TALLINN UNIVERSITY OF TECHNOLOGY

School of Information Technologies

Nishaant Verma

IVCM165608

COMPARATIVE ANALYSIS OF ONLINE PRIVACY AND SECURITY CONCERNS BETWEEN GENERATION Y AND GENERATION Z IN NORTH INDIA: A PILOT STUDY

Master's Thesis

Supervisors: Dr. Hayretdin Bahsi

Prof. Mare Teichmann

Tallinn 2018

Author’s declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Nishaant Verma

21.04.2018

Abstract

The purpose of this thesis is to conduct a pilot study to compare the underlying online privacy and security issues that are of concern to Generation Y (Gen Y) and Generation Z (Gen Z) in North India. Sudden growth in public access to affordable internet brings inexperienced users to the market. In prior literature, research provides comparative analysis of Generation Y and Z from different countries, but none specifically to online privacy and security concerns of these generations. The methodology includes collecting data to measure the online privacy and security concerns of mid-high school students for Generation Z and non-IT working people for Generation Y living in North Indian states. The quantitative analysis of online survey featured 18 close ended questions shows that Gen Y are more concerned about online privacy compared to Gen Z. Whereas when it comes to online security, there is no difference in concern between the two generations, except where Gen Z are more concerned about 'having a strong password' compared to Gen Y. Correlation of gender, household income, education and device usage mapped to concerns show weak correlation. This thesis is written in English and is 81 pages long, including 7 chapters, 10 figures and 15 tables.

Annotatsioon

Võrdlev analüüs privaatsuse ja turvalisuse murede kohta internetis generatsiooni Y ja Z vahel Põhja-Indias: Pilootuuring

Uuringu eesmärgiks oli läbi viia pilootuuring, et võrrelda interneti privaatsuse ja turvalisusega seotud probleeme, mis on murettekitavad generatsioonile Y (Gen Y) ja generatsioonile Z (Gen Z) Põhja-Indias. Järsk kasv avalikus juurdepääsus taskukohase hinnaga võrguteenustele on toonud turule mittekogenud kasutajad. Eelnevad teadustööd on teinud võrdlevaid analüüse generatsioonide Y ja Z vahel eri riikides, kuid ükski ei ole uurinud täpsemalt nende generatsioonide interneti privaatsuse ja turvalisusega seotud muresid. Metodoloogia koosnes andmete kogumisest, et hinnata interneti privaatsuse ja turvalisusega seotud muresid keskkooliõpilaste seas generatsiooni Z uurimiseks ning mitte-IT alal töötavate inimeste seas generatsiooni Y uurimiseks, kes elavad Põhja-India osariikides. Kvantitatiivne analüüs veebikeskkonna küsitlusest hõlmas 18 suletud küsimust ning tulemused näitasid, et generatsioon Y on rohkem mures interneti privaatsuse pärast, kui generatsioon Z. Internetiturvalisuse teemadel muid erinevusi ei olnud peale selle, et generatsioon Z on rohkem mures küsimuse pärast “omad tugevat salasõna” võrreldes generatsiooniga Y. Sugu, leibkonna sissetulek, haridustase ning nutiseadmete kasutamine näitasid puuduvat kuni nõrka korrelatsiooni muretsemise tasemega. Lõputöö on kirjutatud inglise keeles ning sisaldab teksti 81 leheküljel, 7 peatükki, 10 joonist, 15 tabelit.

List of abbreviations and terms

BRICS	Brazil, Russia, India, China and South Africa
CBAM	Concerns-Based Adoption Model
CERT-In	Indian Computer Emergency Response Team
df	Degrees of Freedom
GSDP	Gross Domestic State Product
Gen Y	Generation Y
Gen Z	Generation Z
GDP PPP	Purchasing Power Parity
IM	Instant messaging
M	Mean
N	Sample size
NSA	National Security Agency
RQ	Research Question
SD	Standard Deviation
Sig.	Significance (p-value)
Std.	Standard
t	T-test statistic result
TTPs	Tactics, Techniques, and Procedures
TTU	Tallinn University of Technology

Table of contents

List of figures

List of tables

1 Introduction

Computing, the internet and mobile technology have enabled people to access the internet and share information easily at incredible speeds. As people share more about their lives and perform more everyday tasks online, questions arise about online privacy and security and how user data is stored and shared.

The International Organization for Standardization and International Electrotechnical Commission (IEC) defines cyber security as the "preservation of confidentiality, integrity and availability of information in the Cyberspace" [1]. A subset of cyber security is internet privacy, which can be termed as the "right or mandate of personal privacy concerning the storing, repurposing, provision to third parties, and displaying of information pertaining to oneself via of the Internet" [1].

Worldwide, over 4 billion people currently use the internet and India is ranked as the second largest internet user in the world, with a penetration rate of 47.4%[2]. With this extended usage and increased penetration there is a definite higher exposure to online security breaches and internet privacy issues in India.

This chapter provides an overview of this thesis. The first section gives a brief introduction to the definitions used in this thesis and examines the problem and objectives. This is followed by research questions, plus the importance and limitations of the study.

[1] ISO/IEC 27032:2012 Information technology —Security techniques — Guidelines for cybersecurity (https://www.iso.org/standard/44375.html)

[2] http://www.internetworldstats.com/stats.htm

1.1 Problem

In this thesis we conduct a pilot study on online privacy and security concerns of Gen Y and Gen Z in North India. The country has seen a recent dramatic growth in public access to affordable internet infrastructure. This is due to extensive economic reform leading to a growing purchasing power and since 2015 India is ranked third in the world in terms of Gross Domestic Product (GDP) [2].

India is a very vast and diverse country, consisting of 29 states and 7 union territories; divided into six main zones: North India, South India, East India, West India, Central India and North-East India [3]. North India was chosen as the focus of this pilot study due to the access to existing networks and with consideration of the limitation of time. It is hoped future studies can be conducted on other zones too.

The two generations considered, make up the future of India's growing economy. Gen Y (also known as Millennials) consists of those born between the years 1984 to 1996, prior to the internet age and therefore were first exposed to the internet in their teenage years [4]. Whereas Gen Z, born after 1997, were born in the internet age and have thus been exposed to the internet and technology since infancy [5]. The age ranges for Gen Y and Z are derived from a Harvard business review from a study of different generations across 19 countries [6].

A recent analysis of internet trends in India taken from KPCB [7] shows sudden exponential growth in monthly wireless data consumption between June to December 2016 as seen in Figure 1. Two major factors have caused this rapid growth. The first is the reduction in smartphone prices and the second the reduction in internet data prices. This was mainly caused by the Indian telecommunication company Reliance Jio that promoted free mobile data and heavily discounted mobile internet packages [8].

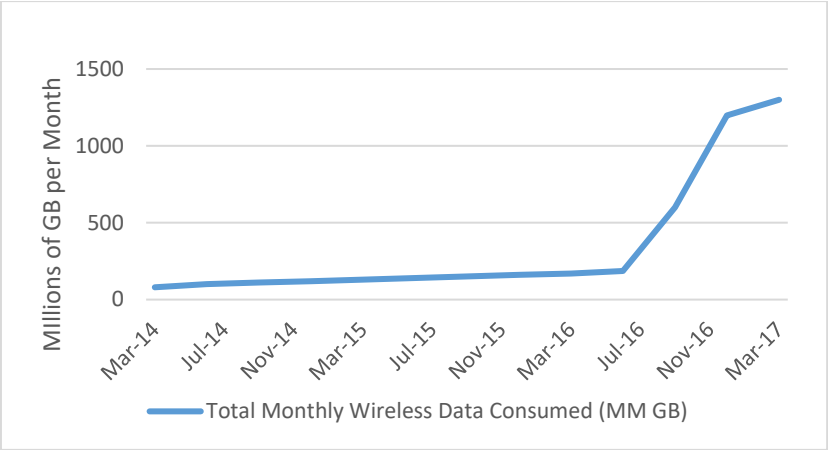


Figure 1: India's wireless internet data usage between March 2014 – March 2017. Total Monthly Wireless Data Consumed (MM GB) by KPCB [7].

This sudden growth brings forward a large amount of new inexperienced users to the market who may not have the awareness or knowledge about online privacy and security. This is where this pilot study will benefit in exposing the underlying online privacy and security issues that are of concern to Gen Y and Gen Z in North India.

1.2 Relevance

This pilot study can be used as a basis for further research for many interested parties. Not only will it indicate whether one generation is more concerned about online privacy and security than the other, but will also highlight their concerns based on different internet aspects.

The breadth of the scope allows this topic to be analysed only at an introductory level, thus classifying it a pilot study. The sample data presented is preliminary and does not give complete insights of the level of concern of Gen Y and Gen Z in North India.

The findings from this study can serve to initiate discussions for policy makers in India to identify areas of improvement and priority areas to focus on. It can also assist to better understand where the generations are lacking in concern and how best for policy makers to use the findings as a starting point to design policies to enable youth to use the internet safely.

It can also serve as a launch pad to start ideas on designing personalised user awareness tools to educate specific generations to protect their data and rights through technology and the internet. The results from this preliminary study could also provide a very basic introduction for the government, schools or universities to partly assist in design cyber security user awareness programs or even generation specific programs.

The findings can part of a larger, more inclusive sample size and the results be disseminated among the public to raise awareness on the problems of potentially alarming low levels of concern. This could enable parents to educate their Generation Z children on effectives steps to protect themselves while using the internet. It will also enable the users in Generation Y to be more conscious of their internet utilisation patterns with regards to online privacy and security.

Overall, there are numerous benefits of utilising the findings arising from this pilot study. These are mostly relevant to public-sector institutions such as governments, schools, universities and the users themselves. Not only will the findings serve to start discussion and raise awareness from a generational perspective in internet security, but will also provide a starting point to expand their knowledge to protect themselves in the future.

1.3 Objectives

The purpose of this thesis is to compare online privacy and security concerns of users from two generations: Generation Y and Generation Z to determine if there are any common concerns or differences. The objectives of this thesis can be further broken down into the following:

a. Analyse the concerns of Generations Y's online privacy and security concerns in comparison to Generations Z in North India; and

b. Determine commonalities or differences between the concerns of Generations Y and Generations Z in North India.

1.3.1 Research Questions

Through this introductory study, the aim is to analyse insights in North India on the following research questions:

1. Are Gen Y users more concerned about their online privacy and security than Gen Z users?

2. What are the differences in concerns between Gen Y and Gen Z users towards online privacy and security?

3. Is there a correlation between gender to the level of concern that Gen Y and Gen Z have for online privacy and security issues?

4. Is there a correlation between the number of hours a device is used to the level of concern that Gen Y and Gen Z have for online privacy and security issues?

5. Is there a correlation between household income and higher education level to the level of concern that Gen Y and Gen Z have for online privacy and security issues?

1.4 Limitations

There are a number of main limitations that could not be controlled. Listed below are the limitations and reasons why they are encountered in the study.

Time frame

Due to the limitation of time, an extensive survey could not be performed to sample a wider range of users from different demographic locations. This study was done in short time frame and it was not possible to collect data from the all the North Indian states due to its size and population. This study represents the views of two generations covering a total sample size of 150 - 200 participants.

Lack of resources

Quantitative research usually requires large sample sizes to have more accurate outcomes. Since there is only one dedicated full-time resource for this thesis, the scale of the study is relatively limited. However, it is hoped that the results will serve as a starting point to simulate discussion on online privacy and security concerns of Gen Y and Gen Z.

Limited outcomes

Quantitative research methods usually comprise of very structured and close ended questions. This restricts the outcome range and potentially may limit the result. Since the questions are designed quantitative statistical analysis and selection limited by the researcher, so the true opinions may be obscured.

Quality Sample Data

There is a possibility that Gen Z are not fully able to self-analyse their online privacy and security concerns since some may be too young. To avoid this potential ambiguity, the Gen Z age group analysed will only be 13 – 18-year olds. In essence, this age group represents middle and high school students.

1.5 Delimitations & Assumptions

The delimitations and assumptions listed below are boundaries set to ensure there is a clear scope that encapsulates the study. This is to ensure that the sample data obtained is targeted and the findings can be justified. This study is conducted based on the following premises:

a. Participants that complete the survey for each generation are from non-IT security background.

b. The survey is completed with genuine and true feedback from participants.

c. The participants of the survey have genuine interest in the research and have no other motive for completing the survey.

d. Extreme outliers will not be considered since they could negatively affect the statistical analysis and hence lead to misleading results.

1.6 Overview of the Thesis

This chapter provides an introduction to the thesis topic, definitions to key words, background to the problem to be studied, relevance, objectives, limitations and assumptions.

Chapter 2 covers additional information on India and online privacy and security. Chapter 3 contains the literature review, followed by Chapter 4, which discusses the methodology in detail.

The statistical analysis of the survey data is examined in Chapter 5, while Chapter 6 covers the evaluation of the findings and conclusion.

2 Background

In this chapter, an overarching background is provided that consists of the definition and difference between online privacy and security. A general overview of the cyber security landscape in India is provided, in addition to an overview of Indian Gen Y and Gen Z in India.

2.1 Defining Concerns

According to the Cambridge Dictionary[1] a 'concern' can be defined as, "a worried or nervous feeling about something, or something that makes you feel worried". In this context, the study examines how "worried or nervous" a user may feel towards specific online privacy and security situational question.

The study questionnaire used in this study has been specifically developed based on the Concerns-Based Adoption Model (CBAM), which applies to users who are experiencing change [9]. Table 1 lists the CBAM model as the seven stages[2] of concern for an innovation:

Table 1: CBAM Model - Typical expressions of concerns in innovation[2].

Stages of Concern	**Expression of Concern**
6. Refocusing	I have ideas that could work even better.
5. Collaboration	How can I relate what I am doing to others?
4. Consequence	How can I refine my activities to have more impact?
3. Management	How to manage my time on activities?
2. Personal	How will using it affect me?
1. Informational	I would like to know more about it.
0. Awareness	I am not concerned about it.

[1] https://dictionary.cambridge.org/dictionary/english/concern

[2] http://www.nas.edu/rise/backg4a.htm

This study utilises the first three stages, namely: 0: Awareness, 1: Informational and 2: Personal, as this accurately aligns with the definition of "worried or nervous" as per the Cambridge Dictionary.

2.2 Threat Actors

Alexander Klimburg defines *threat actors* as parties who conduct cyber-attacks with different behaviours and motives [10]. He also segregates threat actors into three divisions: State Actors, Organized Non-State Actors, and Non-Organized Non-State Actors.

Within this study, four main threat actors are considered: cybercriminals, the Indian Government, other governments and online companies. Each actor's behaviour, tactics or processes are termed as Tactics, Techniques, and Procedures (TTPs) and could include activities such as phishing, using malware or exploiting [11]. In this section we discuss the four applicable threat actors and their motives.

2.2.1 Cyber criminals

Least skilled among all the threat actors, cyber criminals are predominately interested in generating profit by selling sensitive company information or large amounts of personal data [12]. They are usually individuals or work as a team using technology and networks to defraud their victims [13].

Cyber criminals can utilise several different TTPs [14] to access private data as shown in Table 2 below:

Table 2: Types of TTPs of cyber criminals [14].

TTPs	**Description**
Phishing	Fake email can retrieve security information and personal details.
File hijacker	Files are hijacked and held ransom.
Webcam manager	Cyber criminals access and record webcams.
Keylogging	Cyber criminals access and record keyboard typing.
Screenshot manager	Cyber criminals access and take screenshots of your device screen.
Ad clicker	Unknown click on an ad, directs to a malicious link.

2.2.2 Online companies

Nothing in the world comes for free and this can be seen obviously from the cyber activities of online companies. Users who use free tools think that they may be obtaining free service or information, however there is a deep network of back end operations that works hard to track, save and even sell the publics' information.

There are number of ways online companies can use users' private data for profit. Firstly, online companies can track users' online activities like browsing history, search history, social network tracking, geo-location through cookies, supercookies, evercookies [15] and other smart applications to aggregate the data to deduce trends, buying habits and interests [16]. This is then sold to data brokers, who then sell onto other commercial parties [17].

Advertising is another way online companies generate profit from user data. Using the concept of 'behavioural profiling', online companies analyse large amounts of user data to enable specialised targeting of advertisements throughout the user's online journey [18]. The danger with dealing with large amounts of data and profiling is the misuse or leakage of the data. This not only causes distrust for online companies, but also could cause mental or physical harm to the users such as embarrassment, family discord or loss of employment [19].

Experts argue that this kind of profiling is advantageous to the user as they can obtain information that is of interest to them, however that may become a double-edged sword to influence users towards one path of thinking [19]. An obvious example of behavioural profiling is the Cambridge Analytica and Facebook data breach, reported in March 2018 [20]. The scandal highlights how systematic profiling may potentially have swayed an election outcome by specially profiling users and subconsciously influencing.

However, all these TTP methods are highly questionable since users usually authorise permission knowingly or unknowing to the online company to harvest data from their devices. It could be argued that permissions are written in the fine print and not usually read by the layman.

2.2.3 Indian Government

Only in the last 10 years has the Indian Government developed and accelerated their capabilities in intelligence and civilian monitoring through the internet. The Indian Telegraph Act 1985, Rule 419(A) and other related legislation as ruled by the Supreme Court allows for the government to intercept online activity such as voice calls, email, short text messaging and online chatting without going through service providers [21].

There are nine government agencies [22] that are authorised to intercept civilian communication and online activity; two spy agencies and seven ministries/military:

- Intelligence Bureau (IB)
- Research and Analysis Wing (RAW)
- Central Bureau of Investigation (CBI)
- Narcotics Control Bureau
- Directorate of Revenue Intelligence (DRI)
- National Intelligence Agency
- Central Board of Direct Taxes
- Military Intelligence of Assam
- JK and Home Ministry

The technology used by these organisations to intercept information is called the Central Monitoring System (CMS), which gives access to all communication data. All service providers in India were required to install servers called the Interception Store and Forward (ISF) that permits the government agencies to circumvent around needing to gain access from the service providers [22] [23].

Not only does the Indian Government monitor, read and listen in on data in real time, but they also have the capability to analyse data with artificial intelligence. Since 2014, voice traffic shared through Skype, social media and GoogleTalk and text communication such as tweets, forum messages or even Facebook status updates can all be scanned artificially for keywords. This system called the Network Traffic Analysis (NETRA) was developed by the defence ministry [23].

2.2.4 Other governments

There are many nations who have powerful spy networks and are using cyber espionage to access information of citizens of rival countries. According to the whistle-blower Edward Snowden, the American National Security Agency (NSA) gathers online and telephone data on India and its citizens through a data-mining system called Boundless Informant and a network intercepted program called PRISM [24]. India ranked the number one target among BRICS nations [24].

Reports state that China, who has been conducting cyber espionage since a decade ago, do not have priority status on India. China first monitors Taiwan and Hong Kong, followed by USA and Japan [25]. The report also mentions India on the other hand performs cyber espionage on Pakistan and then internal terrorist activity.

A recent incident in India, increased the awareness of cyber threats and the possibility that external nation states could be acting on behalf of the Chinese government [25] [26]. In December 2017, the Indian Government issued a directive to all Army personnel and paramilitary forces to delete certain apps developed in China or having any connection with the Chinese [27]. This is in addition to all Xiaomi branded phones have been banned from being sold as they have multiple apps with spyware capabilities [26].

2.3 Overview of Cybersecurity in India

In 2018, globally India is placed 4th and makes up 5.11% of the global cyber-crime based on Symantec's Internet Security Threat Report [28] . find the source of those attacks, and compiled this ranking of countries, sorted by number of threats that originated from them. In the first 6 months of 2017, sources mention that one cyber-attack was reported every 10 minutes [29] [30].

There were three main initiatives set up by the Indian Government to protect its public and infrastructure. The Information Technology Act, 2000 [30] deals with cyber-crime and electronic commerce and is applicable to citizens of India and even non-citizen, as long as the cybercrime involves a computer or network located in India.

To protect the public and private infrastructure of India, a National Cyber Security Policy was drafted in 2013. This policy framework set up by the Indian Government's

Department of Electronics and Information Technology also protects all user's financial, banking and personal information [31].

Another initiative set up by the Indian Government is the Indian Computer Emergency Response Team (CERT-In). This team deals with cyber-attacks such as hacking, malware attacks (including digital finance) and phishing[1]. The government reported that between 8 months, CERT-In managed 50,000 cyber security incidents [32].

However, laws and initiatives are only valuable if they are known to exist. A recent Indian study [33] conducted in 2015 of youth in India showed that 46.66% did not know that the Indian Information Technology Act even existed.

2.4 Cybersecurity Awareness in India

According to the EMC Privacy Index[2], out of 15 countries surveyed, Indians were the most willing to choose convenience over online privacy. This highlights an alarming issue with awareness and concern of the Indian society over the potential dangers of cyber security. It is therefore imperative to delve into practical and societal causes for the lack of concern in this field.

A 2012 analysis of cyber security education in Indian high schools uncovered that there was no exposure to any cyber security for students aged 15 years and below [34]. The analysis also observed that, cyber security that is actually taught in the old children are purely textbook based, without practical application. The analysis concludes with, "the curriculum should ensure that cyber education is instilled at an early age in the most applicable manner".

However, the Indian Government has started investing into educating their youth about cyber security. The Information Security Education and Awareness (ISEA) program[3] will

[1] www.cert-in.org.in

[2] https://www.emc.com/campaign/privacy-index/index.htm

[3] https://www.isea-pmu.in/home/I_Education

train teachers, industry professionals and faculty members on how to teach and spread cyber security awareness to their students.

To add to that, the Indian state of Kerala will launch a cyber security awareness program called Kids Glove [35] organised by the Police Department, a Cyber division called Cyberdome[1] and the Kerala State Council. The program will train teachers to execute a cyber security digital module through their curriculum.

It seems to be early days for India and it is imperative that the government and the public understand the dangers associated with not keeping up with cyber security issues.

2.5 Generation Y & Z in India

As of the 2011 national Indian Census[2], Generation Y and Z make up 66% of the total population. This large collective of individuals have grown up in an ever-changing globalised world and will have a profound effect on the Indian cultural and cyber security environment in the coming years.

Studies show that Generation Y, have embraced western values, however they are still a transitional generation where decisions are largely influenced by conservative family, national and societal values [36]. Their basic exposure to the advent of personal computers and the internet only started in the 1990s and has enabled them to become digital natives compared to their parents [37].

Comparatively, Generation Z does not know a life without technology. Almost 30 million Gen Zs personally own a mobile phone and 11 million share one with another family member [5]. According to a Bloomberg analysis, Generation Z would be willing to give away a small amount of personal information to make online decisions convenient, but are cautious about the level of information shared [38].

[1] http://www.cyberdome.kerala.gov.in/

[2] http://www.censusindia.gov.in/2011census/population_enumeration.html

2.6 North India Demographics

North India lays sandwiched between Pakistan, China, the Himalayas and the Thar dessert. It consists of the following states: Jammu and Kashmir, Himachal Pradesh, Punjab, Chandigarh, Haryana and Delhi. A map is provided in Appendix 5.

In North India, the percentage of Gen Y aged between 22 – 34 years as of the 2011 national Indian census is 23% of the total population; whereas Gen Z aged between 9 – 21 years stands at 26%[1]. Note in this study, data was not collected from Jammu and Kashmir due to the limitation of time.

By analysing the urban population, literacy rate and Gross Domestic State Product (GSDP) per capita of the North Indian states (Table 3), it is observable that all the five Indian states and union territories rank higher in GSDP per capita compared to the Indian national average. In addition to that, Delhi and Chandigarh, which are both union territories have a very high urban population and a high literacy rate. Haryana, Himachal Pradesh and Punjab have very low urban population since most of the land is used for agriculture [39].

Table 3: North Indian demographics and economic indicators.

State	GSDP per capita[1]	Total Population[2]	Urban Population[11]	Literacy[11]
Delhi	€3,700 (₹303,000)	16,700,000	97%	76%
Chandigarh	€3,000 (₹242,000)	1,060,000	97%	76%
Haryana	€2,300 (₹180,000)	25,300,000	35%	65%
Himachal Pradesh	€1,800 (₹147,000)	6,800,000	10%	73%
Punjab	€1,500 (₹114,000)	27,700,000	37%	67%
India	€1,400 (₹112,000)	1,210,000,000	31%	63%

[1] https://rbidocs.rbi.org.in/rdocs/Publications/PDFs/13SDP_240617EE2A8970184542E895DCE89D75A02259.PDF

[2] http://www.censusindia.gov.in/2011census/population_enumeration.html

3 Literature Review

Up to my knowledge, specific empirical analysis of online privacy and security concerns of two generations Y and Z have not been conducted before. However, extensive research has been made on Generation Y and Z, including other previous generations, on their social media levels [40] and online shopping behaviours [41].

Country specific comparisons and analysis between baby boomers, X, Y and Z generations have been made in marketing media communication [42], generational traits observed in schools [43] and in the workplace [44] [45] [46], among other topics.

In 2017, Sharma and Maidullah present an overview of internet users and online consumers in India of multiple generations ranging from Silver, X, Y and Z [41]. The study conducted of users from the ages of 17 – 66 and the results showed that Generation Z were the most active online shoppers compared to Gen Y, however the latter bought a larger variety of goods than Gen Z.

Other authors [47] [48] [49] also present their privacy findings based on online shopping of different generations or consumers in general. As Jukariya and Singhvi point out that 91% of Indian students surveyed agreed that when shopping on the internet, online privacy and security plays a major role in whether to proceed with the sale. All three literatures are based in India and analyse privacy issues, however they do not perform comparisons on two generations as done in this survey.

The same goes for a robust 2012 study conducted in India evaluated the responses of 10,427 Indian citizens and provided an umbrella summary of the perceptions of online security and privacy in India [50]. Kumaraguru and Sachdeva surveyed the Indian population as a whole, but did not place a generational perspective to their analysis. The authors discovered that approximately 40% of survey participants “would never save / share personal information in / through emails” and a majority of the participants had privacy concerns with social networks such as Facebook and Twitter, where they felt photographs to be “the most privacy invasive data” on social media.

Kamaraguru also conducted a study in 2005 comparing user's online privacy attitudes from India and the United States [51], which showed that Indians showed less concern with regards to online privacy than their American counterparts. A similar empirical study [52] was also conducted comparing concerns and user awareness between selected East African and the United States.

A similar study was conducted to investigate the Croatian publics' attitudes towards privacy, but focuses more on data protection and surveillance on the citizens by the government and private organisations [53]. Here, we consider literature from other parts of the world and discuss literature focusing on privacy and security attitudes towards government surveillance.

Another literature with a multinational perspective surveyed 1261 users from five global cities to understand their perceptions and behaviours regarding online privacy [54]. The study highlights demographics and national culture level factors, in addition to showing that female internet users were more concerned about online privacy than males.

A similar behavioural study was performed by Hoy and Milne, investigating privacy and personal sharing behaviours of college students and social network [55]. The researchers found that women were more proactive about privacy protection compared to men. An interesting statistic from another study also shows similarities to prior research where 83% of women compared to 77% of men valued privacy over convenience [56].

From an online security and education perspective, an American study showed that survey respondents with a diploma or degree did not value privacy over convenience, compared to respondents with a lower education level [56]. Comparatively to this, a study by Sheehan noted respondents with higher levels of education were more concerned about online privacy compared to respondents with lower education level [57].

Research published by Pew Research Center highlights the concerns of American citizens towards privacy perceptions and behaviours since the Edward Snowden leaked documents revealed surveillance by the government [58]. Europe on the other side of the coin, is leading the way in online data privacy reforms and this paper [59] analyses the relationship of privacy and security in European politics. Once again in all these literature, sample the population as one demographic. No comparisons were made between generations, like as performed in this study.

4 Methodology

In this chapter, a description of the research approach and rationale of the study are provided. This is then followed by a description of the population and sample characteristics, along with validity and reliability computations. Later in the chapter, the process of developing the instruments, data collection and analysis is discussion. The chapter is concluded with ethical considerations.

4.1 Research Approach and Rationale

The methodology used in this thesis utilises a quantitative, self-administered questionnaire, close-ended approach using a combination of descriptive statistics and independent variables. An online survey was used via social media to collect data to measure the online privacy and security concerns of the two generations.

Once the survey data were received from participants, the first step is to check each questionnaire to potentially eliminate incomplete questionnaires or ones where participants are not qualified [60] [61], e.g. they do not fall within the age range, are not from North India or have an IT security background.

After which, the data is then coded to allocate numeric codes to answers to allow for application towards statistical techniques [61]. This sample data is then transcribed and converted into Excel format to allow for further data analysis. Cleaning the data then follows [60] [61] by observing and identifying any outliers or errors that could skew the overall results. Finally, the analysis on the data is conducted using the data analysis strategy prescribed in this study and the results are represented in the forms of statistical charts, graphs and tables [61].

4.2 Population and Sample

The population consisted of Indian nationals living in North India, between the ages of 13 – 34 with non-IT security background. Overall there were 190 eligible participants in the survey, out of which 115 were from Generation Y and 75 were from Generation Z. The sample was randomly selected and contacted through the mobile app Whatsapp.

Due to the limitation of time, the sample was obtained from high schools, offices and universities through existing networks in North India. The contacted individuals who completed the survey also sent the survey to others who fell into the sample characteristics. This chain-referral sampling method is called 'exponential non-discriminative snowball sampling' [62].

This sampling method is a sub set of convenience sampling and hence can be classified as non-probabilistic [63]. However, since the sample size can be considered statistically large (>30), hence sample data can be deemed normally distributed and probabilistic in nature [63].

There are many advantages of using snowball sampling, among which is the ability to collect data in a cost and time effective way since the population can easily identify other participants through their peer networks. However, the biggest disadvantage is bias that occurs due to oversampling in one network of peers [62]. This study has aimed to avoid this by ensuring that mutually independent networks in each state were contacted.

4.3 Instrumentation

Through this approach, an online survey was used via social media to collect data to measure the online privacy and security concerns of the two generations. The participants are mid-high school students for Generation Z and non-IT working people for Generation Y. Surveys were targeted to sample population from North India for the two generations.

The benefits of using a web-based survey is the speed and breath at which the survey can be sent out and received [64]. Furthermore, the obtained data can be quickly transformed into analysis [65] [66]. Web-based surveys are also extremely cost effective [64] [66] [67]; with Google Forms used in this study being free and it also has capability to develop graphs and export raw data to excel.

However, there are also some challenges associated with web-based surveys, for example due to the anonymity of the surveys, it is hard to follow up [66] and discuss the answers face to face with participants. Noting all the pros and cons, as long as the survey is conducted diligently, a quantitative approach will save time and still provide credible outcomes.

The online survey is attached in Appendix 3 and focused on the user's online privacy and security concerns with relation to personal data privacy, device concerns, threat actors and communication channel privacy. The questionnaire was customised from surveys performed by the Pew Research Center[1] in 2013. This survey not only provided guidance on language and terminology, but it also provided a validation of the literature.

Two online surveys were developed, one for Gen Y and the other for Gen Z. This was because some questions in the demographics section asked about career for Gen Y and school study stream for Gen Z. Apart from the demographics section all other questions were the same for both generations. The survey for Gen Y can be found in Appendix 3 and for Gen Z in Appendix 4. The online survey was conducted in English and incorporated 18 closed-end questions.

4.4 Data Collection Procedure

The Gen Y and Gen Z surveys were sent to participants by Whatsapp messages. Existing networks in North India, consisting of teachers and employees from high schools, offices and universities. These individuals then sent out the survey to eligible candidates. It was found that students were better connected to their teachers by mobile technology through dedicated Whatsapp groups for each class year.

The participants were sent the Whatsapp message (Appendix 1) that listed the study name, the objectives and the requirements to participate in the study. Each message had a link to the Google Form questionnaire (Appendix 3 & Appendix 4) that can be completed on any device. The survey consisted of 18 close-ended questions and took approximately 5 minutes to complete.

[1] http://www.pewinternet.org/2013/09/05/anonymity-privacy-and-security-online/

4.5 Data Analysis Plan

To obtain the results of the survey, data analysis is performed to systematically analyse the quantitative data and extract valuable conclusions. The survey consists of 18 questions comprising of different types of variables. All variables in this survey can be classified as categorical variables, except Question 6 is classified as a continuous-interval variable. The categorical variables have the following three sub categories: binary, nominal and ordinal.

Choosing the type of analysis depends on how the research and survey was designed. There are two parts to this analysis: descriptive and inferential. The first section, descriptive analysis includes explains the data in descriptive form using minimum, maximum, frequency and measures of central tendency. The second section, inferential analysis employs statistical tests to evaluate the pattern that forms from the data.

Datasets obtained through Google Forms were directly exported to Microsoft Excel where the data was cleaned and outliers removed [52] [53]. Statistical analysis was performed on Microsoft Excel with the Data Analysis add-in tool.

There are many different statistical tests and the right one is chosen based on the design of the research, types of variables and distribution of the data. Considering all the parameters, four different statistical tests are performed to answer the research questions presented in this study: Levene's Test, Student's t-test, Cohen's *d* and Pearson Product-Moment Correlation.

Hypothesis and Significance Level

In statistical analysis, the null hypothesis, H_0 is a theory that has not been proved, but is believed to be true. Whereas, the alternative hypothesis, H_1 is proven by the outcome of the statistical test [68]. If the difference in mean between the two samples are substantial then the null hypothesis, H_0 is rejected [69].

The significance level α, is a pre-chosen probability that compares the calculated significance value (*p*-value) of the hypothesis test to a statistically significant value [68]. Typical α values are 0.1, 0.05, and 0.01; in this study $\alpha = 0.05$.

Therefore, if the *p*-value of the sample statistic is less than or equal to 0.05 then the decision is to reject the null hypothesis, else we fail to reject the null hypothesis [70]. Choosing a significance level is purely arbitrary and in the case of $\alpha = 0.05$, signifies a 95% level of confidence in the result[1].

Levene's Test

Before an independent-samples *t*-test can be performed, it is necessary to test the '*assumption of homogeneity of variance'* using Levene's Test [71]. The *t*-test statistical analysis method is selected based of the resultant of the Levene's Test.

The significance level, $\alpha = 0.05$ is used as a measure to classify whether the samples have violated the assumption of homogeneity of variances [72]. The theory shows that, if the significance value (*p*-value) of the Levene's Test is greater than $\alpha = 0.05$, then variances are considered equal. If the p-value is lesser than $\alpha = 0.05$, then the assumption of homogeneity has been violated with unequal variances [73].

Student's t-test

The Student's *t*-test is *"an inferential statistical test that determines whether there is a statistically significant difference between the means in two unrelated groups* [73] ".

To accept or reject the null hypothesis H_0, the calculated P-value is compared to the chosen $\alpha = 0.05$ [59]. Since this is a two-tailed test with $\alpha = 0.05$, therefore the critical values of the test statistic are: $z_{0.025} = \pm 1.96$. Therefore, the null hypothesis will be rejected if the resultant test statistic is less than -1.96 or greater than 1.96 [63].

This test is used for the following research questions:

RQ1: Are Gen Y users more concerned about their online privacy and security than Gen Z users?

RQ2: What are the differences in concerns between Gen Y and Gen Z users towards online privacy and security?

[1] https://www.statsdirect.com/help/basics/p_values.htm

Cohen's d

This statistical test completes the *t*-test by computing the Cohen's *d* measure. It is an unitless effect size and shows the strength of difference between two sample means [74]. In other words, this represents the distance between the mean of the observations compared to the mean of the null hypothesis. The Cohen's *d* can be calculated using Equation (1) [75]:

$$d = \left|\frac{\bar{x}-\mu}{\sigma}\right| \quad (1)$$

x = sample mean

μ = null hypothesis population mean

σ = null hypothesis population standard deviation

According to Cohen's definition of measurements resultants of small effect has a value between 0 to 0.20, medium effect for values between 0.20 to 0.50 and finally the large effect has values larger than 0.50 [76].

Pearson Product-Moment Correlation

This test is used to "measure of the strength of a linear association between two variables [77]." This strength of the relationship between two variables is represented through *correlation coefficient, r*. To interpret the resultant *r value*, if the value is near zero, there is no correlation. If the value is between 0 to ±0.25 signifies weak correlation and a value between ±0.75 to ±1 signifies a strong negative or positive strength in the relationship [69].

Another useful descriptor is the significance of the relationship that is calculated through the *p-value* and compared to the chosen $\alpha = 0.05$. The correlation is statistically significant if the *p-value* is less than α [68].

This test is used for the following research questions:

> **RQ3:** Is there a correlation between gender to the level of concern that Gen Y and Gen Z have for online privacy and security issues?

RQ4: Is there a correlation between the number of hours a device is used to the level of concern that Gen Y and Gen Z have for online privacy and security issues?

RQ5: Is there a correlation between household income and higher education level to the level of concern that Gen Y and Gen Z have for online privacy and security?

Data coding

To enable statistical analysis of the data, ordinal data collected from the survey is coded to a numerical form as seen in Table 4. Each level on the scale of concerns is assigned a number or code, starting from 1, with and equal increment to 5.

Table 4: Coded ordinal data.

Question 2	**Question 4**	**Question 6**	**Questions 9 -16 & 18**
1 Female 2 Male	1 Middle & High school 2 Diploma & Bachelor's degree 3 Master's degree	1 Below 2,000 2 2,000 - 8,000 3 8,000 -16,000 4 16,000 - 21,000 5 21,000+	0 Not applicable 1 I didn't know this could happen 2 I should be concerned, but I'm not 3 Not concerned 4 Slightly concerned 5 Very concerned

4.6 Ethical Considerations

From the participants side, it was obligatory to confirm consent before the start of the survey. Participants were informed that partaking in the study is purely voluntary and that questions can be skipped at any time if they do not feel comfortable answering. They were also informed that the study is completely anonymous, that no personal information is stored and that participants cannot be identified from results of this study in anyway.

Even though the survey is online based, no geographical locations or IP addresses were retrieved. No identification or personal information was requested from the participants.

5 Results & Analysis

This chapter presents the findings and statistical analysis from sample data collected from the online survey distributed to a Gen Y and Gen Z sample. The analysis first provides an overview of the demographics of the sample, followed by t-tests conducted to statistically answer research questions 1 and 2. This section also provides statistical results for research questions 3 to 5 through Pearson's Correlation Test.

5.1 Demographics

The population consisted of all Indian nationals living in North India, between the ages of 13 – 34 with non-IT security background. Overall the sample includes 190 eligible participants in the survey, out of which 115 were from Generation Y and 75 were from Generation Z. Table 5 represents the gender, education and household income demographic data

Overall, 52 participants were female with 138 males. The ratio of females to males were marginally similar at 27% female to 73% males for Gen Y and 28% females to 72% males for Gen Z. The age range was normally distributed with 75% of the sample being between 16 to 25 years old (M = 3.22, SD = 1.36).

Table 5: Socio-demographic characteristics of Gen Y and Gen Z participants of the study, N = 190.

		Gen Y		**Gen Z**		**Total**	
		n	***%***	***n***	***%***	***n***	***%***
Gender	Female	31	27%	21	28%	52	27%
	Male	84	73%	54	72%	138	73%
Education	Middle & High school	21	18%	75	100%	96	51%
	Diploma & Bachelor's degree	77	67%			77	41%
	Master's degree	17	15%			17	8%
Household Income	below 2,000	71	62%	24	32%	95	50%
	2,000 - 8,000	23	20%	37	49%	60	32%
	8,000 -16,000	12	10%	12	16%	24	13%
	16,000 - 21,000	6	5%	2	3%	8	4%
	21,000+	3	3%			3	1%

Education level of the participants were varied from primary school to Masters level. The highest level of education for the Gen Z participants is primary & high school. On the other hand, only 18% of the Gen Y sample had a high school degree, 67% had a diploma or bachelor's degree and 15% held a master's degree.

5.2 Online Privacy and Security Concerns

To answer the first research question, *'Are Gen Y users more concerned about their online privacy and security than Gen Z users?*', Student's t-tests were performed on questions related to online privacy and security separately. From the survey, sample data from Question 9 is representative of online privacy concern and sample data from Question 11 is representative of online security concern of Gen Y and Gen Z.

Online Privacy Concerns

Question 9 of the survey (Do you care at all that your online activities are being tracked, watched or saved?) was used to test the equality of the means. To statistically prove that Gen Y users are more concerned about their online privacy than Gen Z users, we define the null H_0 and alternative hypothesis, H_1:

H_0: There is no difference between Gen Y and Gen Z online privacy concern level.
H_1: There is a difference between Gen Y and Gen Z online privacy concern level.
Level of significance $\alpha = 0.05$

Table 6 shows the descriptive analysis of Gen Y sample (N = 115) and shows that this sample has higher concern for online privacy (M = 3.83, SD = 1.04) compared to Gen Z (N = 75) with a lower level of concern (M = 3.33, SD = 1.45).

To test the hypothesis, an independent samples t-test was conducted. As shown in Table 7, Levene's test for equality of variances, p=0.00004 validates unequal variances, since the computed p-value is less than $\alpha = 0.05$. This denotes that the homogeneity of variances was not satisfied and thus the independent samples *t*-test was performed by assuming unequal variances. The results of the t-test indicated a rejection of the null hypothesis, H_0 since $t(123.47) = 2.60$, $p = 0.01$. Thus, the level of online privacy concern levels of Gen Y is proven to be statistically significantly larger than that of Gen Z. Cohen's d was calculated to be 0.41, which shows medium effect as listed in Cohen's guidelines [76].

Table 6: Descriptive analysis conducted on survey Question 9 relating to online privacy concern levels.

Online Privacy Concern	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**
Gen Y	115	3.83	1.04	0.10
Gen Z	75	3.33	1.45	0.17

Table 7: Results of independent two-sample t-test assuming unequal variances conducted on survey Question 9 relating to online privacy concern levels.

Q9. Do you care at all that your online activities are being tracked, watched or saved?'.	**Levene's Test**	**t-test**			**Cohen's *d***
	Sig.	t	df	Sig.	
	0.00004	2.60	123.47	0.01	0.41

The data presented in Figure 2 shows the results of Question 9 from the survey: 'Do you care at all that your online activities are being tracked, watched or saved?'. Both Gen Y (28%) and Gen Z (29%) were equally 'very concerned', however 41% of Gen Y were 'slightly concerned' compared to only 23% of Gen Z.

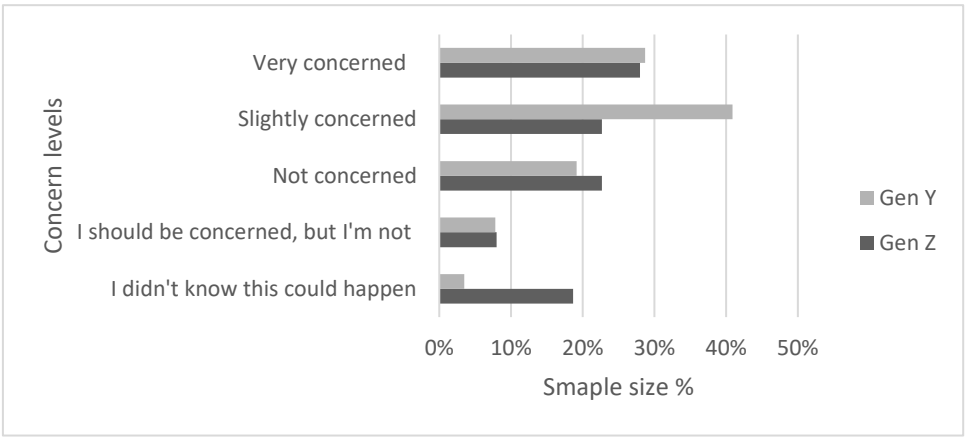


Figure 2: Responses to Question 9: Do you care at all that your online activities are being tracked, watched or saved?

An important observation to note is the almost 20% or a fifth of Generation Z did not know that their online activities could be tracked, monitored or saved. Out of this fifth, 80% was male, this relates with data in Section 5.4 below showing that females are more concerned about their online privacy compared to males.

Online Security Concerns

To answer the second element of the research question, survey Question 11 (How concerned are you about [online security issue]?) was used to test the equality of the means. Independent two-sample t-tests were conducted on different online security concerns and the sample results are as seen in Table 8.

To statistically prove that Gen Y users are more concerned about their online security than Gen Z users, we define the null H_0 and alternative hypothesis, H_1:

H_0: There is no difference between Gen Y and Gen Z online security concern level.
H_1: There is a difference between Gen Y and Gen Z online security concern level.
Level of significance $\alpha = 0.05$

To establish the equality of the means, Levene's test is first conducted and showed no violation of the homogeneity of variances. Hence, a t-test for equal variances was performed on all concerns in Table 8 and sample results show that all online security concerns fail to reject the null hypothesis H_0. This indicates no statistically significant difference in equality of means, coupled with very low Cohen's *d* values suggests a weak practical significance.

The only exception is the online security concern 'having a strong password'. This t-test sample results indicates a rejection of the null hypothesis H_0 since $t(188) = 2.44$, $p = 0.02$. Therefore, based on the sample data, it can be deduced that Gen Y's level of concern about strong passwords is statistically different than Gen Z's. In addition to that, Cohen's *d* was calculated to be 0.36 signifying medium effect.

Table 8: Independent Gen Y and Gen Z two-sample t-test conducted on four survey questions relating to security.

11. How concerned are you about...?	**Levene's Test**	**t-test**			**Cohen's *d***
	Sig.	t	df	Sig.	
malware infecting your phone	0.12	0.09	188	0.93	0.01
online banking	0.48	0.40	188	0.69	0.06
accessing an open Wi-Fi network	0.32	0.66	188	0.51	0.10
having a strong password	0.11	2.44	188	0.02	0.36

5.3 Differences in Concerns

To answer the second research question, '*What are the differences in concerns between Gen Y and Gen Z users towards online privacy and security?*', different aspects such as data privacy, device security and the dangers of threat actors have been considered to evaluate differences in online security and privacy.

Participants were asked which two privacy issues were they most worried about, the responses are illustrated in Figure 3. Nearly half of Gen Z were most concerned about phishing emails. Almost 30% of Gen Y R equally concerned about phishing emails and location tagging themselves on social media. It is interesting to note that only 10% of Gen Y and Gen Z feel concerned about clicking unknown links.

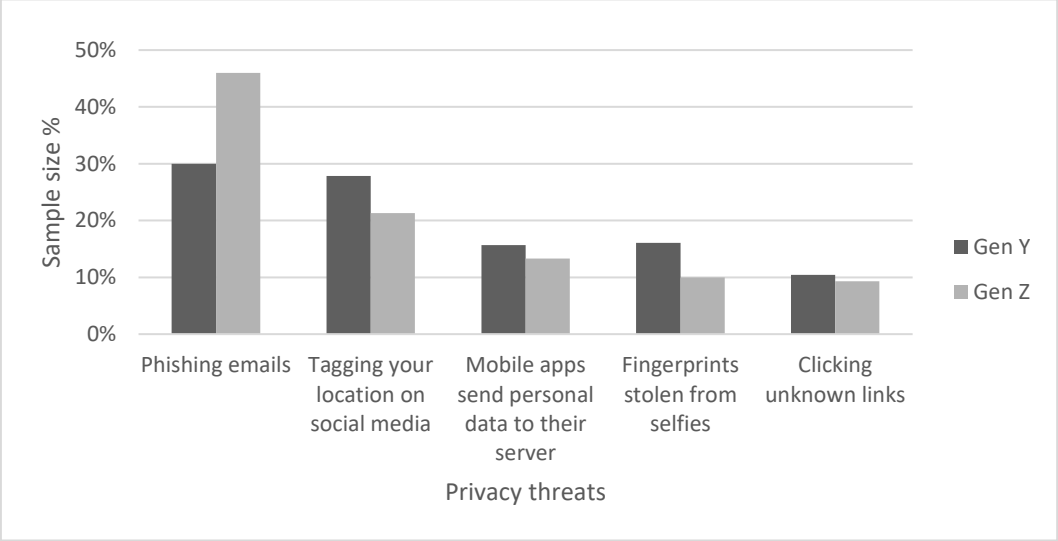


Figure 3: Responses to Question 17: Which two privacy issues are you most concerned about?

To analyse the participants concerns toward device and online security, Question 11 queried how concerned are the participant about keeping their devices and online accounts protected. The results are presented in Table 9.

Table 9: Response to Question 11: Mean level of concern for online security.

11: How concerned are you about…?	Mean level of concern		
	Gen Y	Gen Z	% change
keeping your anti-virus software up-to-date	4.18	4.16	2%
keeping your operating systems up-to-date	4.21	4.17	4%
malware infecting your phone	4.05	4.04	1%
online banking	4.06	4.00	6%
accessing an open Wi-Fi network	3.89	3.97	-9%
buying things online	3.97	4.05	-9%
having a strong password	3.96	4.29	-34%

The mean level of concern for all the variables show a mean range between 3.89 – 4.29 and a difference in mean of 0.40. Therefore, it can be deduced that both Gen Y and Gen Z feel very similar with an average 'slightly concerned' sentiment. The most important concern for both generations was keeping their anti-virus and operating systems updated and the least important was having access to an open Wi-Fi network. A point to note is the 34% difference in concern on having a strong password between Gen Y and Gen Z.

To evaluate how Gen Y and Gen Z perceives data privacy, Question 10 from the survey, extracts the concern levels of the participant if their personal information can be found freely on the internet. The sample results of the mean level of concerns for Gen Y and Gen Z are presented in Figure 6.

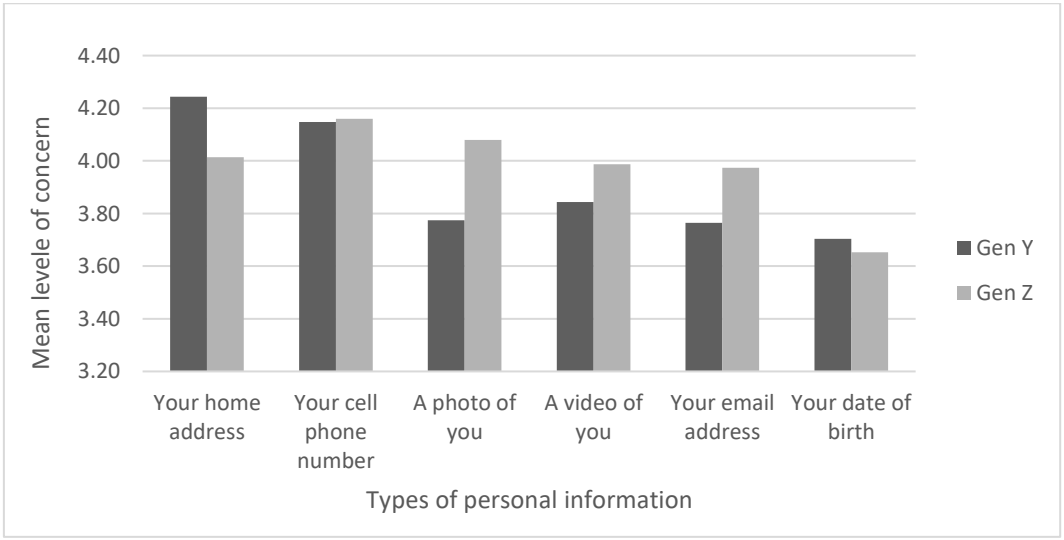


Figure 4: Responses to Question 10: How concerned are you that the following personal information can be found freely on the internet?

Overall, the Gen Z participants were more concerned about their cell phone number (M = 4.16, SD = 0.97) than their home address (M = 4.01, SD = 1.29) found freely online. This is compared to Gen Y participants were most concerned about their home address (M = 4.24, SD = 0.97) being found freely. Data also shows that Gen Ys were not very concerned about their photo, video or email address being freely found online compared to Gen Z.

Participants were also queried about their concerns on four specific threat actors, namely cyber criminals, online companies, the Indian Government and other governments accessing their private online data. Question 12 – 14 and 16 of the survey collates user concerns on the following private data types:

- email contents
- online chat content
- website browsing history
- downloaded files
- GPS location
- webcam and microphone

The mean results were aggregated and represents all data types as seen in Figure 6 below. Both Gen Y (M = 4.12, SD = 0.89) and Gen Z (M = 4.12, SD = 0.93) are equally concerned about cyber criminals accessing their private data. The data also reveals that Gen Y has very consistent level of concerns over all the three remaining threat actors with an average difference in mean of 0.01.

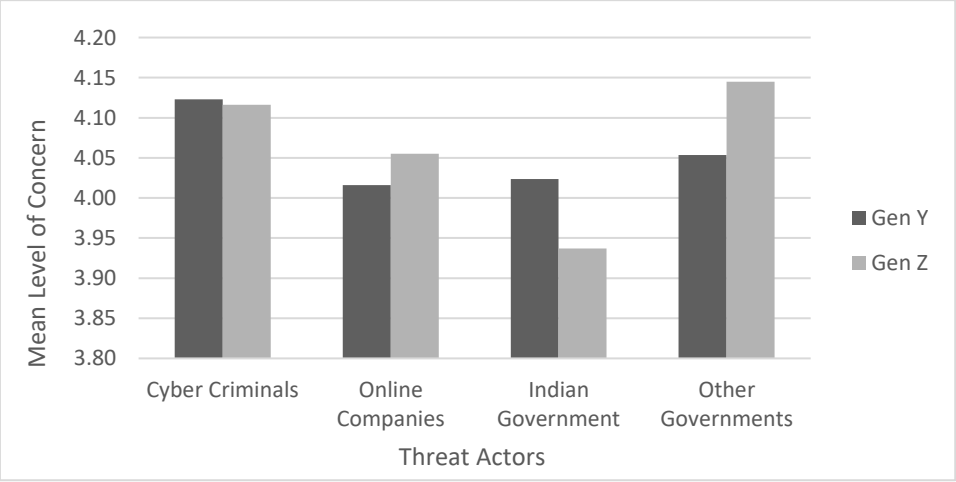


Figure 5: Responses to Questions 12-14,16: How concerned are you that the following threat actors are accessing your data?

Inversely, Gen Z has large differences in mean between level of concerns between threat actors. Gen Z has a very high level of concern for other governments accessing their private data (M = 4.14, SD = 0.91), but very low level for their own Indian Government (M = 4.02, 0.90).

The final question requests survey participants to select which mode of communication did they feel most concerned when sharing private information. Sample results show that the landline was deemed of least concern for both Gen Y (M = 3.92, SD = 0.79) and Gen Z (M = 3.84, SD = 0.85). The communication mode that was made both Gen Y (M = 4.19, SD = 0.78) and Gen Z (M = 4.25, SD = 0.86) participants feel most concerned was using chat or instant messaging (IM) platforms such as Whatsapp or Facebook Messenger.

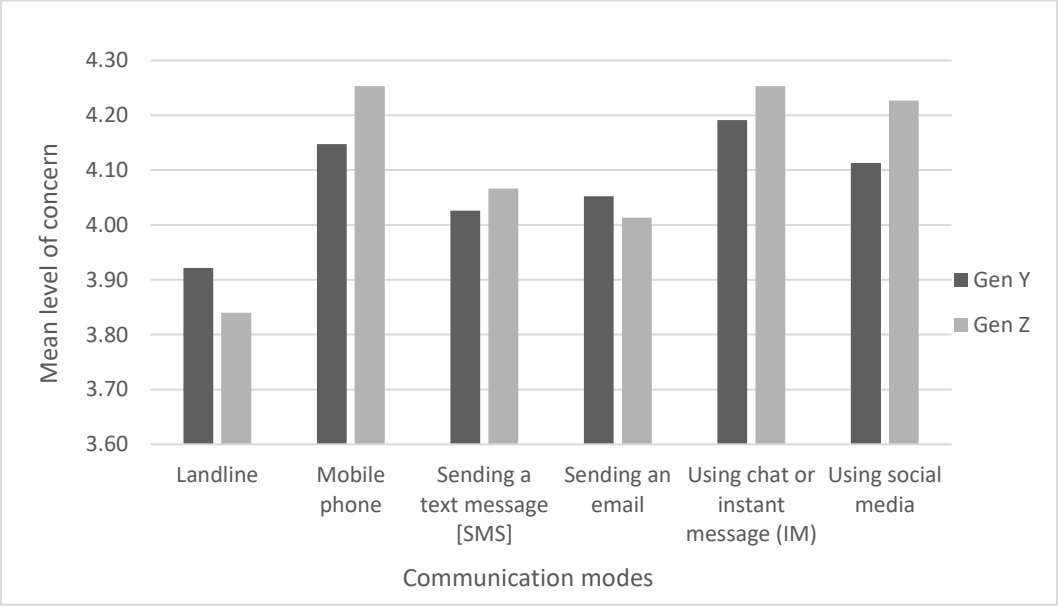


Figure 6: Responses to Question 18: Which mode of communication do you feel most concern when sharing private information?

Communication through social media platforms or a mobile phone was also a concern for both the generations, however for Gen Z had higher mean levels compared to Gen Y for all modes except the landline and email.

5.4 Gender and Concern

To answer the third research question, '*Is there a correlation between gender to the level of concern that Gen Y and Gen Z have for online privacy and security issues?*', and to understand further the correlation between gender to the level of concern that Gen Y and Gen Z have for online privacy and security issues, Pearson's Product-Moment Correlation is employed.

There were in total 31 (27%) females and 84 (73%) males in Gen Y and 21 (28%) females and 54 (72%) males in Gen Z. This shows that the ratio of female to male is almost equal comparatively between Gen Y and Gen Z.

For this exercise to understand the correlation between gender and concern, three questions are selected as seen in Table 10 below to represent the overarching online privacy and security concerns.

For Question 9, Pearson's coefficient shows that gender and the concern level variables for Gen Z have very weak correlation, in addition to a low *p*-value signifying negligible statistical significance, $r(75) = 0.12$, $p = 0.29$.

Table 10: Pearson's correlation results conducted on three survey questions relating to gender and online privacy and security levels of concern.

		Gen Y	**Gen Z**
Q9. Do you care at all that your online activities are being tracked, watched or saved?	Pearson Correlation	0.23	0.12
	Sig. (2-tailed)	0.01	0.29
	N	115.00	75.00
Q10. How concerned are you that the following personal information can be found freely on the internet? [Your home address]	Pearson Correlation	0.24	0.27
	Sig. (2-tailed)	0.01	0.02
	N	115.00	75.00
Q11. How concerned are you about malware infecting your phone?	Pearson Correlation	0.26	0.15
	Sig. (2-tailed)	0.004	0.19
	N	115.00	75.00

Comparatively for the same question, Gen Y receives a stronger Pearson's coefficient of $r(113) = .23$, $p = 0.01$. This coefficient value still shows a weak correlation, however due to a p-value lower than 0.05, this makes the correlation statistically significant. Similar

results can also be seen in Gen Y and Gen Z in question 10. In essence, there is positive, but weak correlation between gender in Gen Y with level of concern.

Using Question 9 as a benchmark, the level of online privacy and security concern is plotted against gender in Figure 7. The graph clearly illustrates that females in both Gen Y and Gen Z have higher levels of concern compared to their male counterparts.

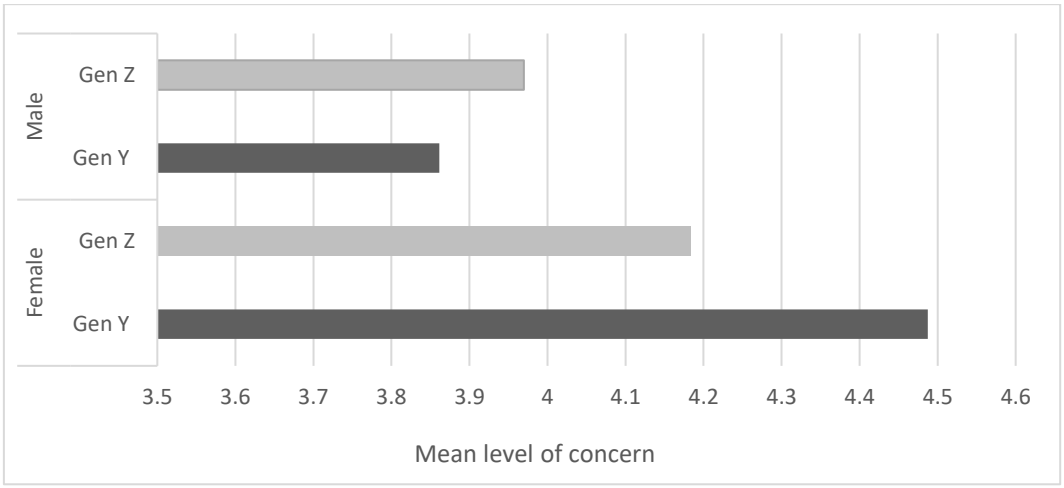


Figure 7: The level of concern for online privacy and security plotted against gender.

5.5 Usage and Concern

To answer the fourth research question, '*Is there a correlation between the number of hours a device is used to the level of concern that Gen Y and Gen Z have for online privacy and security issues*?', a Pearson's Product-Moment Correlation is used to understand the correlation between device usage (hours) and concern.

Firstly, we examine the variables surveyed as seen in Figure 8. Smartphones were most widely used by both Gen Y (M = 3.12, SD = 2.25) and Gen Z (M = 3.32, SD = 2.27) with average usage of 3 hours a day. This is followed by laptops with half the daily usage time. Note that 14 entries were omitted due to more than 90% incomplete answers for this question.

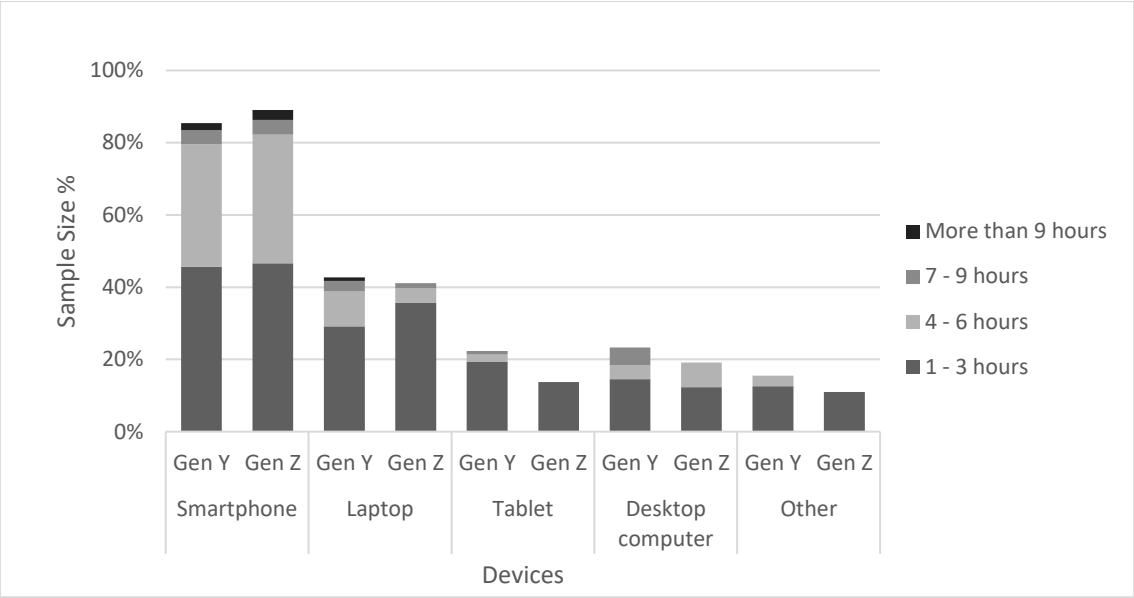


Figure 8: Responses to Question 7: How many hours a day do you use each internet based device?

To test the correlation of hours of device usage to level of concern, Pearson's Product-Moment Correlation was employed. The Pearson's correlation as shown in Table 11 for Gen Y, $r(101) = 0.19$, $p = 0.06$ shows a weak correlation and coupled with a p-value higher than 0.05, making the correlation statistically not significant. These resultants are the same for Gen Z's correlation results.

Table 11: Pearson's correlation results conducted on survey Question 7 relating to device usage (hours) and online privacy and security levels of concern.

		Gen Y	**Gen Z**
Q7. How many hours a day do you use each internet based device?	Pearson Correlation	0.19	0.03
	Sig. (2-tailed)	0.06	0.77
	N	103	73

5.6 Income & Education and Concern

The final research question, *'Is there a correlation between household income and higher education level to the level of concern that Gen Y and Gen Z have for online privacy and security issues?',* revolves around testing whether there is a correlation between household income and higher education level to the level of concern that Gen Y and Gen Z have for online privacy and security issues.

Annual Household Income: Participants were asked to provide their annual household income, which includes all family members living under the same roof. For Gen Z participants this includes their parent's income.

There are differences between household income levels of Gen Y (M = 1.67, SD = 1.03) and Gen Z (M = 1.89, SD = 0.76) as seen in Table 12. More than 60 % of the Gen Y sample have the lowest income, earning below €2000 annually. Figure 9 also shows, that half of the Gen Z participants household income fell into the €2000 - €8000 income bracket.

Table 12: Descriptive analysis conducted on survey Question 6 relating to annual household income.

Income	N	Mean	Std. Deviation	Std. Error Mean
Gen Y	115	1.67	1.03	0.10
Gen Z	75.00	1.89	0.76	0.09

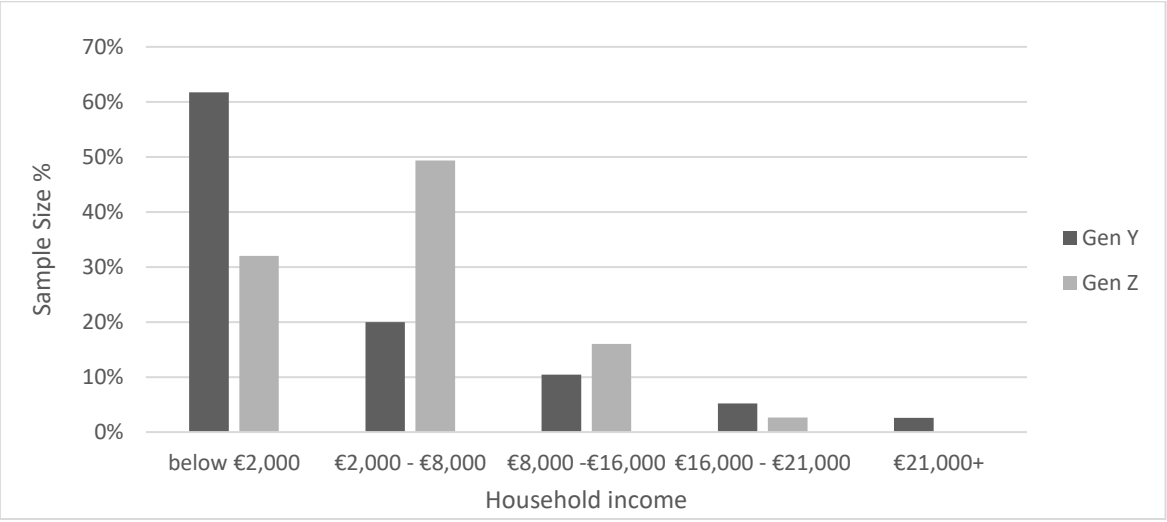


Figure 9: Responses to Question 6: Select the range that best reflects yours or both your parent's total annual household income.

To see if there is a correlation between annual household income and online privacy and security concern levels, Pearson's Product-Moment Correlation is utilised. Table 13 shows a weak correlation for both Gen Y, $r(113) = .10$, $p = .27$ and Gen Z, $r(73) = .20$, $p = .08$. This is in addition to a p-value higher than 0.05, making the correlations statistically not significant.

Table 13: Pearson's correlation results conducted on survey Question 6 relating to annual household income and online privacy and security levels of concern.

		Gen Y	Gen Z
Q6. Select the range that best reflects yours or both your parent's total annual household income.	Pearson Correlation	0.10	0.20
	Sig. (2-tailed)	0.27	0.08
	N	115.00	75.00

Education: Descriptive statistics shown in Table 14 and Figure 10, shows that 67% of Gen Y (M = 1.97, SD = .58) have a diploma or a bachelor's degree. Whereas through the Gen Z (M = 1, SD = 0) descriptive statistics, the mean has a value of 1 and standard deviation since the whole sample consists of primary and high school students.

Table 14: Descriptive analysis conducted on survey Question 4 relating to education.

Income	N	Mean	Std. Deviation	Std. Error Mean
Gen Y	115	1.97	0.58	0.05
Gen Z	75.00	1.00	0.00	0.00

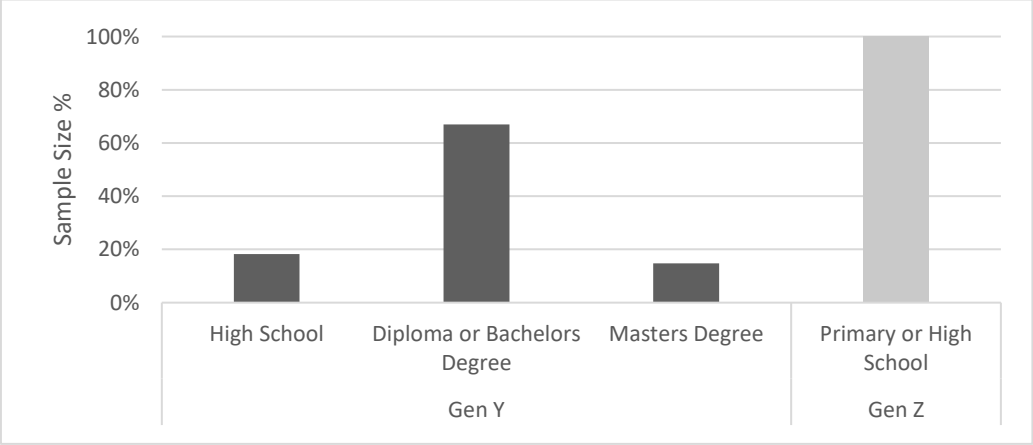


Figure 10: Responses to Question 4: What is your highest level of education?

Correlation as tabulated in Table 15, between education and level of concern is statistically analysed for Gen Y and Gen Z respectively. Pearson's coefficient shows that education and the concern levels for Gen Y, $r(113) = .0.005$, $p = .96$ and Gen Z, $r(73) = .05$, $p = .68$ almost negligible correlation, in addition to a low p-value signifying no statistical significance.

Table 15: Pearson's correlation results conducted on survey Question 6 relating to education level and online privacy and security levels of concern.

		Gen Y	**Gen Z**
4. What is your highest level of education?	Pearson Correlation	0.005	0.05
	Sig. (2-tailed)	0.96	0.68
	N	115.00	75.00

6 Discussion

The aim of this pilot study was to understand the level of concern of the sample set Gen Y and Gen Z regarding their online security and privacy in North India. The study used quantitative statistical analysis to examine data derived from online surveys. Two objectives are discussed using the statistical results listed in Chapter 5.

Objective 1: Analyse the concerns of Generations Y's online privacy and security concerns in comparison to Generations Z

The first objective of this study is to explore the online privacy and security concerns that are the most significant for the two generations. Analysis from Section 5.2 show that Gen Y have a higher level of online privacy concern than their younger Gen Z counterparts.

The sample results also show no difference in level of online security concern between Gen Y and Gen Z, with the exception of online security concern for 'having a strong password'.

Observing overall demographics, there is a difference in sample numbers with having more Gen Y (N = 115) participants than Gen Z (N = 75). One possible reason for this was that North India was going through an exam period, hence many school students did not make this survey a priority.

It could also be due to maturity levels and a lack of appreciation of the applicability of online privacy and security in their lives. Whereas, the excellent completion rates from Gen Y participants could be seen through comments stating that they learned a lot about their own concerns and attitudes through the survey process.

Another obvious reason to the disparity of concern levels is due to the difference in age and education level between Gen Y and Gen Z. As the quote goes, 'with age, come wisdom' and at many points through this survey this difference in thinking can be seen. Majority of the Gen Y sample were degree holders between 19-21 years old, whereas the Gen Z sample were predominately high school students.

When asked to select the top two privacy concerns both Gen Y and Gen Z selected phishing emails and tagging their location on social media. This shows that both generations understand the risk and implications of responding to a scam email or data they put out into social media. Only 21% of Gen Z think that location tagging on social media is a concern. Gen Z participants also commented that they were concerned about script kiddies hacking their social media accounts.

However, an interesting twist to the data shows 30% of Gen Y and 45% of Gen Z selected phishing emails to be of their major concern, but only 10% of Gen Y and Gen Z feel concerned about clicking unknown links. This high concern for phishing emails but low concern for clicking unknown links shows a lack of deep understanding of threats. This means once a cybercriminal is successful in deceiving the participant into opening phishing email, the probability that the participant clicks the unknown link is high.

Another privacy question dealt with how concerned would a participant feel if their personal data could be found freely online. Gen Y was most concerned about their house address, whilst Gen Z was most concerned about their cell phone number being found freely online. This could be due to numerous factors, but as social teenagers, Gen Z's focus predominately revolves around mobile technology and social media, allowing their mobile phones to be very central to their everyday life. This is in addition to their feeling of ownership towards their phone, compared to the lack of ownership for the home they live in.

Hence, having their mobile number freely on the internet could expose them to cyber bullying. This relates directly to Gen Y and Gen Z's top concern with communication, which is using chat or instant messaging (IM) platforms such as Whatsapp or Facebook Messenger and posting information on social media.

Finally, examining correlations between concern levels with gender, household income, education and device usage respectively, all four variables had weak correlation. This could be due to several factors one of which could be due to the sample size not being large or variable enough to enable efficient deduction of the correlation. Also, the distributions of Gen Y and Gen Z are dissimilarity and can cause weakness in the correlation *r* values.

Objective 2: Determine commonalities or differences between the concerns of Generations Y and Generations Z.

To address the second objective, results from the survey are analysed to highlight similarities or variances in concern levels between Gen Y and Gen Z samples. There are many commonalities evident from the survey results, the first being the almost equal ratio of female (27%) to male (73%) percentages in both Gen Y and Gen Z samples. Ideally, an equal representation of both genders would be preferred, however, having the same ratio allows for ease of comparison between generations.

Another demographic to point out are the differences between annual household income brackets of Gen Y and Gen Z samples. More than 60 % of the Gen Y sample have the lowest income, earning below €2000 annually compared to half of the Gen Z participants household income falling into the €2000 - €8000 income bracket.

The statistical test employed was to prove if there was a correlation between higher income families and higher concern levels. Meaning that a family in a higher income bracket would most probably more educated and would have greater exposure to the potential dangers of online privacy and security.

However, the Pearson's statistical test proved that there was weak correlation. Upon analysis, perhaps this is due to the majority of the Gen Y sample are aged between 19 – 25 years old, living as a single entity and have just started their careers. This is contrasted against the more established Gen Z parents who may have more than 10 years of work experience and obtaining higher income.

From a device usage point of view, both Gen Y and Gen Z spent similar hours on their smartphones, where 40% of respective generations spent 1-3 hours and the other 40% spent 4-6 hours on smartphones. In this wired world, Gen Z students in India are getting mobile phones at an average age of 13 years old [5], which enables teenagers to remain in contact with their friends all day. So, it is not surprising to see 40% of the Gen Z sample spend 4-6 hours a day on their mobile phones.

A clear difference in device usage is seen with the Gen Y sample spending longer hours on laptops and desktop computers daily compared to Gen Z. This is most probably since majority of the Gen Y sample will be working professionals or university students, where

completing tasks with a computer is a requirement. However, correlation tests show that there is weak correlation between hours spent on a device and concern levels of both generations.

A clear similarity between the participants concern toward device and software security can be seen through their concern to keep their anti-virus and operating systems updated. This awareness most probably stems from the recent Wannacry ransomware attacks on India [78].

This could be due to the sample population being more aware and exposed to the dangers and hence know that they must patch any vulnerabilities through updates for their anti-virus or operating system. This leads on to provide a reason why both Gen Y and Gen Z samples have a high and equal level of concern towards cyber criminals stealing and selling their personal data.

An encouraging observation can be seen with Gen Z having a 34% increased mean level of concern in having a strong password compared to Gen Y. This could be linked to Gen Z's notable concern towards script kiddies hacking into social media accounts.

However, even though Gen Y are equally aware about protecting their devices and software, they but seem slightly disconnected than Gen Z from the tracking practices of online technology companies. This can be seen from the survey data and could draw upon the fact that Gen Y grew up in an era where a breach of online personal data was unheard of, hence changing the psychology could be a challenge.

Another reason could be due to the discounting nature of the Gen Y sample who understand that free platforms and software need to create revenue and are nonchalant about access to their personal data. Furthering on from this point, the data also illustrates how the Gen Y sample are not very concerned about their photo, video or email address being freely found online compared to Gen Z.

Gen Y, being working professionals, could be opting-in to the public availability of such info, through professional networking sites such as LinkedIn. Additionally, Gen Y had lower mean levels of concern compared to Gen Z for communicating through social media platforms or using chat or IM.

Since Gen Z was born into the social media age and undoubtedly more active in messaging, sharing images and videos with their friends, they may have more exposure and understanding of the dangers of cyber bullying or viral images. Additionally, the recent introduction of IT education initiatives by the public and private sector in North India could be a contributing factor to Gen Z's increased awareness and exposure to the cyber world and its latest threats.

Gen Y's could still be accustomed to using their traditional and more conservative modes of communication such as talking on the mobile phone or through email. Given the comparatively less formal IT education received by Gen Y newer technologies like IM, social media and smartphones create greater concern for them.

Another large dissimilarity between Gen Y and Gen Z's concern level towards threat actors is their trust for their own Indian Government. Gen Z showed lower levels of concern, implying high trust in their government.

This is compared to Gen Y who have comparable concern levels for both the Indian Government and other governments. This could show that Gen Z lacked maturity or limited understanding of the consequence for potential breaches of their personal online data and lack of freedom of rights.

At a macro level, this dissimilarity could be a result of differing perceptions between the generations of the Indian government; North India has seen a growing number of protests conducted by left-wing university student groups against the restrictive policies of the right-wing Modi administration [79]. This could be the reason for the lower trust level Gen Y has towards the Indian government as compared to Gen Z. Gen Z may not be as exposed as Gen Y to such political issues yet due to their age.

The sample data also showed startling information about the innocence and unawareness of the Gen Z sample towards online privacy and security threats. A fifth of Generation Z did not know that their online activities could be tracked, monitored or saved.

From this number, 80% was male, which corresponds to data in Section 5.4 indicating that females are more concerned about their online privacy and security compared to males. The findings of this sample data, showing females having higher levels of concern

compared to males in Gen Z, justify results from prior studies [54] [55] as shown in Chapter 3.

This could be due to multiple factors, but it simply could revolve around the psychological fact that females are generally less risk averse, hence having higher concern levels towards their online privacy and security. India is also a multifaceted country with conservative social norms especially in women's rights, whilst transitioning into a globalised country. Harassment and disrespect for women is prevalent in the online space, hence women may be wary of their online privacy [80].

6.1 Recommendation

The findings and discussion of this pilot study can be used by policy makers, designers of user awareness programs, government, parents and teachers as a starting point of discussion, educational purposes or as a start to a larger study that could include the following points:

- How to be respectful and cautious on social media, especially posting messages, images and videos, sharing geo-location, hacking and being hacked.
- Password generation and sharing.
- The dangers of sharing personal data online such as home addresses, birthdates, images, videos.
- Recognising phishing sites, malicious URLs and apps and unsecured Wi-Fi connections.

6.2 Future Work

In this report we only touch the tip of the iceberg with the topic of online privacy and security concerns in North India. There were a number of questions that surfaced during the study and the following points would be a continuation of this study:

1. It will be also interesting to perform a comparative analysis between other countries and Indian generations to understand how societal, governmental or educational environments can affect the generations level of concerns towards online privacy and security.

2. Future research could be performed through a more structured sample with a consideration of a larger population, i.e. all Indian states.

3. In addition to online privacy and security concerns, it may be beneficial to evaluate the behaviours and actions of the two generations in facing privacy and security issues online.

4. It will be also valuable to discover gaps to why Gen Y and Gen Z do not have the skills and understanding online privacy and security concerns.

5. Finally, another avenue that can be explored is the perspective of gender in online privacy and security concerns.

7 Conclusion

In this thesis we explored the online privacy and security concerns of two generations Y and Z in North India. The objective of the study was to understand and determine commonalities or differences between the samples' online privacy and security concerns.

In the literature review, extensive research has been made on Generation Y and Z, including other previous generations, on their social media levels, privacy awareness and online shopping behaviours. However, specific empirical analysis of online privacy and security concerns of two generations Y and Z have not been performed.

The findings from this study can be utilised as an introductory discussion point by many different parties, such as the Indian Government, schools, universities, parents and teachers to educate and develop personalised online privacy and security content.

The study used quantitative statistical analysis to examine data derived from online surveys. Four different statistical tests are conducted to answer the research questions presented in this study: Levene's Test, Student's t-test, Cohen's d and Pearson Product-Moment Correlation.

The sample results and analysis highlight many privacy concerns important to both Gen Y and Gen Z, such as high concern levels for phishing emails, geo location tags on social media, and allowing personal data such as house addresses and cell phone numbers to be freely available online.

Gen Z sample was most concerned about other governments accessing their private data and substantially reduced levels of concern for their own Indian Government. Other findings include Gen Y not being very concerned about their photo, video or email address being freely found online compared to Gen Z.

Analysis also shows that females have a higher online privacy and security concern levels compared to males, in both Generation Y and Z and that nearly a fifth of Generation Z did not know that their online activities could be tracked, monitored or saved.

Overall, the sample results show that Gen Y are more concerned about online privacy compared to Gen Z. Whereas when it comes to online security, there is no difference in concern between the two generations, except where Gen Z are more concerned about ‘having a strong password’ compared to Gen Y. Correlation of gender, household income, education and device usage mapped to concerns show weak correlation.

References

[1] S. Kumar, Web Usage Mining Techniques and Applications Across Industries, India: IGI Global, 2016.

[2] S. Aiyar, "Twenty-Five Years of Indian Economic Reform -POLICY ANALYSIS NO. 803," 26 October 2016. [Online]. Available: https://www.cato.org/publications/policy-analysis/twenty-five-years-indian-economic-reform#full. [Accessed 21 January 208].

[3] Maps Of India, "Zonal Maps of India," Maps Of India, 13 April 2012. [Online]. Available: https://www.mapsofindia.com/zonal/. [Accessed 9 April 2018].

[4] J. Urban, "How Growing Up With the Internet Made Millennials Different," 3 July 2015. [Online]. Available: https://www.entrepreneur.com/article/247886. [Accessed 14 January 2018].

[5] Ericsson, "Generation Z: Understanding the digital lives of India's young mobile users," Ericsson , Stockholm, 2012.

[6] H. Bresman and V. D. Rao, "A Survey of 19 Countries Shows How Generations X, Y, and Z Are — and Aren't — Different," 25 August 2017. [Online]. Available: https://hbr.org/2017/08/a-survey-of-19-countries-shows-how-generations-x-y-and-z-are-and-arent-different. [Accessed 14 January 2018].

[7] M. Meeker, "Internet Trends 2017," Kleiner Perkins, Menlo Park, 2017.

[8] I. Mehta, "Reliance Jio Is Driving Indian Internet Growth, Says The Mary Meeker Report," Huffpost, 1 June 2017. [Online]. Available: https://www.huffingtonpost.in/2017/06/01/reliance-jio-is-driving-indian-internet-growth-says-the-mary-me_a_22120777/. [Accessed 21 January 2018].

[9] S. Loucks-Horsley, "The Concerns-Based Adoption Model (CBAM): A model for change in individuals," in *Professional Development for Science Education: A Critical and Immediate Challenge*, Dubuque, Kendall/Hunt Publishing Co, 2001.

[10] A. Klimburg, "Strategic Goals & Stakeholders," in *National Cyber Security Framework Manual*, Tallinn, NATO Cooperative Cyber Defence Centre of Excellence , 2012, pp. 68-70.

[11] C. Johnson, L. Badger and D. Waltermire, Guide to Cyber Threat Information Sharing, Gaithersburg: National Institute of Standards and Technology, 2016.

[12] J. Sigholm, "Non-State Actors in Cyberspace Operations," *Journal of Military Studies,* vol. 4, p. 37, 2013.

[13] Trend Micro, "Definition: Cybercriminals," 2018. [Online]. Available: https://www.trendmicro.com/vinfo/us/security/definition/cybercriminals. [Accessed 1 February 2018].

[14] UK National Crime Agency, "Cyber crime - Common cyber threats," UK National Crime Agency, 2018. [Online]. Available:

http://www.nationalcrimeagency.gov.uk/crime-threats/cyber-crime. [Accessed 1 February 2018].

[15] G. Skouma and L. Léonard, "On-line Behavioral Tracking: What May Change After the Legal Reform on Personal Data Protection," in *Reforming European Data Protection Law*, Dordrecht , Springer Science+Business Media , 2015 , pp. 36-60.

[16] S. Hill, "How much do online advertisers really know about you? We asked an expert," Digital Trends, 25 June 2015. [Online]. Available: https://www.digitaltrends.com/computing/how-do-advertisers-track-you-online-we-found-out/. [Accessed 12 February 2018].

[17] B. Naylor, "Firms Are Buying, Sharing Your Online Info. What Can You Do About It?," NPR, 11 July 2016. [Online]. Available: https://www.npr.org/sections/alltechconsidered/2016/07/11/485571291/firms-are-buying-sharing-your-online-info-what-can-you-do-about-it. [Accessed 12 February 2018].

[18] C. Castelluccia, "BehaviouralTracking on the Internet: ATechnical Perspective," in *European Data Protection: In Good Health?*, Springer Science+Business , 2012, pp. 21-33.

[19] M.-D. Tran, Privacy Challenges in Online Targeted Advertising, Grenoble: Computers and Society, 2014.

[20] M. Rosenberg, N. Confessore and C. Cadwalladr, "How Trump Consultants Exploited the Facebook Data of Millions," The New York Times, 17 March 2018. [Online]. Available: https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html. [Accessed 25 March 2018].

[21] I. S. Rubinstein, G. T. Nojeim and R. D. Lee, "Systematic government access to personal data: a comparative analysis," *International Data Privacy Law*, vol. 4, no. 2, p. 96–119, 2014.

[22] S. Singh, "India's surveillance project may be as lethal as PRISM," The Hindu, 7 June 2016 . [Online]. Available: http://www.thehindu.com/news/national/indias-surveillance-project-may-be-as-lethal-as-prism/article4834619.ece. [Accessed 25 February 2018].

[23] D. Balasubramanian, "Privacy in Internet Era: Four Government Surveillance Programs You Must Know About," News18, 17 August 2017. [Online]. Available: https://www.news18.com/news/tech/privacy-in-internet-era-four-government-surveillance-programs-you-must-know-about-1493541.html. [Accessed 23 February 2018].

[24] G. Greenwald and S. Saxena, "India among top targets of spying by NSA," The Hindu, 4 June 2016. [Online]. Available: http://www.thehindu.com/news/national/india-among-top-targets-of-spying-by-nsa/article5157526.ece. [Accessed 28 February 2018].

[25] S. Curtis, "Chinese hackers 'have been spying on Asian governments for a decade'," The Telegraph, 13 April 2015. [Online]. Available: https://www.telegraph.co.uk/technology/internet-security/11532477/China-accused-of-decade-long-cyber-espionage-campaign-in-Asia.html. [Accessed 28 February 2018].

[26] S. Singh, "Indian Army terms Xiaomi apps security hazard, Xiaomi says it is investigating govt advisory," India Today, 1 December 2017. [Online]. Available:

https://www.indiatoday.in/technology/news/story/we-take-security-and-privacy-very-seriously-xiaomi-on-ib-advisory-that-put-popular-mi-apps-under-scanner-for-spying-1098000-2017-12-01. [Accessed 1 March 2018].

[27] N. Groffman, "Indian and Chinese espionage," *Defense & Security Analysis,* vol. 32, no. 2, pp. 144-162, 2016.

[28] Symantec, "2018 Internet Security Threat Report," Symantec, 2018.

[29] S. Chauhan, "27,482 Cases of Cybercrimes Reported in 2017, One Attack in India Every 10 Minutes," India.com, 22 July 2017. [Online]. Available: http://www.india.com/news/india/27482-cases-of-cybercrimes-reported-in-2017-one-attack-in-india-every-10-minutes-2341055/. [Accessed 20 February 2018].

[30] S. Ghate and P. Agrawal, "A Literature Review on Cyber Security in Indian Context," *Journal of Computer & Information Technology,* vol. 8, no. 5, pp. 30-36, 2017.

[31] O. Singh, P. Gupta and R. Kumar, "A Review of Indian Approach towards Cybersecurity," *International Journal of Current Engineering and Technology ,* vol. 6, no. 2, pp. 644-648, 2016.

[32] K. Gupta, "Govt working to set up financial CERT to tackle cyber threats," Livemint, 16 November 2017. [Online]. Available: http://www.livemint.com/Industry/KMK5eQsbcJpYvEMPfp5MHI/Govt-working-to-set-up-financial-CERT-to-tackle-cyber-threat.html. [Accessed 23 February 2018].

[33] S. Narula and N. Jindal, "Social Media, Indian Youth and Cyber Terrorism Awareness: A Comparative Analysis," *J Mass Communication & Journalism,* vol. 5, no. 2, 2015.

[34] S. Salujal, .. D. Bansal and S. Saluja, "Cyber Safety Education in High Schools," *International Conference on Computer Technology and Science,* vol. 47, pp. 107-112, 2012.

[35] The Times of India, "Cyber-security awareness plan for children launched," 14 November 2017. [Online]. Available: https://timesofindia.indiatimes.com/city/thiruvananthapuram/cyber-security-awareness-plan-for-children-launched/articleshow/61639962.cms. [Accessed 25 February 2018].

[36] A. Khare, "“Impact of Indian cultural values and lifestyles on meaning of branded products: Study of university students in India," *Journal of International Consumer Marketing,* vol. 23, no. 5, pp. 365-379, 2011.

[37] R. Iyer, J. Eastman, H. Monteiro, H. Rottier and S. Singh, "Perception of Millennial's Media Attitude and Use: A Comparison of U.S. and Indian Millennials," *The Marketing Management Journal,* vol. 26, no. 2, pp. 69-85 , 2016.

[38] A. Mathew and A. Ailawadi, "How India’s Generation Z Defines Privacy," Bloomberg, 23 July 2017. [Online]. Available: https://www.bloombergquint.com/business/2017/07/22/how-indias-generation-z-defines-privacy. [Accessed 27 February 2018].

[39] Australia Trade Commission, "Punjab, Haryana, Chandigarh and Himachal Pradesh, India," Australian Government, Canberra, 2015 .

[40] R. Levickaite, "Generations x, y, z: How social networks form the concept of the world without borders (the case of Lithuania)," *LIMES: Cultural Regionalistics,* vol. 3, no. 2, pp. 170-183, 2011.

[41] A. Sharma and S. Maidullah, "Generation - Silver, X, Y and Z Internet Users and Consumers of India," in *Strengthing Strategies, Shaping Policies and Empowering Personnel: Key to organizational competitiveness*, New Delhi, Bharti Publication, 2017.

[42] R. Phanthong and W. Settanaranon, "Differences of Consumers' Perception and Attitude towards Marketing Communication through media: comparison generation X, Y, and Z in Thailand," Mälardalen University, Västerås, 2011.

[43] K. Young, "The X, Y and Z of Generations in Schools," *International Journal of Learning,* vol. 16, no. 7, pp. 203-215, 2009.

[44] T. Wiedmer, "Generations Do Differ: Best Practices in Leading Traditionalists, Boomers, and Generations X, Y, and Z," *Delta Kappa Gamma Bulletin,* 2015.

[45] K. Khatri and N. Dixit, "Managing Aspiration of Generation "Y" and Generation "Z" at Work Place," *National Research Conference Special Issue 2016*, 2016.

[46] B. Andrea, H.-C. Gabriella and J. Tímea, "Y and Z Generations at Workplaces," *Journal of Competitiveness,* vol. 8, pp. 90 - 106, 2016.

[47] A. Bhatnagar, S. Misra and H. R. Rao, "Online risk, convenience, and internet shopping behavior," *Communications of the ACM,* vol. 43, pp. 98-1-5, 2004.

[48] S. T. Vijay and M. Balaji, "Influencing the online consumer behavior: the web experience," *Internet Research,* vol. 14, pp. 111-126, 2009.

[49] A. M. Suresh and R. Shashikala, "Identifying Factors of Consumer Perceived Risk towards Online Shopping in India," in *3rd International Conference on Information and Financial Engineering*, Singapore, 2011.

[50] P. Kumaraguru and N. Sachdeva, "Privacy in India: Attitudes and Awareness V 2.0," Indraprastha Institute of Information Technology, New Delhi, 2012.

[51] P. Kumaraguru and L. Cranor, "Privacy in India: Attitudes and Awareness," Carnegie Mellon University, Pittsburgh, 2005.

[52] Z. RUHWANYA, "Attitudes toward, and awareness of, online privacy and security: a quantitative comparision of East Africa and U.S. internet users," Kansas State University, Manhattan, 2015.

[53] J. Budak, I.-D. Anic and E. Rajh, "Public Attitudes towards privacy and surveillance in Croatia," in *Privacy and Security in the Digital Age*, New York, Routledge, 2013, pp. 100-118.

[54] H. Cho, S. S. Lim and M. Rivera-Sánchez, "A multinational study on online privacy: global concerns and local responses," *New Media & Society,* vol. 11, no. 3, pp. 395 - 416, 2009.

[55] M. G. Hoy and G. Milne, "Gender Differences in Privacy-Related Measures for Young Adult Facebook Users," *Journal of Interactive Advertising ,* vol. 10, no. 2, pp. 28-45, 2013.

[56] D. O'Neil, "Analysis of Internet Users' Level of Online Privacy Concerns," *Social Science Computer Review,* vol. 19, no. 1, pp. 17-31, 2001.

[57] K. B. Sheehan, "Toward a Typology of Internet Users and Online Privacy Concerns," *The Information Society ,* vol. 18, no. 1, pp. 21-32, 2011.

[58] M. Madden, "Public Perceptions of Privacy and Security in the Post-Snowden Era," Pew Research Center, 12 November 2014. [Online]. Available: http://www.pewinternet.org/2014/11/12/public-privacy-perceptions/. [Accessed 20 April 2018].

[59] C. Fuchs, "The Privacy & Security - Research Paper Series," *Privacy and Security in Europe* , pp. 1-24, 1 May 2013.

[60] Statistics Canada, Survey Methods and Practises, Ottawa: National Library of Canada, 2010.

[61] G. Peersman, Overview: Data Collection and Analysis Methods in Impact Evaluation, Florence: UNICEF Office of Research, 2014.

[62] Explorable.com , "Snowball Sampling," 2018. [Online]. Available: https://explorable.com/snowball-sampling. [Accessed 11 April 2018].

[63] K. Black, Business Statistics: Contemporary Decision Making, Minneapolis: Wiley, 2009.

[64] E. Heiervang and R. Goodman, "Advantages and limitations of web-based surveys: evidence from a child mental health survey.," *Soc Psychiatry Psychiatr Epidemiol,* vol. 46, no. 1, pp. 69-76, 2011 .

[65] M. v. Gelder, R. W. Bretveld and N. Roeleveld, "Web-based Questionnaires: The Future in Epidemiology," *American Journal of Epidemiology,* vol. 172, no. 11, p. 1292–1298, 2010.

[66] S. Rice, S. R. Winter, S. Doherty and M. Milner, "Advantages and Disadvantages of Using Internet-Based Survey Methods in Aviation-Related Research," *Journal of Aviation Technology and Engineering,* vol. 7, no. 1, p. 58–65, 2017.

[67] C. Greenlaw and S. Brown-Welty, "A Comparison of Web-Based and Paper-Based Survey Methods," *Evaluation Review* , vol. 33, no. 5, pp. 464 - 480, 2009 .

[68] Yale University, "Tests of Significance," Yale University, 1997. [Online]. Available: http://www.stat.yale.edu/Courses/1997-98/101/sigtest.htm. [Accessed 7 April 2018].

[69] Abebe, Daniels and McKean, Statistics and Data Analysis, Kalamazoo: Western Michigan University, 2001.

[70] E. Martz, "Bewildering Things Statisticians Say: "Failure to Reject the Null Hypothesis"," 30 January 2013. [Online]. Available: http://blog.minitab.com/blog/understanding-statistics/things-statisticians-say-failure-to-reject-the-null-hypothesis. [Accessed 9 April 2018].

[71] H. Levene, "Robust Tests for Equality of Variances," in *Contributions to Probability and Statistics*, Stanford, Stanford University Press, 1960, pp. 278-292.

[72] Fundamentals of Statistics, "Two-Sample F-Test," 8 October 2012. [Online]. Available: http://www.statistics4u.com/fundstat_eng/cc_test_2sample_ftest.html. [Accessed 13 April 2018].

[73] Laerd Statistics, "Independent t-test for two samples," Laerd Statistics, 2013. [Online]. Available: https://statistics.laerd.com/statistical-guides/independent-t-test-statistical-guide.php. [Accessed 11 April 2018].

[74] D. Denis, "Understanding Cohen's d," 18 October 2012. [Online]. Available: http://www.bwgriffin.com/gsu/courses/edur9131/content/cohen_d_Denis.pdf. [Accessed 11 April 2018].

[75] D. Groppe, "How to compute p-values and Cohen's d for z-tests," [Online]. Available: http://www.cogsci.ucsd.edu/~dgroppe/STATZ/ztest_pvalue_d.pdf. [Accessed 10 April 2018].

[76] J. Cohen, Statistical power analysis for the behavioral sciences, Hillsdale, 1977.

[77] Laerd Statistics, "Pearson Product-Moment Correlation," Laerd Statistics, 2013. [Online]. Available: https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php. [Accessed 11 April 2018].

[78] ET Bureau, "India third worst hit nation by ransomware Wannacry; over 40,000 computers affected," The Economic Times, 17 May 2017. [Online]. Available: https://economictimes.indiatimes.com/tech/internet/india-third-worst-hit-nation-by-ransomware-wannacry-over-40000-computers-affected/articleshow/58707260.cms. [Accessed 17 April 2018].

[79] S. Vij, "Why Narendra Modi's government is at war with students," DW, 16 February 2016. [Online]. Available: http://www.dw.com/en/why-narendra-modis-government-is-at-war-with-students/a-19051656. [Accessed 19 April 2018].

[80] J. Pasricha, "Keeping the internet safe for women and marginalized communities in India," Access Now, 1 December 2016. [Online]. Available: https://www.accessnow.org/keeping-internet-safe-women-marginalized-communities-india/. [Accessed 20 April 2018].

[81] L. S. Sterling, The Art of Agent-Oriented Modeling, London: The MIT Press, 2009.

[82] P. Golchha, R. Deshmukh and P. Lunia, "A Review on Network Security Threats and Solutions," *International Journal of Scientific Engineering and Research (IJSER)*, vol. 3, no. 4, pp. 2347 - 3878, 2015.

[83] S. D. B. M. a. S. V. Chitrey A., "A Comprehensive Study Of Social Engineering Based Attacks In India To Develop a Conceptual Model, International Journal of Information & Network Security (," vol. 1, no. 1, pp. 45-53, 2015.

[84] J. Burke, "NSA spied on Indian embassy and UN mission," The Guardian, 25 September 2013. [Online]. Available: https://www.theguardian.com/world/2013/sep/25/nsa-surveillance-indian-embassy-un-mission. [Accessed 25 February 2018].

[85] N. Cohen and T. Arieli, "Field research in conflict environments: Methodological challenges and snowball sampling," *Journal of Peace Research*, vol. 48, no. 4, p. 423–435, 2011.

[86] Laerd Statistics, "Descriptive and Inferential Statistics," Laerd Statistics, 2013. [Online]. Available: https://statistics.laerd.com/statistical-guides/descriptive-inferential-statistics.php. [Accessed 11 April 2018].

[87] I. a. t. t. o. s. b. NSA, http://www.thehindu.com/news/national/india-among-top-targets-of-spying-by-nsa/article5157526.ece.

[88] V. K. V and R. Singh, "Latest Face of Cybercrime and Its Prevention In India," *International Journal of Basic and Applied Sciences*, vol. 2, no. 4, pp. 150-156, 2013.

[89] E. Garbarino and M. Strahilevitz, "Gender differences in the perceived risk of buying online and the effects of receiving a site recommendation," *Journal of Business Research*, vol. 57, no. 7, pp. 768-775, 2004.

Appendix 1 – Survey invitation to participants

Hello,

You are invited to participate in a research study to investigate online privacy and security concerns of different generations in India. To be eligible, you must be an Indian between the ages of 13 -34 with no IT Security background.

Participation is purely voluntary and no personal information will be collected.

- If you are between ages 13 - 18 years old, click your survey here: https://goo.gl/forms/4yO0psnNmWNksGvz1
- If you are between ages 19 - 34 years old and NOT working in IT Security jobs, click your survey here: https://goo.gl/forms/DQbvkqvYtEDPBMR93

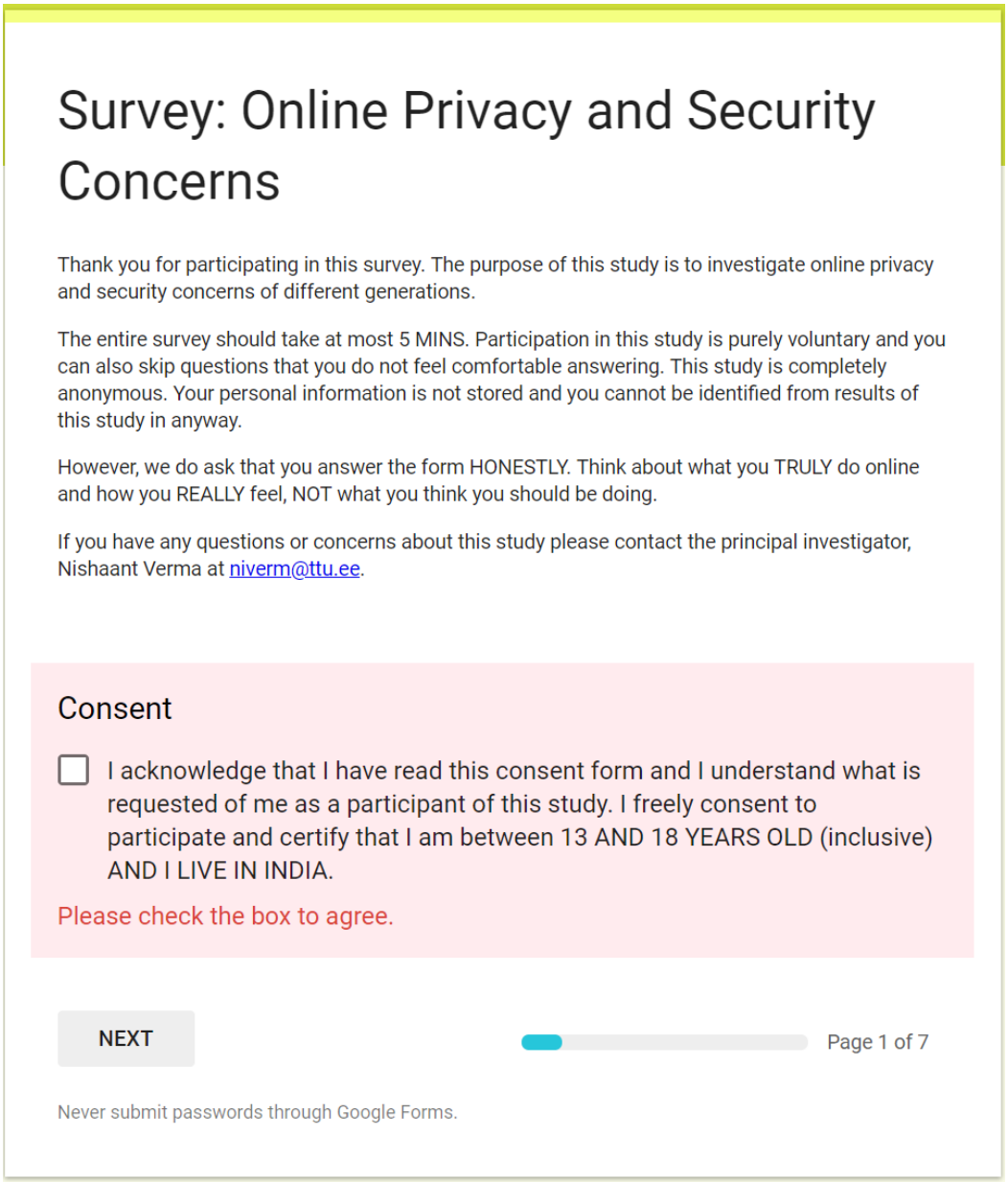
If you have any questions or concerns, please contact the researcher Nishaant Verma at niverm@ttu.ee.

Thank you for your assistance.

Nishaant

Appendix 2 – Consent Form

This section contains two sets of consent forms, for Gen Y (between 19 – 34 years old) and Gen Z (between 13 – 18 years old). The participant can only start the survey when they have consented to the requirements.



Survey: Online Privacy and Security Concerns

Thank you for participating in this survey. The purpose of this study is to investigate online privacy and security concerns of different generations.

The entire survey should take at most 5 MINS. Participation in this study is purely voluntary and you can also skip questions that you do not feel comfortable answering. This study is completely anonymous. Your personal information is not stored and you cannot be identified from results of this study in anyway.

However, we do ask that you answer the form HONESTLY. Think about what you TRULY do online and how you REALLY feel, NOT what you think you should be doing.

If you have any questions or concerns about this study please contact the principal investigator, Nishaant Verma at niverm@ttu.ee.

Consent

I acknowledge that I have read this consent form and I understand what is requested of me as a participant of this study. I freely consent to participate and certify that I am between 13 AND 18 YEARS OLD (inclusive) AND I LIVE IN INDIA.

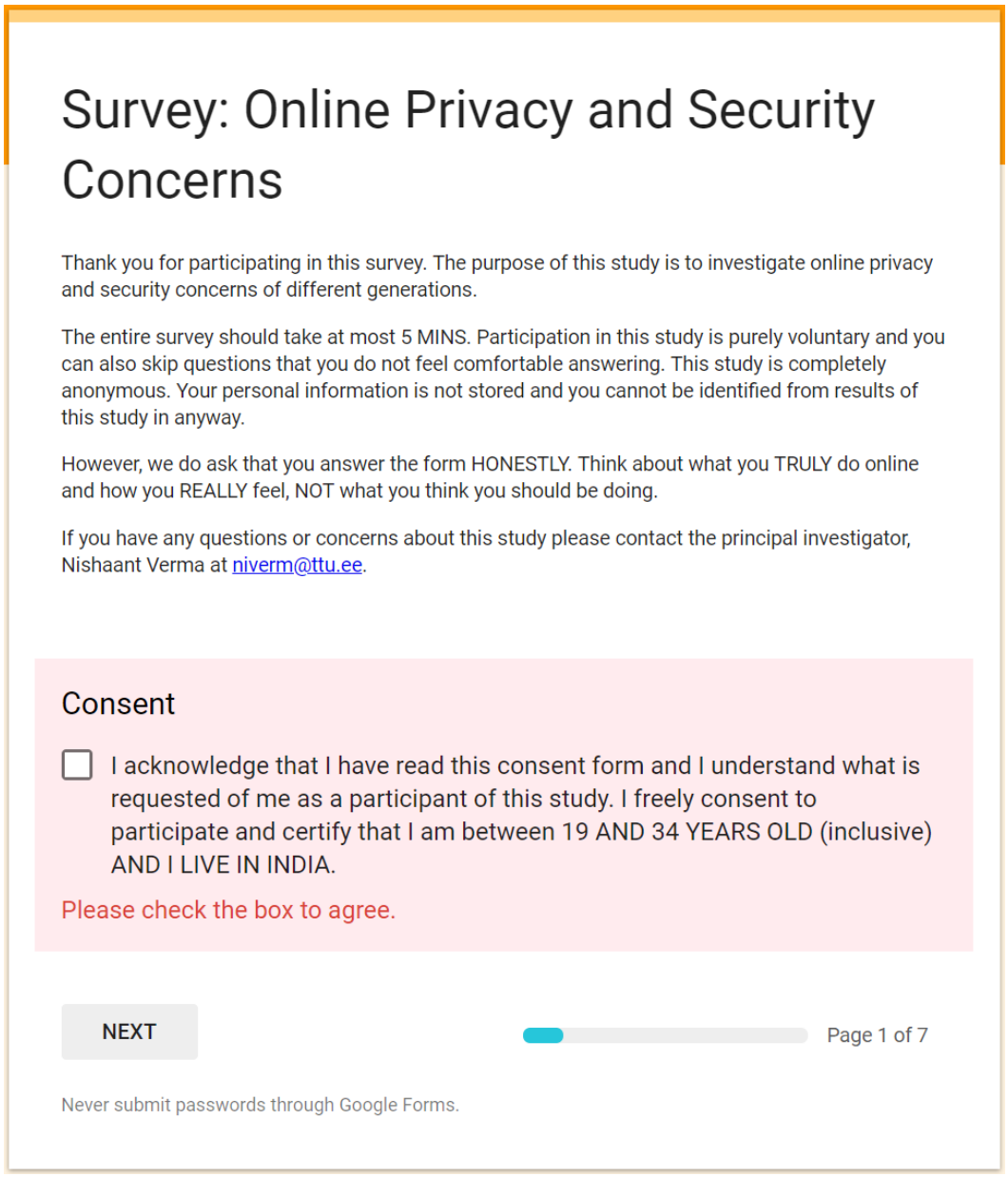
Please check the box to agree.

NEXT

Page 1 of 7

Never submit passwords through Google Forms.

Gen Z consent form.



Survey: Online Privacy and Security Concerns

Thank you for participating in this survey. The purpose of this study is to investigate online privacy and security concerns of different generations.

The entire survey should take at most 5 MINS. Participation in this study is purely voluntary and you can also skip questions that you do not feel comfortable answering. This study is completely anonymous. Your personal information is not stored and you cannot be identified from results of this study in anyway.

However, we do ask that you answer the form HONESTLY. Think about what you TRULY do online and how you REALLY feel, NOT what you think you should be doing.

If you have any questions or concerns about this study please contact the principal investigator, Nishaant Verma at niverm@ttu.ee.

Consent

☐ I acknowledge that I have read this consent form and I understand what is requested of me as a participant of this study. I freely consent to participate and certify that I am between 19 AND 34 YEARS OLD (inclusive) AND I LIVE IN INDIA.

Please check the box to agree.

NEXT

Page 1 of 7

Never submit passwords through Google Forms.

Gen Z consent form.

Appendix 3 – Questionnaire for Generation Z

YOUR BACKGROUND

1. What is your age? *

Choose

If you are NOT between 13 - 18 years old, click here: https://goo.gl/forms/DQbvkqvYtEDPBMR93

2. What is your self-identified gender?

- ○ Female
- ○ Male
- ○ Other: _____

3. In which state do you reside? *

Choose

4. What is your highest level of education?

Choose

5. If you are a school student, please select your stream:

Choose

6. Select the range that best reflects yours or both your parents total ANNUAL household income.

Choose

ONLINE ACTIVITY

7. How many hours a day do you use each internet based device?

	0	1 – 3 hours	4 - 6 hours	7 - 9 hours	More than 9 hours	Not applicable
Desktop computer	○	○	○	○	○	○
Laptop computers	○	○	○	○	○	○
Tablet	○	○	○	○	○	○
Smartphone	○	○	○	○	○	○
Other	○	○	○	○	○	○

8. Were you a victim of a cybercrime-related incident between 1 Jan 2017 – 31 Dec 2017?

- ○ Yes
- ○ No
- ○ Maybe

9. Do you care at all that your online activities are being tracked, watched or saved?

- ○ Very concerned
- ○ Slightly concerned
- ○ Not concerned
- ○ I should be concerned, but I'm not
- ○ I didn't know this could happen

10. How concerned are you that the following personal information can be FOUND FREELY on the internet?

	Very concerned	Slightly concerned	Not concerned	I should be concerned, but I'm not	Not applicable
The company you work for	○	○	○	○	○
Your school or university	○	○	○	○	○
Your home address	○	○	○	○	○
Your home phone number	○	○	○	○	○
Your cell phone number	○	○	○	○	○
Your date of birth	○	○	○	○	○
A video of you	○	○	○	○	○
A photo of you	○	○	○	○	○
Your email address	○	○	○	○	○

DEVICE

11. How concerned are you about....

*Malware is software written specifically to harm and infect the host system. Malware includes viruses along with other types of software such as trojan horses, worms, spyware, and adware. (https://antivirus.comodo.com/blog/computer-safety/malware-vs-viruses-whats-difference/)

	Very concerned	Slightly concerned	Not concerned	I should be concerned, but I'm not	Don't know about it
keeping your operating systems up-to-date	○	○	○	○	○
keeping your anti-virus software up-to-date	○	○	○	○	○
malware* infecting your phone	○	○	○	○	○
accessing an open Wi-Fi network	○	○	○	○	○
online banking	○	○	○	○	○
buying things online	○	○	○	○	○
having a strong password (a unique word more than 7 characters with letters, numbers and symbols)	○	○	○	○	○

ONLINE DATA

12. How concerned are you that the CYBER CRIMINALS* are accessing and selling....

*Cybercriminals nuse technology to commit malicious activities on digital systems with the intention of stealing sensitive company information or personal data, and generating profit.

	Very concerned	Slightly concerned	Not concerned	I should be concerned, but I'm not
your email contacts	○	○	○	○
your email contents	○	○	○	○
your online chat content	○	○	○	○
your website browsing history	○	○	○	○
your downloaded files	○	○	○	○
your GPS location	○	○	○	○
your webcam and microphone	○	○	○	○

13. How concerned are you that the ONLINE COMPANIES - like Google, Facebook & Apple are accessing and using....

	Very concerned	Slightly concerned	Not concerned	I should be concerned, but I'm not
your email contacts	○	○	○	○
your email contents	○	○	○	○
your online chat content	○	○	○	○
your website browsing history	○	○	○	○
your downloaded files	○	○	○	○
your GPS location	○	○	○	○
your webcam and microphone	○	○	○	○

14. How concerned are you that the INDIAN GOVERNMENT is accessing....

	Very concerned	Slightly concerned	Not concerned	I should be concerned, but I'm not
your email contacts	○	○	○	○
your email contents	○	○	○	○
your online chat content	○	○	○	○
your website browsing history	○	○	○	○
your downloaded files	○	○	○	○
your GPS location	○	○	○	○
your webcam and microphone	○	○	○	○

15. How concerned are you about the paragraph below:

The India government has built a Central Monitoring System (CMS) capable of accessing all communication data (telephone calls, both mobile and landline, VoIP calls, emails, and other communication on social media). All this without intervention of the service providers. (http://www.thehindu.com/news/national/indias-surveillance-project-may-be-as-lethal-as-prism/article4834619.ece)

- ◯ Very concerned
- ◯ Slightly concerned
- ◯ Not concerned
- ◯ I should be concerned, but I'm not

16. How concerned are you that OTHER GOVERNMENTS - like China, Pakistan or USA are accessing and monitoring....

	Very concerned	Slightly concerned	Not concerned	I should be concerned, but I'm not
your email contacts	◯	◯	◯	◯
your email contents	◯	◯	◯	◯
your online chat content	◯	◯	◯	◯
your website browsing history	◯	◯	◯	◯
your downloaded files	◯	◯	◯	◯
your GPS location	◯	◯	◯	◯
your webcam and microphone	◯	◯	◯	◯

COMMUNICATION

17. Which TWO (2) privacy issues are you MOST concerned about?

*Phishing is when a scammer uses fraudulent emails or texts, or copycat websites to get you to share valuable personal information. (https://www.consumer.ftc.gov/articles/0003-phishing)

- [ ] Phishing* emails - auto install of virus/steal passwords
- [ ] Tagging your location on social media
- [ ] Fingerprints stolen from selfies
- [ ] Mobile apps send personal data to their server
- [ ] Clicking unknown links
- [ ] Other: _____

18. How concerned are you when sharing private information over the following communication channels?

	Very concerned	Slightly concerned	Not concerned
Using a landline	○	○	○
Using a mobile phone	○	○	○
Sending a text message [SMS]	○	○	○
Sending an email	○	○	○
Using chat or instant message (IM) [Whatsapp, Facebook Messenger]	○	○	○
Using social media	○	○	○

Appendix 4 – Questionnaire for Generation Y

YOUR BACKGROUND

1. What is your age? *

Choose

If you are NOT between 19 - 34 years old, click here: https://goo.gl/forms/4yO0psnNmWNksGvz1

2. In which state do you reside? *

Choose

3. What is your self-identified gender?

- ○ Female
- ○ Male
- ○ Other: _____

4. What is your highest level of education?

Choose

5. If you are in/have graduated from university, please select your major in university:

Choose

6. Select the range that best reflects yours or both your parents total ANNUAL household income.

Choose

ONLINE ACTIVITY

7. How many hours a day do you use each internet based device?

	0	1 – 3 hours	4 - 6 hours	7 - 9 hours	More than 9 hours	Not applicable
Desktop computer	○	○	○	○	○	○
Laptop computers	○	○	○	○	○	○
Tablet	○	○	○	○	○	○
Smartphone	○	○	○	○	○	○
Other	○	○	○	○	○	○

8. Were you a victim of a cybercrime-related incident between 1 Jan 2017 – 31 Dec 2017?

- ○ Yes
- ○ No
- ○ Maybe

9. Do you care at all that your online activities are being tracked, watched or saved?

- ○ Very concerned
- ○ Slightly concerned
- ○ Not concerned
- ○ I should be concerned, but I'm not
- ○ I didn't know this could happen

10. How concerned are you that the following personal information can be FOUND FREELY on the internet?

	Very concerned	Slightly concerned	Not concerned	I should be concerned, but I'm not	Not applicable
The company you work for	○	○	○	○	○
Your school or university	○	○	○	○	○
Your home address	○	○	○	○	○
Your home phone number	○	○	○	○	○
Your cell phone number	○	○	○	○	○
Your date of birth	○	○	○	○	○
A video of you	○	○	○	○	○
A photo of you	○	○	○	○	○
Your email address	○	○	○	○	○

DEVICE

11. How concerned are you about....

*Malware is software written specifically to harm and infect the host system. Malware includes viruses along with other types of software such as trojan horses, worms, spyware, and adware. (https://antivirus.comodo.com/blog/computer-safety/malware-vs-viruses-whats-difference/)

	Very concerned	Slightly concerned	Not concerned	I should be concerned, but I'm not	Don't know about it
keeping your operating systems up-to-date	○	○	○	○	○
keeping your anti-virus software up-to-date	○	○	○	○	○
malware infecting your phone	○	○	○	○	○
accessing an open Wi-Fi network	○	○	○	○	○
online banking	○	○	○	○	○
buying things online	○	○	○	○	○
having a strong password (a unique word more than 7 characters with letters, numbers and symbols)	○	○	○	○	○

ONLINE DATA

12. How concerned are you that the CYBER CRIMINALS* are accessing and selling....

*Cybercriminals nuse technology to commit malicious activities on digital systems with the intention of stealing sensitive company information or personal data, and generating profit.

	Very concerned	Slightly concerned	Not concerned	I should be concerned, but I'm not
your email contacts	○	○	○	○
your email contents	○	○	○	○
your online chat content	○	○	○	○
your website browsing history	○	○	○	○
your downloaded files	○	○	○	○
your GPS location	○	○	○	○
your webcam and microphone	○	○	○	○

13. How concerned are you that the ONLINE COMPANIES - like Google, Facebook & Apple are accessing and using....

	Very concerned	Slightly concerned	Not concerned	I should be concerned, but I'm not
your email contacts	○	○	○	○
your email contents	○	○	○	○
your online chat content	○	○	○	○
your website browsing history	○	○	○	○
your downloaded files	○	○	○	○
your GPS location	○	○	○	○
your webcam and microphone	○	○	○	○

14. How concerned are you that the INDIAN GOVERNMENT is accessing and monitoring....

	Very concerned	Slightly concerned	Not concerned	I should be concerned, but I'm not
your email contacts	○	○	○	○
your email contents	○	○	○	○
your online chat content	○	○	○	○
your website browsing history	○	○	○	○
your downloaded files	○	○	○	○
your GPS location	○	○	○	○
your webcam and microphone	○	○	○	○

15. How concerned are you about the paragraph below:

The India government has built a Central Monitoring System (CMS) capable of accessing all communication data (telephone calls, both mobile and landline, VoIP calls, emails, and other communication on social media). All this without intervention of the service providers. (http://www.thehindu.com/news/national/indias-surveillance-project-may-be-as-lethal-as-prism/article4834619.ece)

- ◯ Very concerned
- ◯ Slightly concerned
- ◯ Not concerned
- ◯ I should be concerned, but I'm not

16. How concerned are you that OTHER GOVERNMENTS - like China, Pakistan or USA are accessing and monitoring....

	Very concerned	Slightly concerned	Not concerned	I should be concerned, but I'm not
your email contacts	◯	◯	◯	◯
your email contents	◯	◯	◯	◯
your online chat content	◯	◯	◯	◯
your website browsing history	◯	◯	◯	◯
your downloaded files	◯	◯	◯	◯
your GPS location	◯	◯	◯	◯
your webcam and microphone	◯	◯	◯	◯

COMMUNICATION

17. Which TWO (2) privacy issues are you MOST concerned about?

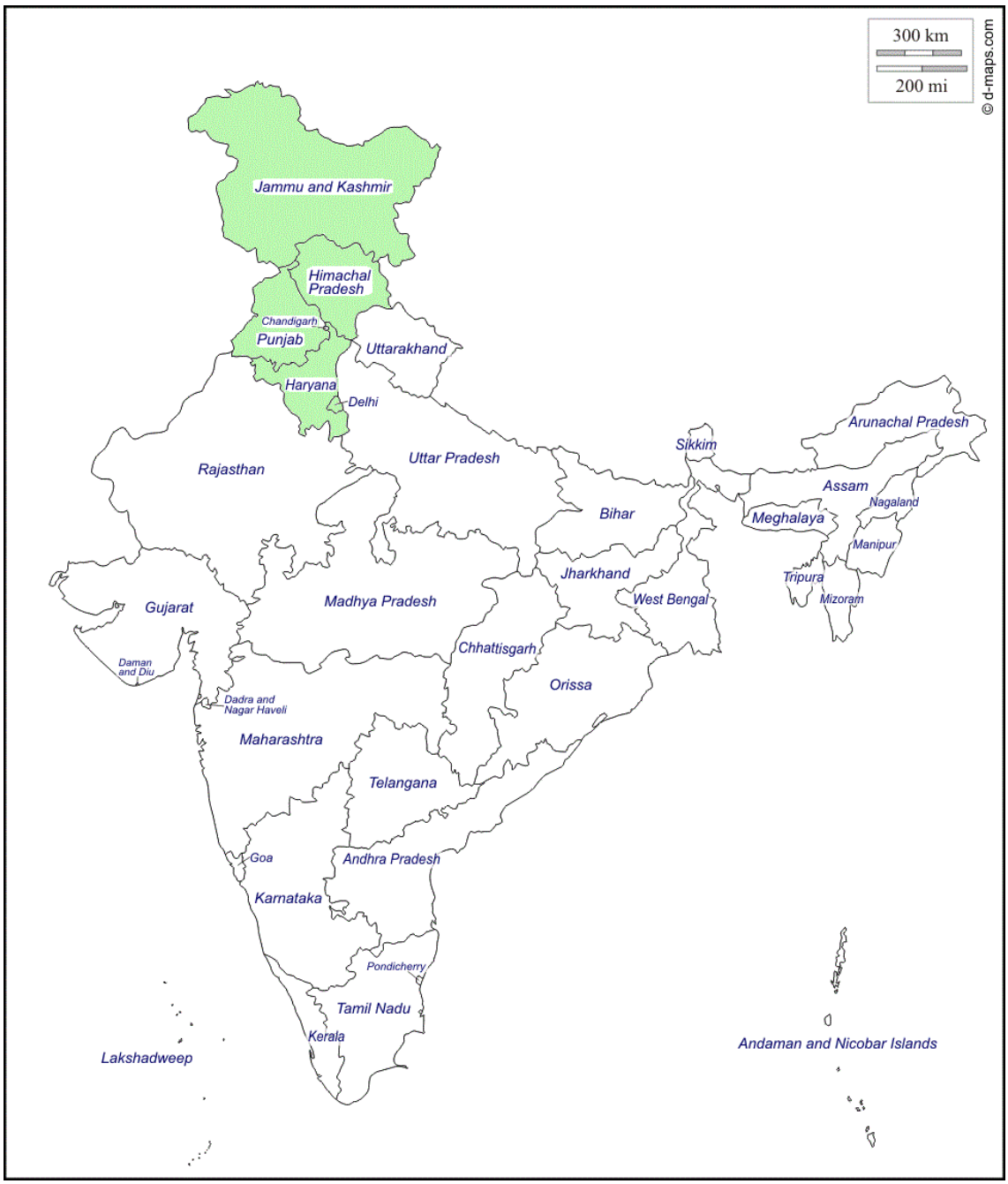
*Phishing is when a scammer uses fraudulent emails or texts, or copycat websites to get you to share valuable personal information. (https://www.consumer.ftc.gov/articles/0003-phishing)

- [ ] Phishing* emails - auto install of virus/steal passwords
- [ ] Tagging your location on social media
- [ ] Fingerprints stolen from selfies
- [ ] Mobile apps send personal data to their server
- [ ] Clicking unknown links
- [ ] Other: _____

18. Which mode of communication do you feel most concern when sharing private information?

	Very concerned	Slightly concerned	Not concerned
Using a landline	○	○	○
Using a mobile phone	○	○	○
Sending a text message [SMS]	○	○	○
Sending an email	○	○	○
Using chat or instant message (IM) [Whatsapp, Facebook Messenger]	○	○	○
Using social media	○	○	○

Appendix 5 – Map of North India



Source: d-maps.com