

Canadian Household Internet Productivity and Recreation Trends 2018*

Brief Overview of Canadian Internet Use Trends

Haocong Wu

27 April 2022

Abstract

Internet access is critical to many users. This paper examines the quality of internet access that users have and how they are impacted by subpar service. With data collected by Statistics Canada in the Canadian Internet Use Survey, and modelled with linear regression, age has a negative correlation with user household internet usage, whilst education level has a strong positive correlation.

Keywords: internet use, internet infrastructure, statistics canada, canadian residents, linear regression

1 Introduction

Internet access is becoming, if not already considered, a human right. It is a standard point of communication used by the majority of the Canadian population. The platform that the internet provides can be used to host a variety of activities, ranging from those of productive nature, to others of recreational and leisurely type. The growing importance of internet access necessitates a revision of the bare minimums of providing internet service.

This paper examines the causal relations between individual internet usage amounts to factors such as connection speed, nature of use, and necessity of use.

The following sections of the paper is organized as stated: the Data section discusses the source of the data and its surveying methodology, the Modelling section attempts to formalize a model incorporating various factors to explain the response variable of user household internet use duration, the Results section discusses the impact of the selected factors, and the Discussion section comments on the results and its validity, potential bias, and weaknesses of the paper.

2 Data

2.1 Data Source

The Canadian Internet Use Survey (CIUS) replaced the Household Internet Use Survey (HIUS) in 2005 (Statistics Canada 2019). The 2018 of the CIUS aims to record and track the quickly increasing Canadian internet utilization. Aspects of internet technologies that Canadians interact with such as online government services, social networks, smartphone use, and online commercial platforms are presented to respondents for their sentiments.

*Code and data are available at: <https://github.com/hcgw0318/2018-Canadian-Internet-Use-Analysis>

Survey data was provided by the portal, maintained by the Ontario Council of University Libraries. The particular set of data used in this report is a subset of the topics covered by the 2018 CIUS.

2.2 Survey Methodology

Based on the 2018 CIUS documentation (Ontario Data Documentation 2020), the survey methodology can be summarized by the following.

The CIUS attempts to gauge the population habits regarding various facets of internet use. The CIUS' stated target population includes all individuals at and above the age of 15 in Canada, excluding those in the Canadian Territories and excluding institutionalized individuals.

The sampling frame of the GSS relies on the availability of telephone access. Households without telephones, as well as households with telephone services not covered by the current frame, were excluded from the sampling frame.

The sampling process consisted of location based random sampling without replacement. A target minimum sample size for each province was determined, and this minimum sample size was allocated to the strata within the province.

Each respondent was contacted to complete either an electronic questionnaire, or participate in a computer assisted telephone interview. If the respondent refused or was absent, further contact attempts were made through telephone contact until reasonable confirmation of non-response, and the respondent is removed from the pool.

The 2018 CIUS had a target sample size of 15,000. The actual number of respondents was 13,810, the total households contacted was 33,248, and the overall response rate was 41.6%.

2.3 Data Processing Tools and Methodology

Data processing and analysis in this project is done in the R statistical programming language (R Core Team 2020). The following R packages were imported and used for their corresponding purpose:

- **tidyverse** (Wickham et al. 2019) : Used for general logical commands, file importing, and data manipulation
- **knitr** (Xie 2021) : Used to generate and knit the Markdown document to PDF
- **ggplot2** (Wickham 2016) : Used to generate graphs and apply styling
- **kableExtra** (Zhu 2020) : Used to assist in table formatting
- **bookdown** (Xie 2020) : Used to improve the Markdown document formatting
- **here** (Müller 2020) : Used to simplify file access in a project environment
- **forcats** (Wickham 2022) : Used to sort data tables by frequency
- **modelsummary** (Arel-Bundock 2022) : Used to display regression model information

2.4 Relevant variables

Variables such as whether the respondent used the internet and had an internet connection at home were used to identify the subset of the sample whose internet usage time is the response of interest. Variables likely to explain variations in internet usage time contained in the survey were age, education level, and purpose of internet usage.

Table 1: Location Demographics

Province	Count
Alberta	1142
British Columbia	1178
Manitoba	847
New Brunswick	918
Newfoundland and Labrador	624
Nova Scotia	860
Ontario	3240
Prince Edward Island	662
Quebec	3618
Saskatchewan	721

Table 2: Age and Gender Demographics

	Female	Male
15 to 24 years	363	343
25 to 34 years	827	694
35 to 44 years	1001	1016
45 to 54 years	1142	1075
55 to 64 years	1523	1463
65 years and over	2365	1998

In Table 1 and Table 2, the survey demographics are shown. It can be seen that the sample population is skewed towards more female respondents, and skewed towards older respondents.

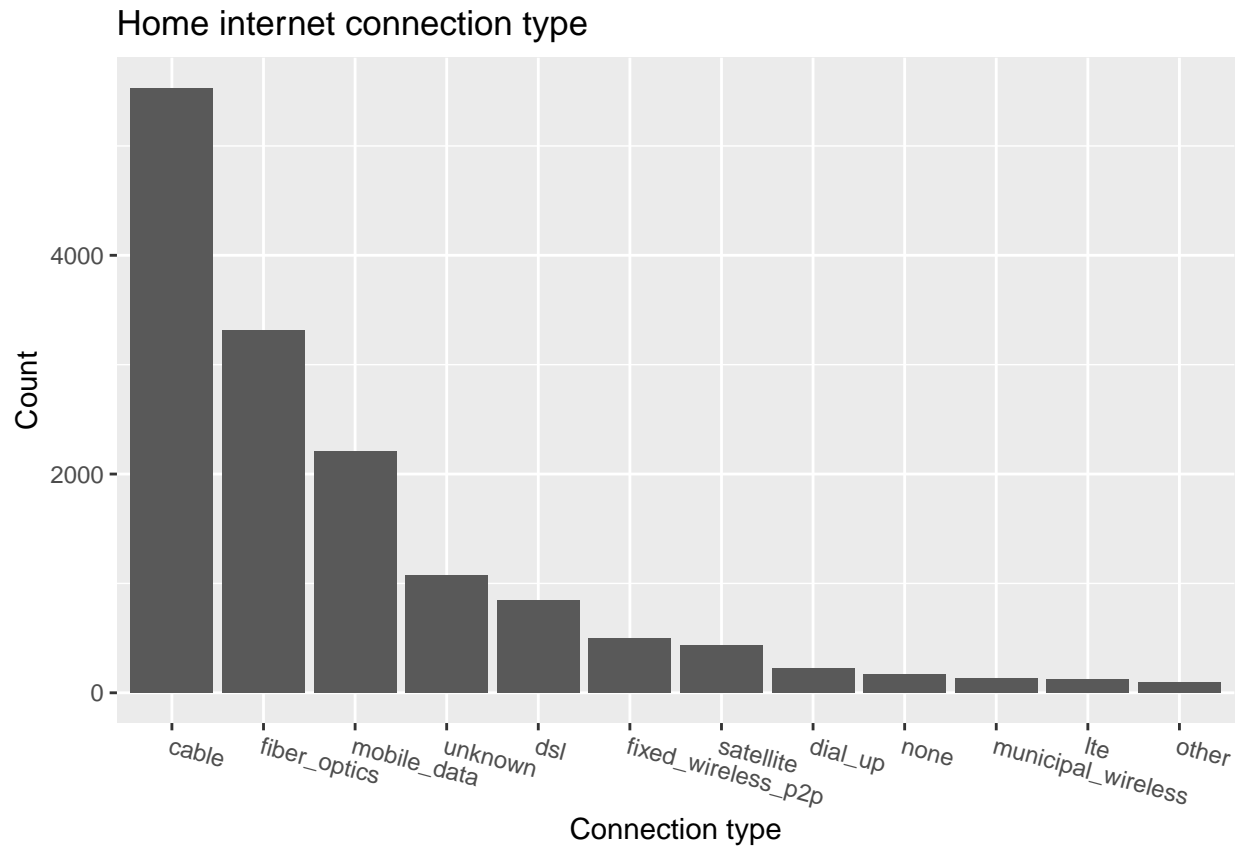


Figure 1: Respondent home connection type

In Figure 1, the three most popular types of home internet connections are cable, fiber optics, and mobile data. After that, approximately 7% of respondents are not sure what type of connection they have.

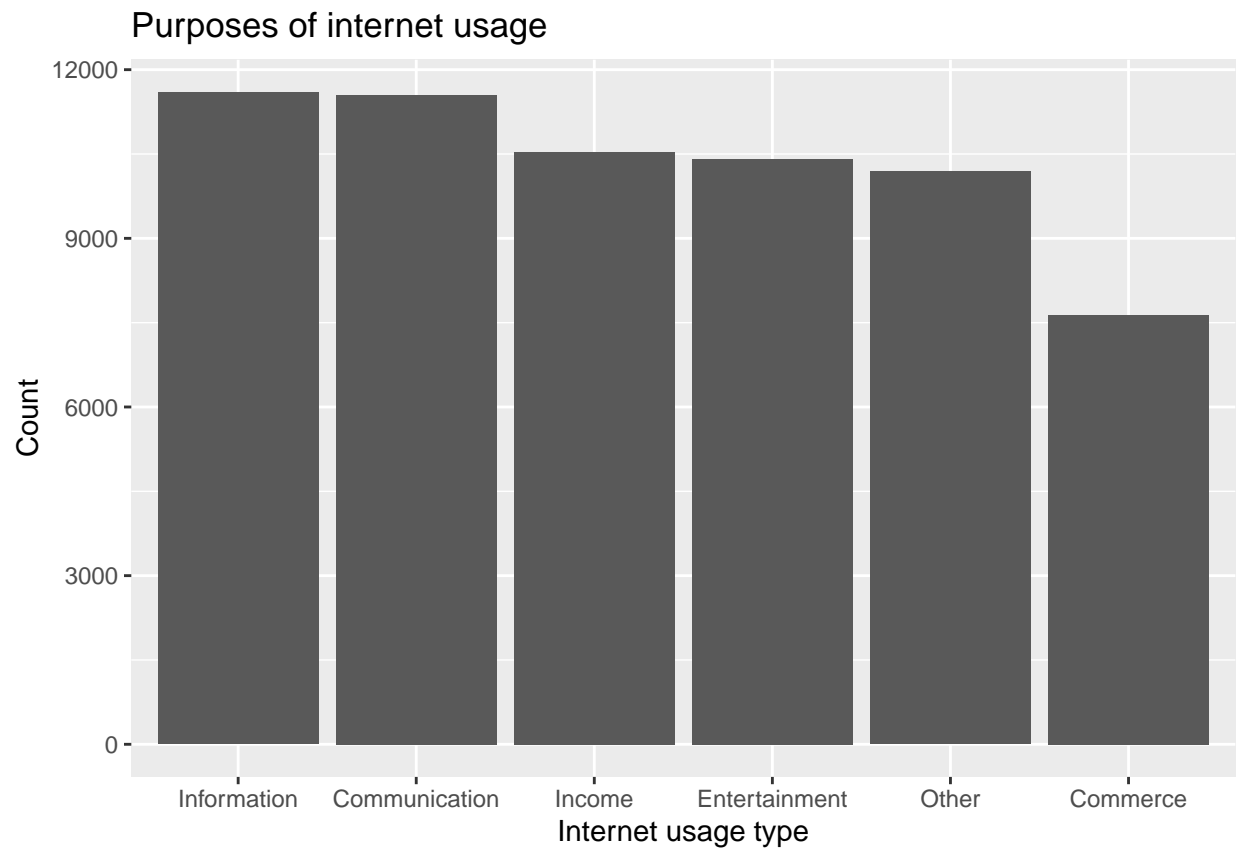


Figure 2: Respondent internet use purpose

In Figure 2, common uses of the internet are displayed.

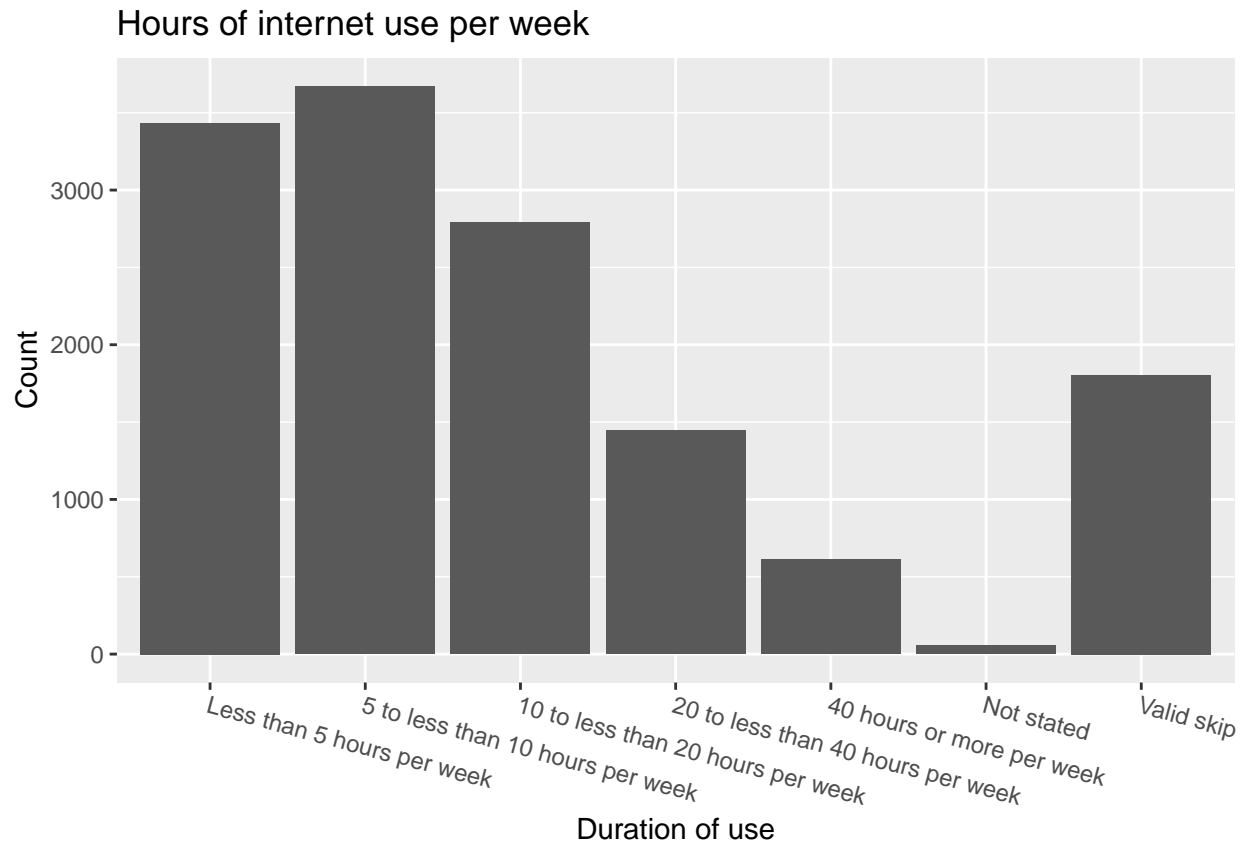


Figure 3: Respondent weekly internet use

In Figure 3, the distribution of weekly internet usage of the respondents is shown. The most common amount of time spent on the internet per week is between 5 to 10 hours.

3 Results

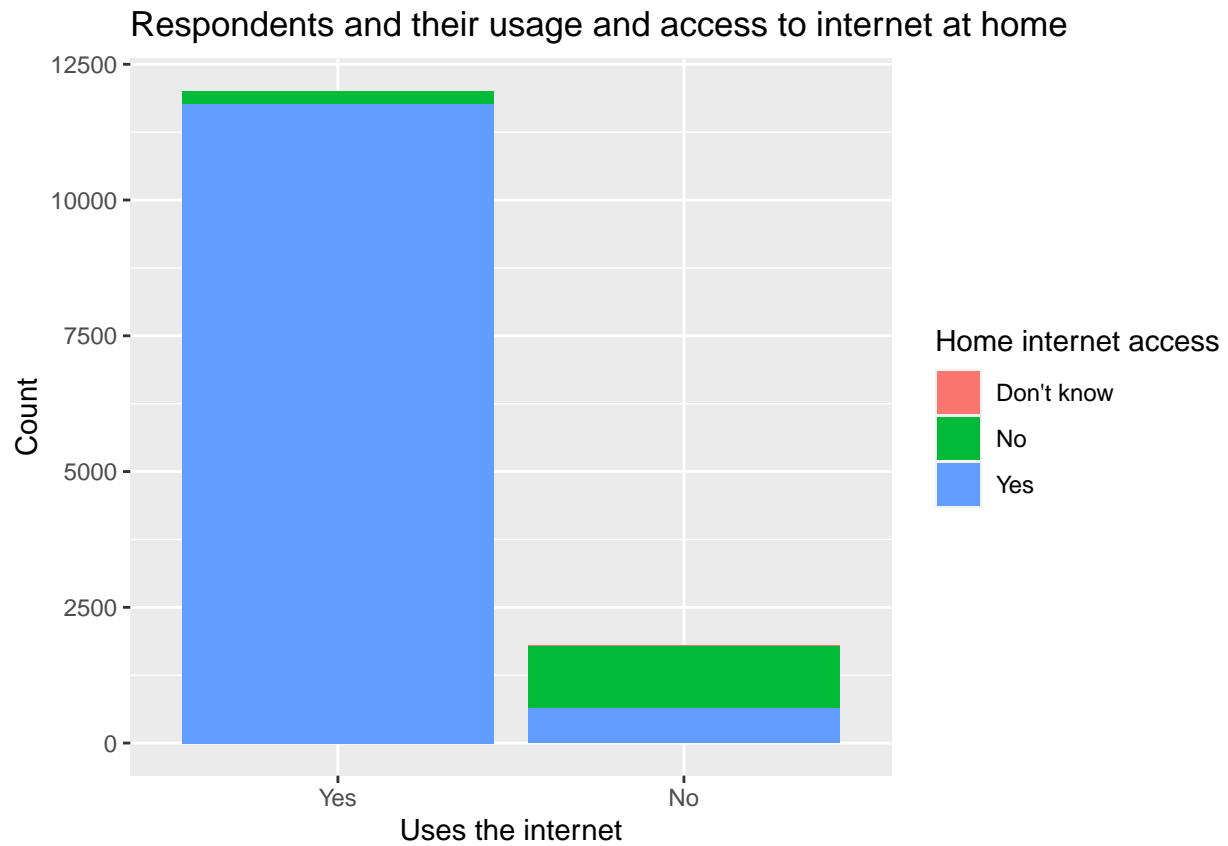


Figure 4: Respondent internet use and home internet access

In Figure 4, it is shown that most users use the internet in some way. The majority of respondents that do not use the internet also do not have an internet connection at home.

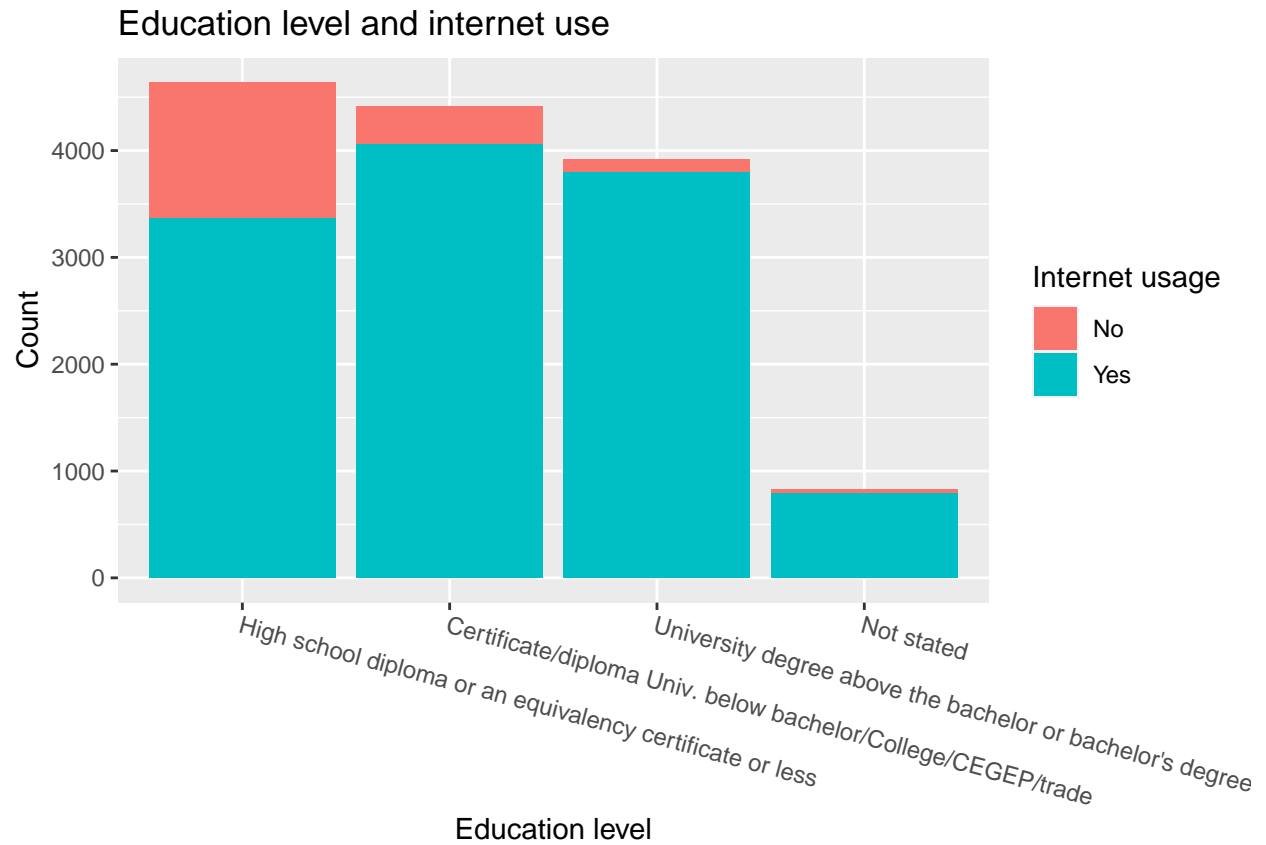


Figure 5: Respondent education level and internet use

In Figure 5, it can be seen that internet usage proportion increases with respondent education level.

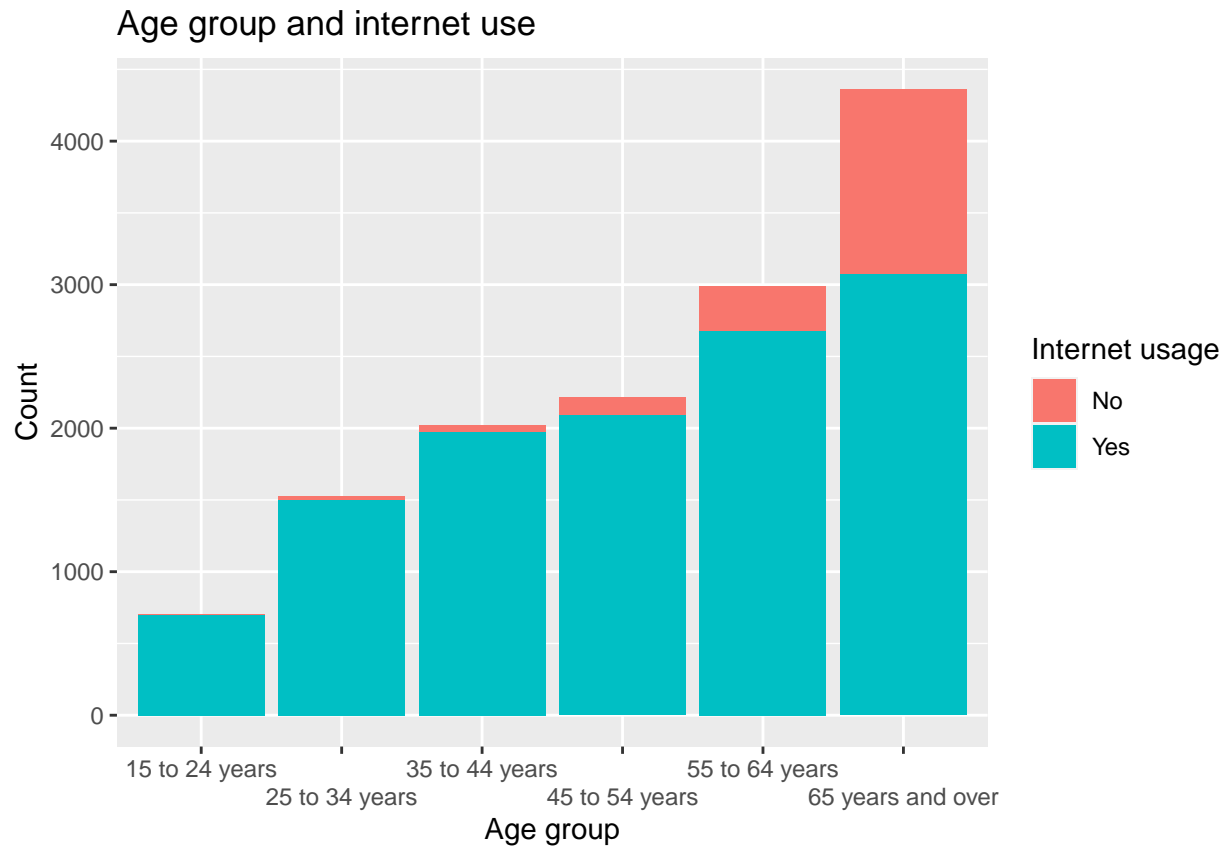


Figure 6: Respondent age group and internet use

In Figure 6, as age group increases, the proportion of the group that does not use the internet increases.

Table 3: Model of weekly internet use quantity

	Model 1
(Intercept)	18.5 (0.5)
genderMale	0.5 (0.2)
education_levelCertificate/diploma Univ. below bachelor/College/CEGEP/trade	-0.2 (0.3)
education_levelUniversity degree above the bachelor or bachelor's degree	1.5 (0.3)
education_levelNot stated	-1.0 (0.5)
age_group25 to 34 years	-3.6 (0.6)
age_group35 to 44 years	-4.8 (0.6)
age_group45 to 54 years	-5.7 (0.5)
age_group55 to 64 years	-6.2 (0.5)
age_group65 years and over	-6.5 (0.5)
no_communicationTRUE	-2.2 (0.7)
no_commerceTRUE	-2.7 (0.3)
no_entertainmentTRUE	-4.2 (0.4)
no_informationTRUE	-1.3 (0.8)
no_other_activitiesTRUE	-1.4 (0.4)
internet_incomeTRUE	3.6 (0.5)
Num.Obs.	11 718
R2	0.084
R2 Adj.	0.083
AIC	91 732.1
BIC	91 857.4
Log.Lik.	-45 849.046
F	71.874
RMSE	12.12

In Table 3, the regression model of internet use hours as a function of gender, education level, age group, and internet activities is displayed.

4 Discussion

4.1 Greater proportion of female respondents

From data displayed in Table 1 and graphed, the respondent demographic can be seen to have a larger female representation. This observation could be caused by a larger proportion of Canadian females being available to pick up the phone due to childcare and household work.

4.2 Greater proportion of older respondents

Despite the decreasing proportion of individuals at each age group that uses the internet shown in Figure 6, the number of respondents increases with each older age group. This at first glance seems conflicting with the fact that the CIUS is conducted primarily on an online platform. A reason for this occurrence may be that the older age groups have more leisure time, and are willing to use their leisure time to respond to this survey.

4.3 Education and age trends

As education level increases, the proportion of respondents that use the internet increases, as shown in Figure 5. In addition, the amount of time that one uses the internet per week increases with education level as well. This result likely arises due to the utility if not necessity of utilizing online resources to progress with further education, as well as the fact that individuals who undertake further education would be more inclined to adapt and utilize new technology.

The converse is true with age, the younger the respondent, the more likely they are to both use the internet and use it for a longer duration per week, this is depicted by Figure 6. This occurrence is explainable by the attraction of online entertainment of social media, video games, and media consumption platforms to younger audiences. The younger the individual, the more accustomed they are to seek entertainment through online means.

4.4 Weaknesses and improvements

This paper attempts to address the current populations internet usage tendencies, and focuses on the duration of internet use. The respondents' demographics, connection type, and usage purposes were variables that were analyzed. Given that the goal of the paper is to find correlation between these factors and internet usage hours, the analysis that consisted of the factors of gender, age, education, and usage nature may be lacking. The analysis can be supplemented with the consideration of factors such as household economy, urban and rural living distinctions, and whether anyone in their household utilized the internet. These factors were not included in the survey, and can be collected from a follow-up supplemental survey.

Another weakness of the analysis is making the assumption of generating uniform random ages for each age group to perform linear regression. This assumption was made due to the lack of more granular internet usage data. To improve on this aspect, perhaps ordinal regression may be a better fit to generate a model.

References

- Arel-Bundock, Vincent. 2022. *Modelsummary: Summary Tables and Plots for Statistical Models and Data: Beautiful, Customizable, and Publication-Ready*. <https://CRAN.R-project.org/package=modelsummary>.
- Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. <https://CRAN.R-project.org/package=here>.
- Ontario Data Documentation, Scholars Portal, Extraction Service and Infrastructure (ODESI). 2020. *Canadian Internet Use Survey, 2018 Study Documentation*.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Statistics Canada. 2019. *Canadian Internet Use Survey (CIUS)*. <https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=1196799>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- . 2022. *Forcats: Tools for Working with Categorical Variables (Factors)*. <https://forcats.tidyverse.org>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Xie, Yihui. 2020. *Bookdown: Authoring Books and Technical Documents with r Markdown*. <https://github.com/rstudio/bookdown>.
- . 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.
- Zhu, Hao. 2020. *kableExtra: Construct Complex Table with ‘Kable’ and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>.