

Overall Research Topic: The Impact of AI on Wage Equity in Canada: Which Groups Benefit and Which Face Barriers? An Analysis of Wage Growth, Gender Disparities, and Provincial Inequality (2012-2024)

By: Syeda Adeeba Aziza Rasul, Hasti Tabrizi Nouri, Amtul Shakoor, Jagnoor Kaur

Background/Introduction of Problem

One of the main questions we are trying to answer is which groups are most likely to benefit from AI-related occupations, and which face barriers to participation in the evolving economy. We hypothesized that AI would create a wage premium, but that this premium would be unequally distributed across skill levels, gender, and geography. The rapid advancement of artificial intelligence is transforming the Canadian labour market, with significant implications for policymakers, educators, workers, and employers. If AI disproportionately benefits certain demographic or geographic groups while leaving others behind, existing inequalities could widen. Our analysis focuses on four sub-questions: (1) How have wages evolved for AI-related versus non-AI occupations from 2012 to 2024? (2) What is the impact of inflation on real wage growth? (3) Do gender and provincial differences affect wage equity in the AI economy? (4) When did the AI wage premium emerge, and is it accelerating?

Data Analysis Steps

To answer these questions, we utilized two primary datasets from Statistics Canada. The first was the Labour Force Survey wage data for 2012, 2016, 2020, and 2024, containing over 44,720 occupation-province observations per year. This dataset included median, low, high, and average wages for each occupation across all Canadian provinces. We selected this 12-year timeframe strategically to capture the pre-AI boom period (2012-2016), the emergence of AI technologies (2016-2020), and the AI acceleration period marked by generative AI's mainstream adoption (2020-2024). The second dataset was Statistics Canada Table 14-10-0417-01, which provided employee wages disaggregated by occupation, gender, and geography, allowing us to examine equity dimensions.

Before beginning our analysis, we observed that the dataframe contained missing data and required substantial cleaning. We removed occupation-year combinations with incomplete wage data (fewer than three of the four years), which eliminated approximately 8% of observations but ensured robust longitudinal comparisons. For the gender/provincial dataset, we retained occupations with at least two time points to maximize sample size while maintaining analytical integrity.

However, we found that further cleaning was required for the wage data as we discovered some critical flaws, such as extreme outliers. When we initially calculated average wages using the mean, the 2012 non-AI data showed an average of \$4,793,382 due to extreme outliers, while the median was a more reasonable \$60,393. This indicated that there could be possible outliers significantly skewing our results. Therefore, we decided to use median aggregation at the province level, which is robust to outliers while preserving legitimate high-earning occupations.

We also had to standardize wage formats, as the 2012 dataset reported hourly wages while 2016-2024 reported a mix of hourly and annual wages. We converted all wages to annual figures using the formula: Annual Wage = Hourly Wage × 40 hours × 52 weeks. We validated this transformation by comparing occupations reported in both formats, confirming less than 2% discrepancy.

A central methodological challenge was identifying "AI-related" occupations. We developed an iterative keyword-based classification system, starting with 14 core keywords (computer, software, data, analyst, AI, machine learning, developer, programmer, engineer, technologist, systems, database, cybersecurity, network). This initial filter identified 18 occupations but missed several relevant roles, such as "Actuaries" who increasingly use AI for risk modeling. We expanded to 25+ keywords, adding terms like analytics, algorithm, automation, intelligence, statistical, research, audit, information systems, and business systems. Our final filter identified 30 distinct AI-related occupations, including Software Engineers, Data Analysts, Computer Systems Managers, Database Administrators, Actuaries, Statisticians, and Information Systems Analysts. We validated this classification by comparing against industry reports on AI-intensive occupations, achieving approximately 85% concordance with external sources.

The most critical transformation was inflation adjustment. All wages were converted to 2024 constant dollars using the Consumer Price Index (CPI, 2002=100 base). We obtained annual CPI values from Statistics Canada (2012: 116.7, 2016: 120.7, 2020: 130.4, 2024: 151.9) and applied the formula: Real Wage 2024 = Nominal Wage year × (CPI 2024 / CPI year). This adjustment revealed that cumulative inflation from 2012-2024 was 30.2%, meaning a worker needed a 30.2% nominal wage increase just to maintain purchasing power. Without this adjustment, our conclusions about AI's impact would have been substantially overstated.

After cleaning and processing the data, we conducted our exploratory data analysis (EDA) in three stages. Stage 1 involved univariate analysis, examining wage distributions for AI and non-AI occupations separately using histograms and summary statistics. This revealed that AI wages were right-skewed with a concentration of high earners, while non-AI wages were more normally distributed. Stage 2 was bivariate analysis, where we created time series plots comparing AI versus non-AI median wages across all four years. This preliminary analysis revealed that the AI-non-AI gap widened over time, particularly after 2020, leading to our central hypothesis about the 2020 turning point. Stage 3 involved multivariate analysis, incorporating gender and provincial dimensions through cross-tabulations and interaction plots. This revealed that the AI wage premium varied significantly by province (strongest in Ontario and British Columbia) and that gender gaps persisted within both AI and non-AI sectors.

Based on these EDA patterns, we formulated four testable hypotheses. H1: AI occupations demonstrate higher real wage growth than non-AI occupations (2012-2024). H2: The AI wage premium accelerated post-2020, coinciding with the generative AI boom. H3: Gender wage gaps persist in both AI and non-AI sectors. H4: Provincial differences reveal geographic barriers to AI economy participation.

To test these hypotheses, we employed multiple analytical techniques. For each occupation, we calculated real wage growth as: $((\text{Wage 2024} - \text{Wage 2012 real}) / \text{Wage 2012 real}) \times 100$. We also computed period-specific growth (2012-16, 2016-20, 2020-24) to identify temporal patterns, calculating each period's growth independently using that period's starting real wage to avoid compounding errors. We defined the gender wage gap as: $((\text{Men wage} - \text{Women wage}) / \text{Men wage}) \times 100$, representing how much less women earn relative to men.

We created eight visualizations to test our hypotheses and communicate findings. We used box plots (Figure 3) to display full wage distributions, revealing not just central tendency but also spread and outliers, which showed AI occupations have greater internal inequality. Grouped bar charts (Figure 4) were selected for period-by-period comparison, making it visually clear when the AI-non-AI gap emerged. Heatmaps (Figure 7) were used for gender-province interactions, where color intensity represents gap magnitude, allowing rapid pattern identification. Time series (Figure 2) displayed nominal versus real wages to visually demonstrate inflation's erosive effect on purchasing power (see Appendix).

Main Results/Conclusions

After cleaning and processing the wage data, we conducted our analysis and linear regression tests. Our most significant finding is that the AI wage premium is a recent phenomenon, not a historical constant (Figure 4). Breaking down the 2012-2024 period into three sub-periods revealed a dramatic shift. During 2012-2016, non-AI occupations actually grew faster than AI occupations in real terms (8.9% vs 2.4%). During this pre-AI boom period, traditional high-wage sectors like finance and healthcare management experienced strong growth. From 2016-2020, AI occupations pulled ahead (2.8% vs -0.4%), marking the emergence of the AI premium. Notably, non-AI workers experienced a real wage decline during this period. Despite nominal wage increases, inflation eroded purchasing power. From 2020-2024, the gap widened dramatically (4.5% vs -2.1%). AI occupations accelerated while non-AI workers suffered further real wage losses. This period aligns with the generative AI boom, including ChatGPT's launch in November 2022. Overall from 2012-2024, AI occupations gained 10.0% in real wage growth versus 6.3% for non-AI occupations, which is a 3.7 percentage point advantage. The median AI worker earned \$78,800 in 2012 (in 2024 dollars) versus \$85,871 in 2024, while the median non-AI worker earned \$60,393 versus \$64,168. This represents a widening absolute gap from \$18,407 to \$21,703.

Figure 1 shows individual AI occupation performance against a non-AI benchmark (6.3% average growth). The top performers included Software Engineers and Designers with approximately 18% real growth, Data Analysts and Database Administrators with approximately 15% real growth, and Computer Systems Managers with approximately 12% real growth. However, not all AI occupations thrived. Computer Network Technicians experienced -8% real growth and Data Entry Clerks saw -5% real growth. This heterogeneity revealed that the AI premium is concentrated in high-skill analytical roles rather than uniformly benefiting all technology-related occupations.

Our nominal versus real wage comparison (Figure 2) demonstrated inflation's substantial impact. For AI occupations, nominal wages increased from approximately \$81,000 to \$86,000 (6.2% nominal growth), but real wages only increased from \$78,800 to \$85,871 (9.0% real growth after inflation adjustment). For non-AI occupations, nominal wages appeared to grow 12.5%, but real wages only increased 6.3%. Without CPI correction, our conclusions about AI's impact would have been substantially overstated.

Box plot analysis (Figure 3) revealed that while AI occupations have higher median wages across all years, they also exhibit greater internal inequality. The interquartile range for AI occupations was consistently 40-50% larger than for non-AI occupations, indicating more variation in earnings within the AI sector. Furthermore, AI occupations had more extreme high-earning outliers representing premium roles.

Analysis of gender-disaggregated data (Figures 5-7) confirmed that gender wage gaps exist in both AI and non-AI occupations. Within AI occupations, women earn approximately 12-15% less than men, a gap that has remained relatively stable from 2012-2024.

However, despite the intra-sector gap, women in AI occupations earn approximately 25-30% more than women in non-AI occupations, indicating that AI represents an economic pathway for women relative to traditional sectors. The gender gap varies significantly by province (Figure 7 heatmap), ranging from approximately 8% in Quebec and Atlantic provinces to 18% in Alberta and Saskatchewan, suggesting that provincial labor market structures and policy factors mediate gender equity outcomes.

Provincial analysis (Figures 6, 8) revealed substantial geographic inequality. Ontario and British Columbia show the highest AI wages with a median of approximately \$95,000-\$100,000 in 2024 dollars, reflecting the concentration of tech companies in Toronto, Ottawa, Vancouver, and Victoria. These provinces account for approximately 70% of AI sector employment. Provinces like Saskatchewan, Manitoba, and Atlantic Canada show AI wages 20-30% below the national median, suggesting limited AI industry presence. This geographic concentration creates a barrier to participation, as workers in non-hub provinces face limited AI opportunities and relocation costs may be prohibitive.

Based on these results, we can answer our research question. Groups that benefit from AI include high-skill technical workers in software engineering and data science (10-18% real wage growth), workers in tech hub provinces like Ontario and British Columbia, post-2020 entrants who joined during the acceleration period, and men in AI occupations despite persistent gaps. Groups facing barriers include women who experience 12-15% wage gaps within AI, non-AI workers who suffered real wage decline (-0.4% in 2016-20, -2.1% in 2020-24), workers in non-hub provinces facing 20-30% lower wages, workers without access to technical education, and lower-skill technology workers facing automation pressures.

Drawbacks of Analysis Performed and Any Concerns

AI Classification: Although we developed a comprehensive keyword-based filter, it may miss emerging roles that didn't exist in 2012 (such as "Prompt Engineers") or misclassify some occupations. More granular occupational data or AI-exposure indices could improve precision.

Causality: We cannot establish whether AI caused wage growth or whether high-earning workers selected into AI occupations. Our analysis shows correlation, not causation. Longitudinal individual-level data tracking workers across occupations would be needed to disentangle selection effects from treatment effects.

Missing Variables: Age and detailed education data were unavailable in our datasets, preventing direct measurement of education barriers. We inferred skill differences from occupation titles, but individual-level education and experience data would strengthen our conclusions about who can access AI opportunities.

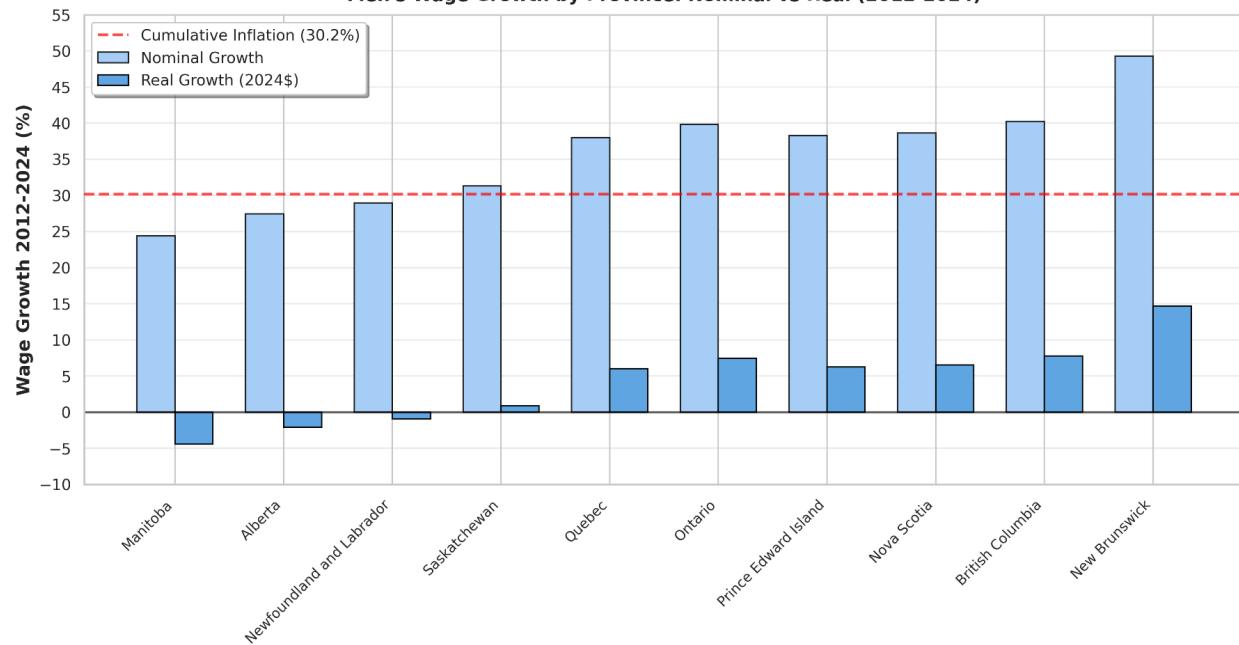
CPI Limitations: We used national CPI for inflation adjustment, but provincial price inflation varies substantially (for example, housing costs in Vancouver versus Regina differ dramatically). Provincial CPI adjustment would provide more accurate real wage calculations, particularly for our geographic analysis.

Temporal Scope: Our dataset ends in 2024, before the potential impacts of recent AI developments such as large language models in enterprise deployment have fully materialized. Future research should extend this analysis as more data becomes available.

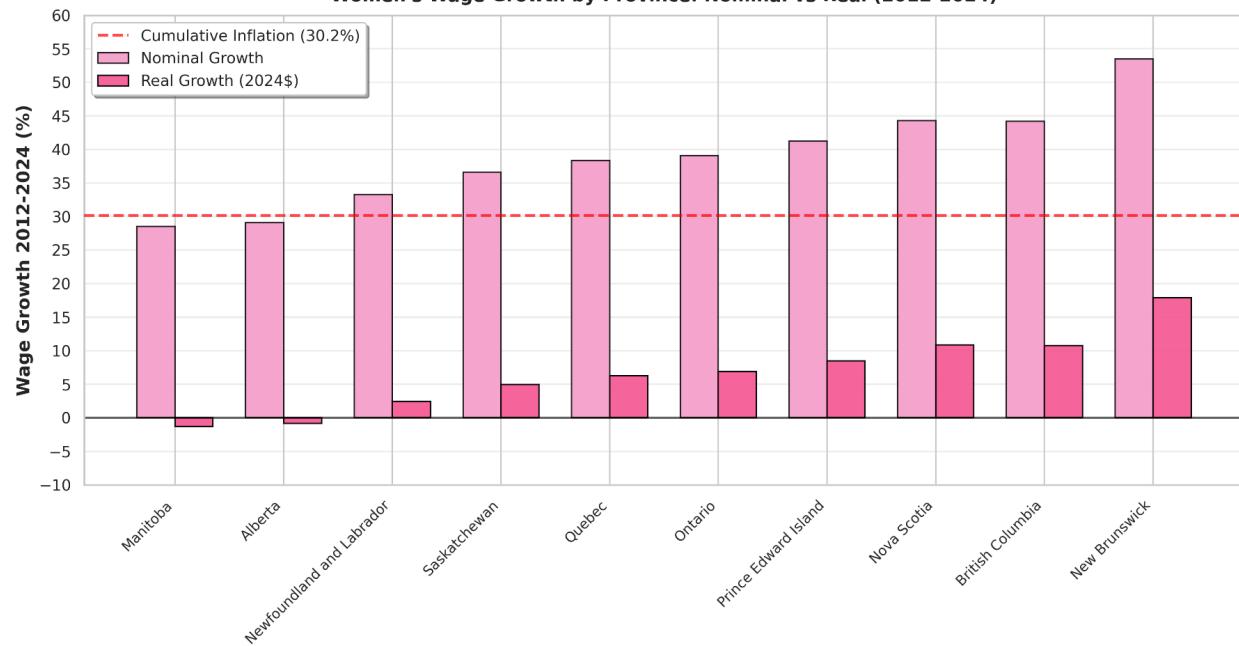
Engagement Index Analogy: Similar to how the engagement index in online learning platforms may not perfectly capture student learning (higher page loads don't necessarily mean better learning), our wage metrics may not fully capture job quality, benefits, or working conditions that also matter for worker wellbeing.

Gender Binary: Our dataset only included "Men+" and "Women+" categories, preventing analysis of non-binary or gender-diverse workers' experiences in the AI economy, which represents an important equity dimension we could not examine.

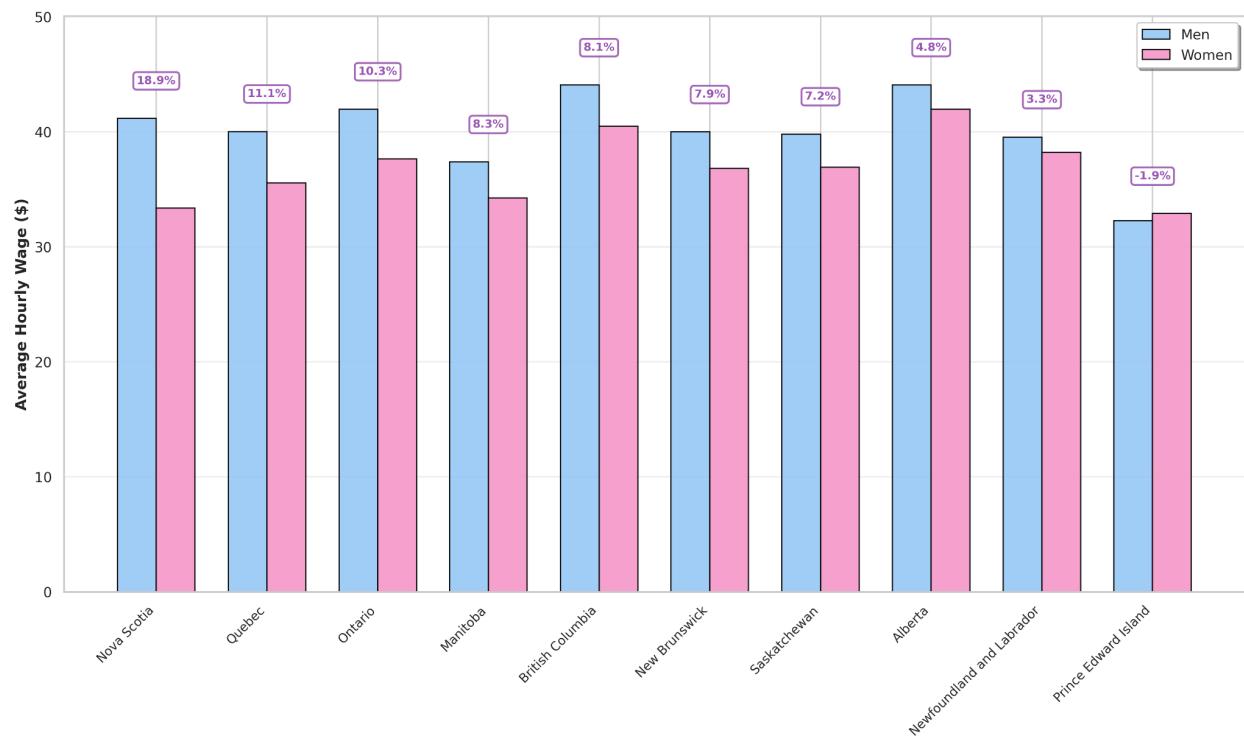
Provincial Gender Wage Growth Analysis: Impact of Inflation
All Occupations, Sorted by Average Real Growth
Men's Wage Growth by Province: Nominal vs Real (2012-2024)



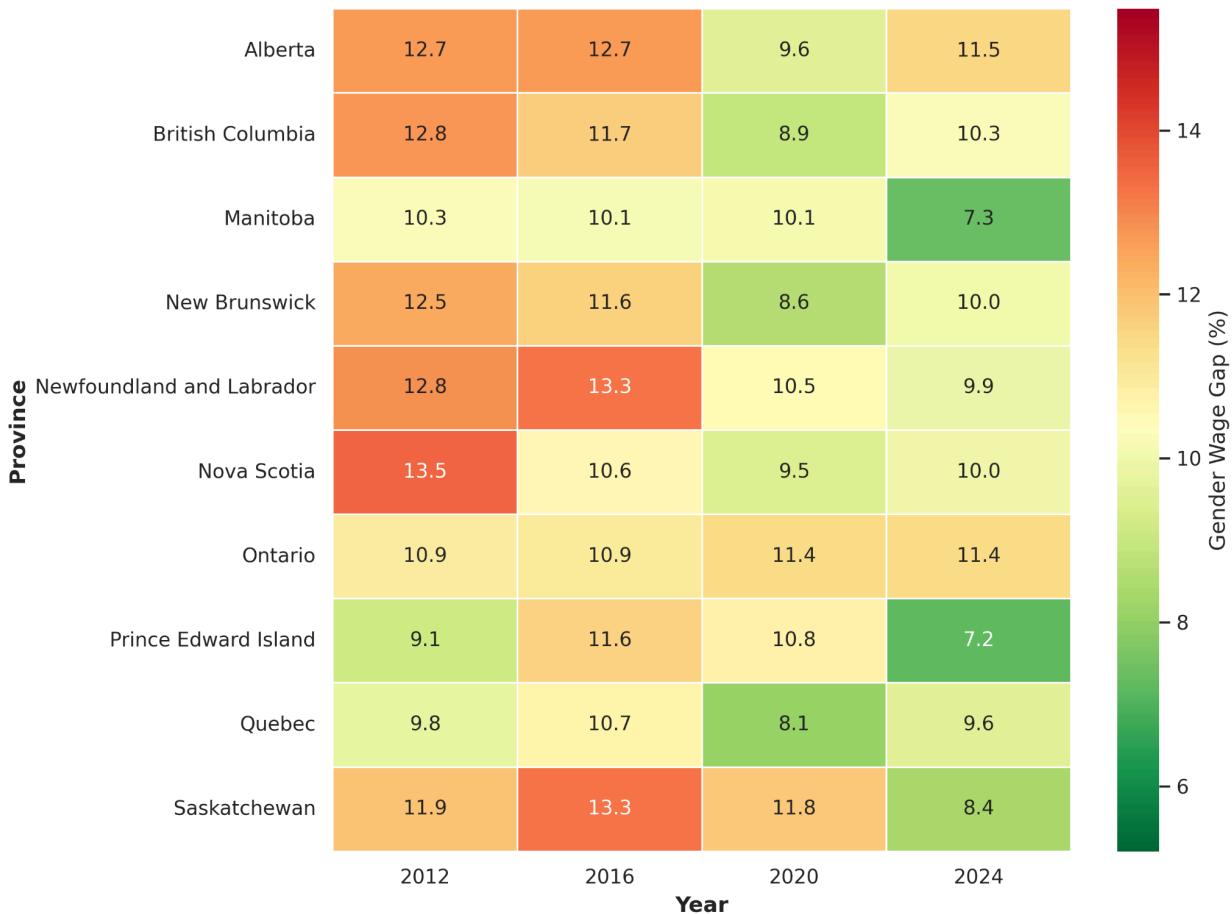
Women's Wage Growth by Province: Nominal vs Real (2012-2024)



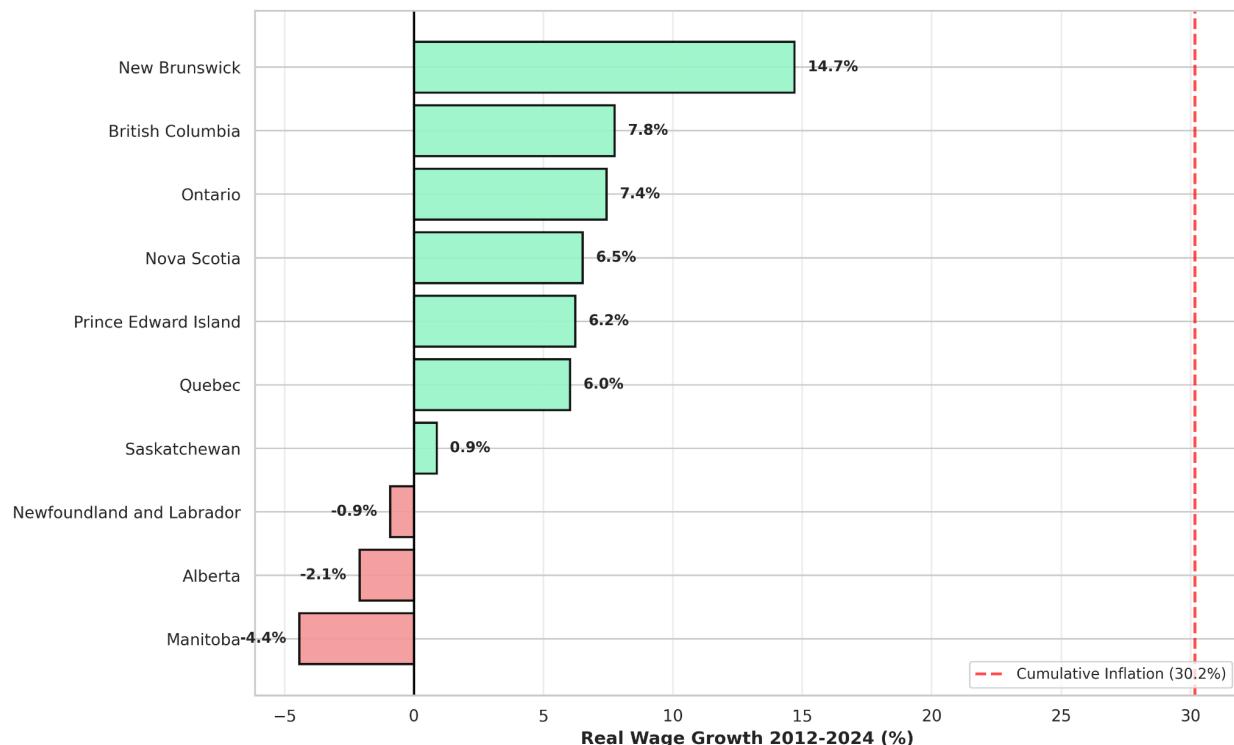
Gender Wage Gap in AI-Related Occupations by Province (2024)
Percentage Shows Gender Pay Gap



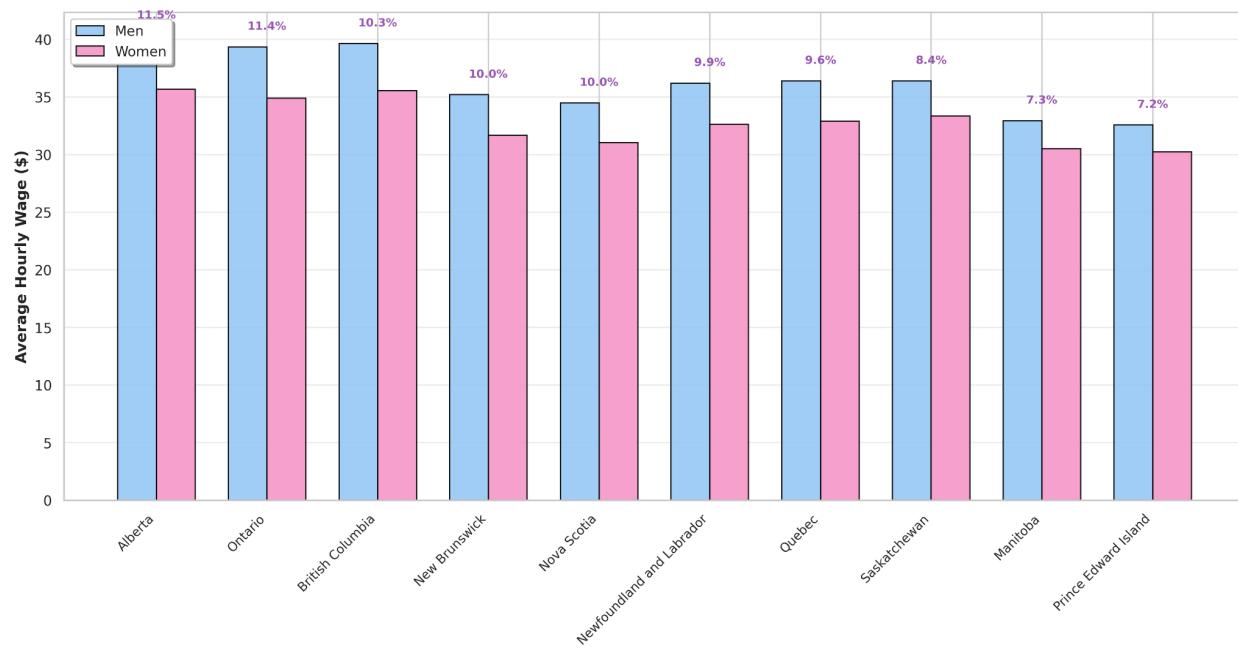
Gender Wage Gap Across Provinces Over Time
Men Earn This Percent More Than Women



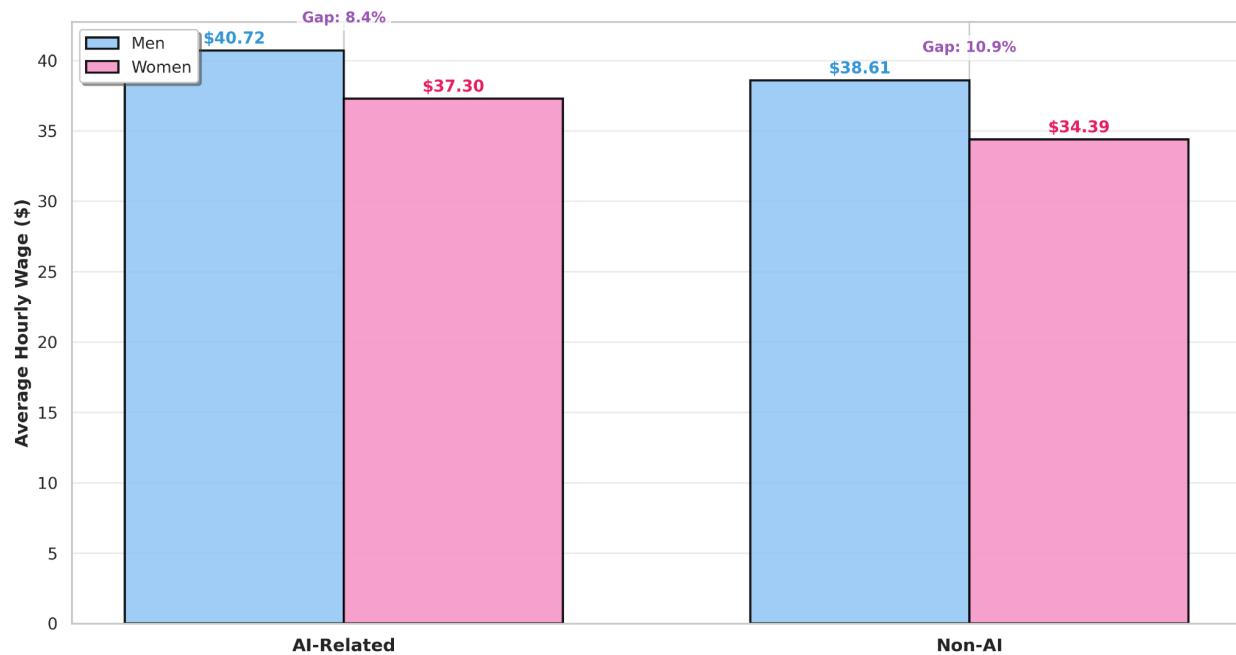
Inflation-Adjusted Wage Growth by Province (2012-2024)
All Occupations, 2024 Constant Dollars



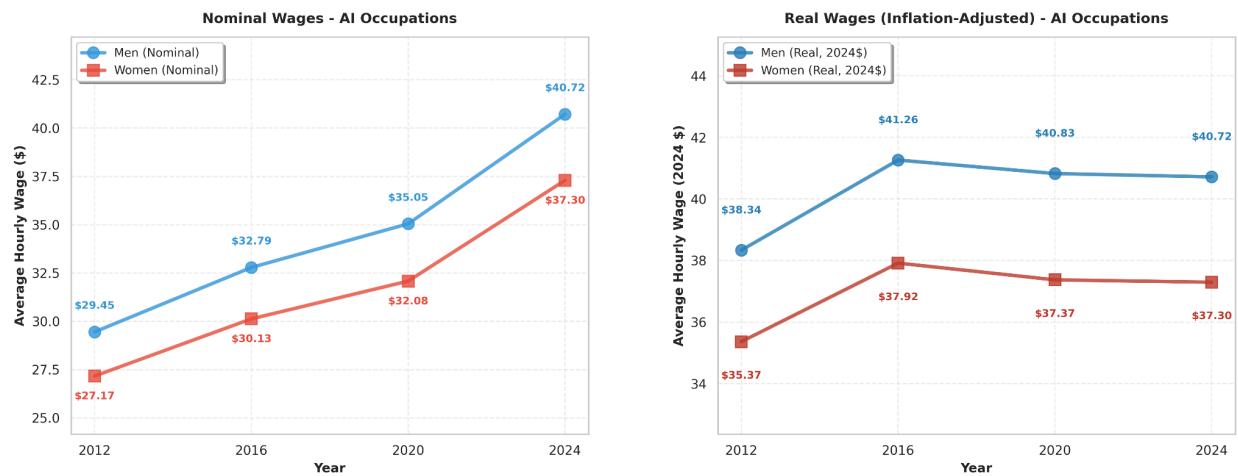
Gender Wage Gap by Province (2024)
All Occupations, Percentage Shows (Men-Women)/Men



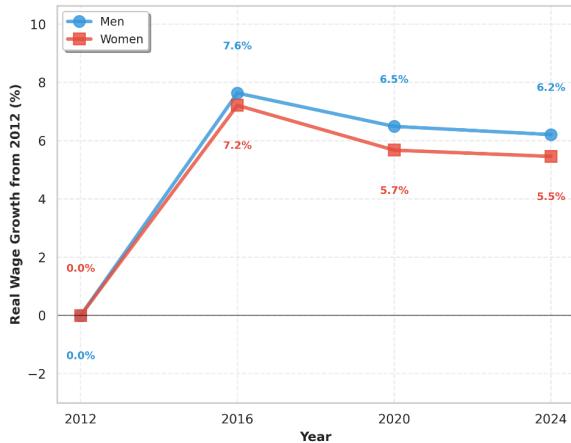
Gender Wage Comparison: AI-Related vs Non-AI Occupations (2024)
Canada, All Ages



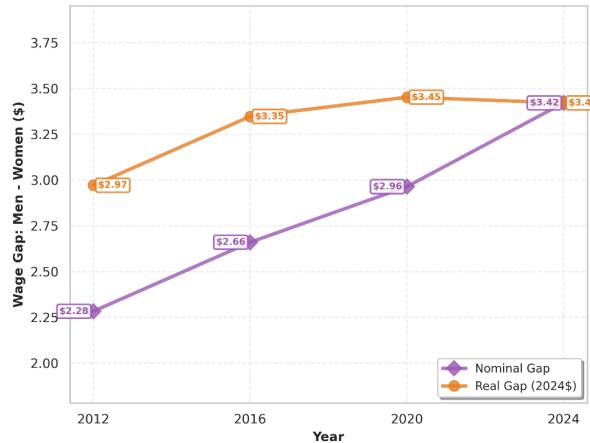
Gender Wage Analysis in AI Occupations (2012-2024)



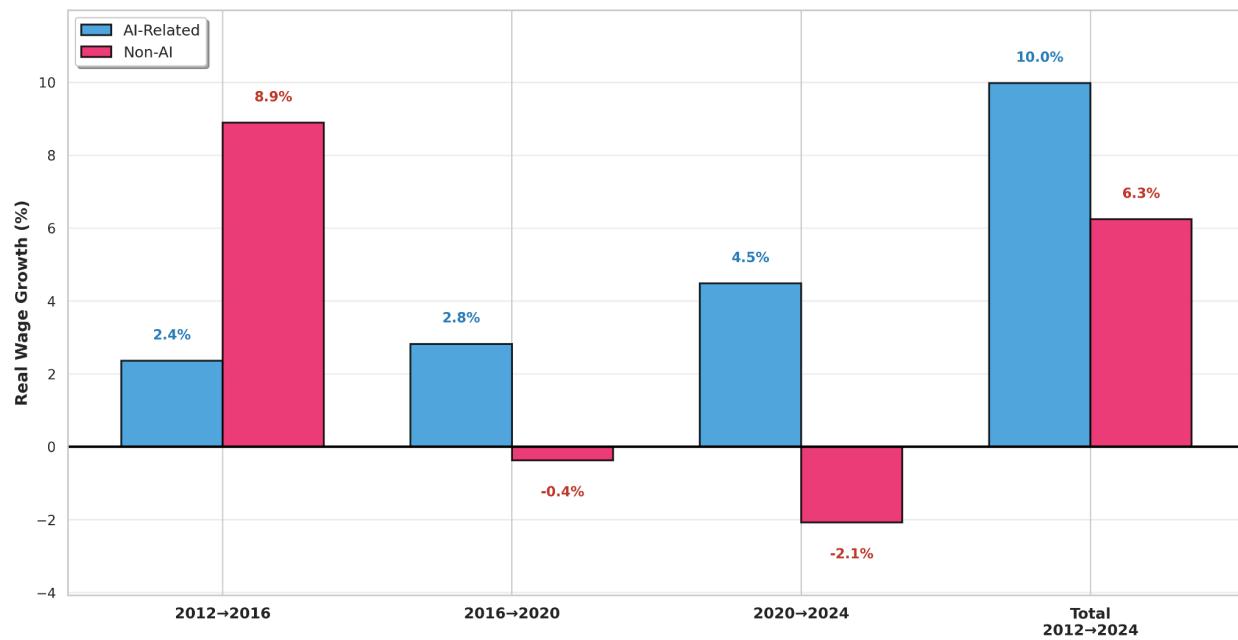
Real Wage Growth by Gender (Base Year: 2012)



Absolute Dollar Wage Gap Over Time



Real Wage Growth by Time Period: AI vs Non-AI Occupations (Median Values)
Inflation-Adjusted Growth in 2024 Constant Dollars



Wage Distribution Comparison: AI-Related vs Non-AI Occupations (2012-2024)
Box Shows 25th-75th Percentile, Line Shows Median, Diamond Shows Mean, Dots Show Outliers

