Exploring People's Perceptions to the Ban on Women's Education in Afghanistan

Abdu Saboor Hamedi
abdul.hamedi@student.uiii.ac.id

Exploring People's Perceptions to the Ban on Women's Education in Afghanistan

By:

Abdul Saboor Hamedi

A thesis submitted in partial fulfillment of the requirement for the

degree of

Master Degree

In

Social Sciences

Indonesian International Islamic University

2023

Approved by

Chairperson of Supervisory Committee

Supervisors:

Supervisor 1   : Sirojudin Arif, Ph.D

Supervisor 2   : Nia Deliana, Ph.D

Table of Contents

List of Figures

List of Appendix

Abstract

This study, based on feminist theory, explores how different categories of people perceive hurdles to women's education. The study illustrates public opinion on this restriction and how it limits women's educational opportunities. This study examined tweets from Afghanistan's four most populous ethnic groups: Pashtuns, Tajiks, Uzbeks, and Hazaras. We have collected 56,820 tweets to find common themes and distinct perceptions of the restriction on women's education. Women respondents were more likely to disagree with the ban on female education in Afghanistan than their male counterparts. The study found there are places where gender disparity is more pronounced. Findings stress the need to remove barriers prohibiting Afghan women from receiving an education due to power dynamics and underlying gender inequities. The study sheds light on how the restriction on female education in Afghanistan has affected the country and how people feel about it. Furthermore, it stresses the significance of gender equality in education and the necessity of addressing power dynamics and underlying gender inequities that continue to restrict women's education.

Keywords: Girl, Women, Education, Afghanistan, Ban, School, University, Taliban

CHAPTER I: INTRODUCTION

Introduction

Without education, especially without equal educational experiences, skills, or qualifications, men, and women of specific classes and social groups have long been doomed to inferior lives in terms of their capacity to grow personally, their choice of employment, their standing as citizens, and their ability to influence leadership and decisions at the national level that have an impact on their local lives (Weiner, 1986).

Million women and girls practically prevent from receiving education around the globe, especially in countries with patriarchal and conservative government forms. Afghanistan is one of those countries where women suffered for decades. Today, women and girls are prevented from educating themselves, or the government has restricted their education system. This restriction has a substantial negative impact on women's daily activities. It limits their lives and keeps them from the crowd and the public.

Study shows that educated women are more likely to have more appotunities, "better educated women that do work are more likely to work in non-farm activities (Anderson, Reynolds, Biscaye, Patwardhan, & Schmidt, 2021, p. 200)." Education for women has been cited as an important factor in the development of feminism by a number of authors. Women's higher education institutions sprung up at the same time as the

first waves of feminism in the United States, as noted by (Plutzer, 1993, p. 149).

The denial of education for women can have severe consequences for society and individuals (Orfan, 2021). Taliban restrictions on women's education in Afghanistan have led to high illiteracy rates among Afghan girls and women, which has limited their participation in society (Mashwani, 2017). The Taliban's return has raised fears that women's rights, including access to education, will be further diminished (Women, 2022). Despite Taliban efforts to limit women's education, evidence shows that there is opposition among Afghans to this action. A survey by Asia Foundation 2021 showed that 87% believed women should have equal access to education (Rieger, 2019). Hodes' 2019 study found that Afghan women who have received an education are more supportive of gender equality and rights.

This study is necessary because it sheds light upon the ban's effects on women's education in Afghanistan and gives insight into Afghan society's attitudes towards this topic. We use a feminist framework to show the power dynamics and gender inequalities that lead to the ban's enforcement and implementation.

We have collected data from Twitter throughout the four largest Afghan ethnicities: Pashtuns, Tajiks, Uzbeks, and Hazaras. These four

ethnicities are the largest group in Afghanistan. We used a qualitative and quantitative mix-method to analyse tweets and identify patterns in the difference between people's attitudes regarding women's education restriction or ban. Further, this study has used python computer language and a machine learning approach (ML) to collect and analyse data. Python is one of the most efficient and compatible computer languages for analysing data and has thousands of libraries. In this study, we have used libraries such as Pandas, NumPy, scikit-learn, etc.

In addition, between August 15, 2021, and March 30, 2023, we collected tweets from 400 different accounts, for a total of over 500,000 tweets; this is raw data; to transform these tweets into information, there is a lot to accomplish, such as eliminating noise, duplicates, and null values.

A comprehensive literature analysis addresses the restriction on Afghan women regarding education. This study covers the historical and cultural backdrop, the prohibition's impact on women's lives, and the role that gender and ethnicity play in determining views regarding women's educational opportunities. After this, we offer our results and detail the process through which we gathered and evaluated data from Twitter. We analyse the ramifications of these results and provide suggestions for further study and policy changes to advance gender equality and women's rights in Afghanistan.

Research Problem

The current study investigates, people's perceptions of women's access to education in Afghanistan. Along with the factors that shape these perceptions, the barriers that women face during access to education, and the impact of gender on individual attitudes and behaviors toward women's education. Also, we seek to understand the complexity of social, cultural, and political factors that influences the perceptions of people regarding women's access to education with a focus on the effect of gender norms.

The sutdy exploreres the barriers that prevent women's access to education, including social norms, cultural practices, and structural challenges. Further, we explorer the role of "gender stereotypes", discrimination, and gender based violence. In nutshell, this study g through the people's perception regarding women's access education, and the impact of gender on individual attitudes toward women's education.

Research Question

There is only single question that this study carries, "To what extent does gender affect people's perception of women's access to education?" we investigate how gender influences individuals' perceptions on the educational opportunities available to women. This question acknowledges that individuals' attitudes and beliefs on particular topics, including

5

education for girls, can be significantly influenced by social and cultural influences to a significant degree.

This question investigates the role that gender plays in people's opinions of women's educational opportunities. It is intended to uncover any prejudices and discriminatory views. The purpose of this research question is to identify potential obstacles to the education of women that are connected to gender-based ideas and attitudes.

This question is an attempt to gain a deeper comprehension of the intricate dynamic that exists between gender, culture, and educational attainment. It focuses on the schooling opportunities available to Afghan women. This study will contribute to continuing conversations on gender equality and human rights. Also, it may provide insights that may be utilised to inform programs and policies that encourage women's empowerment and education.

Research Objective

Researchers and experts have been focused on women's subjects regarding education equality, politics and other issues that distinguish women from men or make men superior. In contrast, women are inferior, and women's issues have dramatically increased in recent years. When it comes to education, Afghan women suppressed and confronted numerous

challenges, particularly during the Taliban. Despite so many attempts by other countries, the Afghan government, and local authorities, all failed.

This study examines the effect of gender on perceptions of Afghanistan's women's education access. The study will examine the views and experiences of Afghans, with particular attention to the role of gender in shaping their opinions and how gender influences people's perceptions of women's education access. The study focuses on the factors that affect attitudes toward women's education in Afghanistan. These factors include cultural, economic, and social aspects.

The study will help better understand Afghan women's difficulties when accessing education. It will also examine the relationship between gender perceptions and people's perceptions. The study promotes gender equality, improves women's education access in Afghanistan, and helps promote human rights and social justice in Afghanistan.

Significant of study

In many places around the globe, the ban on women's education remains a problem. It has long-lasting consequences for women's empowerment and gender equality. This study is critical because it helps us understand women's attitudes to the ban on women's education and the role education plays in promoting gender equality.

The first finding, women are more likely to disagree with the ban on women's education than men. It is important because women appreciate education values, stand for education, and confront cultures and social norms to achieve education.

The second finding, women are more likely to oppose the ban on women's education in different cultures. It is significant as it shows that the same factors affect women's attitudes to education regardless of their cultural backgrounds. These factors could include the desire to achieve personal fulfilment and self-actualization, recognition of the value of education for their children, and belief that education is an essential human right.

The third finding is that education helps women become aware of their society and stand against inequalities; education empowers women to achieve gender quality. Research regarding women's education has shown that education improves women's political, economic, and social status.

Study shows that women value education and highlights the importance of ensuring all women have equal right regardless of political position or educational area. According to the study, efforts to promote gender equality in education and remove barriers to women's education should consider women's views and opinions. Programs and policies that seek to increase women's education access should be sensitive to women's

preferences and cultural contexts. This study highlights the importance of awareness-raising and advocacy campaigns to challenge cultural attitudes and practices behind the ban on women's education.

Also, it provides valuable insight into women's attitudes to the ban on women's education and the role education plays in gender equality. The study shows that women value education and that efforts must be made to remove barriers to women's education as well as promote gender equality and help women reach their full potential.

CHAPTER II: THEORETICAL FOUNDATION

Literature Review

The status of women in any society indicates the progress of that society and the advancement of the countries in the group it is the purposeful use of all human forces and talents, including women. Women plays a prominent role in economic activities and have social and to achieve sustainable development, pay more attention to women who are side by side and men's counterparts have active participation in economic and social

activities (Jütting, 2004, p. 10). Participating human being is considered the center of the development process and endogenous development in the agenda is placed in such development, participation is a fundamental and endogenous variable Being fair calls women to participate as half of society

In this regarding it is important for all human being to access education especially for women, as Aaron says, "Human rights education must be seen as a worldwide movement (Aaron, Braslavsky, & Truong, 2007)" One of our most fundamental freedoms should be the opportunity to acquire the knowledge and skills necessary to thrive as contributing members of society. Yet, cultural and religious beliefs that reinforce gender roles and restrict women's mobility have acted as barriers to women's access to education in Afghanistan (Islam, 2021, p. 1).

The restriction on women's education by the Taliban has only made the situation worse. Reviewing the literature on Afghan women's education is essential to help contextualize the research.

Afghanistan may be the only country where women's rights have made and broken kings and politicians in the last century (Ahmed-Ghosh, 2003, p. 1). Women have been oppressed throughout history, Mujahideen (1992-1996)   (Fluri J. L., 2009, p. 260) Thus, one must analyze Afghanistan's women's situation within the larger historical context of Afghanistan, not through the ideological formulation of "before and after"

the Taliban. Only such a perspective can ensure that women are seen as essential to rebuilding Afghanistan.

Deniz Kandiyoti (2005) says regarding women's education there has been number contentious for several years with number of barriers preventing women from education. The primary research regarding the author is the frequent war for decades (Yapp, 2001) that dramatically impact on women's education. Regarding Deniz apart from conflict Afghanistan is a traditional country, and in many parts of the country women are expected to stay at home and take care children and not pursue education or work, and women education often seen as threat to traditional gender roles and values (Kandiyoti, 2005).

Karlsson and Mansory (2008) regarding these two respectful authors, Afghanistan is a country which has the lowest enrolment of education and "adult literacy". The author says, Afghanistan has long tradition Islamic education (Karlsson & Mansory, 2008). Also, Roozbeh Shirazi (2008) confirms that for seventh century Islam strong impact on social norms and political activities, and "Islam served strong basis for Afghan cultures identity and servers as a powerful reference point for Afghan social mores rights, and obligations regardless of ethnicity (Shirazi, 2008, p. 212)."

The elimination of women from education and limiting the public sphere began during the Soviet Union and continues until today. The violence against women drastically increased when the Taliban seized power. Women have always made sacrifices in the country. Afghan women's status has changed significantly over the past four decades due to instability and fragility. Many competing local ideologies are within its borders, including liberal, moderate, and radical Islam. Also, a diverse population under-educated rural Afghans (who account for three-quarters) and educated, wealthy urban residents (Bank, 2016). The central state's inability to find an equilibrium to meet the needs of these diverse groups has led to the ongoing conflicts in Afghanistan.

Jennifer L. Fluri (2008) social and cultural views towards women's education, a lack of financial resources, security concerns, and a lack of available educational opportunities are all discussed in this study as obstacles to women's education in Afghanistan. According to the authors, these challenges are magnified for women seeking a postsecondary degree since they encounter more roadblocks than their elementary and secondary school counterparts (Fluri J. L., 2008). The effects of the Taliban's policies on Afghan women's access to education are also discussed. Women were not allowed to receive an education under the Taliban's control, as stated by the writers. Even though the restriction was repealed following the Taliban's defeat in 2001, the authors maintain that traditional societal views in rural

regions of Afghanistan have a lasting impact on women's access to higher education.

Zubeda Jalazai and David Jefferess (2011) have compiled a collection of essays on the intersections between gender, education, and development in Afghanistan. The book contains several chapters discussing the difficulties girls and women face in Afghanistan regarding accessing education. They also highlight how political instability, cultural attitudes, and poverty have made it difficult for girls and women to obtain educational opportunities in Afghanistan (Jefferess, 2011). Education opportunities for girls in Afghanistan have been severely affected by political instability and conflict. From 1996 to 2001, the Taliban regime in Afghanistan banned girls from going to school (Telesetskyt, 1998). Even after the Taliban's fall, violence and instability continue threatening girls' education. Insurgent groups have targeted schools, and many families fear their daughters will be kidnapped or killed (Yousufi, 2021).

Further, Hadi Ahmad (2022), during their rule in Afghanistan from 1996 to 2001, the Taliban prohibited education for women. The Taliban's interpretations of Islamic law meant that women were to remain at home and not be educated. Taliban viewed women's education as a threat and believed educated women would challenge their authority (Ahmadi, 2022). Further, the author argues that the Taliban are against education.

David J. Roof (2015) In the last two decades, education in Afghanistan has seen significant changes, with the establishment of new universities, expansion of existing institutions, and increased number of female students. This paper examines the problems facing Afghanistan's higher education system, which include a lack of resources, limited technology access, inadequate infrastructure, and security concerns. According to the authors, these challenges have negatively affected student outcomes and quality education (Roof, 2015).

The paper highlights some positive developments in Afghanistan's higher education. The authors point out that there are now over forty universities in Afghanistan. Also, female students have increased significantly, from almost none in the Taliban era to 40% in some institutions. While there have been positive developments in Afghanistan's higher education system, the paper states that many challenges exist to improving education quality and expanding access. Increasing infrastructure investment and ongoing security efforts are necessary to continue the positive developments in Afghanistan's higher education.

Herz and Sperling (2004) discovered that educated women could have better health and social well-being. It could lead to economic growth (Gene Sperling, 2015). Also, girls' education can help reduce poverty, improve health outcomes, and promote gender equality. Afghan women

have faced significant obstacles in accessing education despite the many benefits of education. The Taliban's ban on women's education significantly impacted Afghanistan's educational opportunities. Girls not allowed to attend school due to the prohibition of women's education had a considerably lower attendance rate.

Valentine M. Moghadam (2003) addresses the significance of education in fostering gender equality in the Middle East. She contends that education is essential for women's empowerment and for questioning established gender roles and prejudices. Moghadam observes that women's access to education has increased dramatically in many Middle Eastern countries in recent decades, positively impacting women's lives. She argues that education may give women the skills and information they need to enter the labor field, engage in public life, and oppose gender discrimination (Moghadam, 2003).

Theoretical Framework

Utilizing a feminist lens enables the exploration of how individuals interact with institutions and may provide suggestions for addressing and dismantling oppressive structures and systems. In addition to focusing on oppression, feminist theory considers the lived experiences of any individual or group, not just women. Disrupting oppression is a fundamental tenet of feminist work, although there may not be agreement on where

feminist thought fits as a theory or paradigm. As hooks (2000) says, "Feminism is a movement to end sexism, sexist exploitation, and oppression (Hooks, 2000, p. 5)."

Marxism and socialism are important pillars of feminist theory's history. Feminist theory has its roots in the 18th century and developed through the equality movements of the 1970s and 1980s. Based on Clare Burton (2013), "Origin of the Family and commentaries on it were central texts to the feminist movement in its early years because of the felt need to understand the origins and subsequent development of the subordination of the female sex (Burton, 2014, p. 2)." Research on gender equality is one aspect of current work in feminist theory.

Today gender inequality is still controversial topic around the globe, especially in the field of women's education and schooling, for instance, Kate Pincock (2018) discussing the issues of women education in Tanzania, Pincock says, "considerations of what empowerment looks like in relation to one's sexuality are particularly important in relation to schooling for teenage girls as a route to expanding their agency (Pincock, 2018, p. 909)." This is not a line of research limited to the oppression of simply female students, but rather may be extended to any oppressed group within an educational system. For instance, non-binary students experience discrimination in educational systems, and even male students may

encounter difficulties. Despite this, students are frequently encouraged to pursue what are deemed "gender acceptable" fields of study. As a result, an oppressive system is created that calls for deliberate effort to overthrow.

Education is another aspect of feminist's inquiry research Jennifer Earles (2016) centered on using physical classroom settings to study relationships, "between gendered literary characters and stories and the normative and marginal responses produced by children (Earles, 2016, p. 369)." According to the author, educators should support the ongoing eradication of oppressive gender norms by carefully choosing the books they use and creating settings that provide students the chance to have meaningful conversations about these contradictions.

Sexuality, gender, race, prejudice, discrimination, equality, diversity, and choice are the fundamental ideas of feminist thought. There are mechanisms and processes in place that act against people based on these attributes as well as against equality and equity. Critical paradigm research necessitates the conviction that truths can be discovered by exploring these conditions already present in the current social order. What's more significant, though, is that this investigation can promote awareness of repressive structures and provide up platforms for different perspectives to be heard (Egbert & Sanden, 2019).

Feminist theory is a crucial theoretical framework for understanding and overcoming the challenges that women experience in education. It offers a perspective through which to examine the structural discrimination against women that persists in academic settings. Despite being a basic human right, girl's and women's access to education is often thwarted for a variety of reasons, including but not limited to cultural norms, financial constraints, preconceived notions about what they should learn, and outright prejudice. Feminist theory provides a scathing analysis of the unequal power relations and systems that keep these barriers to education in place.

Feminist theory relies heavily on the idea of "intersectionality" to explain patterns of oppression. Many other elements, including but not limited to ethnicity, class, sexual orientation, ability, and nationality, impact women's experiences of oppression. Feminist theory, therefore, provides a comprehensive comprehension of the various and often overlapping obstacles women confront in pursuing education. When it comes to educational opportunities, women may suffer both gender and racial discrimination. Furthermore, feminists' theory sheds light on the needs and challenges of patriarchal structures and ideologies in educational systems in some countries such as Afghanistan. Women are given fewer chances to advance their education because patriarchy promotes and reinforces sexist views about women's inherent inferiority to males.

Hypothesis

H1: Women are more likely to disagree with the ban on woman's education. The right to education is one of the fundamental human rights. Traditional societies often deny girls the opportunity to reach their full potential. Their education is also affected by school fees and safety concerns.

These barriers are slowly being removed due to the overwhelming evidence that educated girls can benefit society and the international focus on universal schooling. It's helpful to encourage girls to pursue careers in science and technology, which have a majority of men. But this is still a problem. While there have been some improvements in sending girls to secondary and primary schools, less than half are enrolled in university. In this regard, feminists further highlight the rhetoric of equal opportunities is almost the sine qua nu by which liberal-feminist perspectives are recognized "Equal means the same. (Sandra, 1987)"

Education is considered a basic necessity and right for the citizens of any nation, which is a powerful tool to reduce inequalities in society, especially for women who face discrimination in many fields and have a particular need for this. Education is considered a turning point in empowering women because it enables them to face challenges and change their lives. Still, many women in our country are illiterate, backward, weak,

and exploited, which can be changed with proper education. Education is one of the ways to spread the message of empowering women.

No goals and aspirations of the citizens of a country can be achieved without effective education. Education educates the individual and helps them understand that women are essential to society. Career achievement, self-awareness, and satisfaction are many things that are guaranteed by the effective use of training. Guidance and counseling also help women choose careers and follow career paths through training. Education helps women become empowered through science and technology and face the challenges of today's technological age.

Education educates women and enables them to make decisions and take responsibility in their homes and the outside world. Education is a goal and a means to achieve other desirable destinations. Based on the information above, women are more likely to disagree with a ban on women's education: Education is a tool for empowerment, and educated women are more likely to be aware of their rights and empowered to take action to improve their lives. Social change: Women who receive an education are more likely to challenge traditional gender roles and stereotypes. Education can lead to broader social changes that benefit women, such as greater participation in the workforce and increased political representation. Economic benefits: Education can lead to higher-

paying jobs and better economic opportunities for women. Also, it helps women to achieve greater financial independence and improve their quality of life. Health benefits: Educated women are more likely to have better health outcomes for themselves and their families. They are more likely to seek medical care and make healthier lifestyle choices.

Methodology

This study investigates Afghan people's perceptions regarding women's ban on education. Afghan women have been suffering for a long time with various issues, such as gathering in public, voting, freedom of speech, more importantly, education; this study focuses on education. Education is one of the essential elements of our daily lives; without education, human beings are blind that cannot see anything. Afghan women have been limited from gaining education or banned from attending school. We use mixed-methods research to address a research issue by combining quantitative and qualitative techniques.

While quantitative researchers focus on numbers, qualitative researchers examine text and image data (such as transcripts, field notes from observations, journals, and photographs). There is some numerical analysis done in qualitative research (for example, when ethnographers or case study authors offer a descriptive table of demographic information), and there is also some numerical analysis done in quantitative research (for

example, when text data are translated into frequencies), but these two types of data are treated differently. Also different are the methods used to analyse the data, which in quantitative studies may include things like descriptive and inferential statistics as well as checks of effect sizes, and in qualitative studies may involve things like coding text segments, developing themes or categories from the codes, and establishing relationships between the themes (as in grounded theory or a narrative study) (Swanson & Holton III, 2005, p. 317).

More information can be gleaned via a mixed-methods study than from a quantitative or qualitative approach alone (Creswell, 1999, p. 455). Comprehension is the second aim of mixed methods research, which integrates qualitative and quantitative techniques to comprehend better the phenomenon being studied or account for discrepancies in the data. Some see understanding as speeding up the learning process for a researcher (Morse, 2003, p. 191). Mixed-method studies have numerous potential advantages but also some potential limitations. One advantage is that it boosts a researcher's faith in their results. By drawing parallels between different aspects of the phenomenon being studied, new insights may be gained or old ones revised. (Dunning, Williams, Abonyi, & Crooks, 2006, p. 147).

Quantitative research	Qualitative research
The goal is to try to explain the observed phenomena by classifying features, counting them, and building statistical models.	The intent is to provide a comprehensive analysis.
The researcher has a precise goal in mind before beginning the search.	The researcher may have only a vague idea of what needs to be found going in.
Suggested for use in the concluding stages of scientific investigations.	Advised for use in the preliminary stages of studies.
Before any data is collected, every facet of the study is meticulously planned.	The pattern develops as the research progresses.
Quantitative information is gathered using questionnaires and other research instruments.	The researcher acts as a tool for collecting information.
Information is represented numerically and statistically.	Words, images, and physical artefacts all constitute data.
Attempts to acquire an accurate read on the issues at hand by employing techniques like surveys and questionnaires.	Individuals' perceptions of events are given weight, therefore methods like participant observation and in-depth interviews are used.
While quantitative data saves time and allows for hypothesis testing, it sometimes lacks nuance because of its lack of context.	Qualitative information is more detailed, time-consuming, and limited in its applicability.
In most cases, the researcher keeps an objective distance from the topic being studied.	The researcher often develops a personal interest in the study's topic.

Table 1.Quantitative vs Qualitative, *(Castellan, 2010, p. 75)*

Argument

We must address the issue of women's education right away since it's so important. Every woman has the right to obtain an education, a fundamental human right. Unfortunately, this right is denied to many women worldwide, and it is time to alter this.

Traditional civilizations that uphold antiquated beliefs and traditions are one of the main causes of women being refused education. Many of these civilizations discourage girls from going to school and expect them to stay home and take care of the family. It's essential to tackle these attitudes and behaviors because they significantly impede women's education.

School fees are a significant barrier to women's education. It is especially true for females since many families cannot afford to pay for their children's education. As a result, many girls are compelled to leave school, reducing their education length, regardless of a person's gender or financial situation.

Education is a fundamental human right and an effective tool for decreasing inequality and empowering women. But numerous women all over the world experience prejudice in a variety of settings and have restricted access to education. This thesis contends that as education is vital for empowering women, bringing about social change, fostering the

economy, and promoting women's health, women are more inclined to disagree with a ban on women's education.

Education is a potent weapon for the empowerment of women. Women who have had an education are more likely to know their rights and feel empowered to take action to better their life. Women who pursue education get the knowledge, abilities, and confidence they need to exercise their agency and make informed decisions. Women who seek education are more likely to question prevailing gender norms and prejudices, which may result in broader social changes that are advantageous to women, such as increasing employment rates and political representation.

CHAPTER III: RESEARCH DESIGN

The Place and Time of the Study

1- The study explores people's perceptions regarding women's education and whether women agree with the ban on women's education.

2- The study occurred on January 01 2023, and data collection began on December 01 2022, until March 30 2023. The data is based on Afghan ethnicity, more specifically, the four largest ethnics, which are Pashtuns, Tajiks, Hazaras, and Uzbeks

Research Material and Tools

Here are the primary items that we needed for our research:

- Python is a high-level programming language for web development, data analysis, artificial intelligence and scientific computing. It is an open-source language with many developers contributing to its development.

- Pandas is a Python library that is used for data manipulation and analysis. It provides data structures for efficiently storing and manipulating large datasets. It also includes data cleaning, filtering, aggregation, and visualization functions.

- Natural Language Processing (NLP) is a subfield of artificial intelligence that focuses on the interaction between computers and humans using natural language. It involves using algorithms to analyses, understand and generate human language. NLP is used in various applications such as chatbots, sentiment analysis, speech recognition and machine translation.

- Visual Studio Code (VSCode) is a source code editor that Microsoft develops. It provides debugging, syntax highlighting, code completion and Git integration. VSCode supports various programming languages, including Python, and provides extensions for multiple libraries, such as Pandas.

Data Analysis

The analysis of data is a crucial aspect of any study because it reveals patterns in the actions of a sample or population. We have gathered 179,122 rows of data from the four most populous groups in Afghanistan: Pashtun, Tajik, Uzbek, and Hazara. Tweets was gathered over three months, beginning on January 1, 2023, and ended on March 30, 2023. The data collection procedure began with a thorough cleaning and data organization phase.

As part of this process, we eliminated any data that was redundant, incorrect, or otherwise problematic. After this was completed, the data could be analyzed in greater depth. The next step was to dive into the data and see what insights could be gleaned from it. This process involved analyzing the demographics and content of tweets from all four groups of people. Descriptive data were used in this analysis, including means and standard deviations. The data was ready for additional analysis when the descriptive analysis was finished.

This involved looking for patterns and studying the connections between the various ethnic groups. This was accomplished with the use of inferential statistics like regressions and correlations. The findings were then presented straightforwardly and succinctly. The results were illustrated with charts and tables for clear comprehension. The goal was to give a

thorough summary of the data and any conclusions that could be drawn from it.

Overall, the data analysis method shed light on the four biggest ethnic groups in Afghanistan and their patterns of behavior over the course of the study's three-month duration. Interesting trends and patterns, as well as connections between the various ethnic groups, were uncovered by the investigation.

Data Cleaning

It's the procedure of finding wrong information, fixing it by eliminating unnecessary bits, and re-entering the right ones. Data cleansing include eliminating mistakes and verifying information. Cross-checking data is a viable option for fixing this problem. Problems usually disappear after data is checked for accuracy.

Model Building

The goal of model building is to define the nature of the relationship between the variables through the use of statistical or machine learning models. In this scenario, the algorithm tries to identify tweets that are in favor of education from dataset as well as recognize the neutral tweets. The program evaluates user sentiment from tweets using a binary score, revealing interesting information about how individuals feel about school.

Research Management

The study challenge at hand is investigating Afghan people's perceptions toward women's education. The study aims to find how gender affects people's perceptions women's access to education. The hypothesis states that women are more likely than men to reject banning women from pursuing education. The research methodology employs both qualitative and quantitative methodologies in a mixed-methods approach.

The data collection was done through Twitter from four largest Afghan ethnicities, namely: Pashtuns, Tajiks, Hazaras, and Uzbeks using machine approach to analyze the data. Using different technics and various libraries, such Pandas, numpy, matplotlib, seaborn and so on. The data collected during January 1st 2023 continued until March 30th 2023.

We have collected tweets from 263 different users the largest tweets come from Pashtuns followed by Tajik, Hazara, and Uzbek. The amount of data at first place was 179,121 after removing noises, duplicates, null data, and other unnecessaries data we have received 56,821 using different keywords such as: women, girl, education, school, ban, university, and Taliban. Keywords helped us to get more required data from the large dataset.

We have labeled our dataset using machine learning approach, we have put certain keywords as mentioned above, if any of those words found

in tweet we label them as 1 if not 0, in this case 1 means positive and 0 means neutral. The study employed cross-tabulation analysis to examine how gender and ethnic designations on tweets differed. According to the data, 96–98% of tweets from people of all other genders and ethnicities supported education, compared to 95% of Pashtun men who tweeted. These findings suggest that gender may influence attitudes toward education among specific ethnic groups and emphasize the need for targeted efforts to reduce gender disparities in education. Overall, the data has provided us the solid answer regarding education and the disagreement of female on education.

Independent Variable Gender: Gender refers to person identifies as male or females or another gender identity. In this study gender refers as independent variable, because we are interested how gender affects people's perceptions of women's education access to education. The dependent variable in this study is education, we are willing whether male and female have different perceptions regarding access education. By studying the relationship between these two variables we can have insights how gender impacts people's perceptions and beliefs regarding education for women.

CHAPTER IV: RESEARCH RESULTS AND DISCUSSIONS

Research Results

We have collected 56,820 tweets from different Afghan ethnicity; the tweets we collected are between January and March of 2023. We have seen that the most significant number of tweets come from Pashtuns, which consists of 51%; the second largest tweets come from Tajiks which is 37%,

followed by Hazara 7% and the last Uzbek, which makes 5% of the tweets

in our dataset.

Pashtun	51%
Tajik	37%
Hazara	7%
Uzbek	5%

Table 2. Largest Tweet

We also found that most tweets are in favor of education denoted as

(label 1), with very few tweets falling into the neutral category (label 0). An

overall accuracy of 98% was achieved by the label prediction model we

trained with a precision of 0.99 and a recall of 1.00 for label 1. We also used

cross-tabulation analysis to look at how labels on tweets varied by ethnic

and gender.

Metric	Value
Accuracy	0.98
Precision	0.84 (0) / 0.99 (1)
Recall	0.66 (0) / 1.00 (1)
F1-score	0.74 (0) / 0.99 (1)

Table 3.Logistic Regression Result

Our data shows that women of all ethnicities are more vocal on

Twitter about the importance of education than men. Meanwhile 95% of

Pashtun male tweets were in support of education, while 96-98% of tweets

were in favor of education across all other gender-ethnicity groupings.

These findings point to the necessity for focused interventions to overcome

gender differences in education and imply that gender may play a role in determining attitudes toward schooling among particular ethnic groups.

Our results highlight the value of gender-sensitive policies and initiatives to overcome educational inequality in Afghanistan. Aside from that, the findings imply that gender and ethnicity may influence how some groups see education. Policymakers and educators wishing to increase educational opportunities and support inclusive growth in Afghanistan will find this information crucial.

Meanwhile, the largest amount of our dataset talks regarding women's education in Afghanistan; these tweets consist of various issues. Mostly the tweets blame the Taliban because they have banned women from education at the moment in Afghanistan. As the table below shows, the largest tweets are related to education, while a few tweets are not talking about education; we call those tweets neutral, denoted as zero (0).

| Positive Tweet | 94.56% |
| Neutral | 3.44% |

Table 4. Tweet Values

It's important to note that a dataset must consist of two values, 0-1; otherwise, we would encounter an error during data analysis. Thus we have neural data in our dataset.

Discussion

Our study sheds light on people's perceptions regarding women's access to education and the role of gender in shaping those perceptions. Those perceptions come from the four largest Afghan ethnicities: Pashtuns, Tajik, Hazaras, and Uzbeks. The data has been collected from Twitter based on education through several keywords using a machine learning approach. Most tweets come from Pashtun ethnic, with 51% followed by Tajiks, which consists of 31% Hazara, and Uzbek, which makes 7% and 5%; the finding shows that Pashtuns are essentially a presence on the Twitter platform.

Regarding the people's perceptions towards women's education, we found that most tweets supported education, with very few neutral tweets. The finding shows that significant tweets regarding women's education that helps women to proceed with their educations, despite cultural challenges.

Interestingly, our data showed that women from the four largest ethnicities are more willing to tweet about education than men. Also, the data showed that several barriers and challenges prevent Afghan women from access to education, such as extremist ideologies. A need for focused interventions to overcome gender inequalities in education may be indicated by the fact that women are more vociferous on Twitter than males about the value of education. Furthermore, our results stress the significance of

incorporating multiple viewpoints and voices into Afghanistan's education policy and practice dialogues. To put our results in perspective, we looked at how they stacked up against international studies of the same kind. Notwithstanding Afghanistan's unique cultural, political, and economic circumstances, our research finds that global support for women's education is growing. This indicates that there is cause for optimism and hope despite the persistent obstacles women in Afghanistan face in gaining access to education.

In conclusion, our research sheds light on how different Afghan ethnic groups view women's education and how gender plays a role in influencing those perspectives. Our research shows the necessity of engaging several viewpoints and voices in debates about education policy and practice in Afghanistan and the need for specific initiatives to address the gender gap in schooling. Help create a more fair and just society in Afghanistan by building on these insights and taking action to alleviate women's difficulties in accessing education.

CHAPTER V: CONCLUSION AND SUGGESTIONS

Conclusion

This study aimed to examine how Afghans think about the country's educational prohibition for women. In other words, what are people's perceptions regarding women's access to education? We utilized a mixed-methods strategy that included both quantitative and qualitative techniques to get a complete picture of the problem. Twitter users from Afghanistan's four largest ethnic groups Pashtuns, Tajiks, Hazaras, and Uzbeks provided the data. Descriptive statistics were used for the quantitative data, as well as machine learning methods were employed to analyze the data.

According to the finding of the study, we can say that women are more likely to disagree with the ban on women's education in the country.

The result shows that women are more regularly tweets in support of education, and more likely stands with education and other women which has been banned from going to school and university. Also regarding the result of the study, education is essential for developing country and a nation.

Similarly, we found that women's and men's attitudes towards women's education are different, women are more supportive of education while men are not as supportive as women, further the study highlights the importantance of promoting gender equality and empowering women in the country. The research, the Pashtun ethnic produced the most influential tweets, followed by the Tajiks, the Hazaras, and the Uzbeks. This finding accords with Afghanistan's ethnic composition, where the Pashtun people predominate. Women of all ethnic backgrounds are more likely to tweet on the support of education than men, according to the study. This finding may indicate that women in Afghanistan are more likely to be vocal advocates for women's education than men.

The trained label prediction model achieved an overall accuracy of 98%, with a precision of 0.99 and a recall of 1.00 for label 1. This finding suggests that machine learning algorithms can effectively analyze large datasets and provide accurate results. The study also used cross-tabulation analysis to examine how labels on tweets varied by ethnicity and gender.

The findings of this analysis point to the necessity for focused interventions to overcome gender differences in education and imply that gender may play a role in determining attitudes toward schooling among particular ethnic groups.

In conclusion, the study's results provide light on how Afghans view the government's decision to outlaw girls' access to higher education. The study's findings have substantial implications for anyone working to improve education and gender equality in Afghanistan at all government and civil society levels. The study's findings inform the following recommendations:

References

Aaron, B., Braslavsky, C., & Truong, N. (2007). *School Knowledge in Comparative and Historical Perspective Changing Curricula in Primary and Secondary Education.* New York City: Springer.

Ahmadi, H. (2022). Higher education of Afghanistan under the Taliban rule-Review and analysis of past and current impacts.

Ahmed-Ghosh, H. (2003). A History of Women in Af omen in Afghanistan: Lessons Learnt for the F ghanistan: Lessons Learnt for the Future or Yesterdays and T ys and Tomorrow: Women in Af omen in Afghanistan. *Journal of International Wowen's Studies*, 1.

Bank, W. (2016). *World development report 2016: Digital dividends.* Washington, DC 20433: World Bank Publications.

Burton, C. (2014). *Subordination: Feminism and social theory.* Simultaneously published in the USA and Canada: Routledge.

Castellan, C. M. (2010). Quantitative and qualitative research: A view for clarity. *International journal of education*, 75.

Creswell, J. W. (1999). Mixed-method Research: Introduction and Application. *In Handbook of educational policy, pp. 455-472. Academic press*, 455.

Dunning, H., Williams, A., Abonyi, S., & Crooks, V. (2006). A Mixed Method Approach to Quality of Life Research: A Case Study Approach. *Springer*, 147.

Earles, J. (2016). Reading gender: a feminist, queer approach to children's literature and children's discursive agency. *Gender and Education*, 369.

Egbert, J., & Sanden, S. (2019). *Foundations of Education Research.* New York: Routledge.

Fluri, J. L. (2008). feminist-nation building in Afghanistan: an examination of the Revolutionary Association of the Women of Afghanistan (RAWA). *Feminist Review*.

Fluri, J. L. (2009). Geopolitics of gender and violence from below. *Political Geography*, 260.

Gene Sperling, a. R. (2015). *What Works in Girls' Education: Evidence for the World's Best Investment.* Massachusetts Avenue, N.W., Washington, D.C. 20036: The Brookings Institution.

Hooks, B. (2000). *Feminism is for everybody: Passionate politics.* South End Press.

Islam, J. M. (2021). Gender Equality Practices by the Teachers inside the Classrooms: A
        Comparative Study between a Government and a NonGovernment Primary
        School. *BRAC Institute of Educational Development BRAC University*.

Jefferess, Z. J. (2011). *Globalizing Afghanistan: Terrorism, War, and the Rhetoric of
        Nation Building.* Durham, North Carolina, United States.: Duke University Press.

Jütting, C. M. (2004). The impact of social institutions on the economic role of women in
        developing countries. *Oecd Development Centre*, 10.

Kandiyoti, D. (2005). The Politics of Gender and Reconstruction in Afghanistan. *United
        Nations Research Institute For Social Developmen*, 2,18.

Karlsson, P., & Mansory, A. (2008). Islamic and Modern Education in Afghanistan -
        Conflictual or Complementary. *Institute of International Education Stockholm
        University*.

Mashwani, H. U. (2017). Female Education in Afghanistan: Opportunities and
        Challenges. *International Journal for Innovative Research in Multidisciplinary
        Field*.

Moghadam, a. M. (2003). Modernizing Women: Gender and Social Change in the Middle
        East. *Lynne Rienner Publishers*.

Morse, J. M. (2003). *Handbook of Mixed Methods in Social & Behavioral Research.* New
        Delhi: Sage.

Orfan, S. N. (2021). Political participation of Afghan Youths on Facebook: A case study of Northeastern Afghanistan. *Cogent Social Sciences*.

Pincock, K. (2018). School, sexuality and problematic girlhoods: reframing 'empowerment' discourse. *Third World Quarterly*, 909.

Rieger, T. A. (2019). A survey of the Afghan people: Afghanistan in 2019. *The Asia Foundation*.

Roof, D. J. (2015). Day-By-Day: Higher Education in Afghanistan. *FIRE: Forum for International Research in Education. Vol. 1. No. 3.*

Sandra, A. (1987). Feminist Theory and the Study of Gender and Education. *International Review of Education / Internationale Zeitschrift für Erziehungswissenschaft / Revue Internationale de l'Education , 1987, Vol. 33, No. 4, Women and Education*, 6.

Shirazi, R. (2008). Islamic Education in Afghanistan: Revisiting the United States' Role. *CR: The New Centennial Review, 8.*

Swanson, R. A., & Holton III, E. F. (2005). *Research In Organization Foundations and Methods of Inquery.* San Francisco, California: Berrett-Koehler.

Telesetskyt, A. (1998). In the Shadows and Behind the Veil: Women in Afghanistan Under Taliban Rule. 1.

Weiner, G. (1986). Feminist Education and Equal Opportunities: Unity or Discord. *British Journal of Sociology of Education , 1986, Vol. 7, No. 3 (1986), pp. 265-274.*

Women, U. (2022). Women's rights in Afghanistan, one year after the Taliban taken-over.

Yapp, M. (2001). The Legend of The Great Game. *In Proceedings-british Academy (Vol. 111, Pp. 179-198). Oxford University Press Inc..*

Yousufi, F. (2021). The Prospect of W ospect of Women's Rights in the P s Rights in the Post-Taliban-Go aliban-Government ernment. *Journal of international Women's Studies.*

Appendix

Fetch Tweets

The code below fetches tweets from each individual accounts, by their usernames, Twitter allows their users to have a single unique username, we are allowed to fetch maximum 3200 from certain user.

```
1. import tweepy
2. import pandas as pd
3. consumer_key = 'your_consumer_key'
4. consumer_secret = 'your_consumer_secret'
5. access_token = 'your_access_token'
6. access_token_secret = 'your_access_token_secret'
7. auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
8. auth.set_access_token(access_token, access_token_secret)
```

```
 9. api = tweepy.API(auth)
10. username = 'twitter_username'
11. tweets = api.user_timeline(screen_name=username, count=100)
12. tweets_list = []
13. for tweet in tweets:
14.    tweets_list.append({
15.        'created_at': tweet.created_at,
16.        'text': tweet.text,
17.        'retweets': tweet.retweet_count,
18.        'favorites': tweet.favorite_count
19.    })
20. df = pd.DataFrame(tweets_list)
21. print(df.head())
```

Table 5. Fetch tweets by username

Find Related data by Keywords

After fetching large amount of data, it's always good practice to break down the data into small pieces. The code below makes another excel sheets according the given keywords ["education", "women", "woman", "schools", "school", "universities", "university", "closed", "close", "ban", "banned", "from school", "girls", "girl", "work"]. The total amount of data we have received from 179,121 - 56,820, almost half of the data is not related our research.

```
1. import pandas as pd
2. tweets_df = pd.read_excel('../data_processed/complete.xlsx')
3. tweets_df = tweets_df.dropna()
4. tweets_df['text'] = tweets_df['text'].str.lower()
```

45

```
 5. keywords = ["women", "education", "schools", "study", "learn", "right", "ban",
"people",
 6.          "afghan", "educated", "girls", "female", "protest", "pashtun", "tajik", "uzbek",
"hazara",
 7.          "rights", "prohibit", "patriarch", "close schools", "close universities",
"university", "man"]
 8.  matched_tweets = {}
 9. for keyword in keywords:
10.     tweets_df['text'] = tweets_df['text'].fillna('')
11.     matched_tweets[keyword] = tweets_df[tweets_df['text'].str.contains(keyword,
case=False)]
12.
13. for keyword, df in matched_tweets.items():
14.     print(f"{keyword}: {len(df)}")
15.  for keyword, df in matched_tweets.items():
16.     df.to_excel(f"../data_processed/{keyword}_matched_tweets.xlsx", index=False)
```

Table 6. Find related data by keywords

Data pre-processing

Is an important stage in any natural language processing operation. It entails cleaning and translating raw text data into a format that machine learning models can easily understand. The following steps were included in the pre-processing of Twitter data:

Remove ULR, and Mentions

With code below, we are able to remove all URLs, and mentions from the dataset, our dataset has bunch mentions and URLs. We import the regular expression library re in this example, then build a sample tweet with a URL, a mention, and some hashtags. The actual tweet is printed.

Following that, we use the re.sub() function to replace any URL that begins with http and any mention that begins with @ with an empty string, thereby deleting them from the tweet. Lastly, we publish the processed tweet, which is devoid of URLs and mentions.

```
1. def remove_usernames_links(tweet):
2.     tweet = re.sub("@[^\s]+","",str(tweet))
3.     tweet = re.sub("http[^\s]+","",str(tweet))
4.     return tweet
5. df = df.replace(r"\n"," ", regex=True)
6. df["text"] = df["text"].apply(remove_usernames_links)
```

Table 7. Remove URL, and mention

Top Ten Trends

The code below, uses regular expression to extract hashtags from a DataFrame's 'tweet' column and stores them in a 'hashtags' list. The 'Counter' module from the 'collections' library is then used to tally the frequency of each hashtag. The top ten most often hashtags are extracted and saved in the 'top hashtags' list using the 'most_common()' method of the 'Counter' class. Finally, the code outputs the top ten hashtags and their counts as named constants, with the hashtag transformed to title case by eliminating the '#' symbol.

```
1. from collections import Counter
2. hashtags = []
3. for row in df['text']:
4.     hashtags.extend([tag.strip("#") for tag in row.split() if tag.startswith("#")])
5. counts = Counter(hashtags)
6. top_tags = counts.most_common(10)  # modify topn parameter here
```

47

```
 7. fig, ax = plt.subplots(figsize=(12, 8))
 8. ax.bar([tag[0] for tag in top_tags], [len(tag[0]) for tag in top_tags])
 9. ax.set_title("Top 50 Hashtags with Lengths")
10. ax.set_xlabel("Hashtag")
11. ax.set_ylabel("Length")
12. plt.xticks(rotation=90)
13. plt.show()
```

Table 8. Top 10 Trends

Here are the most top 10 or frequent trends in our datasets. Millions of trends made by users everyday around global, as the table 5 shows, Afghanistan is at the top with 5224 times, and followed the Taliban trends.

Hashtag	Count
#Afghanistan	5224
#taliban	3457
#rabieh_sadata	1619
#afghanistansupporttaliban	1033
#kabul	929
#unitedafghanistan	829
#stophazaragenocide	757
#afghan	549
#afghanwomen	509
#pakistan	483

Table 9. Top Trends

Top Ten Trends Chart

The chart below shows the top most frequent or top ten hashtags in our dataset, chart help as to see the trend in a graph.
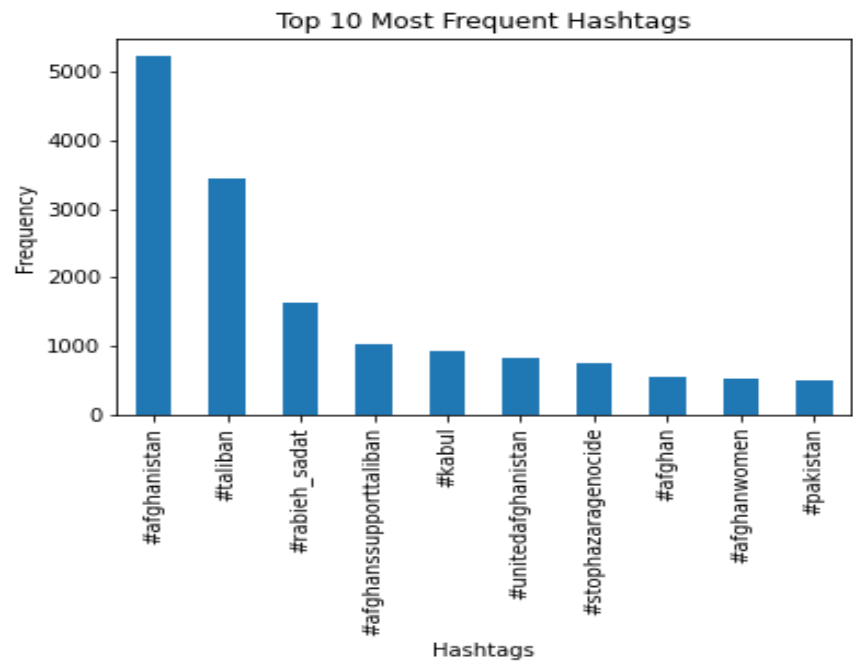
Figure 1 Top Ten Trend Chart

Remove Hashtags

It's import to remove hashtags after we get the results, because it's part data cleansing. Words with hashtags makes no sense in data analysis, since several words sticks together makes the user confuse. And this is still part of data cleansing.

```
1. import pandas as pd
2. import re
3. pattern = r'#\w+'
4. df['hashtags'] = df['text'].str.extractall(pattern)[0].values
5. print(df)
```

Table 10. Remove Hashtags

Here is the new update of our trends, after removing hashtags, it looks much better since we removed the hashtags.

Hashtag	Count
Afghanistan	5224
taliban	3457
rabieh_sadata	1619
afghanistansupporttaliban	1033
kabul	929
unitedafghanistan	829
stophazaragenocide	757
afghan	549
afghanwomen	509
pakistan	483

Table 11. Hashtags Removed

Expand Contractions

The code below, expands contractions in a Pandas DataFrame column named 'text' to do text pre-processing. A list of commonly used contractions and their expanded forms can be found in the contraction map dictionary. The function expand contractions searches for contractions in a text input using a regular expression pattern. When a contraction appears in the text, it is substituted with its extended form. The select dtypes method in Pandas is used to select columns of type 'object,' which often contain textual data. The apply method is used to apply the expand contractions function to all of the DataFrame's text columns. Finally, the first 5 rows of the pre-processed DataFrame are displayed using the head technique.

```
1. import pandas as pd
2. contraction_map={
```

```
3.    "ain't": "is not",
4.    "aren't": "are not",
5.    "can't": "cannot",
6. }
7. def expand_contractions(text):
8.    pattern = re.compile('({})'.format('|'.join(contraction_map.keys())),
flags=re.IGNORECASE|re.DOTALL)
9.    def replace(match):
10.        return contraction_map[match.group(0).lower()]
11.    return pattern.sub(replace, text)
12. text_cols = data.select_dtypes(include=[object]).columns
13. data['text'] = data['text'].apply(lambda x:expand_contractions(str(x)))
14. data.head(5)
```

Table 12. Expand Contractions

The table below, shows that our contractions has changed to complete words. This is

important in data cleansing as well.

No.	Contraction	Fixed
1	ain't	Is not
2	aren't	Are not
3	can't	Cannot

Table 13. Fixed Contractions

Remove ASCII Characters

A character encoding standard for electronic communication is
called ASCII (American Standard Code for Information Interchange). It
was created for the first time in the 1960s and is still in use today. ASCII
can represent up to 128 different characters because it employs a 7-bit code
to encode its characters. These symbols contain capital and lowercase
letters, numbers, punctuation, control, and a few unusual symbols.
Computers and other electrical devices use ASCII to represent text. Because

each character is represented by a distinct code, text may be processed and

displayed by computers in a uniformed manner. Over time, the ASCII

standard has grown to accommodate more characters, including symbols

and those from foreign languages.

```
1. import re
2. sample_string = 'This is a sample string with ASCII characters:
!@#$%^&*()_+={}[]|\:;"<>,.?/`~'
3. clean_string = re.sub(r'[^\x00-\x7F]+', '', sample_string)
4. print('Original string:', sample_string)
5. print('Cleaned string:', clean_string)
```

Table 14.Remove ASCII Characters

As a result, the table below shows that the code above has

successfully run and remove ASCII character our dataset.

No.	ASCII	Removed
1	This is a sample string with ASCII characters: !@#$%^&*()_+={}[]|\:;"<>,.?/`~	This is a sample string with ASCII characters

Table 15. ASCII Removed

Remove White Spaces

The whitespace function is used on the data DataFrame in this
example to remove any empty strings and leading/trailing whitespace in the
'name' and 'tweet' columns. The modified DataFrame is then printed to
demonstrate the changes. The following is the output of the above code:

```
1. def whitespace(columns_name, text):
2.    data[columns_name] = data[columns_name].replace('', '', regex=True) #remove
empty string
3.    data[columns_name] = data[columns_name].str.strip() #remove whitespace
4. whitespace('name',data)
5. whitespace('text',data)
```

Table 16. Remove Whitespace

As expected, the table below shows, all whitespace removes from our dataset. Always good practice to remove whitespace from the dataset.

No	White space	Remove
1	I love Pizza	I love Pizza
2	Its fun to analyze data	Its fun to analyze data

Table 17. Fixed all Whitespaces

Drop Irrelevant Data

Most of the dataset contains irrelevant data, null values, during analyzing the data if we do not remove or drop them, we would encounter error. That Makes the analysis hard, it's always good practice to remove all those null values during the data cleansing. The code below shows how to drop the null values from the dataset.

```
1. df.dropna(subset=['text'], inplace=True)
2. df.dropna(inplace=True)
3. df.drop_duplicates(subset=['text'], inplace=True)
4. df[['text']].sample(5)
```

Table 18. Drop Null Values

Remove Punctuations

Many symbols used in written language to fix and make clear the various sections of a sentence or phrase are referred to as punctuation in a

53

dataset. These can include signs like dashes, commas, periods, question marks, exclamation points, semicolons, colons, parentheses, and more.

```
1. import string
2. def remove_punctuations(text):
3.     for punctuation in string.punctuation:
4.         text = text.replace(punctuation, '')
5.     return text
6. df['text'] = df['text'].apply(remove_punctuations)
```

Table 19. Remove Punctuations

As a result, the table below shows that all punctuation has remove from the dataset. Punctuations are not important to keep them in the dataset and it should be removed all.

No	Paragraph (with punctuation)	Paragraph (without punctuation)
1	"Hello, how are you?"	Hello how are you
2	"I love to read books!"	I love to read books
3	"It's a beautiful day today."	Its a beautiful day today
4	"The cat chased the mouse."	The cat chased the mouse

To sum up, data cleaning is a crucial phase in the data analysis process, and using pandas to clean and preprocess data can be an efficient and effective way to make sure that datasets are prepared for analysis. Data analysts can find and fix missing or inconsistent values, get rid of duplicates, deal with outliers and errors, and change variables to better suit their analysis needs by using panda's functions and methods.

Machine Learning

Python is a preferred programming language because of its extensive capabilities, applicability, and simplicity. Due to its independent platform and widespread use in the programming community, the Python programming language is the most suitable for machine learning. A component of Artificial Intelligence (AI) called machine learning tries to make a machine learn from experience and carry out tasks automatically without necessarily having to be programmed to do so. Contrarily, Artificial Intelligence (AI) is a more general term for machine learning in which computers are made to be sensitive to the human level by perceiving visually, by speaking, by language translation, and thereafter making important decisions.

Create Contingency

A contingency table is a type of frequency distribution table that shows the relationship between two or more categories. The contingency table in our code shows the frequency distribution of the categorical variable "label" for each combination of the categorical variable's "ethnic" and "gender".

```
1. ct = pd.crosstab([df['ethnic'], df['gender']], df['label'], normalize='index')
2. print(ct)
```

Table 20. Contingency Code

The frequency with which each possible value combination appears in our data collection is displayed in the table. Relationships between categorical variables can be examined with the help of contingency tables. The underlying structure of our data can be revealed, and patterns and relationships between variables can be discovered. Predictions and inferences can be made by analyzing the frequencies in the contingency table to see if particular combinations of categorical variables are more likely to occur together.

	label	0	1
ethnic	gender		
hazara	female	0.012%	0.98%
	male	0.035%	0.96%
pashtun	female	0.034%	0.96%
	male	0.047%	0.95%
tajik	female	0.018%	0.98%
	male	0.024%	0.97%
uzbek	female	0.012%	0.98%
	male	0.030%	0.96%

Table 21. Contingency Results

This study evaluated tweets about education from different ethnic groups in Afghanistan to look for differences in perspective based on gender and background. According to the data, women of all backgrounds are more likely to tweet in favor of education than males are. And while Pashtun women are the least likely and Hazara men are the least likely to tweet in favor of education, Tajiks are the most likely of all ethnic groups to do so.

Even while social constraints prevent women from attending school in Afghanistan, these results imply that women still place a high value on education and are more inclined to voice their support for education. However, the study's sample size may not be representative of the whole population, and the tweets' context may have influenced the sentiments they express.

More research could build on this one by examining the broader social, economic, and cultural variables that drive gender inequality in education in Afghanistan, as well as the underlying reasons that influence attitudes towards education in different ethnic groups. Research in this area may take the form of in-depth interviews or surveys designed to reveal the perspectives and experiences of various racial and gendered groups in connection to formal education.
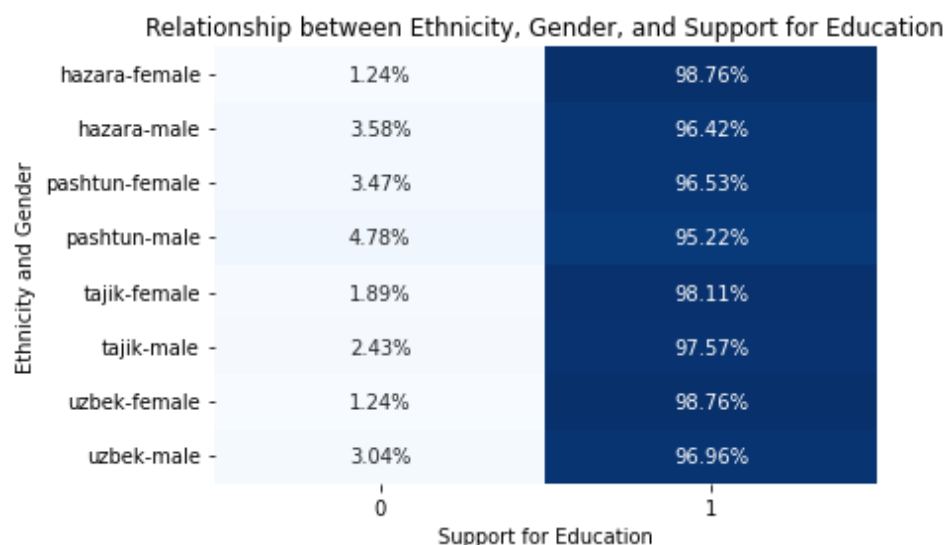
Figure 2. Relationship between ethnicity, Gender and Support for Education

Women of all ethnic and cultural backgrounds tweet more in favor of education, as shown by the heatmap study. This suggests that women are more likely to express their support for education on Twitter. Furthermore, the heatmap reveals that Pashtun males have a different perspective on education than men from other groups since they have the lowest percentage of tweets advocating for education. Uzbek men, meanwhile, have a disproportionately high number of pro-education tweets.

Relationship Between Gender and Label

The table breaks down the sample size according to ethnicity and gender. It shows that, with the exception of the Hazaras, there are more men than women taking part. The largest group is Pashtuns, with the smallest being Uzbeks. The table can be used as a guide for further study by providing an overview of the demographics of the sample population.

label	Female	male
Hazara	1128	3098
Pashtun	10036	18917
Tajik	10010	10899
Uzbek	728	2002

Table 22. Individual Tweet From all Ethnics

Most Common Words on Favor of Education

Here are the top 10 education-related words found in tweets with the label 1, indicating a favorable attitude toward education. Several tweets contained variations on these key phrases.

words	length
women	27471
school	21986
afghanistan	21778
education	17177
taliban	13501
girl	10543
right	7952
university	7059
people	6824

Table 23.Top 10 Positive Words

Most Common Words on Neutral Tweets

Words	Length
afghanistan	1029
people	379
kabul	247
islamic	240
peace	214
situation	206
today	194
country	188
emirate	187

Table 24.Most Common Words on Neutral Tweets

Vectorization

Vectorization is the act of transforming textual input into a numerical format that machine learning algorithms can handle in natural language processing (NLP). Machine learning models can only process numerical data, so this is essential.

```
1. from sklearn.feature_extraction.text import CountVectorizer
2. from sklearn.linear_model import LogisticRegression
```

```python
3. import numpy as np
4. vectorizer = CountVectorizer(stop_words=stopwords.words('english'))
5. X_train_vectorized = vectorizer.fit_transform(X_train)
6. unwanted_words = ['afghanistan', 'taliban', 'afghan', 'kabul']
7. for word in unwanted_words:
8.     try:
9.         idx = vectorizer.vocabulary_[word]
10.        del vectorizer.vocabulary_[word]
11.        vectorizer._validate_vocabulary()
12.        X_train_vectorized = X_train_vectorized[:,
np.arange(X_train_vectorized.shape[1]) != idx]
13.    except KeyError:
14.        pass
15. X_train_vectorized = vectorizer.fit_transform(X_train)
16. lr = LogisticRegression()
17. lr.fit(X_train_vectorized, y_train)
18.
```

Table 25. Transforming Textual to numeric

The table displays the outcomes of vectorizing the preprocessed text data using scikit- vectorization learns object. Each column in the table represents a distinct word from the lexicon, and each row in the table represents a data point. The frequency of each word in the associated data point is shown by the values in the table. For instance, the value "1" in row 0 and column 985 denotes that the first data point has one instance of the term corresponding to column 985. Only the non-zero values are displayed since the values are stored in a sparse matrix format. The table offers a useful method for numerically representing text data, which can be utilized as data for machine learning models. Based on the text input, the generated matrix can be used to train a model to predict the target variable.

row	(i, j) index	count
0	(0, 985)	1
0	(0, 4830)	1

0	(0, 9910)	1
0	(0, 13283)	1
0	(0, 14703)	1
0	(0, 17155)	1
0	(0, 20273)	1
0	(0, 21178)	1
0	(0, 26797)	1
0	(0, 29986)	1
0	(0, 31600)	1
0	(0, 32324)	1
1	(1, 985)	1
1	(1, 2209)	1
1	(1, 3562)	1
1	(1, 5029)	1
1	(1, 7365)	1
1	(1, 10101)	2
1	(1, 16913)	1
1	(1, 16971)	2
1	(1, 18601)	1
1	(1, 19578)	1
1	(1, 19579)	1
1	(1, 25931)	1
1	(1, 26958)	1

Table 26. Bag of Words

The bag-of-words (BoW) model is a popular technique for vectorizing text data. Each document is represented as a vector in the BoW model, where each element in the vector denotes the frequency of a specific word inside the document. For instance, if we have a collection of documents with the terms "apple," "banana," and "orange," we may represent the documents "I enjoy apples," "She likes bananas," and "We eat oranges every day" as the vectors [1, 0, 0], [0, 1, 0], and [0, 0, 1], respectively.

Train Model

The code below, evaluate the performance of our dataset how well the dataset has been trained

```
1. from sklearn.metrics import classification_report
2. y_pred = lr.predict(X_test_vectorized)
3. print(classification_report(y_test, y_pred))
```
Table 27. Evaluate Performance

The logistic regression model was quite effective in labeling tweets about learning. The dataset used to train the algorithm contained mostly pro-education tweets and some neutral ones (label 0). (label 1). Table 1 displays the evaluation criteria.

Metric	Value
Accuracy	0.98
Precision	0.84 (0) / 0.99 (1)
Recall	0.66 (0) / 1.00 (1)
F1-score	0.74 (0) / 0.99 (1)
Table 28.Model

For tweets that are neutral the precision was 0.84, while for tweets that were in favor, it was 0.99, showing a very high degree of accuracy. The algorithm did a better job of recognizing tweets connected to education than tweets not related to education, with a recall of 1.00 for pro-education tweets compared to 0.66 for neutral tweets. Overall, the model seems to have done quite well, with a weighted F1-score of 0.98. Overall, the tweets that we collected and trained them regarding education suggest that tweets are more likely to be in favor education rather than neutral.

63