The world's anxiety of cybercrime in 2019*

Julie Nguyen

19 April 2023

Abstract

The emergence of the internet has altered the way we live, work, and communicate with each other. Despite being a tool to improve life in overall, technology and the internet also pose some potential risks that greatly affect the life of people. This paper aims at identifying the perceptions of risk of using the internet among people around the world by exploring the patterns and the relationship between gender, the number of children, the living location and the pereception of risk towards online bullying. Findings show that females tend to be more worried about experiencing cybercrime than males, people having more children also witness the same pattern. Results also show that gender, the number of children, and the living location affect the anxiety of the prospect of being a victim of cybercrime.

1 Introduction

The emergence of the internet has altered the way we live, work, and communicate with each other. It has become an integral part of our daily lives, offering numerous benefits such as instant access to information, online shopping, and communication with people from all over the world. However, along with these benefits, the Internet also poses various risks to its users.

The dangers of using the internet increase as it develops and expands. Some of the main dangers that people encounter include cybercrime, online scams, identity theft, cyberbullying, and pestering. Regardless of the person's age, gender, or location, these dangers can have significant financial and mental repercussions. Findings from Gray (2017) about emerging technology and its potential risk suggest that the Internet of Things (IoT) fell into the group of technology that had lots of risks, but not so much reward. On the one hand, they gave more access to the life of people. On the other hand, they could also "give cyber-criminals a window to our world".

Therefore, it is essential to recognize these dangers. Understanding and identifying risk is as important as keeping the world healthy. This paper focuses on identifying the perceptions of risk of using the internet among people around the world. The paper also analyzes the correlation between education, income, and living locations, and how these factors affect the perception of risk towards the internet. In exploring the patterns of risk perception, the paper hopes to provide insights into people in different demographics experience and perceive potential risks posed by digital technology, through which will improve digital well-being, raise awareness, and enhance the ways to protect internet users in the world.

This paper will firstly explore the patterns of risk posed by technology in the world and in different global regions from the data of "World Risk Poll" (2019) in 2019 using R Core Team (2022). Various groups (of gender, education, and income) are observed to see how they affect the respondent's perception of risk caused by the internet. Other packages such as tidyverse (Wickham et al. 2019), readr (Wickham, Hester, and Bryan 2022), haven (Wickham, Miller, and Smith 2023), janitor (Firke 2021), ggplot2 (Wickham 2016), tidymodels (Kuhn and Wickham 2020) etc. are also used to aggregate, visualize, and model the data.

The main findings show that females tend to experience more anxiety towards the prospect of cybercrime than males. People living in low-income countries tend to witness the same pattern. People who have less kids tend to be less worried about the potential of risk.

In the following paper, I will delve into the data from "World Risk Poll" (2019), and how it is obtained and arranged for proper analysis. This section will also guide us through the understanding of the initial survey and ethical concerns

*Code and data are available at: https://github.com/julieenguyennn/World-Risk-Analysis

regarding the data collection. Moving on to the Model section, we will discuss our model and its implications for how we interpret our results going forward. In the Results section, graphical data is shown as to provide a clear understanding of the results given. Lastly, the Discussion section will include what has been done, what was learnt as well as identify some weaknesses and propose the next steps for the paper.

2 Data

2.1 Dataset

To attempt an analysis on …., I manipulated the "World Risk Poll" (2019) data from the year 2019. These data are collected by the Llyod's… The dataset was gatherd in… and last updated on… The raw data includes observations. Using R (R Core Team 2022), tidyverse (Wickham et al. 2019), readr (Wickham, Hester, and Bryan 2022), haven (Wickham, Miller, and Smith 2023), janitor (Firke 2021), ggplot2 (Wickham 2016), tidymodels (Kuhn and Wickham 2020), I cleaned and extracted the necessary data to complete an exploratory analysis and modelling.

2.2 Variables

09 variables were selected to conduct analysis for the paper including 06 variables for respondents' general information (gender, education attainment, global region, urbanicity, and the number of children in household) and 03 variables relating to the anxiety towards the risk of being online (cyber bullying, false information and fraudulent activities). I first viewed the presence of anxiety towards different gender, education, and global region to investigate the global trend of online risk and anxiety in the year 2019. Along with these variables, the remainings are used later on to explore whether there is a causal relationship between the respondent's identity and their being worried about risks posed by the internet.

The anxiety towards different types of cybercrime were measure by asking the respondents "When using the Internet or social media, do you worry about any of the following things happening to you?". The answered is scaled as "Yes", "No". Gender of the respondents was recorded as 1 for "Male", and 2 for "Female". For the variable regarding the number of children in household, it is categorized into from 0 to 5 as the number of children accordingly, except for 5 was coded as "5 and above". Urbanicity refers to the place that the respondents were living in at the time of the survey. It is recorded as 1 for "A rural area or on a farm/Small town or village", and 2 for "A large city/Suburb of a large city". Lastly, the education attainment of the respondent is divided into 3 levels "Up to 8 years of basic education", "9-15 years of education", and "Completed four years of education beyond high school and/or received a 4-year college degree". Other types of responses that fell into "Do not Know" and "Refused" in all questions are omitted from analysis.

2.3 Plot

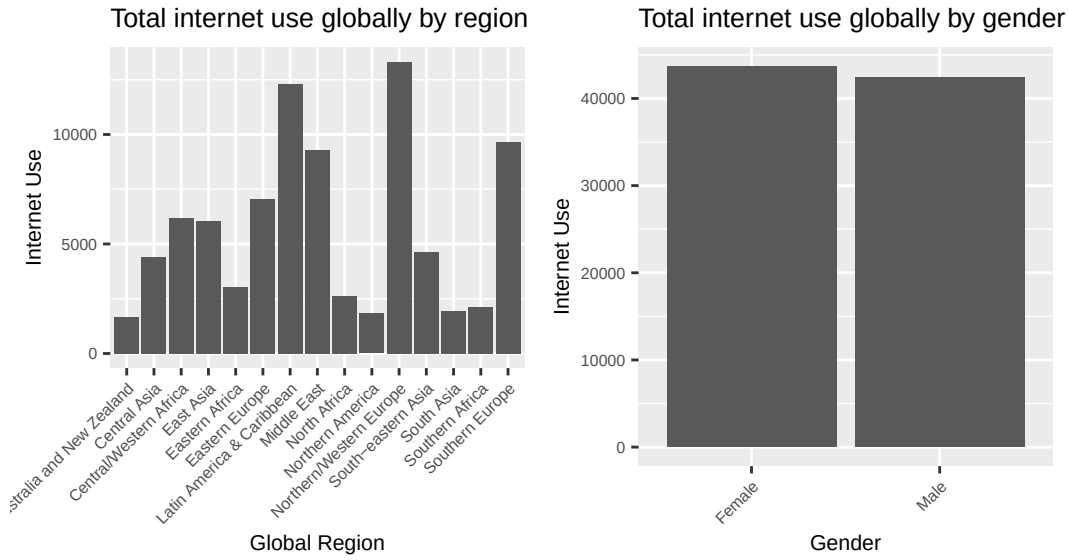


Figure 1: Anxiety towards cyber risks by region

(**internetuse?**) depicts the demographics of total internet use globally by region and gender. A total number of 86,088 observations is recorded as having used the internet in the last 30 days from the time of survey. Among 86,088 observations, it is recorded the the majority comes from Europe, with 15.47% in Northern/Western Europe, 11.23% in Southern Europe, and 8.21% in Eastern Europe. On the other hand, the region of Australia and New Zealand only witnesses a total of 1666 respondents, which takes up 1.94% of the total observations. Regarding the gender, the survey witnesses a quite equal proportion of female and male respondents, 50.75% and 49.25% respectively.

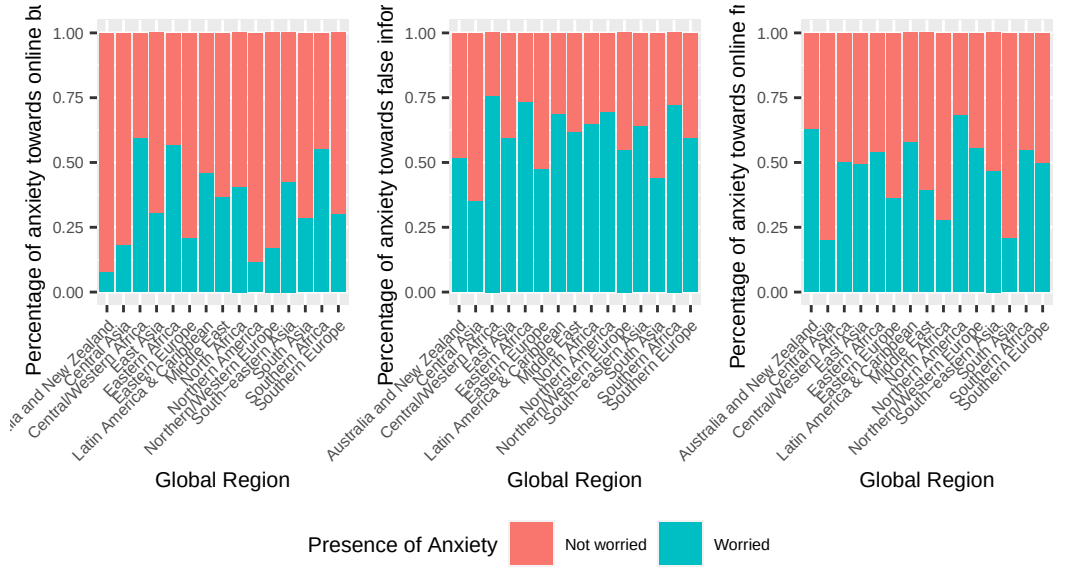


Figure 2: Anxiety towards cyber risks by region

?@fig-regionrisk illustrates the proportion of respondents feeling anxious towards the three types of cybercrime. In general, people were more worried about the prospect of being a victim of false information. A majority of respondents in Africa reported to have been worried about being online bullied, with a percentage of xxx% in Central/Western Africa, xxx% in Eastern Africa, xxx% in North Africa, and xxx% in Southern Africa. People from Australia and New Zealand were deeply concerned with confronting false information and fraudulent activities online. On the other hand,

the region of Central Asia and South Asia recorded similar attitudes of being "not worried" than "worried" for any of the three types of cybercrime.

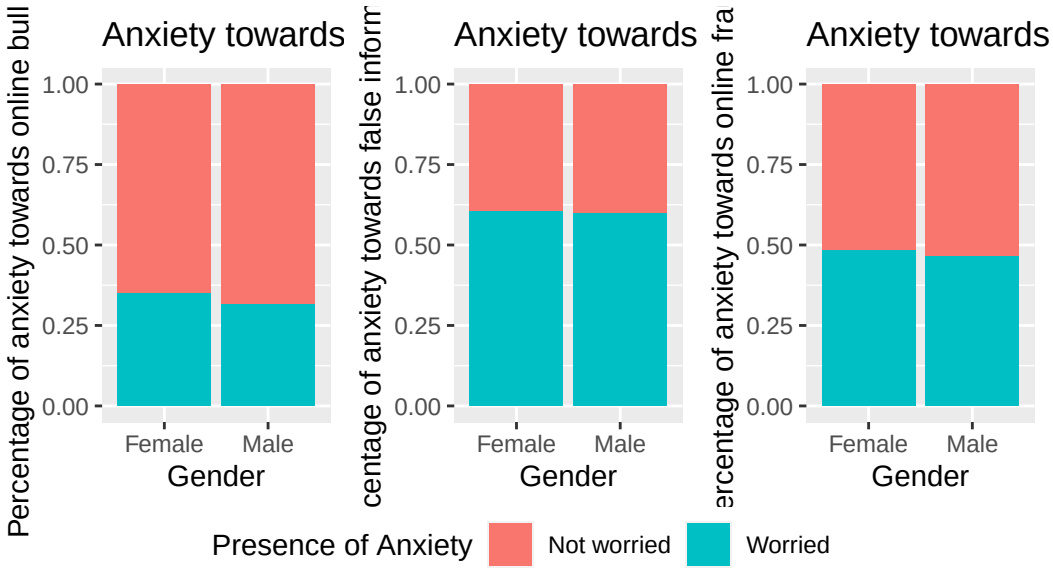


Figure 3: Anxiety towards cyber risks by gender

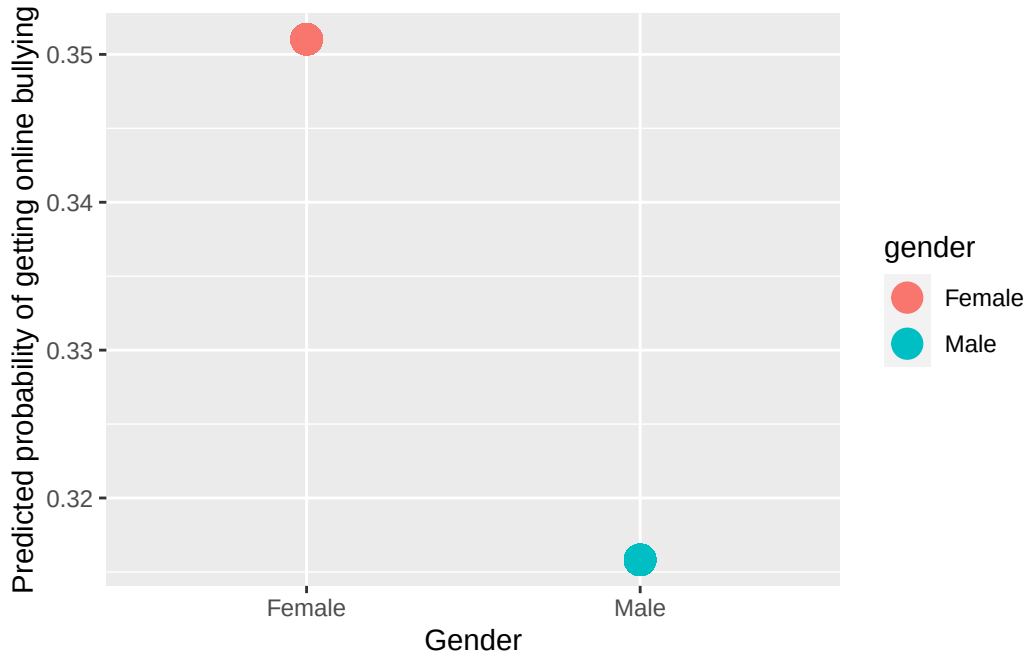
?@fig-genderrisk shows the percentage of people feeling worried about three types of cybercrime by gender. There is an overall trend, although not too distinct, that female respondents feel more anxious than male towards the online risks.

3 Model

Multiple logistic regression is used for analyzing the correlation between gender, the number of children, living location and the anxiety of the prospect of being a victim of cybercrime. $\text{logit}(p) = \beta 0 + \beta 1 * x1 + \beta 2 * x2 + ... + \beta n * xn$

```
               gender
online_bullying Female  Male
    Not worried  28354 29008
    Worried      15336 13390

               children_in_household
online_bullying     0     1     2     3     4     5
    Not worried 32137 10519  8595  3375  1410  1326
    Worried     12642  5855  5183  2506  1146  1394

             gender
false_info    Female  Male
  Not worried  17341 17061
  Worried      26349 25337

             education
false_info        1     2     3     4     5
  Not worried  4523 20405  9327    74    73
  Worried      6250 31426 13850    98    62

             education
fraud             1     2     3     4     5
  Not worried  6442 27458 11285    86    81
  Worried      4331 24373 11892    86    54
```

```
Call:
glm(formula = online_bullying ~ gender, family = "binomial",
    data = data_bullying)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.9299  -0.9299  -0.8712   1.4470   1.5183

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.61457    0.01002  -61.31   <2e-16 ***
genderMale  -0.15850    0.01448  -10.95   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 109634  on 86087  degrees of freedom
Residual deviance: 109514  on 86086  degrees of freedom
AIC: 109518

Number of Fisher Scoring iterations: 4
```




```
                              gender
probability_of_online_bullying Female  Male
             0.315816783810539      0 42398
             0.351018539711486  43690     0
```

```
Call:
glm(formula = online_bullying ~ children_in_household, family = "binomial",
    data = internet_risk)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.1987  -0.9408  -0.8145   1.3983   1.5904
```

```
Coefficients:
                       Estimate Std. Error z value Pr(>|z|)
(Intercept)            -0.93298    0.01050  -88.87   <2e-16 ***
children_in_household1  0.34710    0.01939   17.90   <2e-16 ***
children_in_household2  0.42719    0.02048   20.86   <2e-16 ***
children_in_household3  0.63528    0.02838   22.38   <2e-16 ***
children_in_household4  0.72567    0.04113   17.64   <2e-16 ***
children_in_household5  0.98299    0.03977   24.72   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 109634  on 86087  degrees of freedom
Residual deviance: 108206  on 86082  degrees of freedom
AIC: 108218

Number of Fisher Scoring iterations: 4
```

4 Results

5 Discussion

This gives valuable insights for pre-covid.

6 Appendix

References

Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data.* https://CRAN.R-project.org/package=janitor.

Gray, Alex. 2017. "What New Technologies Carry the Biggest Risks?" World Economic Forum. https://www.weforum.org/agenda/2017/01/what-emerging-technologies-have-the-biggest-negative-consequences/.

Kuhn, Max, and Hadley Wickham. 2020. *Tidymodels: A Collection of Packages for Modeling and Machine Learning Using Tidyverse Principles.* https://www.tidymodels.org.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2022. *Readr: Read Rectangular Text Data.* https://CRAN.R-project.org/package=readr.

Wickham, Hadley, Evan Miller, and Danny Smith. 2023. *Haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files.* https://CRAN.R-project.org/package=haven.

"World Risk Poll." 2019. The Lloyd's Register Foundation. https://wrp.lrfoundation.org.uk/.