

Catalogue no. 11F0019M — No. 465
ISSN 1205-9153
ISBN 978-0-660-40772-2

Analytical Studies Branch Research Paper Series

Internet-use Typology of Canadians: Online Activities and Digital Skills

by David Wavrock, Grant Schellenberg and
Christoph Schimmele

Release date: November 09, 2021



 Statistics Canada Statistique Canada

Canada 

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at STATCAN.infostats-infostats.STATCAN@canada.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Depository Services Program

- Inquiries line 1-800-635-7943
- Fax line 1-800-565-7757

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "[Standards of service to the public](#)."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

© Her Majesty the Queen in Right of Canada as represented by the Minister of Industry, 2021

All rights reserved. Use of this publication is governed by the Statistics Canada [Open Licence Agreement](#).

An [HTML version](#) is also available.

Cette publication est aussi disponible en français.

Internet-use Typology of Canadians: Online Activities and Digital Skills

by

David Wavrock, Grant Schellenberg and Christoph Schimmele

Social Analysis and Modelling Division
Statistics Canada

11F0019M No. 465
2021008
ISSN 1205-9153
ISBN 978-0-660-40772-2

November 2021

Analytical Studies Branch Research Paper Series

The Analytical Studies Branch Research Paper Series provides for the circulation of research conducted by Analytical Studies Branch staff and collaborators. The Series is intended to stimulate discussion on a variety of topics, such as labour, immigration, education and skills, income mobility, well-being, aging, firm dynamics, productivity, economic transitions, and economic geography. Readers of the Series are encouraged to contact the authors with their comments and suggestions.

All the papers in the Analytical Studies Branch Research Paper Series go through institutional and peer review to ensure that they conform to Statistics Canada's mandate as a governmental statistical agency and adhere to generally accepted standards of good professional practice.

Table of contents

Abstract	5
Executive summary	6
1 Introduction	7
2 Literature review	8
3 Data and methods	10
3.1 Data	10
3.2 Methodology	10
3.3 Variable selection.....	11
3.4 Generating the typology	11
4 Results	12
4.1 Internet-user groups—on-line activities and digital skills.....	12
4.2 Sociodemographic characteristics of individuals in user groups	17
4.3 Telework and Internet-user groups.....	21
5 Discussion and conclusions	22
6 Appendix	24
6.1 Supplementary tables	24
6.2 Methodology and robustness checks	26
6.2.1 The k-modes algorithm	26
6.2.2 Establishing number of clusters	27
6.2.3 Variable selection criteria	27
6.2.4 Selecting the optimal specification	28
6.2.5 Summary of steps	28
6.2.6 Robustness checks.....	29
References	33

Abstract

A robust Internet typology is essential for monitoring how individuals are responding to the digital transformation and for assessing the divide between digital 'haves' and 'have nots'. Individuals' capacities to use the Internet and digital technologies are an important aspect of this digital divide. Using data from the 2018 Canadian Internet Use Survey (CIUS), this study presents an Internet-use typology that is based on the range and complexity of online activities and digital skills that Canadians report they perform. Five Internet-user groups are identified. About 9% of Canadians are Internet Non-users and another 16% are 'Basic users'. Together, almost one-in-four Canadians (24%) had either no engagement or only very limited engagement with the Internet and digital technologies in late 2018 or early 2019. Another 20% of Canadians are identified as 'Intermediate users' of the Internet and digital technologies, with a further 22% identified as 'Proficient users' and 34% identified as 'Advanced users'. Age and education are strongly associated with the distribution of Canadians across these groups. About 63% of seniors are either Non-users or Basic users of the Internet and digital technologies, as are 40% of Canadians with a high school diploma. In contrast, 77% of Canadians aged 15 to 34 are either Proficient or Advanced users, as are 76% of Canadians with a Bachelor's degree or higher. Individuals in these latter Internet-user groups are most well-positioned to access online programs and services and adapt in an increasingly online society.

Executive summary

The economic and social changes associated with digital technologies continues to reshape the lives of individuals, communities, and societies. The degree to which individuals are able to operate and adapt in this context is an important issue, as it speaks to their capacity to benefit from the opportunities and avoid the risks associated with the Internet and digital technologies.

This paper presents a typology of Internet users in Canada. The analysis uses the 2018 Canadian Internet Use Survey (CIUS), and specifically, responses to 36 questions regarding respondent's online activities and digital skills. A k-modes cluster algorithm using these variables was run 2000 times and the specification that yielded the highest within-group homogeneity of observations was selected.

The cluster analysis yielded four distinct groups of Internet users that we labelled as Basic, Intermediate, Proficient, and Advanced. A fifth group of Internet Non-users was also identified. Across Canada's 10 provinces, 9% of Canadians aged 15 or older were Non-users of the Internet and just under 16% were Basic users. Together, just under one-in-four Canadians (24%) had either no engagement or very limited engagement with the Internet and digital technologies when surveyed between November 2018 and March 2019. Another 20% of Canadians were categorized as Intermediate users, with a further 22% categorized as Proficient users and 34% as Advanced users.

Of the 36 activities and skills used in the analysis, the average number reported ranged from 5.1 among Basic users to 28.9 among Advanced users. Types of activities and skills varied across the four user groups. Instant messaging, for example, was used by most Intermediate, Proficient, and Advanced users, at 78%, 86%, and 94%, respectively, but by only 28% of Basic users. Other activities set other groups apart. For example, 88% of Advanced users uploaded files to the cloud compared with just 39% of Proficient users, 16% of Intermediate users, and 6% of Basic users.

An ordered logistic regression model was used to estimate the strength and significance of the relationships between eight socioeconomic characteristics—age, education, income, urban/rural status, household size, immigration status, employment status, and sex—and the likelihood of being in each of the five groups. Age and education were the strongest predictors. Net of other characteristics in the model, individuals aged 15 to 34 were 12.3 percentage points more likely to be classified as Advanced users and 3.9 percentage points less likely to be classified as Basic users than the reference group comprised of individuals aged 35 to 49. In contrast, individuals aged 65 or older were 27.6 percentage points less likely to be classified as Advanced users and 12.9 percentage points more likely to be classified as Non-users than 35- to 49-year-olds.

A strong relationship between educational attainment and membership in user groups was also observed. For example, compared with individuals who have a high school diploma, those with a university degree were 22.7 percentage points more likely to be Advanced users and 9.8 percentage points less likely to Basic users, net of other characteristics in the model. Income and, to a much lesser extent, urban/rural status, household size, immigration status and employment status were significantly correlated with membership in user groups. Sex was the only variable in the model that was not statistically significant.

Forthcoming analysis will use both the 2018 and 2020 cycles of the CIUS to estimate the how the distributions of Canadians across Internet-user groups differed prior to and during the COVID-19 pandemic.

1 Introduction

The COVID-19 pandemic has both highlighted and accelerated the roles that digital technologies and Internet connectivity play in our lives. Physical distancing and “stay-at-home” responses have changed how Canadians work, shop, attend school, and spend their leisure time. In September 2020, over 40% of Canadians reported that since the start of the pandemic they were spending more time on social media and messaging services, and about one-third reported they had increased their spending on home and mobile Internet connectivity (Statistics Canada 2020a).

From February to May 2020, total retail sales fell 17.9%, while retail e-commerce doubled, reaching \$3.9 billion (Aston et al. 2020). This accelerated an ongoing shift, with the share of total retail sales comprised of e-commerce increasing from 2.4% to 4.0% from 2016 to 2019, and then surging to 11.4% in April 2020.

In the workforce, the share of paid employees who were teleworking reached 40% in April 2020—comparable to the estimated share of Canadian workers estimated to be in jobs that could possibly be carried out from home (Statistics Canada 2020b). As Deng, Morissette, and Messacar (2020) note, this suggests that “...the Canadian labour market responded very quickly to the onset of the pandemic by increasing its prevalence of telework to the maximum capacity.” The share of businesses using telework arrangements increased from 17% to 33% from February to May 2020, with such arrangements used by over half of firms in some industries (Statistics Canada 2020c). One-in-seven workplaces (14%) said it was likely or very likely to continue to use such arrangements once the pandemic is over.

The transformative impacts of technological change have long been recognized, and the shifts observed through the pandemic are part of the ongoing digital transformation. Nonetheless, the prevalence and impacts of digital technologies have been highlighted through the pandemic. This includes recognition that not all Canadians are equally well-positioned to cope with, adapt, or even thrive in the increasingly online environment (Frenette, Frank, and Deng 2020). The extent to which there is widespread capacity across individuals to benefit from the opportunities and avoid the risks presented by the digital transformation is an on-going public policy concern. This is often presented in terms of “digital divides”—the distinction between digital ‘haves’ and ‘have-nots.’ This divide may stem from various factors, such as the quality and availability of digital infrastructure, access to digital devices, and the skills to use both (OECD 2019).

So how well were Canadians positioned in terms of digital skills and capacities on the eve of the COVID-19 pandemic? As life in Canada was about to go increasingly online, to what extent might they best be characterized as basic, intermediate, or advanced users of the Internet and digital technologies? This paper provides such a perspective, focusing on the online activities and skills of Canadians as reported on the Canadian Internet Use Survey (CIUS) fielded from November 2018 to March 2019. CIUS information is used to classify respondents into five mutually exclusive categories, from those who were highly skilled and adept users of the Internet and digital technologies to those that did not use them at all. This provides a pre-pandemic benchmark of how Canadians were engaging in and adapting to the changing digital environment at that time. Current and future Internet use surveys during and after the pandemic will provide subsequent snapshots.

A typology is offered, comprised of four groups of Internet users differentiated by the online activities and digital skills they exhibit, and a fifth group comprised of non-users of the Internet. Previewing the results, many Canadians engage in a wide range of online activities using a broad range of skills, with over one-half identified as either ‘Proficient users’ or ‘Advanced users’. In contrast, one-quarter of Canadians either do not use the Internet at all or do so in only the most basic way. This activities- and skills-based approach sheds new light on the digital divide in Canada.

The paper is organized into four sections. Section 2 below provides a review of the relevant research literature. Section 3 presents the data sources and methodology used. Section 4 presents the results of the analysis, detailing the distribution of Canadians across the Internet-use typology, the features of each group, and the sociodemographic characteristics associated with inclusion in each. Section 5 discusses the implications and limitations of the study.

2 Literature review

One way to assess Canadians' capabilities in an increasingly digital society is to develop a typology. Internet-user typologies divide people into discrete groups based on tendencies in their online activities and practices. These typologies are useful for understanding Internet use in general and also for examining how differences in Internet use correspond to sociodemographic characteristics and other individual-level attributes (Blank and Groselj 2014; Brandtzæg, Heim, and Karahasanović 2011). Such differences, or digital divides, have been described as a “new form of social inequality” that affects life chances via the activities that people conduct online (Zillien and Hargittai 2009). Previous studies have shown that digital divides follow traditional forms of inequality since characteristics such as education, age, and income associate with the types of activities that people tend to conduct online (Büchi, Just, and Latzer 2016; Helsper and Galácz 2009; Lutz 2019; Robinson, Winborg, and Schulz 2018; Scheerder, van Deursen, and van Dijk 2017).

The Internet is a “multipurpose infrastructure” that can be used for an extensive range of activities and purposes (Büchi, Just and Latzer 2016). With the growth of the Internet, patterns of Internet use have become progressively more complex and differentiated. Developing an Internet-user typology involves condensing this plethora of usage into a smaller set of empirically-derived categories (Blank and Groselj 2014). Typical points of comparison have included the amount of Internet use, variety of use, and content preferences (Blank and Groselj 2014; Brandtzæg, Heim and Karahasanović. 2011; Montagnier 2007; Spiezia and Montagnier 2010). Amount of use compares people on rates of Internet access (users versus non-users) and on frequency (daily, weekly, monthly) and duration (hours per day or week) of Internet use. The variety of Internet use refers to the number of activities an individual performs online, capturing the extensiveness of online engagement. Content preferences refer to the different types of activities performed online, such as social media, entertainment, financial transactions, and searching for information.

In the first decade of the Internet, the key points of differentiation among Canadians were between users and non-users of the Internet and how frequently individuals used the Internet (Middleton and Sorensen 2005; Singh 2004). This refers to the “first-level” digital divide, which stemmed largely from the uneven distribution of telecommunications infrastructure across regions and the affordability of technology across households. The first-level divide has disappeared for most Canadians. In 2018, about 91% of Canadians reported using the Internet for personal use (this excludes work- and school-related use) in the previous three months (Statistics Canada 2021). However, gaps in Internet access still persist. According to a recent CRTC (2020) report, rural areas¹ and the territories trail well behind the rest of Canada in access to high-speed Internet service. About 59% of households in rural areas do not have access to broadband services that meet the CRTC benchmark (50 Mbps download and 10 Mbps upload speeds and unlimited data transfer) and these services are unavailable throughout the Yukon, Northwest Territories, and Nunavut. In contrast, these broadband services are available to almost all Canadian households in places that have populations of 100,000 persons or more.

Beyond access to infrastructure, attention regarding the digital divide has turned to *how* the Internet is accessed (Frenette, Frank and Deng 2020; Hargittai 2002; Napoli and Obar 2014; van

1. Population of less than 1,000 or a population density of fewer than 400 persons per square kilometre.

Deursen and van Dijk 2019). Napoli and Obar (2014) argue that there are qualitative differences in Internet access between people who use multiple devices and those who depend on mobile devices alone. A smartphone may allow students to check class schedules or grades but is less optimal than a laptop or desktop computer for writing a term paper or other tasks that are easier to perform with a full keyboard and larger screen. Among Canadian households with children under age 18, 24% of those in the lowest income quartile depend on mobile devices to access the Internet compared with 8% of those in the highest income quartile (Frenette, Frank and Deng 2020). One implication is that some Canadian children, particularly those from lower income households, may not have devices best suited to participation in online learning activities during COVID-19 lockdowns.

Overall, in Canada and other countries where Internet use is at near-saturation levels, the distinction between haves and have-nots is no longer primarily a matter of barriers to Internet access or frequency of use (Borg and Smith 2018; Scheerder, van Deursen, and van Dijk 2017). More recent studies distinguish Internet-user groups based on a combination of amount of use and the extensiveness and types of activities that people conduct online (Blank and Groselj 2014; Haight, Quan-Haase, and Corbett 2014; Middleton, Veenhof, and Leith 2010; Montagnier and Wirthmann 2010). On the latter, online behaviours have been differentiated in terms of capital-enhancing versus leisure-oriented activities (Borg and Smith 2017; van Deursen and van Dijk 2015). Capital-enhancing activities are those that provide offline benefits, such as using the Internet for job searches, e-learning, financial transactions, or professional networking. Research indicates that individuals in disadvantaged groups tend to use the Internet comparatively less for capital-enhancing activities and comparatively more for leisure (Helsper and Galácz 2009; van Deursen and van Dijk 2014; Zillien and Hargittai 2009). In other words, having equal access to the Internet is not equivalent to having equal opportunities to benefit from it (Brandtzæg, Heim, and Karahasanović 2011; Büchi, Just, and Latzer. 2016). Among people connected to the Internet, digital divides arise from differences in human capital and social capital needed to adopt new technologies and leverage Internet-enabled opportunities into benefits (Chen 2013; Büchi, Just and Latzer 2016; Korupp and Szydlik 2005). Skills are critical in this regard.

Van Dijk and van Deursen (2014) define six types of Internet skills.² To begin, the *operational skills* that are needed to use computer hardware, software programs, and Internet applications and the *formal skills* that are needed to navigate the hypermedia structures of the Internet. But digital literacy involves more than technical skills. Van Dijk and van Deursen describe four content-related competencies needed to benefit from Internet use, including *information skills*, *communication skills*, *content-creation skills*, and *strategic skills*, and argue that the ability to exercise one type of skill is contingent on competencies on other types of skills. For example, conducting communication-based activities requires basic technical proficiencies to navigate websites, the ability to find and use information, and perhaps even the ability to create content. The breadth of skills involved raises the prospect that deficits in one area will impede progress in others, diminishing overall capacity to engage in, and derive benefit from, the Internet and digital technologies.

Rapid change in Internet use, particularly the constant evolution of online activities, underscore the need to continually update the metrics used to differentiate Internet users (Büchi, Just, and Latzer 2016). The amount of “know-how” needed for digital literacy continues to advance with technological innovations and the shift of both routine activities and essential services into digital environments. This places the emphasis on the skills needed to use Internet-enabled devices, to establish and manage online connectivity, to use software and applications, and effectively navigate the Internet (Hargittai and Micheli 2019; Scheerder, van Deursen, and van Dijk 2017). These elements, broadly defined as both activities and skills, are reflected in the approach to Internet-user groups outlined next.

2. For a complementary review of Information and Communication Technologies (ICT) skills see OECD (2016).

3 Data and methods

3.1 Data

Data for the Internet-user typology presented in this study are from the 2018 Canadian Internet Use Survey (CIUS). The 2018 CIUS was completed by a nationally representative sample of almost 14,000 Canadians aged 15 years and older living in the 10 provinces, excluding full-time residents of institutions. The survey was fielded from November 2018 to March 2019, just over a year prior to the onset of the COVID-19 pandemic. The CIUS collected information on individuals' use and experiences with the Internet and digital technologies, including the types of activities and tasks they perform, their encounters with security threats and online harassment, and their digital and software skills.

The CIUS includes variables identifying whether or not respondents had home access to the Internet and whether or not they used the Internet in the past three months. The typology presented here treats 'Non-users' as a distinct group, being one of the five groups identified in the analysis. This group is set aside in generating the typology as described below, but is included in the presentation of results.

In generating the typology, we focused on variables that gave a profile of the typical online activities respondents engaged in, as well as those variables that demonstrated proficiency with digital technologies. Activity variables related to communication, accessing information, entertainment, and e-commerce, as well as skill variables related to learning, software use, privacy, and personal device features were all included in generating the typology. Other variables related to where and how respondents accessed the Internet were not used to generate the typology, but are included as supplementary information.

3.2 Methodology

We use a cluster algorithm approach to generate our typology of Internet users in Canada. Cluster algorithms are designed to group a dataset in such a way that within each cluster group, observations are more 'similar' to or 'near' one another than they are with observations in other groups. When appropriately applied, the cluster groups can be interpreted as distinct typological groups of users. Furthermore, this approach reduces bias in the results by making the process data-driven and less dependent on preconceived ideas of which activities or skills are more or less important than others.

Our typology utilizes the k-modes algorithm developed in Huang (1998) to generate typology groups on Internet users, omitting non-users and assigning them to a separate group *ex ante*. The k-modes algorithm is related to the popular k-means algorithm, which generates k cluster groups on a dataset given a pre-defined number k . The k-means algorithm has been known to produce arbitrary cluster results on discrete data, whereas the k-modes algorithm does not. Given the nature of the CIUS data and Internet surveys in general, the k-modes seems well suited for generating our user typology, and is employed as the primary algorithm in this study for generating the final typology.³

Given a dataset of N observations with P variables and a desired number of groups k , the k-modes algorithm produces a vector of length N that assigns each observation to one of k cluster groups, as well as a $k \times P$ matrix of 'modes'. Each row in this matrix of modes represents a mode

3. A cluster approach, as opposed to other dimensionality-reducing methods such as principal component analysis (PCA), has the advantage of compressing data points themselves into groups as opposed to compressing variables, which is the case for PCA. This results in a typology based on user profiles as opposed to correlations between activities and skills, which was not particularly informative given the volume and diversity of CIUS variables.

in the data space⁴ that has iteratively been determined to be central to each cluster group. That is to say, an observation belongs to group K if and only if the observation is nearest to the mode corresponding to group K . In the k-modes algorithm, 'nearness' is based on simple matching between two records for each of the P variables. Two observations have distance zero if and only if they are identical in each variable, and the distance is otherwise equal to the number of variables in which they differ.

3.3 Variable selection

The CIUS contains questions on a wide variety of activities and skills related to Internet use. In total, 64 variables within the CIUS data are potentially relevant for generating our typology. Following recent literature (e.g., Blank and Groselj 2014; Borg and Smith 2017; van Deursen and van Dijk 2015), we limit the generation of the typology to variables pertaining to activities and skills, and leave other variables related to the type or frequency of use for subsequent analysis.

Many of the activity and skill variables in the data have relatively low uptake. Some variables, such as online gambling, may have low uptake due to low demand or low confidence in engaging with that activity. Others, such as searching for employment or receiving free training from a community centre or seniors care facility are dependent on other factors and may not be widely applicable to the whole population. Additionally, there are some variables that are nearly ubiquitous across Internet users, such as the use of email and researching information.

Because the k-modes algorithm determines distances between observations using a simple matching metric, it is particularly sensitive to differences in any variable that is used to generate cluster groups. In particular, low-uptake variables will have an impact on the final clustering when a respondent's value for that variable is one. Similarly, high-uptake variables will have an impact if a respondent's value is zero. As such, it is desirable to remove these variables from consideration in generating the analysis, as they add 'noise' to the clustering that will impact final results.

A selection criteria was developed to eliminate variables with high or low uptake so that the typology was generated only on those variables with sufficient variation. (Details of the selection criteria can be found in Section 6.2). In total, 25 variables were removed due to low uptake and 3 were removed due to high uptake, leaving 36 variables to be used in generating the typology. The three variables removed due to high uptake or near ubiquity in the data corresponded to using email, searching for information online, and checking the weather. Although these variables are not included in generating the typology, they are included in parts of the results and analysis section.

3.4 Generating the typology

The k-modes cluster algorithm initializes by randomly selecting k modal observations from the data and iteratively generating cluster groups. At each iteration, it determines which points in the data are closest to each of the k modes, and redefines the mode to be the observation that minimizes within-group distances for all points relative to the mode. Because the algorithm begins by randomly selecting k observations, it does not generate identical cluster groups between applications. As such, an individual application of the k-modes algorithm does not generate a stable typology.

To overcome this limitation, we ran the k-modes cluster algorithm a large number of times (2000) and selected the specification that yielded the highest within-group homogeneity of observations.

4. Modes are not guaranteed to correspond to observations in the data; however, each entry in each mode does correspond to a category present in the data. For example, if a variable has observed values $\{a,b,c,d\}$, then each mode's value at that variable must be one of $\{a,b,c,d\}$.

Cluster algorithms seek to find groupings where observations within a group are more similar to one another than to observations in other groups, so choosing the best fit out of a large sample of applications of the algorithm produces the optimal result for the typology. (Details on determining best fit, as well as greater detail about the approach taken to generate the 2000 specifications can be found in Section 6.2.)

Although our final typology is based on the best fit from a large number of specifications, it is important to compare it against all other specifications generated to ensure our results are not coincidental. With this in mind, we present our results for the best fit cluster typology in the upcoming sections, and include a more detailed robustness analysis in Section 6.2.

4 Results

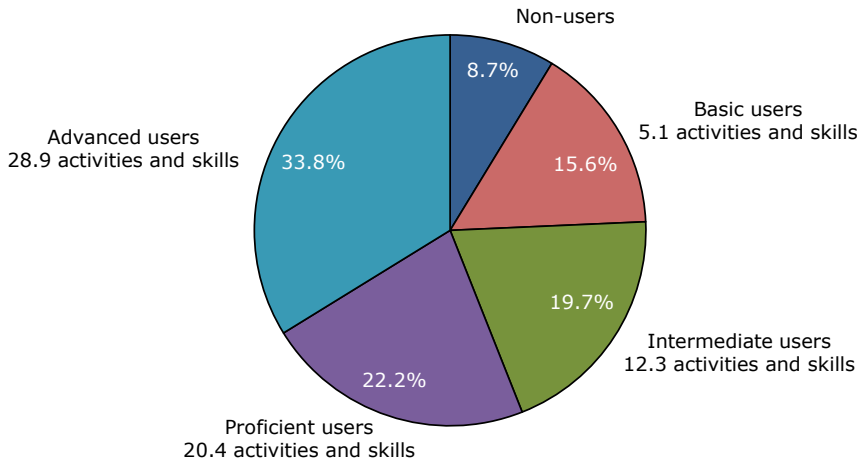
4.1 Internet-user groups—on-line activities and digital skills

The cluster analysis yielded a typology composed of four distinct groups of Internet users, to which is added a fifth group of Internet non-users. The four user groups are ordered from ‘Basic’ to ‘Advanced’, reflecting the activities and skills of their constituents. Almost 9% of Canadians were Non-users of the Internet and just under 16% were ‘Basic users’ (Chart 1). Of the 36 activities and skills used to define the cluster groups, Basic users engaged in 5.1 on average.⁵ Considering both groups together, just under one-in-four Canadians (24%) had either no engagement with the Internet or very limited engagement when surveyed in late 2018 and early 2019. These groups could be argued to be on the have-not side of the digital divide.

Another 20% of Canadians were identified as ‘Intermediate users’ of the Internet and digital technologies, with a further 22% identified as ‘Proficient users’. Individuals in these groups on average did, respectively, 12.3 and 20.4 of the 36 skills and activities used in the cluster analysis, with differences also observed in terms of the sophistication of skills. Finally, 34% of Canadians were identified as ‘Advanced users’ on the basis of the activities and skills they exhibit. On average, they did 28.9 of the 36 activities and skills used in the cluster analysis.

5. The average number of activities across all CIUS respondents for the 36 activities (variables) included in the cluster specification was 17.5.

Chart 1
Distribution of Canadians aged 15 or older across Internet-use typology, Canada, 2018



Note: Activities and skills: average number reported by group; maximum = 36.
Source: Statistics Canada, 2018 Canadian Internet Use Survey.

The shares of individuals in each Internet-user group exhibiting the 36 activities and skills underlying the typology are presented in Table 1. A selection of those activities and skills are shown graphically in Chart 2. Most Basic users, like Internet users overall, use the Internet for email (74%), checking the weather (60%), and finding basic information (60%).⁶ Aside from these high-uptake activities, Basic users engaged in other common activities to a much lesser extent than other user groups. For example, less than 40% of Basic users use the Internet to access the news, obtain directions, or do their banking and less than 30% use instant messaging or social media. The vast majority of individuals in the other Internet-user groups do these things. It is plausible that keeping up in a rapidly changing digital environment could be a challenge for individuals in the Basic-user group.

In contrast to Basic users, far larger shares of individuals in the Intermediate-user group engaged in online communications, with about three-quarters of them using instant messaging and social media. Intermediate users were also more likely to consume online entertainment, with almost two-thirds of them listening to music and using video-sharing websites such as YouTube and almost half (45%) watching streamed content. Intermediate users also exhibited significant uptake of e-commerce activities such as online banking (81%) and buying new goods and services online (58%).

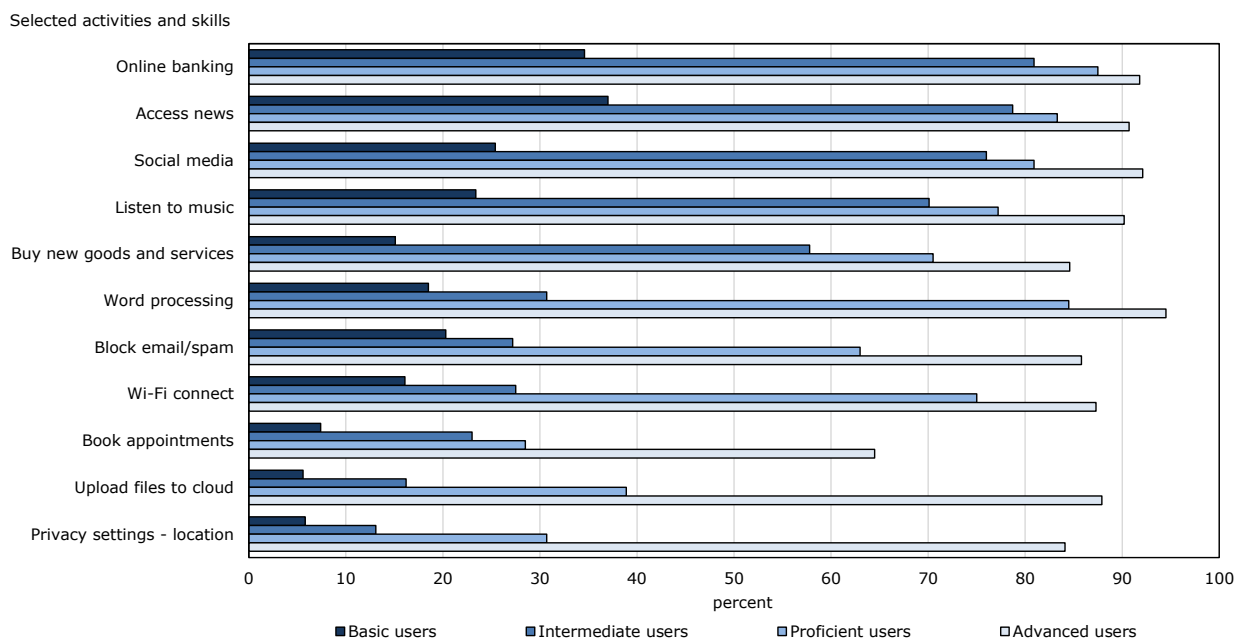
The Proficient-user group is generally more adept with technical skills related to basic office skills, as well as uploading files to the cloud, browser and email management, and wireless connectivity. A large majority (84%) of them use word processing software, move folders and files (82%), use spreadsheet applications (61%), and download files (62%). Relative to Intermediate users, Proficient users were more likely to be familiar with Wi-Fi (75%) and Bluetooth (71%) connectivity, and to consume paid streaming services.

The Advanced-user group exhibits the most activities and skills. Some 80% or more of Advanced users adjust their security settings to limit personal and location information, and most (88%) upload files to a cloud storage service of some kind, many doing so for the purpose of sharing

6. These activities were not used in the cluster analysis, as their high uptake limits their utility in differentiating internet-user groups. See Appendix Table 1.

(73%) or backing up files (75%). Online technologies appear to be part of Advanced users' daily lives, with most booking appointments (65%), registering for courses or checking class schedules (67%), or conducting video/audio calls online (74%).

Chart 2
Percentage of Internet users engaged in selected activities and skills, by Internet-user group, Canada, 2018



Source: Statistics Canada, 2018 Canadian Internet Use Survey.

In addition to the 36 activities and skills used in the cluster analysis, information can be gleaned from the 28 activities and skills that were considered but not used. The shares of individuals in each Internet-user group exhibiting these excluded activities and skills are shown in Appendix Table 1. Although many of the activities and skills presented in this table are generally less common than the variables used in the cluster analysis, the magnitude of the difference between Basic users and other users and the difference Advanced users and other users are both notable. On most items presented in Appendix Table 1, there is a fairly large increase in the uptake when moving from a Basic user to an Intermediate user. While the difference between other user groups tends to be more modest and is small on some items (e.g., emailing, checking the weather), there is still a general increase between adjacent user groups on all items. Advanced users have the highest uptake on all items, particularly on more sophisticated activities such as uploading content, blogging, online training, and software skills.

Table 1
Percent of Canadian Internet users who reported doing on-line activities and skills, by Internet-user group, Canada, 2018

	Basic user	Intermediate user	Proficient user	Advanced user
	percent			
Communication				
Instant messaging	28.4	77.9	85.5	94.2
Contact government	27.3	67.0	80.2	90.2
Social media	25.4	76.0	80.9	92.1
Online voice/video call	13.1	35.7	57.5	73.6
Book appointments	7.4	23.0	28.5	64.5
Accessing info				
Access new s	37.0	78.7	83.3	90.7
Obtain locations/directions	37.9	88.6	94.2	97.6
Entertainment				
Listen to music	23.4	70.1	77.2	90.2
Watch video-sharing (e.g., YouTube)	18.8	68.4	79.5	91.8
Watch streamed content	16.2	45.2	71.6	82.9
Read books and magazines	5.0	16.2	25.0	44.3
Listen to podcast	3.6	18.7	25.7	45.4
E-commerce				
Online banking	34.6	80.9	87.5	91.8
Buy new goods and services	15.1	57.8	70.5	84.6
Training/education				
Check schedules / register for classes	5.1	23.6	29.5	67.1
Informal training/learning	3.3	16.2	23.3	49.7
Software skills				
Word processing	18.5	30.7	84.5	94.5
Copy/move files/folders	14.8	26.8	81.9	92.9
Spreadsheets	9.0	14.7	61.2	82.7
Upload files to cloud	5.6	16.2	38.9	87.9
Photo editing	4.3	10.2	31.8	72.7
PowerPoint, etc.	4.4	11.0	40.9	78.0
Internet skills				
Delete browser history	24.0	32.8	69.5	80.1
Block email/spam	20.3	27.2	63.0	85.8
File download	8.5	17.4	62.3	86.7
Privacy settings - personal info	6.5	13.2	27.3	78.3
Block other messages	5.9	11.9	22.5	61.5
Privacy settings - location	5.8	13.1	30.7	84.1
File backup via cloud	4.3	10.0	22.9	75.4
File share via cloud	1.6	5.9	18.1	73.0
Device skills				
OS updates	18.0	21.2	60.4	78.1
Smartphone as GPS	17.4	37.8	83.9	90.8
Wi-Fi connect	16.1	27.5	75.0	87.3
Photo/video transfer between devices	11.3	20.1	65.7	82.4
Bluetooth connect	9.6	22.0	70.8	84.0
Device settings - location	4.8	12.2	33.8	81.6
	number			
Average number of activities and skills reported (Maximum = 36)	5.1	12.3	20.4	28.9

Note: OS: Operating system; GPS: Global Positioning System.

Source: Statistics Canada, 2018 Canadian Internet Use Survey.

Three batteries of additional questions from the CIUS provide additional insight on the Internet-user groups. The first set of questions asked individuals about whether they use each of six types of Internet-enabled devices. As shown in Table 2, Smartphones were most prevalent, used by 44% of Basic users and by virtually all Proficient and Advanced users. Laptops and netbooks are also widely used, with prevalence ranging from 40% of Basic users to 83% of Advanced users. Desktop computers and tablets were used by smaller shares of individuals in each group, but were quite prevalent. The average number of devices used by individuals in each group is shown at the bottom of Table 2. This increases steadily from an average of 1.6 devices used by Basic users to 3.5 devices used by Advanced users.⁷ Looking more closely, 12.4% of Basic users accessed the Internet only using a Smartphone compared with 9.9% of Intermediate, 3.3% of Proficient, and 1.6% of Advanced users.

The second battery of questions asked respondents about the locations where they access the Internet. Most individuals accessed it from home, with this share ranging from 86% among Basic users to 97% among Advanced users. Very few Basic users accessed the Internet elsewhere, such as in business establishments or other public locations. The preponderance of seniors and near-seniors in this group is an important consideration, resulting in fewer in the Basic-user group accessing the Internet at work or school, and perhaps less general mobility associated with health or access issues. Still, it is important to note that digital engagement does not occur outside the home for most Basic users. Individuals in this group accessed the Internet in an average of 1.4 locations, compared with 3.9 among Advanced users.⁸ And while 62% of Basic users only accessed the Internet at home, just 7% of Advanced users did so.⁹ For the latter, more wide ranging access is consistent with a profile in which digital technologies and Internet connectivity are a greater part of their daily lives.

The third set of questions aims to capture the digital nature of daily lives is the use of Internet-enabled home technologies. As shown in Table 2, the shares reporting that they use a Smart TV ranged from 19% among Basic users to 48% among Advanced users. One interpretation is that, although Smart TVs are increasingly prevalent, the differences above reflect variation in the use of their Internet connectivity.¹⁰ Given the dominance of smart TVs in the market as of 2018, it may also simply reflect who most recently bought a TV. Smart speakers and other Internet-enabled devices are perhaps clearer measure of home technology adoption. The use of Smart speakers ranges from 4% among Basic users to almost 22% of Advanced users, while the use of other Internet-enabled devices derived from six CIUS questions, ranged from 7% to 31%. So even among the most advanced group, the adoption of such technologies was not widespread in late 2018 or early 2019, but followed the expected pattern of increasing prevalence across the four user groups.

7. On a range of 1 to 6.

8. On a range of 1 to 7.

9. In addition, 33% of Intermediate users and 20% of Proficient users only accessed the internet from home.

10. The CIUS defined Smart TVs as those "...with integrated Internet capabilities or a set-top box for television that offers computing ability and Internet connectivity." (Statistics Canada n.d., Question UI_23).

Table 2
Selected Internet-related characteristics, by Internet-use group, Canada, 2018

	Basic user	Intermediate user	Proficient user	Advanced user
	percent			
Access Internet using ...				
Smartphone	44.0	79.7	92.9	97.5
Laptop or netbook	40.0	56.1	69.9	82.8
Desktop computer	35.3	37.7	49.5	56.1
Tablet	31.8	44.5	51.4	52.2
Media streaming device	4.3	15.1	27.1	42.1
Other mobile devices	1.9	5.8	7.8	16.6
Access Internet ...				
At home	86.3	92.7	94.8	97.3
At work	18.3	42.6	55.3	66.5
At someone else's home	15.5	41.3	52.8	71.9
At business establishment	8.6	30.1	41.3	62.2
At another public location	7.4	24.1	34.0	53.0
At school	4.2	10.5	13.8	26.7
At public library	4.2	8.0	9.8	17.3
Internet-enabled home technologies				
Smart TV	18.6	34.8	46.3	48.0
Smart speaker	3.9	10.4	15.0	21.5
Other Internet-enabled device ¹	6.8	13.4	20.0	31.3
	number			
Mean number of Internet-enabled devices	1.6	2.4	3.0	3.5
Mean number of access locations	1.4	2.5	3.0	3.9
Mean number of Internet-enabled home devices	0.3	0.7	0.9	1.2

1. This includes the following: Video cameras connected to the Internet; Smart door or window locks; Smart thermostats; Smart switches or lights; Smart large appliances; and other devices not including smart TVs or speakers.

Source: Statistics Canada, 2018 Canadian Internet Use Survey.

4.2 Sociodemographic characteristics of individuals in user groups

Thus far, the analysis has focused on the activities and skills exhibited by individuals in each user group. Sociodemographic characteristics are now incorporated into the analysis. Table 3 shows how Canadians with specific sociodemographic characteristics are distributed across the five groups in the typology.

Age, educational attainment, and income¹¹ are strongly associated with the distribution of individuals across the Internet-user groups. As expected, the shares of individuals in the Advanced-user and Proficient-user groups decline across age groups. Over one-half (54%) of Canadians aged 15 to 34 are Advanced users and another 23% are Proficient users of the Internet and digital technologies. Together, almost 8-in-10 individuals in this age range are in the two most tech-savvy groups. About two-thirds of individuals aged 35 to 49 are in these two groups, with 42% classified as Advanced users and 26% as Proficient users. A smaller share of individuals aged 50 to 64 are classified as Advanced users (23%), but still, almost half of individuals in this age group are either Advanced or Proficient users. In contrast, over 60% of Canadians aged 65 or older are either Non-users or Basic users of the Internet and digital technologies. Age profiles *within* Internet-user groups (i.e., compositional characteristics) offer an additional perspective and are shown in Appendix Table 2. For example, of all individuals in the Non-user group, 68% are aged 65 or older and 88% are aged 50 or older.

11. Incomes in the CIUS are based on a combination of self-reported income and imputation based on demographic variables.

Table 3
Distribution of individuals across Internet-user groups, by socioeconomic characteristics,
Canada, 2018

	Non- users	Basic users	Intermediate users	Proficient users	Advanced users	Total
	percent					
Age group						
15 to 34 years	1.2	5.8	16.0	23.2	53.7	100.0
35 to 49 years	2.7	8.2	21.2	25.9	42.0	100.0
50 to 64 years	7.1	20.0	24.8	24.8	23.2	100.0
65 years or older	28.8	33.8	17.6	13.1	6.7	100.0
Educational attainment						
Less than high school	40.9	29.2	16.4	8.4	5.0	100.0
High school diploma	16.2	23.6	21.9	19.3	19.0	100.0
Non-university postsecondary	6.1	16.7	20.9	27.0	29.4	100.0
Bachelor or higher	2.1	8.0	13.9	27.6	48.4	100.0
Student	0.9	5.3	14.2	20.9	58.7	100.0
Household income						
\$24,999 or lower	29.5	22.6	17.7	13.9	16.4	100.0
\$25,000 to \$49,999	14.9	22.6	20.4	18.5	23.5	100.0
\$50,000 to \$74,999	5.8	18.8	23.1	21.8	30.5	100.0
\$75,000 to \$99,999	2.9	13.4	22.8	25.5	35.4	100.0
\$100,000 or more	1.1	7.9	16.8	26.1	48.0	100.0
Place of residence						
Urban	8.3	15.1	19.4	21.9	35.4	100.0
Rural	10.6	18.5	21.7	24.0	25.3	100.0
Household size						
Multi-person household	6.0	14.6	19.9	23.5	36.0	100.0
Single-person household	23.3	21.1	18.8	15.0	21.8	100.0
Immigration status						
Canadian-born	9.0	14.7	16.9	23.2	36.1	100.0
Immigrant	9.6	15.3	18.5	23.9	32.6	100.0
Sex						
Male	7.8	15.3	18.5	24.0	34.4	100.0
Female	9.5	15.9	20.9	20.5	33.2	100.0
Employment status						
Employed	3.0	10.8	20.3	25.2	40.8	100.0
Not employed	18.1	23.5	18.9	17.2	22.2	100.0

Note: Percentages may not add up to 100.0% because of rounding.

Source: Statistics Canada, 2018 Canadian Internet Use Survey.

Educational attainment is strongly correlated with the distribution of individuals across user groups. Almost half (48%) of individuals with a Bachelor's degree or higher were Advanced users, while this was the case for 29% of those with a non-university post-secondary credential, and 19% of those with a high school diploma (Table 3). This may reflect the confounding effect of age as, on average, older Canadians have lower levels of educational attainment than younger Canadians (Ferguson and Zhao 2013). This is partially shown in Appendix Table 2, which shows that 76% of Basic users are aged 50 and older and 79% of Basic users have an educational attainment of less than a Bachelor's degree.

The shares of individuals identified as Advanced users increased steadily across categories of household income, from 16% among those with incomes under \$25,000 to 48% among those with incomes of \$100,000 or more. Similarly, individuals who were employed rather than non-employed were about twice as likely to be Advanced users (at 41% and 22% respectively). Again, these results likely reflect the confounding effects of age and education, two characteristics correlated with income and employment.

Smaller differences in distributions across user groups are observed across other socioeconomic characteristics. An urban/rural difference is observed, with the share of rural residents identified as Advanced users 10 percentage points smaller than the share of urban residents (at 25% and 35%, respectively). Individuals living alone are more likely to be Non-users or Basic users than those living with others.

Finally, differences in the distributions of women and men across user groups were generally small, as were the distributions of immigrants and Canadian-born individuals.

To more precisely estimate the strength and significance of the relationships between each socioeconomic characteristic and the likelihood of being in each user group, an ordered logistic regression model was run. The results are presented as marginal effects in Table 4, interpreted as the percentage point difference in the likelihood of individuals in one category (e.g., female) being in an Internet-user group relative to a reference category (e.g., male), net of other variables in the model.

The multivariate results confirm that age and education are the strongest predictors of being in each of the five user groups. Net of other characteristics in the analysis, individuals aged 15 to 34 were 12.3 percentage points more likely to be classified as Advanced users and 3.9 percentage points less likely to be classified as Basic users than the reference group composed of individuals aged 35 to 49. Seniors, in contrast, were 27.6 percentage points less likely to be classified as Advanced users. Being a senior was also the strongest predictor of being a Non-user or a Basic user, after accounting for education and other characteristics that could confound the relationship between age and Internet use. Persons aged 65 and older were 12.9 percentage points more likely to be classified as Non-users and 16.4 percentage points more likely to be classified as Basic users than the reference group.

Table 4
Estimated percentage point difference in likelihood of being in each Internet-user group relative to reference group

	Non-users	Basic users	Intermediate users	Proficient users	Advanced users
estimated percentage points					
Age group					
Age 15 to 34	-1.69 ***	-3.90 ***	-4.32 ***	-2.37 ***	12.28 ***
Age 35 to 49 (reference group)
Age 50 to 64	3.89 ***	6.79 ***	5.01 ***	-1.02 ***	-14.67 ***
Age 65 or older	12.90 ***	16.35 ***	6.69 ***	-8.37 ***	-27.56 ***
Educational attainment					
Less than high school	11.41 ***	8.06 ***	0.51	-7.43 ***	-12.55 ***
High school (reference group)
Non-university postsecondary	-3.54 ***	-4.25 ***	-2.19 ***	1.39 ***	8.60 ***
University degree or higher	-6.84 ***	-9.81 ***	-6.41 ***	0.35	22.70 ***
Enrolled as student	-6.57 ***	-9.27 ***	-5.93 ***	0.62	21.14 ***
Household income					
Income less than \$25,000	6.54 ***	6.22 ***	2.83 ***	-2.08 ***	-13.50 ***
Income \$25,000 to \$49,999	1.47 **	1.67 **	0.97 **	-0.16 *	-3.95 **
Income \$50,000 to \$74,999 (reference group)
Income \$75,000 to \$99,999	-0.67	-0.83	-0.54	-0.03	2.06
Income \$100,000 or more	-2.52 ***	-3.39 ***	-2.41 ***	-0.57 ***	8.90 ***
Place of residence					
Urban (reference group)
Rural	0.89 *	0.86 *	0.57 *	0.06 **	-2.38 *
Household size					
Multi-person household
Single-person household	1.41 ***	1.37 ***	0.88 ***	0.05	-3.72 ***
Immigration status					
Non-immigrant (reference group)
Immigrant	1.29 ***	1.24 ***	0.81 ***	0.10 **	-3.45 ***
Sex					
Male (reference group)
Female	0.42	0.41	0.28	0.05	-1.16
Employment status					
Employed (reference group)
Not employed	1.69 ***	1.76 ***	1.13 ***	0.06	-4.63 ***

... not applicable

* significantly different from reference category ($p < 0.05$)

** significantly different from reference category ($p < 0.01$)

*** significantly different from reference category ($p < 0.001$)

Source: Statistics Canada, 2018 Canadian Internet Use Survey.

The strong relationship between educational attainment and membership in user groups also remains evident in the multivariate results. For example, compared with individuals who have a high school diploma, those with a university degree were 22.7 percentage points more likely to be Advanced users and 9.8 percentage points less likely to be Basic users. Non-university post-secondary credentials, as well as student enrollment, were also positively correlated with being a Proficient or Advanced user and negative correlated with being a Non-user, Basic user, or Intermediate user of the Internet and digital technologies. The same patterns hold across household income categories.

Individuals with annual household incomes of less than \$25,000 were 13.5 percentage points less likely to be in the Advanced-user group than individuals with household incomes ranging from \$50,000 to \$74,999 (the reference group), while those with incomes of \$100,000 or more were

8.9 percentage points more likely to be in the Advanced-user group, net of education, employment, and other characteristics.

Smaller differences in the likelihood of being in each user group are observed across other variables in the multivariate models. Urban/rural place of residence, immigration status, household size, and employment status are significantly correlated with the likelihood of being in most user groups, although the estimated differences generally range from 1 to 4 percentage points and are far smaller than the correlations associated with age, educational attainment, and household income.

No significant differences in the likelihood of being in each of the five user groups are observed between women and men when other characteristics are taken into account.

4.3 Telework and Internet-user groups

As highlighted in the introduction, the COVID-19 pandemic has had a large impact on employment, including a sharp increase in the use of telework arrangements. The relationships between telework and digital activities and skills likely runs in both directions. Individuals with stronger Internet and digital skills may be more likely than others to telework, while teleworking itself may strengthen the Internet and digital skills of those engaged in it. While the 2018 CIUS does not allow conclusions to be drawn about the direction of causation, it does provide an opportunity to explore the correlation between telework and the Internet-user groups presented above.

CIUS respondents were asked if they had done any telework during the past 12 months.¹² This is quite a broad definition, and no information was collected regarding the frequency, regularity, or intensity of telework over the year. Across all CIUS respondents, 13% of employed individuals reported that they teleworked during the prior year, a rate comparable to estimates from previous years (Turcotte 2010). Among employed individuals, 70% of those who had teleworked were classified as Advanced users compared with 33% of those who had not teleworked.

Evidence from the late 2000s shows that age, education, and income¹³ are among the characteristics positively associated with the likelihood of teleworking. One question this raises is whether the strong relationship between telework and being an Advanced user remains when these characteristics are taken into account. To assess this, the ordered logistic regression model presented above was modified and re-run. Specifically, the analysis was limited to individuals aged 15 to 64, an employment status / telework variable was included, and the Non-user and Basic-user groups were combined into a single category because no teleworkers were in the Non-user group. The rest of the model remains unchanged.

Net of other observed characteristics, employed individuals who teleworked were 26.4 percentage points more likely to be classified as Advanced users than employed individuals who had not teleworked (Table 5). The CIUS does not include information on other important correlates of teleworking, such as occupation and industry, and hence this relationship warrants further scrutiny as additional information becomes available. That said, the result in Table 5 is consistent with the view that digital skills enable individuals to benefit from the opportunities provided by the digital transformation, and highlights the possibility that elevated rates of telework during the COVID-19 pandemic strengthened the digital skills of those able to work in this way.

12. Telework was defined as "...a flexible work arrangement whereby employees have approval to carry out some or all of their work duties from a place other than the employer's office." (Statistics Canada n.d., Question OW_71).

13. In 2008, the incidence of telework was higher among employees in management and professional occupations—two occupational groups with relatively high earnings—than among employees in all other occupational groups (Turcotte 2010).

Table 5**Marginal effects—estimated percentage point difference in likelihood of being in each user group relative to reference category among Canadians aged 15 to 64**

	Non-users and basic users	Intermediate users	Proficient users	Advanced users
	estimated percentage points			
Employed—yes telework	-9.77 ***	-9.67 ***	-6.98 ***	26.42 ***
Employed—no telework (reference)
Not employed	0.48	0.33	0.05	-0.87

... not applicable

*** significantly different from reference category ($p < 0.001$)**Note:** Estimated percentage points calculated as marginal effects. Model also includes age group, educational attainment, household income, urban/rural status, household size, immigration status, and sex.**Source:** Statistics Canada, 2018 Canadian Internet Use Survey.

5 Discussion and conclusions

The economic and social changes associated with digital technologies continues to reshape the lives of individuals, communities, and societies. The breadth of the digital transformation is immense, affecting virtually all aspects of peoples' lives, and the pace of change is quick. The degree to which individuals are able to operate and adapt in this context is an important issue, as it speaks to their capacity to benefit from the opportunities offered by digital technologies and Internet connectivity, while avoiding the risks they pose.

Gauging such capacity is a challenging task given the scope and pace of change, but new tools are needed. The OECD Working Party on Measurement and Analysis of the Digital Economy (WPMADe) recently recommended the development of a technology user typology "...based on the clusters of digital technologies used, purpose and intensity of usage, and personal characteristics to see how those affect well-being outcomes." (Hatem and Ker 2021, p. 24). This project offers such a tool.

Some observations about the typology process itself are warranted. We found that some online activities, such as email and checking the weather are now so ubiquitous as to be unhelpful in differentiating technology-user groups. Other online activities also provided little analytical leverage, mainly because they were less prevalent overall. Overall, the questions needed to construct a robust typology were certainly not clear at the outset of the process. The questions needed for a typology may also differ across countries because of different patterns of use, and would need to be updated from time to time to keep pace with change. Creating a typology that could be operationalized with fewer than 36 questions warrants consideration, if a more parsimonious approach could allow technology-user groups to be identified on a broader range of surveys.

The typology presented above is composed of four groups of Internet users differentiated by the online activities and digital skills, and a fifth group comprised of Non-users of the Internet. About 9% of Canadians are Internet Non-users and another 16% are 'Basic users'. Together, almost one-in-four Canadians (24%) had either no engagement or only very limited engagement with the Internet and digital technologies in late 2018 or early 2019. If both groups are considered to be digital have nots, there may be more people on the wrong side of the digital divide in Canada than previously thought. The majority of these individuals are seniors and near-seniors, raising a host of questions regarding the impacts that their lack of digital engagement has in their lives. What challenges does this create? What opportunities are foregone? The same questions apply to the 11% of 35 to 49 year olds who are Non-users or Basic users. When age is taken into account, the likelihood of being a Non-user or Basic user of the Internet remained significantly correlated with educational attainment and income. As

other analysts have noted, differentiation between digital haves and have nots follows well-established dimensions of vulnerability and marginalization (Büchi, Lutz and Latzer 2016; Reisdorf and Groselj 2017).

For organizations seeking to engage people online, the limited prospects of reaching some populations may matter. It certainly does for e-government and the online delivery of publicly administered programs and services. Over 90% of Canadian youth aged 15 to 34 are at least Intermediate users and many (54%) are Advanced users. Similarly, about 90% of Canadians with a Bachelor's degree or higher are at least Intermediate users and about half (48%) are Advanced users. These demographic groups are well-positioned to access online programs and services.¹⁴ Lessons derived from e-learning through the pandemic will be instructive in this regard. In contrast, almost two-thirds of seniors (63%) are Non-users or Basic users of the Internet and seem poorly-positioned to access programs and services online. The same is true for Canadians with a high school education or less.

The fact that almost one-in-four Canadians have limited or no digital engagement has implications for national statistical offices as well. Household surveys fielded online offer opportunities to reduce the cost and increase the timeliness of surveys. The extent to which Basic users participate in such surveys, and any unobserved characteristics that differentiate them from the rest of the population, are important considerations in survey design.

Considering other user groups, 22% of Canadians were Proficient users of the Internet and another 34% were Advanced users. Individuals in the latter group exhibited the broadest and most in-depth digital engagement. This group demonstrated a high rate of adoption for many activities and appeared to be well-positioned to make the shift on-line that occurred through the COVID-19 pandemic. In particular, the high rate of Advanced users who had experience with telework prior to the pandemic relative to other groups positioned them well for the sudden spike in work from home.

In terms of next steps, a subset of at least the 36 questions used to create the typology should be included in the final cycle of the Canadian Internet Use Survey (CIUS) scheduled to go into the field in late 2022 and early 2023. With this widespread increases in telework, e-commerce, e-learning, and other online activities, one might expect that the online activities and digital skills exhibited by Canadians have increased, perhaps markedly. The CIUS can offer to estimate this in quite precise terms.

14. This is not to diminish access issues associated with access to devices, skills and even infrastructure, as highlighted by Frenette, M., K. Frank, and Z. Deng (2020).

6 Appendix

6.1 Supplementary tables

Appendix Table 1

Percent of Canadian Internet users who reported doing on-line activities and skills that were not used in the cluster analysis, by Internet-user groups, Canada, 2018

	Basic user	Intermediate user	Proficient user	Advanced user
	percent			
Communication				
Email	73.6	93.3	97.4	99.0
Use dating sites and apps	1.1	3.9	5.0	9.5
Upload content	2.3	9.0	8.9	19.7
Blogging	0.6	3.2	4.4	14.4
Accessing info				
Check weather	57.8	91.8	94.7	97.0
Research information	56.5	88.0	95.0	97.8
Search for employment	5.2	18.3	22.6	33.0
Entertainment				
Play video games	16.9	27.6	31.4	43.7
Watch other streamed live TV	4.2	19.8	27.1	40.9
Watch live-streamed sports	1.6	9.1	10.2	16.1
Gambling	0.7	1.9	2.4	2.5
E-commerce				
Donate to registered charity	3.8	10.4	15.5	24.5
Buy used goods	4.3	16.7	21.6	33.1
Homeshare	1.5	4.8	13.2	24.5
Rideshare	2.0	5.6	17.5	35.1
Buy/sell stocks	1.1	5.0	7.3	12.5
Trade goods or services via app/site	0.8	5.2	7.0	14.8
Use virtual wallet	0.6	6.3	10.9	26.7
Training/education				
Receive instructions from friends/family	25.1	22.7	27.9	34.5
Free online training	4.5	9.6	18.2	39.5
Formal training via organization/institution	2.2	9.1	13.5	25.2
Employer-provided training	1.9	3.0	7.7	15.2
Free training via community/senior centre	1.2	1.5	1.2	1.8
Self-paid training	0.9	1.2	3.4	8.0
Other public free training	0.6	1.3	2.2	6.4
Software skills				
Advanced spreadsheet functions	1.5	2.8	15.7	42.2
Coding	0.5	1.4	4.5	22.4
Device skills				
Set router security settings	2.0	2.7	11.7	29.7

Source: Statistics Canada, 2018 Canadian Internet Use Survey.

Appendix Table 2**Compositional characteristics of Internet-user groups, by socioeconomic characteristics, Canada, 2018**

	Non-users	Basic users	Intermediate users	Proficient users	Advanced users	Total
	percent					
Age group						
15 to 34 years	4.5	11.7	25.4	32.7	49.6	31.2
35 to 49 years	7.4	12.4	25.4	27.7	29.4	23.7
50 to 64 years	20.3	31.7	31.0	27.6	17.0	24.7
65 years or older	67.8	44.2	18.1	12.1	4.1	20.4
Total	100.0	100.0	100.0	100.0	100.0	100.0
Educational attainment						
Less than high school	37.5	16.3	7.6	3.0	1.2	8.2
High school diploma	34.3	30.2	23.5	15.7	10.2	19.0
Non-university postsecondary	19.9	32.9	34.5	34.0	24.4	29.3
Bachelor or higher	6.7	15.3	22.3	33.7	39.0	28.4
Student	1.5	5.4	12.1	13.6	25.2	15.1
Total	100.0	100.0	100.0	100.0	100.0	100.0
Household income						
\$24,999 or lower	41.0	17.4	10.8	7.5	5.8	12.0
\$25,000 to \$49,999	38.5	32.3	23.1	18.6	15.5	22.3
\$50,000 to \$74,999	10.2	18.4	17.9	15.0	13.8	15.3
\$75,000 to \$99,999	6.1	15.3	20.6	20.5	18.7	17.9
\$100,000 or more	4.3	16.5	27.7	38.4	46.2	32.5
Total	100.0	100.0	100.0	100.0	100.0	100.0
Place of residence						
Urban	81.2	81.8	83.1	83.4	88.5	84.6
Rural	18.8	18.2	16.9	16.6	11.5	15.4
Total	100.0	100.0	100.0	100.0	100.0	100.0
Household size						
Multi-person household	59.2	79.5	85.6	89.8	90.2	84.8
Single-person household	40.8	20.5	14.4	10.2	9.8	15.2
Total	100.0	100.0	100.0	100.0	100.0	100.0
Immigration status						
Canadian-born	73.8	66.8	60.6	74.1	75.6	70.8
Immigrant	25.4	22.4	21.4	24.6	22.0	22.8
Total	100.0	100.0	100.0	100.0	100.0	100.0
Sex						
Male	44.6	48.3	46.4	53.3	50.3	49.4
Female	55.4	51.7	53.6	46.7	49.7	50.6
Total	100.0	100.0	100.0	100.0	100.0	100.0
Employment status						
Employed	21.7	43.8	64.4	71.2	75.6	62.8
Not employed	78.3	56.2	35.6	28.8	24.4	37.2
Total	100.0	100.0	100.0	100.0	100.0	100.0

Note: Percentages may not add up to 100.0% because of rounding.

Source: Statistics Canada, 2018 Canadian Internet Use Survey.

6.2 Methodology and robustness checks

This appendix provides additional detail on the procedures used to generate the typology and results from robustness checks from the large set of cluster specifications generated.

6.2.1 The k-modes algorithm

The k-means algorithm is one of the most well-known algorithms in unsupervised machine learning, and is commonly used for the purposes of clustering data due to its quick runtime and wide applicability. Given a pre-defined value k , the k-means algorithm generates cluster groups by iteratively calculating the Euclidean distance between all points in the data and each of k 'means', assigning each data point to the cluster corresponding to the closest mean, and re-calculating the mean of each cluster group to feed into the next iterative step. The algorithm terminates after a pre-specified number of iterations, or when data points no longer move between groups.

The k-means algorithm is well-suited for continuous datasets, and although it produces results that are sensitive to initial (random) conditions, it generally performs well in segregating data. However, it has been known to produce arbitrary results on discrete datasets, in particular because discrete data are congregated in a finite number of points within a continuous interval. Furthermore, Euclidean distance is not an intuitive measure of distance on discrete data, and is in fact not applicable for variables with more than 2 categories and no ordinal properties.

The k-modes algorithm developed in Huang (1998) is specifically designed to overcome the limitations of the k-means algorithm with respect to discrete datasets. It does so by replacing the concept of a 'mean' (the average value for each variable within a cluster group) with a mode (based on the most commonly observed values for each variable in a cluster group), and by replacing Euclidean distance with simple matching distance.¹⁵ Huang (1998) further demonstrates that the k-modes algorithm performs well with real-world datasets and accurately classifies discrete data within a reasonable bound of error for a wide range of data.

The limitations of the k-modes algorithm as it regards generating the user typology should be noted. First, distances between observations can only take on integer values, hence the likelihood of observations being equidistant to two different modes is relatively higher than for the k-means algorithm. In such a circumstance, observations are randomly assigned to a cluster group. Secondly, modes within cluster groups need not be unique, especially if each variable in the parameter space is uniformly distributed.¹⁶ Third, distances are bounded above by the number of variables used in generating cluster groups—if P variables are used to generate the clusters, then the largest number of variables for which two observations can differ is P . A result of this is that the k-modes algorithm on binary data appears to generate cluster groups that are 'ordinal' and roughly order the data from observations with lower average positive responses to higher average positive responses.

Caution should therefore be exercised in interpreting the meaning of cluster groups in this typology. Cluster algorithms generate groups that are 'predictive' in the sense that given a set of responses to the Canadian Internet Use Survey (CIUS), a respondent can be placed into a single cluster group based on their distance to each group's corresponding mode. However, given a cluster group, one cannot necessarily generate a representative user profile corresponding to that

15. It's worth noting that, when applied to discrete binary data, the simple distance metric is equivalent to the Hamming distance on binary strings of length P for P binary variables, and that the set of binary strings of length P combined with the Hamming distance form a metric space. Simple matching distance on binary data is therefore an intuitive and 'natural' way of measuring distance between any two observations.

16. This issue is relatively minor in our case, as the binary nature of the data combined with the large number of observations reduces the probability that distribution within a given cluster group is exactly 50% for positive and negative responses.

cluster group that is representative of the entire group. The typology therefore does not reveal any underlying fact corresponding to each group, but rather serves to label respondents with regards to their similarity to one another.

6.2.2 Establishing number of clusters

The k-modes algorithm requires that the value k representing the number of clusters to be generated must be specified beforehand. Many of the criteria for establishing the optimal number k that are often applied to k-means specifications can also be applied to the k-modes algorithm. Here, the value of k that is optimal is such that within-group similarity of clusters is maximized while out-of-group similarity is minimized.

We use the elbow method and the gap statistic to determine the optimal number of clusters for our typology. The elbow method determines the optimal number of clusters based on the within-group sum of squared distances for a range of values of k , and yields the optimal k where the magnitude of the second order change in within-group sum of squared distances is highest. The gap statistic compares the actual distribution of the data with a null distribution, calculating within-group variation of clusters relative to a uniform random distribution. The optimal k is the value where the gap statistic is maximized, implying the within-group similarity is not random.

Using these two methods, we tested for the optimal number of clusters on the CIUS data excluding non-users (that is, those who responded that they did not use the Internet in the past 3 months) for values of k between 2 and 10. The elbow method resulted in an optimal number of $k=3$ clusters, whereas the gap statistic returned $k=5$ as the optimal number. We therefore chose $k=4$ for the number of user clusters, as it lay between these two values. Adding the Non-user group back into the dataset resulted in a total of five cluster groups.

6.2.3 Variable selection criteria

Because the simple matching distance used in the k-modes algorithm only yields integer values, it is particularly sensitive to differences in any variable used in the algorithm. For this reason, variables with relatively low or high uptake across the full set of respondents will have a larger impact on final cluster results than they would with other cluster algorithms. Furthermore, many of these variables represent activities that are carried out conditional on other factors, including demographic factors.

We establish a criteria to eliminate high and low uptake variables from being included in the data used to generate the typology. Recall that the k-modes algorithm run on P variables produces a $k \times P$ matrix of modes representing the k modes corresponding to each cluster group. Because our specification is generated on binary data, we can assign positive responses a value of 1 and negative responses a value of 0, and compute averages across cluster specifications that represent the frequency that each entry in the mode matrix is a positive response (represented by a 1).

In order to select our final set of variables, we run the k-modes algorithm 100 times and record the mode matrix produced by each iteration. These can each be organized from low intensity user groups to high intensity user groups based on the number of positive responses (1) in each mode. Taking the average across all specifications gives us a matrix where each entry is the percentage of positive responses present for each variable in each mode across the full set of cluster specifications. Taking the average over the four modes yields a vector of length P that contains 'uptake scores' for each variable in the data which range between 0 and 1. A value of 0 means that no mode in any specification contained a positive response for that variable, whereas a value of 1 means that every mode in every specification contained a positive response for that variable.

Using this vector, we assign a threshold of 0.05 and eliminate all variables whose uptake score falls below that value. Similarly, variables whose uptake score exceeds $1 - 0.05 = 0.95$ are also removed and classified as 'ubiquitous' variables. These variables are not included in the data used to generate cluster groups, but are retained for analysis after the typology has been generated. The value 0.05 was chosen because it appeared to eliminate most redundant variables from the data without eliminating too many variables that had limited uptake among Advanced users.

6.2.4 Selecting the optimal specification

Having eliminated variables with high or low uptake, we are now in a position to generate the typology. Because the k-modes does not generate a unique cluster specification between iterations, we run the algorithm 2000 times and select the optimal specification as described below. Although the k-modes algorithm terminates when the mode is stationary between iterations, we add an additional check of stability for each iteration by feeding the resulting mode matrix back into the k-modes algorithm until the output mode matrix is equal to the input mode matrix. This ensures that each iterative step of the algorithm finds a set of modes that locally minimizes the sum of within differences across groups.

For each specification, we save information on the sum of within differences, a vector of length k containing the sum of distances between all points in each cluster. The sum of these values across all four clusters is the total sum of within differences for that cluster specification. This metric is low when within differences across cluster groups are low (that is, observations are similar within groups) and high when they are not, hence serves as a useful simple metric for evaluating the homogeneity of each of the cluster groups.

We selected cluster specifications that minimized the total sum of within differences. This resulted in a set of cluster specifications, many of which were permutations of one another, but many of which were not. From this group, we selected those specifications where the total *average* of within differences was lowest. That is to say, we divided each value in the vector of sum of within differences by the total number of observations in the group to produce the average within difference, and took the sum across clusters of these values.

This resulted in a set of cluster specifications that were identical to one another, except with regards to the arbitrary numeric labels assigned to them. Ordering the modes in the mode matrix from each specification from lowest to highest number of positive responses showed that the mode matrix for these specifications were also identical in this regard, as was the N-vector assigning each observation in the data to each group. As these specifications all produced the same cluster groups, the first specification from this set was selected and appropriately labelled to serve as our final typology on the CIUS data.

6.2.5 Summary of steps

What follows is a summary of the steps taken to generate the final user typology using the k-modes algorithm.

1. Determine the optimal number of clusters on Internet users using the elbow method and gap statistic.
2. Establish a selection threshold between 0 and 1. Eliminate variables where the average of modes across test specifications and cluster groups does not exceed the threshold, and classify variables that exceed one minus the threshold as ubiquitous and similarly eliminate them from the set of variables used to generate the typology.
3. Given the limited set of variables, run the k-modes algorithm a large number of times. For each run of the algorithm, do the following:

- a. Initialize the algorithm with randomly generated modes; the algorithm will return the final mode matrix.
 - b. Using the mode matrix output from Step a, re-run the algorithm, this time inputting the matrix of modes into the algorithm instead of a value k .
 - c. Repeat Step b until the k-modes algorithm outputs the same mode matrix as the matrix that was input.
 - d. Save the vector of length N assigning each observation in the data to clusters for further analysis, as well as the within-group sum of differences for selecting the optimal specification.
4. From the large number of k-mode results, select the cluster specification that minimizes the sum of total within differences across all clusters. If there are multiple specifications, select the one with the lowest *average* within differences across all clusters. (For each cluster, divide the total sum of within differences and divide it by the number of observations in the cluster, and take the average across all clusters).
 5. Verify that the results of Step 4 only produce cluster specifications that are permutations of one another. Select any of these permutations as the final cluster specification.
 6. Rank the clusters based on the number of positive responses in each mode. The lowest user group has the fewest number of positive responses in the mode, and the highest user group has the highest number of positive responses in the mode.

6.2.6 Robustness checks

Because the k-modes clustering algorithm used in generating the typology does not generate a unique typology with the same dataset, it is pertinent to check the other cluster groups generated by the process in the methodology section to see how the typology could differ and how the differences impact our main demographic results. The selected typology represents the best fit in the sense that it minimizes within-group differences between individuals, and although the typology was generated numerous times in the data, it is still not the only typology that resulted from this process.

Using the full set of 2000 k-modes results on the CIUS data, we can compare how our base typology differs from other possible typologies in a relatively large set of cluster specifications. In particular, we look at how our base typology compares to other typologies with regards to demographic composition between clusters, and how this impacts the results of the ordinal regression analysis. It should be noted that each cluster specification classifies Non-users the same by design, and so demographics of this group are identical between all specifications. Marginal effects in the ordinal regression may differ between specifications however, as the coefficients and the marginal effects for the Non-user group may be impacted by changes in other groups.

Appendix Table 3 presents the average marginal effects for each demographic category and cluster group, with averages taken across all 2000 cluster specifications. Indicators of statistical significance are also taken from the average p-values across all 2000 cluster specifications. As such, they do not represent the p-value of the estimate itself, but rather serve as an indicator for overall confidence between models.

Appendix Table 3
Average marginal effects across cluster specifications

	Non-Users	Basic Users	Intermediate Users	Proficient Users	Advanced Users
marginal effects					
Age group					
15 to 34 years	-0.016 ***	-0.052 ***	-0.035 ***	-0.013 **	0.116 ***
35 to 49 years
50 to 64 years	0.040 ***	0.090 ***	0.034 **	-0.018	-0.146 ***
65 years or older	0.131 ***	0.197 ***	0.023 *	-0.083 **	-0.267 ***
Educational attainment					
Less than high school	0.110 ***	0.084 ***	-0.017 *	-0.062 ***	-0.115 ***
High school diploma
Non-university postsecondary	-0.037 ***	-0.054 ***	-0.012 **	0.017 **	0.086 ***
Bachelor or higher	-0.070 ***	-0.125 ***	-0.044 ***	0.016 *	0.222 ***
Student	-0.067 ***	-0.117 ***	-0.039 **	0.018 *	0.206 ***
Household income					
\$24,999 or lower	0.065 ***	0.073 ***	0.015	-0.023 *	-0.130 ***
\$25,000 to \$49,999	0.015 **	0.021 **	0.007 *	-0.003	-0.039 **
\$50,000 to \$74,999
\$75,000 to \$99,999	-0.007	-0.010	-0.004	0.001	0.020
\$100,000 or more	-0.024 ***	-0.039 ***	-0.017 ***	-0.001 *	0.081 ***
Place of residence					
Urban
Rural	0.006	0.007	0.003	0.000	-0.015
Household size					
Multi-person household
Single-person household	0.013 **	0.016 **	0.006 **	-0.001	-0.033 ***
Immigration status					
Canadian-born
Immigrant	0.012 **	0.014 **	0.006 **	-0.001	-0.031 **
Sex					
Male
Female	0.005	0.007	0.003	0.000	-0.014
Employment status					
Employed
Not employed	0.016 ***	0.021 ***	0.008 **	-0.002	-0.044 ***

... not applicable

* significantly different from reference category ($p < 0.05$)

** significantly different from reference category ($p < 0.01$)

*** significantly different from reference category ($p < 0.001$)

Source: Statistics Canada, 2018 Canadian Internet Use Survey.

Comparing these results to Table 4, we can see that many of the marginal effects from the main result (using the selected user typology) are close to the average marginal effects across all of the specification generated in our analysis. Patterns observed in the age of respondents for the main result are also observed here, with younger individuals on average being more likely to be placed in the Advanced-user group and older individuals more likely to be placed in the more Basic-user or Non-user groups. Similarly, higher levels of educational attainment and higher annual incomes are correlated with placement in the Advanced-user group, and vice versa for the Basic-user or Non-user groups.

One notable difference in the main versus average marginal effects is with regards to urban and rural respondents. The average marginal effects across specifications are lower in magnitude and are not statistically significant for any cluster group. This is opposed to the main results, where coefficients are relatively small but statistically significant for all cluster groups with 95%

confidence. As these coefficients are small in either case: this result is somewhat marginal. Results for single- or multi-person households, immigrant status, gender and employment status yielded average marginal effects close to the main marginal effects presented in Table 4.

Appendix Table 4 builds on the results from Appendix Table 3 and presents the range (minimum/maximal) of marginal effects for each demographic and cluster group. We observe for age that the marginal effects for those ages 15 to 34 are always positive for the Advanced-user group (that is, individuals in this group were always more likely to be Advanced users relative to the baseline group), while the marginal effects for those aged 50 to 64 and 65 or older are always negative. We observe the opposite patterns for Non-users and Basic users, with those aged 15 to 34 always being less likely to be in either the Non-user or Basic-user group, and those aged 50 or older always being more likely. By contrast, the marginal effect is more balanced (not always positive) for Proficient users aged 15 to 34, as well as for Intermediate and Proficient users aged 50 to 64 and 65 or older.

Individuals with less than high school as their highest level of educational attainment were always found to be more likely to be Non-users or Basic users, and were always found to be less likely to be Proficient or Advanced users. Higher levels of education including current college and university students were always more likely than high school graduates to be Advanced users, and were always less likely to be Non-users or Basic users, with mixed effects for Intermediate and Proficient users at all levels of education.

Among other demographic indicators, rural respondents were always found to be less likely to be Advanced users, although the effect is marginal. Single-person households were always found to be less likely to be Advanced users than multi-person households, and more likely to be Non-users or Basic users. Immigrants were generally more likely to be Non-users, Basic users or Intermediate users, and were always found to be less likely to be Advanced users. Results were ambiguous for women relative to men, although the interval of marginal effects becomes increasingly negative going from the Non-user group to the Advanced-user group. Unemployed individuals were always found to be less likely to be Advanced users, and more likely to be Non-users or Basic users.

Appendix Table 4
Minimum and maximum marginal effects across cluster specifications

	Non-users		Basic users		Intermediate users		Proficient users		Advanced users	
	minimum marginal effects	maximum marginal effects	minimum marginal effects	maximum marginal effects	minimum marginal effects	maximum marginal effects	minimum marginal effects	maximum marginal effects	minimum marginal effects	maximum marginal effects
Age group										
15 to 34 years	-0.020	-0.011	-0.102	-0.019	-0.065	-0.009	-0.060	0.064	0.044	0.146
35 to 49 years
50 to 64 years	0.033	0.045	0.038	0.127	-0.016	0.077	-0.112	0.052	-0.187	-0.043
65 years or older	0.104	0.143	0.091	0.246	-0.100	0.118	-0.281	0.039	-0.360	-0.075
Educational attainment										
Less than high school	0.092	0.128	0.048	0.098	-0.089	0.036	-0.154	-0.008	-0.182	-0.029
High school diploma
Non-university postsecondary	-0.043	-0.032	-0.077	-0.027	-0.039	0.021	-0.018	0.079	0.028	0.115
Bachelor or higher	-0.079	-0.058	-0.190	-0.057	-0.107	0.014	-0.063	0.150	0.079	0.264
Student	-0.077	-0.049	-0.180	-0.053	-0.098	0.017	-0.054	0.144	0.063	0.258
Household income										
\$24,999 or lower	0.047	0.076	0.038	0.100	-0.023	0.048	-0.104	0.023	-0.168	-0.041
\$25,000 to \$49,999	0.008	0.021	0.008	0.035	-0.002	0.019	-0.027	0.010	-0.059	-0.013
\$50,000 to \$74,999
\$75,000 to \$99,999	-0.012	0.001	-0.022	0.001	-0.012	0.000	-0.007	0.016	-0.002	0.035
\$100,000 or more	-0.030	-0.016	-0.058	-0.018	-0.038	-0.001	-0.026	0.046	0.035	0.103
Place of residence										
Urban
Rural	0.000	0.010	0.001	0.015	0.000	0.008	-0.009	0.005	-0.028	-0.001
Household size										
Multi-person household
Single-person household	0.007	0.022	0.007	0.035	-0.001	0.017	-0.029	0.011	-0.053	-0.017
Immigration status										
Canadian-born
Immigrant	0.005	0.018	0.006	0.027	0.000	0.014	-0.023	0.008	-0.049	-0.009
Sex										
Male
Female	-0.002	0.012	-0.002	0.018	-0.001	0.008	-0.013	0.005	-0.034	0.005
Employment status										
Employed
Not employed	0.012	0.022	0.010	0.037	-0.001	0.022	-0.028	0.012	-0.063	-0.016

... not applicable

Source: Statistics Canada, 2018 Canadian Internet Use Survey.

References

- Aston, J., O. Vipond, K. Virgin, and O. Youssouf. 2020. "Retail E-Commerce and COVID-19: How Online Shopping Opened Doors While Many Were Closing." StatCan COVID-19: Data to Insights for a Better Canada. Statistics Canada Catalogue no. 45280001.
- Blank, G., and D. Grosej. 2014. "Dimensions of Internet Use: Amount, Variety, and Types." *Information, Communication, & Society* 17 (4): 417-435.
- Borg, K., and L. Smith. 2018. "Digital Inclusion and Online Behaviour: Five Typologies of Australian Internet Users." *Behaviour & Information Technology* 37 (4): 367-380.
- Brandtzæg, P.B., J. Heim, and A. Karahasanović. 2011. "Understanding the New Digital Divide – A Typology of Internet Users in Europe." *International Journal of Human-Computer Studies* 69 (3): 123-138.
- Büchi, M., N. Just, and M. Latzer. 2016. "Modeling the Second-Level Digital Divide: A Five-Country Study of Social Differences in Internet Use." *New Media & Society* 18 (11): 2703-2722.
- Chen, W. 2013. "The Implications of Social Capital for the Digital Divides in America." *The Information Society* 29 (1): 13-25.
- CRTC (Canadian Radio-television and Telecommunications Commission). 2020. *Communications Monitoring Report*. Catalogue no. BC9-9E-PDF. Ottawa: CRTC.
- Deng, Z., R. Morissette, and D. Messacar. 2020. *Running the Economy Remotely: Potential for Working from Home During and After COVID-19*. StatCan COVID-19: Data to Insights for a Better Canada. Statistics Canada Catalogue no. 45-28-0001.
- Ferguson, S.J., and J. Zhao. 2013. *Education in Canada: Attainment, Field of Study, and Location of Study*. Statistics Canada Catalogue no. 99-012-X2011001.
- Frenette, M., K. Frank, and Z. Deng. 2020. *School Closures and the Online Preparedness of Children during the COVID-19 Pandemic*. Economic Insights, no. 103. Statistics Canada Catalogue no. 11-626-X.
- Haight, M., A. Quan-Haase, and B.A. Corbett. 2014. "Revisiting the Digital Divide in Canada: The Impact of Demographic Factors on Access to the Internet, Level of Online Activity, and Social Networking Site Usage." *Information, Communication, & Society* 17 (4): 503-519.
- Hargittai, E. 2002. "Second-Level Digital Divide: Differences in People's Online Skills." *First Monday* 7 (4).
- Hargittai, E., and M. Micheli. 2019. "Internet Skills and Why They Matter." In *Society and the Internet: How Networks of Information and Communication are Changing our Lives*, Second Edition, ed. M. Graham and W. H. Dutton, p. 109-124. Oxford: Oxford University Press.
- Hatem, L., and D. Ker. 2021. "Measuring Well-Being in the Digital Age." *Going Digital Toolkit Note*, No. 6. Organisation for Economic Co-operation and Development.
- Helsper, E.J., and A. Galácz. 2009. "Understanding the Links Between Social and Digital Exclusion in Europe." In *World Wide Internet: Changing Societies, Economies, and Cultures*, ed. G. Cardoso, A. Cheong, and J. Cole, p. 146-178. Macau: University of Macau.
- Huang, Z. 1998. "Extensions to the k-Means Algorithm for Clustering Large Data Sets with Categorical Variables." *Data Mining and Knowledge Discovery* 2 (3): 283-304.

Korupp, S.E., and M. Szydluk. 2005. "Causes and Trends of the Digital Divide." *European Sociological Review* 21 (4): 409-422.

Lutz, C. 2019. "Digital Inequalities in the Age of Artificial Intelligence and Big Data." *Human Behavior and Emerging Technologies* 1 (2): 141-148.

Middleton, C., and C. Sorensen. 2005. "How Connected are Canadians? Inequities in Canadian Households' Internet Access." *Canadian Journal of Communication* 30 (4): 463-483.

Middleton, C., B. Veenhof, and J. Leith. 2010. *Intensity of Internet Use in Canada: Understanding Different Types of Users*. Business Special Surveys and Technology Statistics Division Working Papers, no. 2. Statistics Canada Catalogue no. 88F0006X.

Montagnier, P. 2007. *Broadband and ICT Access and Use by Households and Individuals. Working Party on the Information Economy*. Organization for Economic Co-operation and Development.

Montagnier, P., and A. Wirthmann. 2010. *Digital Divide: From Computer Access to Online Activities. Working Party on Indicators for the Information Society*. Organization for Economic Co-operation and Development.

Napoli, P.M., and J.A. Obar. 2014. "The Emerging Mobile Internet Underclass: A Critique of Mobile Internet Access." *The Information Society* 30 (5): 323-334.

OECD (Organisation for Economic Co-operation and Development). 2019. *How's Life in the Digital Age: Opportunities and Risks of the Digital Transformation for People's Well-being*. Paris: OECD Publishing.

OECD (Organisation for Economic Co-operation and Development). 2016. [New Skills for the Digital Economy: Measuring the demand and supply of ICT skills at work](https://doi.org/10.1787/5jlwnkm2fc9x-en). OECD Digital Economy Papers, no. 258. Paris: OECD Publishing. <https://doi.org/10.1787/5jlwnkm2fc9x-en>.

Reisdorf, B.C., and D. Groselj. 2017. "Internet (Non-) Use Types and Motivational Access: Implications for Digital Inequalities Research." *New Media & Society* 19 (8): 1157-1176.

Robinson, L., Ø. Winborg, and J. Schulz. 2018. "Interlocking Inequalities: Digital Stratification Meets Academic Stratification." *American Behavioral Scientist* 62 (9): 1251-1272.

Scheerder, A., A. van Deursen, and J. van Dijk. 2017. "Determinants of Internet Skills, Uses, and Outcomes: A Systematic Review of the Second- and Third-Level Divide." *Telematics and Informatics* 34 (8): 1607-1624.

Singh, V. 2004. *Factors Associated with Household Internet Use in Canada, 1998-2000*. Agriculture and Rural Working Paper Series. Working Paper no. 66. Statistics Canada, Agriculture Division. Statistics Canada Catalogue no. 21-601-MIE.

Spiezia, V., and P. Montagnier. 2010. "Measuring ICT Engagement and Dependency: A Statistical Framework." Working Party on Indicators for the Information Society. Organisation for Economic Co-operation and Development.

Statistics Canada. n.d. [Canadian Internet Use Survey 2018](https://www150.statcan.gc.ca/n1/pub/98-646-x2018001/article/00001-eng). Last updated November 23, 2020. Available at: <https://www150.statcan.gc.ca/n1/pub/98-646-x2018001/article/00001-eng> (accessed September 21, 2021).

Statistics Canada. 2020a. "Canadians Spend More Money and Time Online During Pandemic and Over Two-fifths Report a Cyber Incident." *The Daily*. October 14. Statistics Canada Catalogue no. 11-001-X.

Statistics Canada. 2020b. "Canadian Perspectives Survey Series 1: COVID-19 and Working from Home, 2020." *The Daily*. April 17. Statistics Canada Catalogue no. 11-001-X.

Statistics Canada. 2020c. [Table 33-10-0247-01 Percentage of workforce teleworking or working remotely, and percentage of workforce expected to continue teleworking or working remotely after the pandemic, by business characteristics](https://doi.org/10.25318/3310024701-eng). <https://doi.org/10.25318/3310024701-eng>.

Statistics Canada. 2021. [Table 22-10-0082-01 Internet use and intensity of use per week by gender, age group and highest certificate, diploma, or degree completed, inactive](https://doi.org/10.25318/2210008201-eng). <https://doi.org/10.25318/2210008201-eng>.

Turcotte, M. 2010. "Working at Home: An Update." *Canadian Social Trends* 91: 3-11. Statistics Canada Catalogue no. 11-008-X.

van Deursen, A.J.A.M., and J.A.G.M. van Dijk. 2014. "The Digital Divide Shifts to Differences in Usage." *New Media & Society* 16 (3): 507-526.

van Deursen, A.J.A.M., and J.A.G.M. van Dijk. 2015. "Toward a Multifaceted Model of Internet Access for Understanding Digital Divides: An Empirical Investigation." *The Information Society* 31 (5): 379-391.

van Deursen, A.J.A.M., and J.A.G.M. van Dijk. 2019. "The First-Level Digital Divide Shifts from Inequalities in Physical Access to Inequalities in Material Access." *New Media & Society* 21 (2): 345-375.

van Dijk, J.A.G.M., and A.J.A.M. van Deursen. 2014. *Digital Skills: Unlocking the Information Society*. New York: Palgrave Macmillan.

Zillien, N., and E. Hargittai. 2009. "Digital Distinction: Status-Specific Types of Internet Usage." *Social Science Quarterly* 90 (2): 274-291.