www.cbc.ca/news/canada/montreal/legault-electoral-reform-1.6597622>.
- Carmichael, Kelly. 2017. «Broken Trust on Electoral Reform». Policy Options. Consulté le 10 mars 2024. <https://policyoptions.irpp.org/magazines/february-2017/broken-trust-on-electoral-reform/>.
- The Globe and Mail. 2023. «Globe Editorial: Budget 2023: The Growing Generation Gap between What Ottawa Spends on Older and Younger Canadians», 20 mars 2023. Consulté le 10 mars 2024. <https://www.theglobeandmail.com/opinion/editorials/article-mind-the-growing-generation-gap/>.
- Kershaw, Paul, et Lynell Anderson. 2016. «Measuring the age distribution in Canadian social spending». *Canadian Public Administration* 59 (4): 556-79. <https://doi.org/10.1111/capa.12193>.
- Kishchuk, Oksana. 2023. «Gen Z - Top Issues Facing Canada». Abacus Data (blog). 31 août 2023. Consulté le 10 mars 2024. <https://abacusdata.ca/genz-top-issues-facing-canada/>.
- Munger, Kevin M. 2022. *Generation Gap: Why the Baby Boomers Still Dominate American Politics and Culture*. New York: Columbia University Press.
- Parliament of Canada. «Parlinfo». Consulté le 10 mars 2024. [https://lop.parl.ca/sites/ParlInfo/default/en\\_CA/People/parliamentarians](https://lop.parl.ca/sites/ParlInfo/default/en_CA/People/parliamentarians).
- Rands, Chris. 2017. «Do You Think Kellie Leitch Should Have Her Own Party?» *Trudeau Asks Electoral Reform Advocate*. CBC News. 10 février 2017. Consulté le 10 mars 2024. <https://www.cbc.ca/news/politics/electoral-reform-trudeau-leitch-1.3975354>.
- Rivard, Alexandre, Marc André Bodet, Jean-François Godbout, et Éric Montigny. 2024. «Provincial Parliamentary Biographies». Harvard Dataverse. <https://doi.org/10.7910/DVN/LZPEFQ>.

- Vallée-Dubois, Florence. 2023. «Government Spending Preferences over the Life Cycle». *Journal of Public Policy* 43 (3): 468-89. <https://doi.org/10.1017/S0143814X23000065>.
- McClean, Charles T. 2019. “Does it matter that politicians are older than their constituents? Yes.” Manuscript, UC San Diego, downloaded from <https://www.charlesmcclean.com/research> (Accessed 3 February 2021)