Using Machine Learning & Data Science to fight Cyber Crime

Amrita Shalini, Graduate Student, Information Technology
Mentor – Dr. James Wolf

School of Information Technology
Illinois State University

Abstract

In the modern world the technology is rising quite rapidly. In order to operate our industries, government, national defense and other vital functions, society is becoming more dependent on data and networks. Nearly every person with computer knowledge has entered into the use of this fast-growing industry. Cyber crime is a criminal act committed via internet and technology. Most cyber crime is an attack on personal, corporate, or government information. Though the attacks do not occur on a physical body, they do occur on the virtual personal or corporate body, which is the set of information attributes that define people and institutions on the Internet. Today, cyber crime has drawn a great deal of attention around the globe because of the recorded dramatic increase rate of case reported. As a matter of fact, a new form of cyber crime case is being recorded every single day. In reality, cyber crime is the most brutal computer crime in the modern world that has caused a great deal of many people to lose their most confidential information. In other words, our virtual identities are essential elements in the digital age Machine learning & Data Science is one of the defense mechanisms to protect our virtual identities. This not only helps to react to cyber crimes by enhancing the cyber protection systems but also is instrumental in formulating defense strategies proactively by predictive analysis.

Data Explain

We Combine data from a large , representative sample of adult Americans (n=5000) with data from **Have I Been Pwned** to estimate the lower After drawing the bound of the average number of breached online accounts per person. The better educated, the middle aged
sample, YouGov used the emails associated with the accounts to query the haveibeenpwned.com(HIBP) AIP. HIBP is a non-profit clearinghouse of information about online account breaches. HIBP's stated aim is to provide a way for people to check if they are at risk from online breaches.
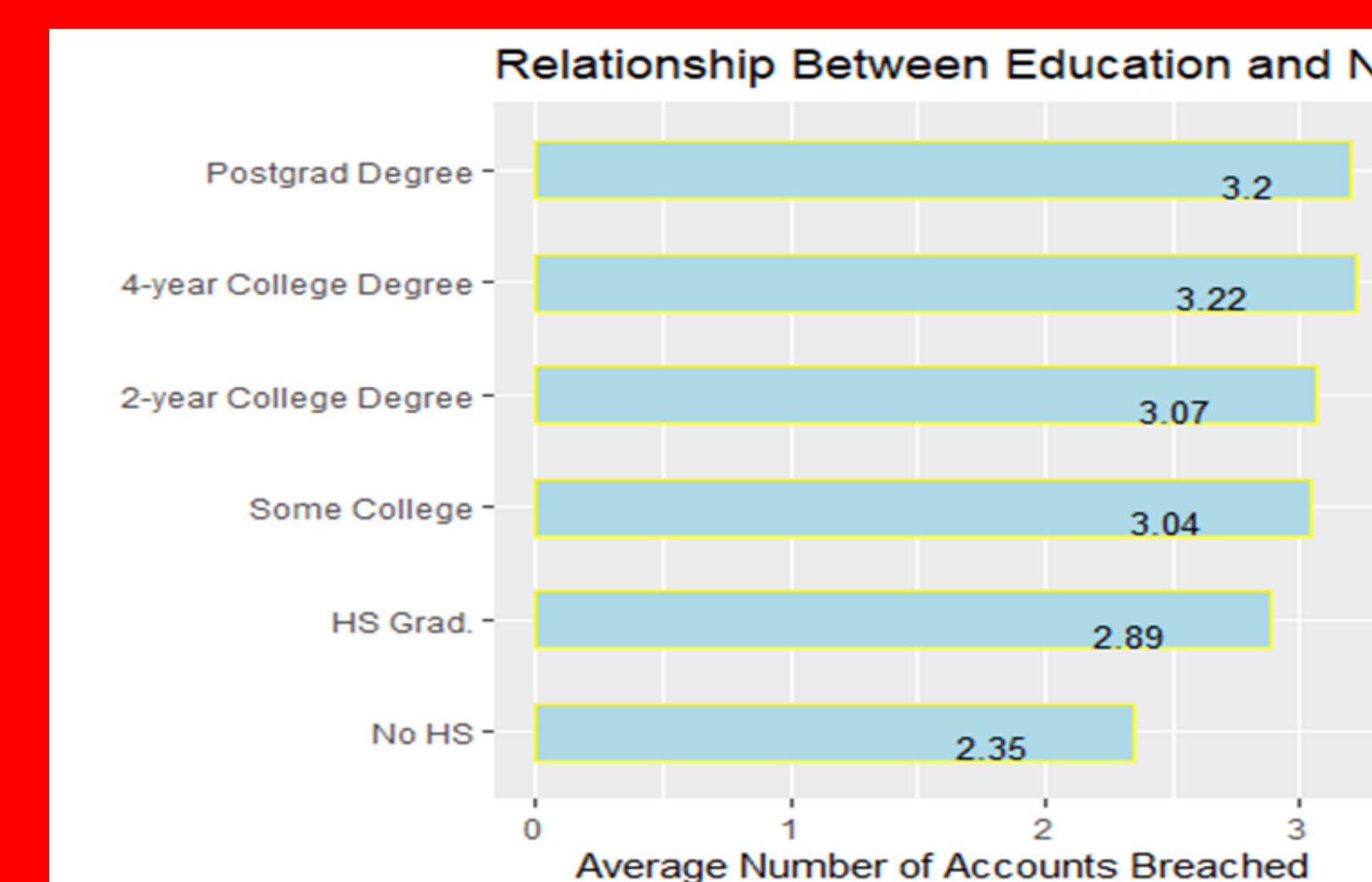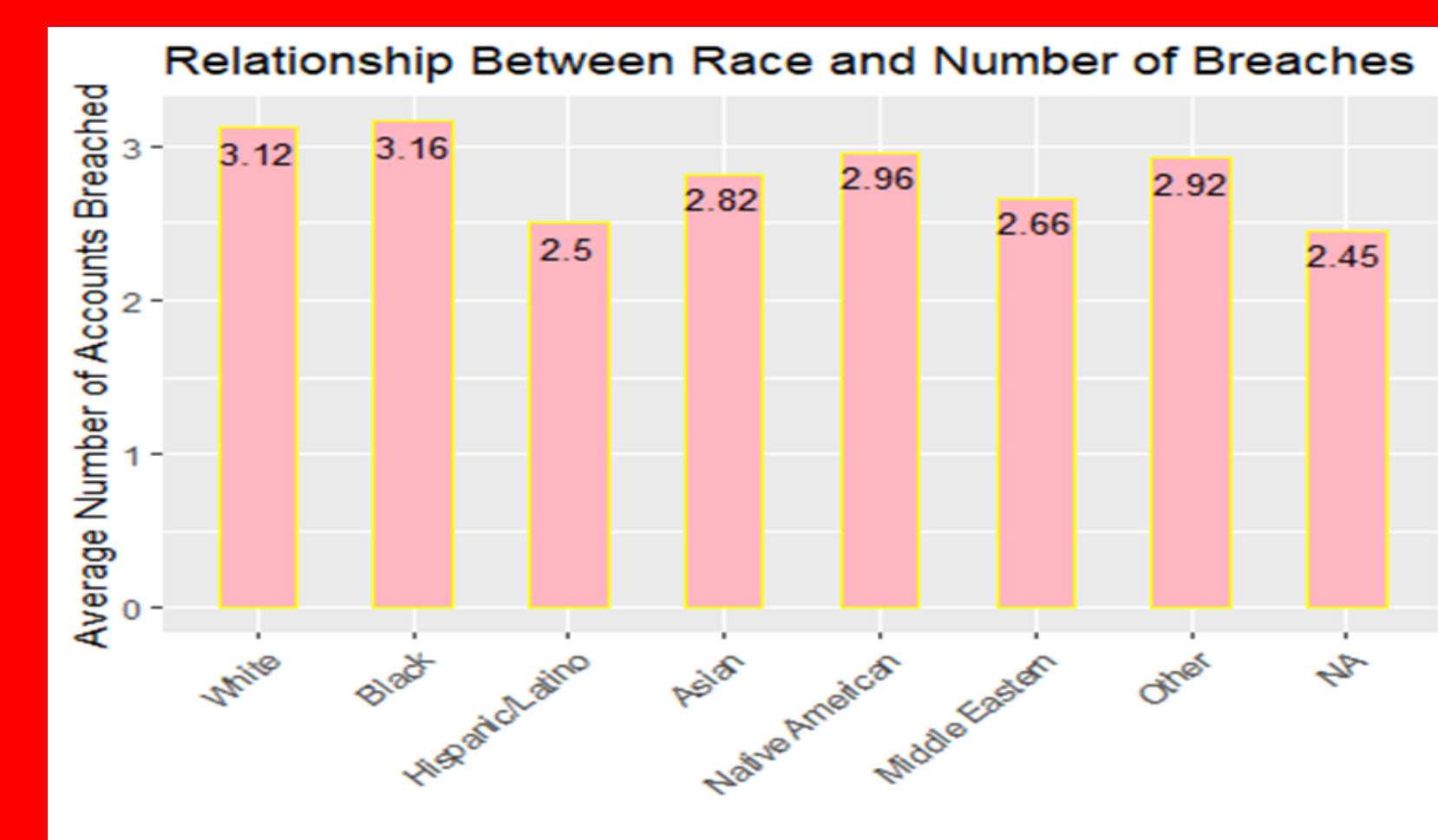
Relationship Between Gender and no. of Breaches
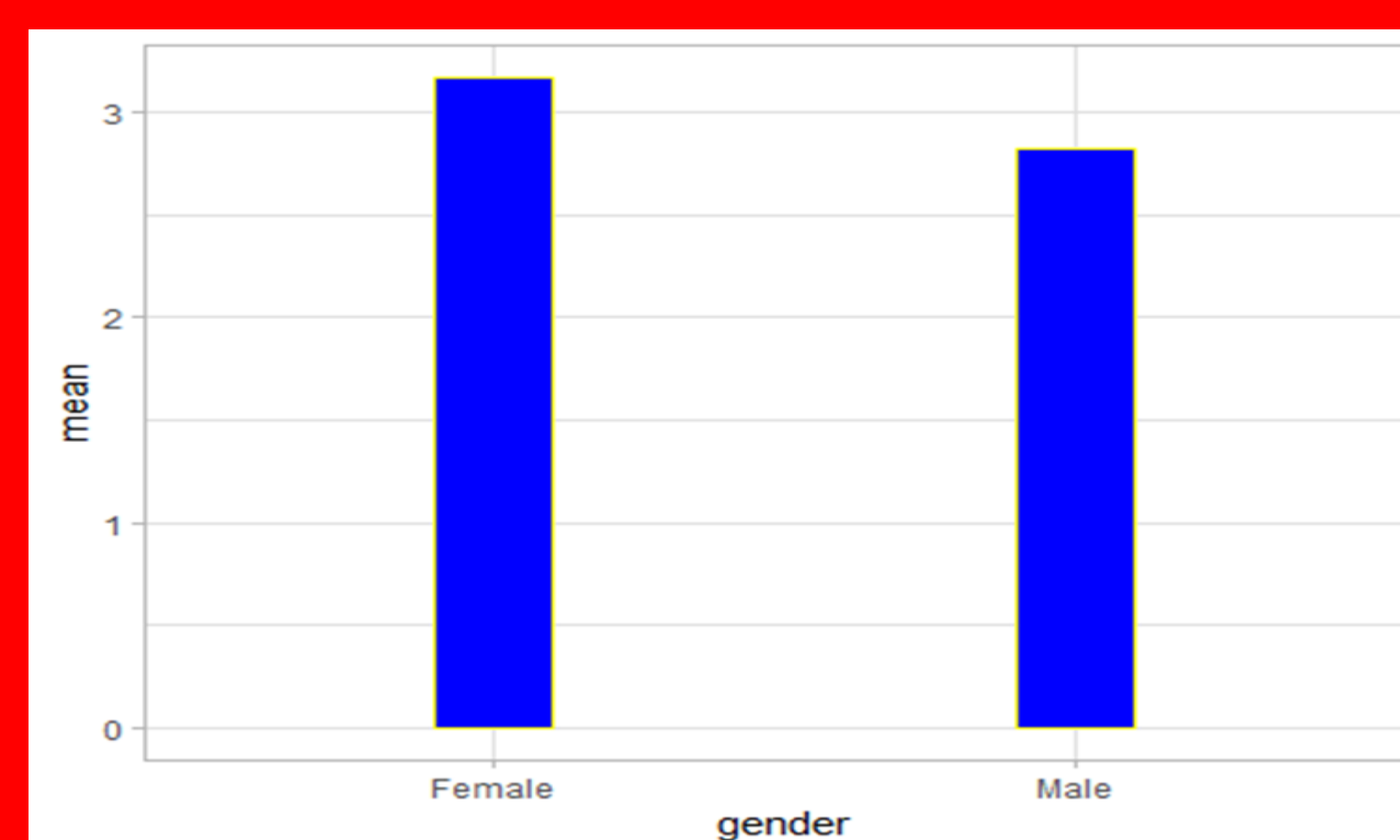
Most Frequently Implicated Domains
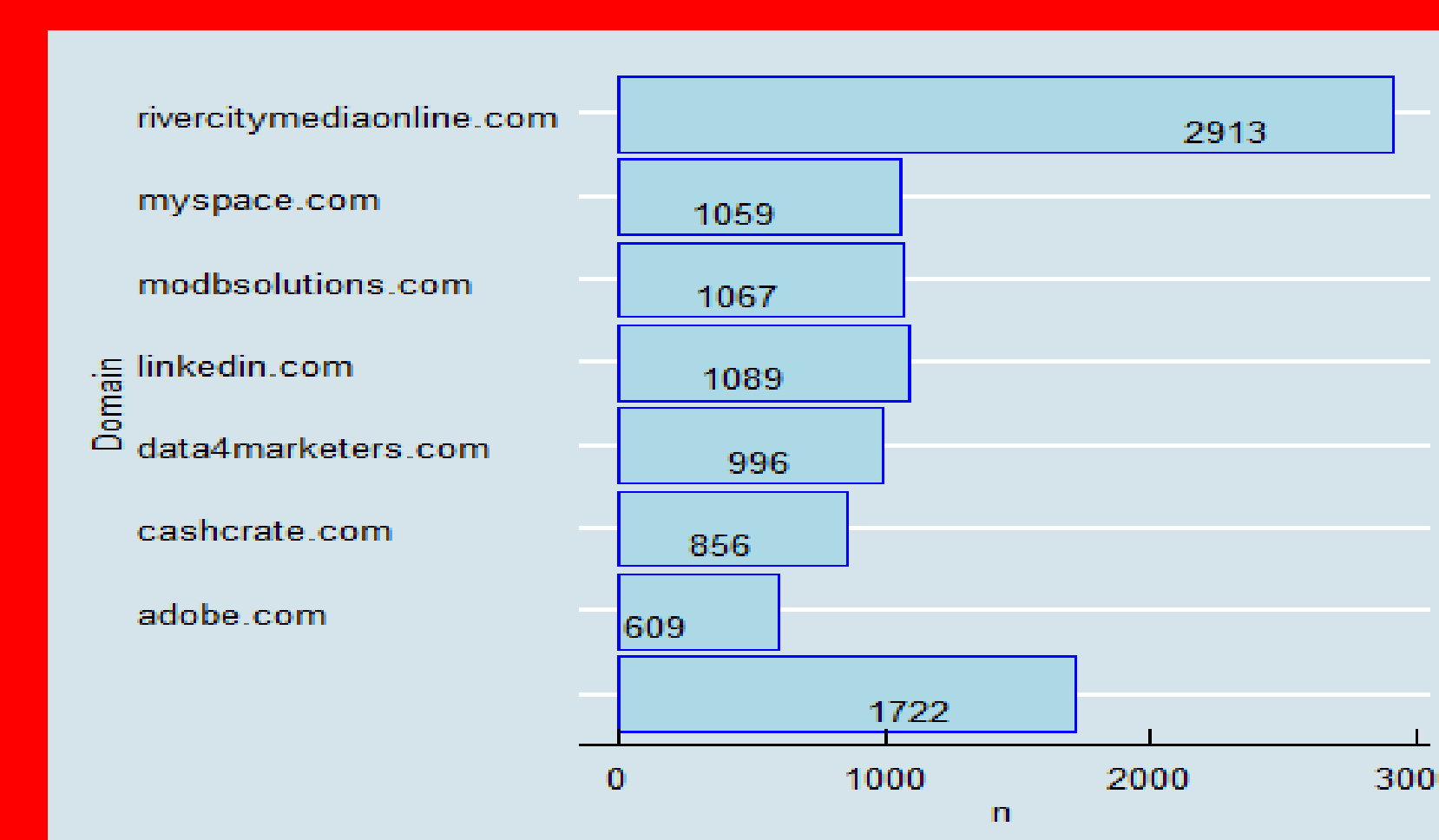
Relationship Between Education and no. of Breaches

Relationship Between Race and no. of Breaches

Conclusion

This Project not only provides me an opportunity to enhance my understanding of the major subject I plan to graduate in , but also is critical in the light of more and more commerce happening electronically. Unfortunately, due to sensitivity of data not much data is easily available for analysis and hence I have picked this research topic for my project to help create additionally analysis. Coming from third world nation, there is a personal perspective too. In India more and more electronic is picking up and through this project, want to analyze the account breach trend and be able to share this back in India as well.