**Data Analysis – You’re Now the Hacker**

# Introduction

Data Hackers Inc. specializes in identifying critical information from aggregate data sources, compiled via phishing attempts or directly from corporate data sources. The most recent heist includes a plethora of data from a popular online shopping platform. As a Principal Hacker, I have been tasked with identifying High Value Targets (HVTs) from the collected data. These are candidates that are best suited for identity and credit card theft, and their data will either be used internally by Data Hackers Inc., or sold to external bad actors.

# Methodology

HVTs, as defined in the introduction, are optimal candidates from the aggregate dataset for identity and credit card theft. Based on this principle, the following evaluation protocol was developed for identifying the theft potential of a given person:

1. Likelihood of existing credit or identity fraud.
   1. This was determined through a complex analysis of social security numbers in the data set. See the Analysis section for more information.
2. Spend-to-income ratio (12 month purchase history divided by the annual income).
   1. High spend-to-income ratio is an attribute of targets that would be unlikely to notice irregular spending.
3. Credit Card Expiration date.
   1. Only targets with valid credit cards will be considered.
4. Travel frequency.
   1. Targets that frequently travel will not receive fraudulent spending alerts from their credit card provider.
5. Active account status.
   1. Targets with active accounts are less likely to identify spending on the platform as fraudulent.

The target criteria, listed above, neglects some common factors for identity theft and scams. Factors such as age and income weren’t directly considered, as they are secondary to the factors above.

# Analysis

Present the findings and insights derived from your data analysis. Use tables, graphs, and other visualizations to highlight key results. Discuss trends, patterns, and any anomalies observed.

# Conclusion

Summarize the main outcomes of the study and their implications. Reflect on the significance of the analysis and suggest potential areas for further research or application.

# Appendix

Include any supplementary materials such as raw data, detailed calculations, or code snippets referenced in the paper. Ensure these materials are properly labeled and easy to navigate.

# Sources

<https://www.newsnationnow.com/business/your-money/american-credit-card-debt/>

<https://www.ftc.gov/news-events/data-visualizations/data-spotlight/2022/12/who-experiences-scams-story-all-ages>