

Charles Malon Silva Rocha - 2021376

**Data Exploration and Preparation**

**CA 1**

**Dublin**

**2024**

**CCT College Dublin**

**Assessment Cover Page**

*To be provided separately as a word doc for students to include with every submission.*

|  |  |
| --- | --- |
| **Module Title:** | Data Exploration & Preparation |
| **Assessment Title:** | CA 1 |
| **Lecturer's Name:** | Muhammad Iqbal |
| **Student Full Name:** | Charles Malon Silva Rocha |
| **Student Number:** | 2021376 |
| **Assessment Due Date:** | 1st October 2024 |
| **Date of Submission:** | 24th November 2024 |

**<https://github.com/CharlesMalonRocha/Data-Exploration-and-Preparation-CA1>**

**Declaration**

|  |
| --- |
| By submitting this assessment, I confirm that I have read the CCT policy on Academic Misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or other source. I declare it to be my own work and that all material from third parties has been appropriately referenced. I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution. |

Contents

[Introduction 4](#_Toc12908)

[Problem Domain and Dataset Preparation 5](#_Toc29540)

[Exploratory Data Analysis 6](#_Toc32272)

[Encoding and PCA 7](#_Toc15045)

[Poster Presentation 7](#_Toc24108)

[Video Presentation 7](#_Toc22608)

[Conclusion 8](#_Toc10761)

[References 9](#_Toc27877)

# Introduction

This assignment analyzes data on violence against women and girls to uncover patterns and trends that can inform prevention and policy efforts. By cleaning and exploring the data, we aim to understand factors linked to violence, identify unusual trends, and highlight areas that need attention. This project combines technical analysis with a socially important focus, aiming to contribute meaningful insights that support community safety and well-being. The findings will be presented through visuals and a summary poster.

# Problem Domain and Dataset Preparation: Violence against women

Violence against women and girls is a huge problem around the world, and it impacts society in so many ways. Looking at the data around this kind of violence really matters for a few reasons. First, it helps us see patterns and trends, which means policymakers, social services, and law enforcement can figure out where to put resources and how to prevent it better. By digging into details—like where this violence happens, who it affects most, and what types are common—we can start understanding why it’s happening and tackle it at the roots.

For this assignment, checking out the data will help us see important trends and factors, which can spark real conversations and lead to practical ideas. The analysis can spotlight areas that need attention right now and guide future research or efforts, which fits with global goals like the United Nations' Sustainable Development Goals, especially Goal 5 on Gender Equality and Goal 16 on Peace, Justice, and Strong Institutions.

Plus, violence against women has serious effects on health, the economy, and mental well-being for both individuals and communities. It can cause long-term health issues, lower productivity, and lead to higher social and medical costs. In a community, a lot of violence can create fear, weaken trust, and break down social bonds. Tackling these problems needs solid data to steer our actions, and analyzing it is a big step toward making our communities safer.

This data was retrieved from the following:

<https://www.kaggle.com/datasets/whenamancodes/violence-against-women-girls/data>

# Description of the Dataset

**Regarding Violence Against Girls and Women**

The goal of the Demographic and Health Surveys (DHS) Program is to improve knowledge of population trends and health in developing nations worldwide.

"One of the most pervasive, enduring, and catastrophic human rights violations in our world today," according to the UN, is violence against women and girls (VAWG). Because of the stigma, shame, quiet, and impunity surrounding it, it goes mostly undetected.

Generally speaking, it can take the following forms: sexual violence and harassment (rape, forced sexual acts, unwanted sexual advances, child sexual abuse, forced marriage, street harassment, stalking, cyber-harassment), intimate partner violence (battering, psychological abuse, marital rape, femicide), and human trafficking (slavery, sexual exploitation), child marriage or female genital mutilation

**About the Data**

The information came from a poll of men and women in South American, Asian, and African nations that examined their perspectives and the reasons they said were offered for harming women. The information also examines the respondents' various sociodemographic categories, such as age group, employment, marital status, and education level.

Therefore, in order to convey the message that violence against women and girls is never acceptable or justified, it is imperative that the nations where these viewpoints are prevalent give public awareness campaigns and educational opportunities for women and girls top priority.

|  |  |
| --- | --- |
| COLUMN | DEFINITION |
| Record Id | Numeric value unique to each question by country |
| Country | Country in which the survey was conducted |
| Gender | Whether the respondents were Male or Female |
| Demographics Question | Refers to the different types of demographic groupings used to segment respondents – marital status, education level, employment status, residence type, or age |
| Demographics Response | Refers to demographic segment into which the respondent falls (the age groupings are split into 15-24, 25-34, and 35-49) |
| Survey Year | Year in which the Demographic and Health Survey (DHS) took place. |
| Value | % of people surveyed in the relevant group who agree with the question (e.g. the percentage of women aged 15-24 in Afghanistan who agree that a husband is justified in hitting or beating his wife if she burns the food) |
| Questions | Respondents were asked if they agreed with the following statements:   * A husband is justified in hitting or beating his wife if she burns the food * A husband is justified in hitting or beating his wife if she argues with him * A husband is justified in hitting or beating his wife if she goes out without telling him * A husband is justified in hitting or beating his wife if she neglects the children * A husband is justified in hitting or beating his wife if she refuses to have sex with him * A husband is justified in hitting or beating his wife for at least one specific reason |

# Encoding and PCA

# Poster Presentation

# Video Presentation

# Conclusion

# References

Chauhan, A. (2022). *Violence Against Women & Girls*. [online] Kaggle.com. Available at: https://www.kaggle.com/datasets/whenamancodes/violence-against-women-girls/data [Accessed 11 Nov. 2024].

‌