

Canada's Labour Force - Analysis

Sherry Luo and Tony Wang

17/12/2019

Part I: Job Tenure in Canada

For this project, we conduct an exploratory analysis on job tenure data collected from Statistics Canada from 1976-2018. Though there are many ways to look at employment data, with job security and the so-called “gig economy” as a topic of discussion in recent years Katz and Krueger (2016), we wish to start by seeing how the length of employment of full-time workers has changed over time in Canada by gender and age. Job tenure measures the length of time an employee has been employed by their employer. For this analysis, job tenure was categorized into three groups: short-term (less than 1 year of employment), intermediate (1 to 5 years) and long-term (more than 5 years).

Back-to-back Pyramid Plot

Story:

We start the analysis with an animated pyramid plot to compare changes in job tenure by gender from 1976-2018. It appears that job tenure has increased for both genders, and a higher number of men appear in each of three groups compared to women. Additionally, changes in job tenure also depends on financial state of Canada. Generally, when the economy is stable, the number of short-term jobs increase and transitional stages (moving from short-term to intermediate and intermediate to long-term positions) follow accordingly. However, when a financial crisis strikes such as the recession in 2008-2012, the number of short-term jobs decrease, but the transition stages in both the intermediate and long-term group remain relatively unchanged (i.e. layoffs affect workers with lower seniority more compared to workers with higher seniority). The pyramid plot is a nice visual but comparing the distributions of men and women is not obvious and difficult to do. But due to the animated component, each year can only be viewed individually not collectively (hence the possibility of being wrong, as we will see).

Method:

The cansim package allowed us to easily obtain the data from Statistics Canada. For cleaning, we filtered the necessary data and relabel/order the age and job tenure brackets for clarity. To create a symmetric

back-to-back plot, the female values were multiplied by negative one. In terms of design, we selected high contrast colors to compare males and females and added direct labelling. However, as mentioned previously, comparing the distribution over time was not easy. Thus it may have been better to overlay the distributions rather than back-to-back as suggested by Dr.Bolker. The resulting plot will be less attractive but will be more informative for comparison. For more control on playback, we saved the plot as a MP4 and if played continuously each frame is played for 2 seconds (fps=2). Finally, for visual purposes we assumed the bin width to be equal for each job tenure bracket.

Average Tenure Plots

Story:

We created three plots based on the average job tenure. In the first plot, we examined the average tenure by gender and more detailed trends are presented. We see a huge gap between men and women but with time this gap narrowed around the early 2000s due to more women entering the work force around the early 1990s. There are more men in the workforce as indicated by similarity in trend for “males” and “both sex”. Comparing men and women, the average tenure for women increased steadily over time but decreased for men after 1996. However, this plot may be misleading as it does not consider the various age groups of each gender. To address this problem, it is best to separate the age groups for each gender. In this second plot, the average job tenure by gender and age are examined. For women, we see that for all age groups except 15 to 24 year old age brackets, there is a slight increase in the average job tenure over time. For men aged 65 years or more, job tenure increases until 1996 and substantially decreases afterwards. We also see a similar trend in 55 to 64 year old males, though the decrease is very marginal compared to 65 years and older. This slightly contradicts our observation from before since job tenure actually decreased for men after 1996. One issue that arises is cohort size for each age group. We tried to indicate cohort size by varying the line width for second plot. We see that majority of the workforce is comprised of workers that are 55 years and older.

Method:

The two interactive plots were created with plotly. For both plots, we opted for high contrast colors and used the same color palette for aesthetic consistency while trying to keep the plots accessible to the visually impaired. In the first plot, average tenure for men was represented in blue and dark pink for females. We felt purple would be more visually obvious to represent “the halfway” since average tenure for both sexes lie between males and females. In the second plot, the coloring scheme highlights the ordering of the age categories. To determine the line width in the third plot, we summed the overall cohort size for each age group and gender. For all the years, 65 years and plus had the largest cohort size, followed by 55 to 64, 25 to

54 years old and then 15 to 24 years old had the smallest cohort size. Then for each cohort size, we manually specify a line width.

Facets were used to allow a side-by-side comparison of trends for each gender and age group. We used a linear axis as the discrepancies between groups were substantial and a log-scale would have been rather deceptive when comparing the differences between groups. Additionally, units were rounded to tenth of the year as years is more comprehensible to the general audience. The interactive elements allowed the data to be explored in better detail without adding excessive clutter to the plots. ## Sunburst Plot

Story:

Finally, we conclude the analysis with a sunburst plot. Initially, cumulatively from 1976-2018, the workforce is 40% female and 60% male. As we move upwards to each ring level, we can isolate and analyze specific combinations of the data. With this ordering, we can explore the age composition for each sex at different job tenure levels for different years. For example, among female workers who have been at their job for 1 to 5 years in 2009, majority are between 25 and 54 years old. This scenario can also be examined for men and we see that in 2009, male workers who have been at their job for 1-5 years are mostly between 25 and 54 years old, mirroring the trend that we saw in female workers.

Method:

Sunburst plot was created with the sunburstR package. As Dr.Bolker has stated, these plots are heavily dependent on the ordering of the variables. To suit our analysis above, we selected the ordering to be gender, job tenure, specific year then age group. We added a tooltip that allows a user to select a specific wedge of the plot and the breakdown can be seen on the right. We did consider rounding the numeric values in the tooltip, however, customizing it required knowledge of javascript. The d3.js categorical color scheme was used here.

Possible Explanations for Observed Trend

Overall, we see that job tenures are increasing for females, whereas job tenure for males is decreasing. Several possible reasons for these trends are:

- Better workplace accommodations for women such as maternity leave, as well as possible social shifts on the role of men in parenting.
- The unstable yet highly lucrative nature of work in the technology sector, which is heavily male-dominated, may partially explain the trend towards shorter job tenures for males, who may hop from one startup company to another for short periods of time before making enough money to exit the traditional full-time workforce.
- Globalization may result in highly experienced workers having greater access

to lucrative job positions in other countries, which would result in an outflux of workers in older age brackets.
-People from the silent generation may have come to retirement age in large numbers from 1997 onwards

Part II: Gender Pay Gap Analysis

For the second part of the analysis, we will look at some specific aspects of the workforce in recent years. One topic that frequently comes up when people discuss labour data is the gender pay gap, which can generally be described as a discrepancy between how much men and women get paid for comparable work. Everyone has heard about the 70 cents to the dollar saying, but considering how people's pay differ depending on fields of study, education levels, region, and many other features, it is beneficial to take a precursory glance at these sorts of variations to see how things actually work.

Geospatial Facet Plot

Story:

We begin by looking at the gender pay gap through the creation of an interactive, multi-faceted geospatial plot, coloured by the predominant field for each workforce cohort. From this plot, we see that:

- At the College-level and below, business is the predominant field of study for all provinces and Yukon, and the trades/natural resources is the most common field in the NorthWest Territories and Nunavut.
- At the University Certificate level, the most populous provinces in Canada (Ontario, Quebec, British Columbia and Alberta) have business as the most common field of study, with the rest of the provinces and territories having Education as the most common field.
- At the Bachelor's degree level, Social sciences become the most common field in Ontario, British Columbia and the Yukon, and Business becomes the most common field of study in New Brunswick, but for all other provinces, the field of study remains the same as the University Certificate level.
- Geographical Industry trends at the Master's degree level largely resemble those at the University Certificate level, with the exception of New Brunswick and Manitoba
- At the Doctorate level, sciences are the single most common field of study in every province and territory; however, it is interesting to note that this is for the physical sciences (i.e. biology, chemistry, and physics), meaning that this does not include mathematics and computer sciences.

Now that we have an idea of what the most common fields of study are in different provinces and territories, we can look at the gender balance within each of these fields as well as any gender pay gaps. They can be briefly summarized as follows:

- Where Business and administration is the predominant field, females generally make up between 40 to 80 percent of that particular cohort, with the proportion of females generally declining as the education level goes up. There appears to be a universal gender pay gap across all provinces that goes anywhere between 7% (minimum value, in Prince Edward Island at the Bachelor’s level) to around 24% (as observed in Alberta at the Master’s degree level).
- In provinces and territories where Education is the most common field of study, females make up at least 60% of that particular cohort. For people with Master’s degrees, Nunavut and Prince Edward Island have slightly greater median pay levels for females than males, though the difference is very small. In other areas where education is the most common field, the median salary of males is greater than that of females, though this difference ranges from slight (Nova Scotia, Master’s degree level) to substantial (e.g. Saskatchewan, University certificate level).
- Science disciplines are the predominant field of study at the doctorate level across all genders. In all provinces, females make up between 30 to 45 percent of the Science PhD cohort, and with the exception of Nova Scotia, males have a greater median salary in all provinces than females. The territories have so few Sciences PhDs that with the exception of males in the Northwest Territories, Statistics Canada actually does not provide wage data. No inferences can be made for these regions given the small number of relevant individuals.
- Social sciences are predominant at the Bachelor’s degree level for only two provinces and one territory. These cohorts are roughly 60% female and there appears to be a consistent and rather sizable gender pay gap in favour of males across all three of these jurisdictions.
- The trades, services, natural resources and conservation fields have very low female representation (roughly 30%), and a pay gap of either 10 percent (Nunavut) or 25 percent (Northwest Territories) in favor of males.

Method:

To look at gender pay gaps, we combined industry and wage data obtained from the 2016 Census into an interactive, multi-faceted geospatial plot, coloured by the predominant field for each workforce cohort. The data we used comes from two datasets of the 2016 Census, one of which covered wages of individuals by gender and field of study (Catalogue No. 98-400-X2016280 Statistics Canada (2019)), and one of which tallied the number of employees by various factors such as field of study (Catalogue No. 98-402-X2016010-T4, Statistics Canada (2017)), gender, and earned qualifications. The wage data from statistics Canada was extracted from the source data, a 6GB CSV file, using a workstation desktop with the necessary system resources to do so, and all plots were used on the reduced dataset to reduce system resource usage. These datasets were combined and used to create a five facet geospatial plot using the simplified digital boundary files. The use of digital boundary files rather than detailed cartographic files accelerates the rendering process

and reduces the final file size, albeit at the cost of some information about the northern territories' islands. For the purposes of this analysis, this appears to be fine.

The viridis “C” or “plasma” color scale was used for brightness and high contrast, not uniformity. Though the levels of the factors are not directional, these colors are still bright enough and distinctive enough such that they are suitable for distinguishing the different fields. It was decided to facet the geospatial data by education levels in order to best see how education levels may be related to the fields and location, and the rest of the information on cohort sizes, female representation, and salaries by field and gender were added into the hover tooltip through the use of several dummy aesthetics.

Advantages and Disadvantages:

The advantage of the spatial plot is that it allows people to immediately identify the largest educational cohort within each province, and see some useful information about demographics and median pay within that cohort should they desire to. That is, this graphic provides a very brief high-level overview for those who are not interested in aggregated national level data, but want to get an idea of what they may expect when they look at regional-level data. However, we do acknowledge that doing things this way, while being a non-invasive way to present large amounts of data on the most common industries within each province or territory, does have limitations:

- First, this plot only shows relevant information for the largest field of study. The issue here is that if there are two educational cohorts that both make up, for example, around 40 percent of the total population for that province, the smaller of the two will not be shown.
- Secondly, there is currently no built-in mechanism to simultaneously pinpoint multiple entries at once. This makes it tedious to make multiple comparisons.
- Third, there is an argument to be made that province-level data may not be as relevant as the national-level difference between different age groups.
- As a technical point, the hover only really works when you go over the edges of the polygon. We suspect this is an issue with how ggplotly interacts with ggplot in the creation of traces, though we were unfortunately unable to correct this behaviour.

Overall, it seems that we do need to take a more in-depth look at the situation given what we have seen.

Facet Dot Plot

Story

Given the considerations we noted above, it was decided that a single static plot, with multiple facets for different industries, age groups, and education levels for national-level data, would allow us to somewhat

address all of the issues at hand from the initial exploratory plot on regional-level data. Data was sourced again from the 2016 census, except this time, age information was retained and since we are no longer interested in cohort sizes, it was possible to source all of the information from the 2016 Canadian Census, Catalogue No. 98-400-X2016280 Statistics Canada (2019). From the facet plot, we see that:

- The gender pay gap seems to be the least prominent for the 15 to 29 year old age cohort. While the differences are slightly in favour of Males at the College level, and somewhat in favour of females at the Doctorate level, the difference at other levels of Education for this cohort, while technically in favour of males, is extremely small.
- For those aged 30 to 59, there is a consistent pay gap in favour of males across virtually every industry and education level up to the Master's degree level. While there are some exceptions where females are paid similarly, the gender pay gap appears fairly persistent across a variety of circumstances. While there is a pay gap at the doctorate level that generally favours males, the picture is not as clear.
- At the 60+ age group, we notice that the gender pay gaps are still present, but are far more erratic than those in other age groups. While pay gaps at different educational levels still seem to favour males, these trends are much less consistent across both educational levels and industries.
- In terms of industries, health care at the University Certificate, Bachelor's degree and Master's degree consistently offers comparable pay to both genders. However, at the College level, there is a persistent pay gap in favour of males at all education levels, while at the doctorate level and above the 15-29 year age brackets, there is a persistent pay gap in favour of females that enlarges at the 60+ age group. Otherwise, there are few consistent trends with respect to industry.

Overall, based on the national level data, we see that there is a persistent gender pay gap across a variety of settings and industries, but the picture is not very clear cut. It seems that the strongest trend appears to be the appearance of a pay gap for those at the university level after the 15 to 29 year bracket that persists well into the working years, but narrows as workers approach the age of 60. It is entirely possible that the observed trends are the result of some sort of "motherhood penalty", where younger women upon having children get treated differently and "lose out" on pay increases and promotion Budig, England, and others (2001), but work experience gained over time may reduce these differences as workers get older. A more detailed analysis will be needed if we are to check the validity of this conjecture.

Method

For the purposes of this analysis, we are simply interested in whether there is a consistent pay gap or not based on the factors that are most likely to cause pay variations. Given that there are a large range of possible salaries for different fields of study and educational levels, a log scale was chosen for the horizontal axes for income. Faceting was done by placing education levels as parallel rows and age groups vertically, as each facet by itself is read left to right and salient comparisons are generally not made between education

levels for this topic, but rather between different age groups within the same education level. Colors were selected on the basis of common associations for each gender and to maximize contrast on the background, which was darkened and spaced out to help the reader distinguish between separate facets. Gridlines were included by default and kept as they seem to help somewhat in tracking which values belong to which fields. Finally, fields were left in reverse alphabetical order as there is no way to consistently and naturally order this particular factor. This plot has the advantage of simplicity and clarity at the expense of losing regional-level information in favour of national-level information, but as region may not be as important of a predictor in the preliminary stages of exploratory analysis, this trade-off may be worthwhile for the detailed information it provides.

References

- Budig, Michelle J, Paula England, and others. 2001. “The Wage Penalty for Motherhood.” *American Sociological Review* 66 (2). American Sociological Association: 204–25.
- Katz, Lawrence F, and Alan B Krueger. 2016. “The Rise and Nature of Alternative Work Arrangements in the United States, 1995-2015.” Working Paper 22667. Working Paper Series. National Bureau of Economic Research. <https://doi.org/10.3386/w22667>.
- Statistics Canada. 2017. “Education Highlight Tables: Major Field of Study Stem and Base (Non-Stem) Groupings for Selected Age Groups, Sex and Selected Highest Levels of Educational Attainment, for Canada, Provinces and Territories, Census Divisions, Census Subdivisions with a 5,000-Plus Population, Census Metropolitan Areas and Census Agglomerations, 2016 Census.” <http://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hltfst/edu-sco/Tables/Files/98-402-X2016010-T4-csv-eng.zip>.
- . 2019. “2016 Census of Population, Statistics Canada Catalogue No. 98-400-X2016280.” <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/dt-td/CompDataDownload.cfm?LANG=E&PID=112449&OFT=CSV>.