

Gender and Women: Educational Perception Among Varied Category Groups in Afghanistan

Abdu Saboor Hamedi

abdul.hamedi@student.uiii.ac.id

Gender and Women: Educational Perception Among Varied Category Groups 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

©COPYRIGHT

By

Abdul Saboor Hamedi

Supervisors:

Supervisor 1 : Sirojudin Arif, Ph.D

Supervisor 2 : Nia Deliana, Ph.D

Table of Contents

[List of Figures vi](#_Toc136007519)

[List of Tables vii](#_Toc136007520)

[Abstract 1](#_Toc136007521)

[CHAPTER I: INTRODUCTION 2](#_Toc136007522)

[Introduction 2](#_Toc136007523)

[Research Problem 6](#_Toc136007524)

[Research Question 7](#_Toc136007525)

[Research Objective 8](#_Toc136007526)

[Significant of study 9](#_Toc136007527)

[CHAPTER II: THEORETICAL FOUNDATION 12](#_Toc136007528)

[Theoretical Framework 12](#_Toc136007529)

[Literature Review 14](#_Toc136007530)

[Hypothesis 22](#_Toc136007531)

[CHAPTER III: RESEARCH DESIGN 25](#_Toc136007532)

[Research Methodology 25](#_Toc136007533)

[Data Collection 26](#_Toc136007534)

[Data Cleaning 28](#_Toc136007535)

[Labeling Data 28](#_Toc136007536)

[The Place and Time of the Study 34](#_Toc136007537)

[Research Material and Tools 34](#_Toc136007538)

[Research Management 35](#_Toc136007539)

[CHAPTER IV: RESEARCH RESULTS AND DISCUSSIONS 37](#_Toc136007540)

[Research Results 37](#_Toc136007541)

[Women’s Education: Activist Perceptions 37](#_Toc136007542)

[Female Activists Top Frequent Words 47](#_Toc136007543)

[Male Activists Top Frequent Words 51](#_Toc136007544)

[Women’s Education: From an Activist, Taliban and Ordinary People Perceptions 54](#_Toc136007545)

[Analyzing the Frequency of Keywords Regarding Women’s Education 80](#_Toc136007546)

[Taliban Top Frequent Words 82](#_Toc136007547)

[Female Ordinary People Top Frequent Words 84](#_Toc136007548)

[CHAPTER V: GENDER AND WOMEN'S EDUCATION: A CLOSER LOOK AT THE ACTIVISTS 87](#_Toc136007549)

[VI Discussion 88](#_Toc136007550)

[HAPTER V: CONCLUSION AND SUGGESTIONS 90](#_Toc136007551)

[Conclusion 90](#_Toc136007552)

[References 94](#_Toc136007553)

[Appendix 102](#_Toc136007554)

# List of Figures

[Figure 1:Sentiment Distribution (activists) 38](#_Toc136007555)

[Figure 2: Sentiment Distribution Male vs Female (activists) 42](#_Toc136007556)

[Figure 3: Sentiment Distribution on Women Education (activists) 43](#_Toc136007557)

[Figure 4: Female Activists Top Frequent Words 48](#_Toc136007558)

[Figure 5: Female Activists Word-Cloud 50](#_Toc136007559)

[Figure 6: Male Activists Top Frequent Words 51](#_Toc136007560)

[Figure 7: Male Activists Word-Cloud 53](#_Toc136007561)

[Figure 8: Sentiment Distribution of (activists, taliban, people) 55](#_Toc136007562)

[Figure 9: Sentiment Distribution of five Keywords (activists, taliban, people) 56](#_Toc136007563)

[Figure 10: Sentiment Distribution on Women Education (activists, taliban, people) 77](#_Toc136007564)

[Figure 11: Percentage Women Education Keyword (activists, taliban, people) 79](#_Toc136007565)

[Figure 12: Frequency of Five Keywords by Group over Time (activists, taliban, people) 81](#_Toc136007566)

[Figure 13: Taliban Top Frequent Words 83](#_Toc136007567)

[Figure 14: Taliban Word-Cloud 84](#_Toc136007568)

[Figure 15: Female Ordinary People Top Frequent Words 85](#_Toc136007569)

[Figure 16: Female Ordinary Word-Cloud 85](#_Toc136007570)

[Figure 17: Male Ordinary Word-Cloud 86](#_Toc136007571)

# List of Tables

[Table 1:Tweets From Activists (Male & Female) 30](#_Toc136007572)

[Table 2: Tweets From Taliban 31](#_Toc136007573)

[Table 3: Tweets From Ordinary People (Female ) 32](#_Toc136007574)

[Table 4: Tweets From Ordinary People (Male ) 33](#_Toc136007575)

[Table 5: Sentiment Analysis of Tweets Containing the Keyword (activists) 40](#_Toc136007576)

[Table 6: Calculate Sentiment for Males and Female on Five Keywords (activists) 41](#_Toc136007577)

[Table 7: Sentiment Distribution on Women Education (activists) 42](#_Toc136007578)

[Table 8: Calculate Sentiment Scores for Entire Dataset (activists) 44](#_Toc136007579)

[Table 9: Calculate Sentiment Scores for Five Keywords (activists) 45](#_Toc136007580)

[Table 10: Calculate Sentiment Scores for Women Education (activists) 46](#_Toc136007581)

[Table 11: Frequent Words (female activists) 49](#_Toc136007582)

[Table 12: Frequent Words (Male activists) 52](#_Toc136007583)

[Table 13: Sentiment Distribution of (activists, taliban, people) 56](#_Toc136007584)

[Table 14: Sentiment Analysis of Tweets Containing the Keyword (activists, taliban, people) 59](#_Toc136007585)

[Table 15: Tweets Related to five Keywords (activists, taliban, people) 59](#_Toc136007586)

[Table 16: Sentiment Distribution on Women Education (activists, taliban, people) 78](#_Toc136007587)

[Table 17: Percentage and Count on Women Education (activists, taliban, people) 79](#_Toc136007588)

[Table 18. Fetch tweets by username 102](#_Toc136007589)

[Table 19. Find related data by keywords 103](#_Toc136007590)

[Table 20. Remove URL, and mention 103](#_Toc136007591)

[Table 21. Top 10 Trends 104](#_Toc136007592)

[Table 22. Remove Hashtags 104](#_Toc136007593)

[Table 23. Expand Contractions 104](#_Toc136007594)

[Table 24.Remove ASCII Characters 105](#_Toc136007595)

[Table 25. Remove Whitespace 105](#_Toc136007596)

[Table 26. Drop Null Values 105](#_Toc136007597)

[Table 27. Remove Punctuations 106](#_Toc136007598)

[Table 28. Contingency Code 106](#_Toc136007599)

[Table 29. Transforming Textual to numeric 107](#_Toc136007600)

[Table 30. Evaluate Performance 107](#_Toc136007601)

[Table 31: Labeling Data 107](#_Toc136007602)

[Table 32: Sentiment Analysis 107](#_Toc136007603)

[Table 33: Sentiment Distribution 108](#_Toc136007604)

[Table 34: Calculate Positive Tweets 108](#_Toc136007605)

[Table 35: Code Scheme 109](#_Toc136007606)

[Table 36: Anlyzing Five Keywords 109](#_Toc136007607)

[Table 37: Average Sentiment Percentages 109](#_Toc136007608)

[Table 38: Comparison of Sentiment Whole Tweets 110](#_Toc136007609)

[Table 39: Comparison of Sentiment Five Keywords 110](#_Toc136007610)

[Table 40: Comparison of Sentiment on Keyword Women Education 110](#_Toc136007611)

[Table 41: Chi-square Test of Independence 110](#_Toc136007612)

[Table 42: Cross-Tabulation of Sentiment and Gender 110](#_Toc136007613)

[Table 43:Frequent Words 111](#_Toc136007614)

[Table 44:Time Series 112](#_Toc136007615)

[Table 45: Word-Cloud 112](#_Toc136007616)

# Abstract

This thesis investigates the topic of women’s education in Afghanistan from multiple angles, including those of activists both male and female, ordinary people, and the Taliban officials. In order to examine the amounts and types of support for women’s education, the study employs content analysis and sentiment analysis of tweets. The study is based on feminist viewpoint theory, which stresses the significance of learning from women’s personal accounts. To compile this data, we mined Twitter for posts from Afghan activists, Taliban officials, and Afghan people. The results show that the various groups have diverse perspectives and feelings. By actively supporting and promoting equal access to education, activists show their great support for women’s education. There is also substantial backing from the people, but to a little smaller amount. In contrast, the Taliban show little interest in or concern for the education of girls and women. To further gender equality in education, the study emphasizes boosting activist voices, addressing societal issues, and advocating legislative reforms. Policymakers, educators, and groups in Afghanistan seeking to better women’s education and empower women will get useful insight from the findings.

**Keywords:** Women, Girl, Education, University, School, Twitter

# CHAPTER I: INTRODUCTION

# Introduction

Millions of women, especially without equal educational experiences, skills, or qualifications, 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).

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 imposed restrictions on their education system. This restriction has a substantial negative impact on women’s daily activities. This limits their access to public gatherings, travel, study and other activities.

Studies show that educated women are more likely to have more apportunities, “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 receive an education are more supportive of gender equality and human rights. This study is necessary because it sheds light on the effects of the ban on women’s education in Afghanistan and gives insight into Afghan society’s attitudes towards this topic.

They study uses a standpoing theory (Allen, 2017) the study conducted from a standpoint of women marginalized in the modern education. “This standpoint theory built from the experience of women will be feminist in nature” (Parson, 2020). Further, the study examines the tones of the tweets posted by three different groups, activists, Taliban and ordinary people. The goal is to analyze the tweets and gauge their level of support for women’s education. In this analysis, we looked at the sentiment distribution across several topics connected to women’s education and compute the percentage of positive tweets from women. Descriptive data are included, as well as a chi-square test of independence and a comparison of sentiment ratings across sexes and those between activists and government officials/employees. The study wraps up with a review of the most popular phrases and terms that were found in the tweets.

The study used content analysis to analyze data and test hypothesis as well as sentiment analysis (Evans & Clark, 2015; Kitzie & Ghosh, 2016). Twitter operates as a network in which individuals rapidly disseminate and exchange thoughts, it possible for individuals to share their thoughts and feelings openly via the use of various forms of media, including text, photographs, videos, etc (Karamouzas, Mademlis, & Pitas, 2022, p. 1). Thus, these mediums can be utilised to track public opinion on a selected topic. Public opinion “represents the views, desires, and wants of the majority of a population concerning a certain issue, whether political, commercial, social, or other (El Barachi, AlKhatib, Mathew, & Oroumchian, 2021, p. 1)”

Further, the researcher uses the Python programming language to analyze the tweets. Python programming language is quickly becoming one of the most widely used for computational science. It is attractive for algorithm creation and exploratory data analysis because of its high-level interactive character and growing ecosystem of scientific libraries (Pedregosa, 2011, p. 2826). Also, Python is one of the most efficient and compatible programming languages for analyzing data and has thousands of libraries, also, the researcher uses several libraries, such as Pandas and NumPy, scikit-learn, these libraries allows us to understand our data and analyze them based on the study’s requirements.

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

Despite some progress made in recent years, Afghan women still remain significantly away from education with barriers that hinders them from going to school and achieving educations. Thesis barriers includes, poverty, cultural norms, conflict, and lack of facalities. The goal of the study is to pay attention on difference in the perspectives of activists, the Taliban and the ordinary Afghan people, and how gender affect male and female activist regarding women’s education.

We explore the factors that shape these perceptions, and the barriers that women face during access to education, also, the impact of gender on individual attitudes and behaviors toward women’s education. We seek to understand the complexity of social, cultural, poverty, conflict and political factors that influences the perceptions of people regarding women’s access to education with a focus on the effect of gender norms.

Barriers to women’s education that prevent them from access to education including cultural practices, poverty, and education infrastructure. In a summary this study looks at how different demographics views access to education and how it differ by gender.

# Research Question

The aims to investigate on the difference in perspectives of activists, Taliban and the ordinary Afghan people, and how gender affect male and female activist regarding women’s education. The research question to be answered is: To what extent does gender, controlling for political inclination, affect people’s perception of women’s access to education?” The study uses Twitter data to discover if there are any gender-based trends or disparities in how people feel about lifting the restriction on women’s education. It stresses the necessity of destroying oppressive systems to realize gender parity.

Data was gathered from Twitter, a widely used social media platform, and analyzed using content analysis techniques to see how people feel about the topic. The study attempts to illuminate the complexity of gender and its role in forming attitudes regarding women’s education by investigating patterns and trends on three groups, activists, Taliban, and ordinary people. The findings will help educated people and those fighting for gender equality in the classroom better understand Afghan women’s obstacles and inequities while trying to get an education.

# Research Objective

In recent years, scholars have paid attention to issues that effect women, including access to education, equality, participation in politics, and other aspects of life that women usually face with difficulties. There are a number of reasons behind unequal education, one of the reasons that most of the scholars quote, is that “higher education research has traditionally been framed within a masculine paradigm, often with man participants (Parson, 2020, p. 515)”. Other than masculinity, conflict effects women’s education, especially countries like Afghanistan, which has experienced civil and international wars for decades.

This study aims to explore how gender shapes activists perceptions of women access to education, and to find what are the perspectives of activists, Taliban and ordinary Afghan people. The researcher dives into the perceptions and personal experiences of individual Afghan people. Paying attention to how gender shapes their vision or ideas and influences their views on women access to educational opportunities. The study focuses on recognizing various facts that contributing to attitudes toward women’s education in Afghanistan, considering culture, ecnomic, conflict, social aspects. We believe, by examining these factors reader can gain a deeper understanding of complexity that prevent women from education in the country.

The study will help better understand Afghan women’s difficulties when accessing education. It will also examine the relationship between gender ­­­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

The significance of this study lies in its focus on gender as a factor in how people see women’s educational opportunities in Afghanistan. Understanding the cultural and societal elements that contribute to gender imbalance in schooling can be gained through the analysis of tweets from Afghan people, and this study aims to do just that. As a result, the study will benefit the academic community by providing new insights into this pressing topic and demonstrating the value of social media as a resource for scholars in Afghanistan and beyond.

Furthermore, this study will be one of the first to extensively investigate the difference on perspectives of activists, Taliban and Afghan people regarding women’s access to education through social media. By giving them a stronger voice through social media, this study will help improve conditions for female students in Afghanistan. In doing so, it will shed light on the achievements and challenges faced by Afghan women and inspire future generations of girls and women to pursue education and careers.

The study seeks to shed light on the power of social media to advance gender equality and elevate underrepresented communities’ voices. The difficulties Afghan women encounter in gaining access to educatoin can be better understood by studying tweets from Afghans both inside and outside of Afghanistan. These results will provide the foundation for future efforts to advance gender equality in the classroom and beyond. This research acknowledges the special challenges faced by Afghan women and sets out to shine a light on such obstacles in the hopes of inspiring change. By sharing our findings, we want to encourage further discussion about how social media may help advance gender and political equality.

This study holds significant importance as it brings attention to the critical issue of women’s education in Afghanistan and the concerning public opinion that tends to hinder their access to education. By employing Twitter as a data collection tool, this research will contribute to the expanding literature on the gender imbalance in education. The resulting findings are expected to inspire further investigations into the intricate relationship between gender dynamics and social media in Afghanistan and other contexts worldwide. By deepening our understanding of these dynamics, we can effectively address the challenges that women face in pursuing education and work towards creating more inclusive educational environments.

# CHAPTER II: THEORETICAL FOUNDATION

## Theoretical Framework

As long as the barriers on women’s education concern, only women themselves can understand these barriers and problems. It’s a mistake to say a man can feel what is happening on women’s education, the pain and sorrow which is on the heart of a women because of education no one can feel it as a women does. Consequently, the researcher proposed that a study that is feminist in character will be one that is framed by viewpoint theory and founded on the experiences of women. I consider the feminist lens specifically, a feminist standpoint lens as crucial in this discussion of research methodologies because there needs to be a connection between theory and methodology (Allan & Tinkler, 2015).

Rebecca Ropers-Huilman and Kelly T. Winters (2011) highlights: “Feminist research, by definition, is committed to considering how gender implicates or is implicated by the phenomenon of interest” (Ropers-Huilman & Winters, 2011, p. 671). However, middle-class women’s perspectives and views have historically been prioritized in studies on women in education. Because of this, the experiences of middle-class women have frequently been taken to be representative of all women’s experiences, and the voices of women from other backgrounds or experiences have been marginalized or ignored. Despite the wide variety of feminist theoretical frameworks, researcher believe feminist viewpoint theory and postcolonial approaches to feminist research are essential when undertaking international education research with women. Researchers should take into account the fact that women’s experiences in the traditionally male-dominated realm of education vary greatly from one another.

Diverse groups of people often have dramatically diverse experiences with marginalizing structures, even if they share common features (such as the promotion and tenure process). Postcolonial researchers start from the vantage point of the women whose lives they are investigating by using feminist standpoint theory as a framework. When studying the educational backgrounds of women, feminist viewpoint theory encourages researchers to do it from the female perspective. Researchers should begin their investigation from the perspective of the population of interest, but this does not imply that interviews with members of the population of interest are a required part of the research process or that researchers will not talk to members of the opposite gender (Harding, 2004). The intersectionality of gender, race, ethnicity, and class is often analyzed in studies that are framed by feminist perspective theory (Hasse-Biber, 2013).

## Literature Review

Inequalities in education and training are only one area where gender plays a role. Two-thirds of the world's adult illiterate population is female see (UNESCO, 2013), and over 65 million females do not attend school. The Education for All (EFA) campaign (launched in Dhaka in 2000) aimed to solve these problems by ensuring that all children throughout the world had access to quality education and identified six goals, one of which was to eliminate gender gaps and inequalities in education by 2015. Progress has been made toward ensuring that girls have equitable access to basic education, according to the newly published EFA 2015 report by UNESCO. However, it is predicted that only 69% of nations have achieved gender parity in basic education as of 2015. This number reduces to 48% in secondary education. This demonstrates that we still have a ways to go before reaching true gender parity in our society (Cin, 2017).

Over the course of decades, a worldwide policy debate has developed around issues of gender and education, specifically surrounding girls' access to formal education. There have been distinct changes in this conversation and the subsequent behavior throughout time. Education is likely the most powerful protective factor in lowering the likelihood of child poverty, particularly for females (Roby, Lambert, & Lambert, 2009). The United Nations defines poverty as a lack of access to goods and services that prevents children from enjoying their rights, reaching their full potential, and participating as full members of society (UNICEF, 2006).

Issues of representation, distribution, and recognition of women’s and girls’ interests are also often overlooked. Questions of, education quality, culture, sustainability, and governance can all be framed through the lens of gender. The themes discussed in this literature review revolve around these essential concerns for development and for the advancement of women and girls in particular, also, how gender shapes people’s perceptions on women access to education,which is main discussion of this study.

Understanding the role that collective mobilizations, and especially women's movements, have played in the reconstitution of women's cultures and the redefinition of femininity is crucial if the devaluation of women is itself a cultural construct, as social constructionist perspectives on gender assume. We are curious about the ways in which women's communities foster cultures that encourage resistance through the cultivation of communal processes, consciousness, and practices.

The phrase "women's communities" is used to highlight the social relations and physical spaces where women engage, as well as the areas where culture is produced and enacted (Hurwitz & Taylor, 2012). The study of women’s alternative communities sheds light on the emergence and development of symbolic systems specific to women’s values, bodies, emotions, and labors in contexts other than those in which men participate in creative activity.

As a result of menstruation, childbirth, and motherhood’s symbolic significance in many cultures, women were often kept physically apart from men, and sex-segregated institutions and the domestic sphere were established to fix the private from the public sphere. Strict gender segregation encouraged the development of a culture exclusive to women. (Freedman, 1979; Taylor & Rupp, 1993).

Today in contemporary world Afghanistan has faced the same satuation where women are deprived from education, participate in political or make decision for their lives. Afghanistan may be the only country where women’s rights have made and broken kings and politicians in the last century (Ahmed-Ghosh, 2003). 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) highlights barriers to Afghan women that affected their education. Deniz, there has been a 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 impacts women’s education. Regarding Deniz, apart from conflict, Afghanistan is a traditional country. In many parts of the country, women are expected to stay at home and take care of children and not pursue education or work, and women's education is often seen as a 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).”

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.

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.

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.

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).

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).

## Hypothesis

**H1:** Women are more likely to have positive sentiment on women’s education or support women education compared to men. The idea behind this hypothesis is that women may have a more nuanced grasp of the negative effects of restrictions on their ability to proceed with their education and, therefore, be more vocal and disagree with the ban on education.

Historically, women in Afghanistan, dominated by male society, face serious obstacles to gaining an education and engaging in society and politics. Due to structural gender inequities, Afghan women have suffered disproportionately from restrictions on their educational ability. They know from personal experience how prejudice and other educational obstacles may limit one’s options in life.

Therefore, it is reasonable to assume that women may be more likely to oppose the ban on women’s education because of their experiences. Poverty and social inequality are both pervasive issues in Afghanistan. Because of these obstacles, the educational opportunities for women are pitiful. There have been great restrictions on women’s education during the past four decades. Women endured unacceptable policies from 1996 – 2001; they were not allowed to proceed with their education and could not participate in politics to decide their future.

The researcher mentioned earlier that Afghan women have historically experienced restrictions on education, they faced discrimination in education because of their gender, and they were considered inferior in the country. Thus, we can say that it’s a group of solidarity today. Women have more positive sentiments and support women’s education.

**H2:** Those who involved in activism are more likely to have more positive sentiment on women education compared to those involved in government. It is expected that activists are more likely to support women education that those affiliated with the government. This hypothesis implies that those who engage in activism, such as those who fight for educational reforms or work to advance women’s rights, are more likely to hold positive views and show significant support for women’s education.

Those who work in government, whether as policymakers or government employees, may be less likely to have a good attitude about women’s education. This hypothesis suggests that taking an activist stance significantly impacts forming and encouraging constructive worldviews and advocacy initiatives aimed squarely at improving educational opportunities for women. It also hints at a potential divide between activists and government employees in worldviews, priorities, and strategies. It is clear that there is a need for more research and study into the complex interplay between activism and government participation in molding attitudes about women’s education. Insight into the myriad factors influencing public support for women’s education can guide policies and activities to expand women’s access to education.

**H3:** Activists are more likely to support women’s education compared to the ordinary people and the Taliban. The hypothesis suggests that activists are more likely to support women’s education compared to ordinary people and the Taliban. Activists, driven by their dedication to social justice and belief in the transformative power of education, actively advocate for gender equality and view education as a fundamental right and tool for empowerment.

They passionately use platforms like social media to raise awareness, generate discussions, and mobilize support for women’s education initiatives. In contrast, ordinary people may exhibit varying levels of engagement and awareness on the issue, while the Taliban’s lesser emphasis on women’s education reflects their different priorities. The active support from activists for women’s education demonstrates their commitment to breaking down barriers, challenging societal norms, and promoting equal opportunities for all.

# CHAPTER III: RESEARCH DESIGN

# Research Methodology

These days, surveys are used more than ever before as a research tool. Personality, IQ, happiness, learning styles, and so on may all be gauged through surveys. There's nothing stopping someone from posting a series of photos they feel best captures their personality on a dating service instead of filling out a questionnaire (Scott & Kosslyn, 2015). Additional, area where content analysis might prove valuable is in restarting the pursuit of social intelligence.

The study of social intelligence fell out of popularity in the late 1990s, mostly owing to technical limitations. Vast availability of new media, especially digital images and films, will be incredibly useful in advancing ideas of social intelligence and cultural competency (Kihlstrom & Cantor, 2000).

To test the hypothesis, the study have used content analysis with sentiment analysis (Evans & Clark, 2015) to check the tones of the tweets related to women’s education. Along with chi-square, which evaluates the consistency between an observed percentage and theoretical predictions; test for statistical significance between two variables often representing nominal or ordinal levels of measurement (Lewis & Burke, 1949), also, it is common practice to use cross tabulation (crosstab) as an analysis method to compare the outcomes of one or more variables with those of another.

## Data Collection

Data collection began between December, 01, 2022, and March, 30, 2023, the used free Twitter application programming interface (API), this API allows individual to retrive up to 3200 maximum rows from a singl account (Kitzie & Ghosh, 2016). The researcher in this study analyzes three different groups of tweets, namely: Activists, both male and female; Taliban officials, since there is no female on the Taliban’s cabinet, thus the study only focuses on the male side, last but not least, ordinary Afghan people, male and female, not by gender but in general, eventhough there are an equal amount of sample for both.

Researcher have individually collected the tweets, and there are three different approaches to recognizing the activists, to fetch their tweets. First, the we have joined the Afghan Twitter spaces, to recognize activists from there. The second approach, finding Afghan activists through google. The third approach is, activists have been frequently invited by Afghanistan’s TV, there are several TV channels which has hundreds of followers and views inviting activists in daily bases. E.g., [TOLOnews](https://twitter.com/TOLOnews), [ArianaNews](https://twitter.com/ArianaNews_).

For the Taliban officials, there is a single approach, which is [Zabehulah\_M33](https://twitter.com/Zabehulah_M33), and “Official Twitter Account of the Spokesman of Islamic Emirate of Afghanistan, Zabihullah Mujahid”. This is the official account of the Taliban spokesman, and Zabihullah has followed other officials; it provides an easy way to recognize the officials and fetch their tweets.

The researcher have done the same approachn on [mobeenkhan1231](https://twitter.com/mobeenkhan1231). Mobeen is a famous person on social media. He is an active member of the Taliban, most of the Taliban has followed him on Twitter, this gives us more ability to access Taliban officials to access to their tweets.

Related to the ordinary people, the researcher have implemented the same approach but this time, on Tolo News and Ariana News “[TOLOnews](https://twitter.com/TOLOnews),  [ArianaNews](https://twitter.com/ArianaNews_)”. It gives us ability to recognize Afghan people from their followers or on the comment section, because mostly people use either Pashtu or Persian (Dari) language to comment on a post. To analyze the data researcher has taken a sample which consists of 54 different Twitter users, whether its activists are the Taliban, but the number of tweets could be different, dependent on the users and how active they are. The duration of data collection starts from 2021, December, 01 to 2023, March, 30.

## 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.

The obtained dataset is cleaned up before analysis to reduce time and effort. All the information is changed to lowercase and any slang or unnecessary terms are eliminated. URLs and textual references are also omitted since they are unnecessary for analysis.

Eliminating frequent, brief function words (such as pronouns, prepositions, conjunctions, abbreviations, interjections, etc.) that are still essential to language syntax is known as "removing stopwords" from a corpus. Is, but, and, which, on, any, and some are all examples of stopwords. A text file containing a list of stopwords may be used to remove them from any language written in the Latin script (Domagoj Margan, 2015).

## Labeling Data

Labeling the tweets gives us more access to those tweets which talks related to women’s education. In another word, one can differentiate which tweets are related to education which one are not. The researcher have labeled data using five different keywords, which are related to the study, and frequently appears on tweets, which are ‘education’, ‘school’, ‘university’, ‘women’, ‘girl’, if any of these keywords appear on tweets it denotes with one (1) if not its zero (0). Basically 1 means the tweet related to one of these keywords if not the tweet about something else, which would be taken as a neutral tweet. Tweets like “beautiful women”, or “good school” has removed from dataset, unless tweets that talks the about current topic, such as “women banned from education”, or “girls cannot go to school”, the labeling process done by machine.

Another keyword, the study used is “women education”. This is a single keyword that consists of two words, this is exactly focusing on women’s education, by breaking down the data into a small piece and analyze only those tweets which has word “women education”. Overall, the data analysis method shed light on the three different groups, which are activsts, Taliban and ordinary people in Afghanistan 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 groups, were uncovered by the investigation

|  |  |  |  |
| --- | --- | --- | --- |
| Male | | Female | |
| Unclean | Clean | Unclean | Clean |
| remember world let taliban get away completely banning women education five years one hundred ninety nine | remember world let taliban get away completely banning women education five years | Please support women education Afghanistan | please support women education afghanistan |
| Women in Afghanistan after the return of the Dark ages—Taliban rule. | women afghanistan return dark ages taliban rule | Almost one Month since surrender Kabul Taliban banned women girl, work, media, sports, and Music Public | almost one month since surrender kabul taliban banned women girl work media sports music public |
| "The Taliban banned women’s education in Afghanistan on the order of Pakistan," claims Hamid Karzai. | Taliban banned women education afghanistan order pakistan claims hamid Karzai | Unfortunately current problem women education work country serious sad twenty first century | unfortunately current problem women education work country serious sad twenty first century |
| The Taliban are afraid of women’s education because they don’t want a bright future for Afghanistan | Taliban afraid women education want bright future afghanistan | Today, all universities reopened in Afghanistan after winter break, but ONLY for men. These brave young women are peacefully protesting outside Kabul University against Taliban BAN on women going to university. Let us amplify their voices! | today university reopened afghanistan winter break men brave young women peacefully protesting outside kabul university taliban ban women going university let us amplify voices |
| Taliban banned girls’ education in Afghanistan while many Taliban leaders send their daughters to study abroad | taliban banned girl education afghanistan many taliban leaders send daughters study abroad | Taliban ban Afghan Women Education Basic Human Right Least World Could Ban Taliban | taliban ban afghan women education basic human rights least world could ban taliban |

Table 1:Tweets From Activists (Male & Female)

Tweets from both male and female activists have been collected in Table 1. The table differentiates between processed and unprocessed tweets and offers a qualitative description of each. Several methods, already detailed, were used to get the purified tweets. The dedication displayed by these activists in their fight to advance women’s education is symbolized by the table. It highlights their individual insights and contributions to the continuing debates on this topic.

Table 2: Tweets From Taliban

|  |  |
| --- | --- |
| Unclean | Clean |
| Respected Maulvi Abdul Kabir said: Now compared to the past, good conditions have been provided for the return of migrants and aid is provided to them in a transparent manner, and we assure that the rights of migrants and returnees and the human rights of all Afghan people are protected. 3/4 | respected maulvi abdul kabir also said islamic emirate want deprive women education allow go homes mentioned examples women girls actually engaged business education work offices hospitals |
| As the majority of the people of Afghanistan are Muslims, the Afghan government considers the observance of the Islamic hijab in accordance with the religious and cultural values ​​of the Afghan society and in accordance with the ideals of the majority of Afghan women, and emphasizes that nothing against the religious and cultural belief of the Islamic society has been imposed on Afghans. . 4/6 | majority people afghanistan muslims afghan government considers observance islamic hijab accordance religious cultural values afghan society accordance ideals majority afghan women emphasizes nothing religious cultural belief islamic society imposed afghans |
| If Europe is working for women’s rights in Afghanistan, but it is a slave, then in a religious and conservative society like ours, who can claim destruction rather than work to secure women’s rights? Europe should not look at the Afghan society/women from the perspective of its own values. | europe working womens rights afghanistan slave religious conservative society like claim destruction rather work secure womens rights europe look afghan society women perspective values |
| These are female protesters harassing security forces who are protecting them. This shows the tolerance good behavior of IEA soldiers but no one will show this picture,no western journalist will tweet this. IEA will always protect the women their rights in Afghanistan. #Peace https://t.co/UJ561d2SEQ | female protesters harassing security forces protecting shows tolerance good behavior iea soldiers one show pictureno western journalist tweet iea always protect women rights afghanistan peace |
| Head of the Ministry of Higher Education: University doors are closed to women for four reasons. 1\_ The existence of women’s dormitories 2\_ They come from one province to another province without mahram 3\_ Students do not observe hijab 4\_ And male and female students continue to live together. https://t.co/C2Nub9QmHr | head ministry higher education university doors closed women four reasons existence womens dormitories come one province another province without mahram students observe hijab male female students continue live together |
| On the basis of the decree of His Highness Amirul Momineen Hufzaullah, former governor of Kabul Province, Sheikh Nada Mohammad Nadeem was appointed as the head of the Ministry of Higher Education. On the basis of the order of the Supreme Commander of the Faithful, the former governor of Kabul, Sheikh Nada Mohammad Nadeem, was appointed as the Acting Minister of Higher Education. https://t.co/BHBGAMOTHh | basis decree highness amirul momineen hufzaullah former governor kabul province sheikh nada mohammad nadeem appointed head ministry higher education basis order supreme commander faithful former governor kabul sheikh nada mohammad nadeem appointed acting minister higher education |

Tweets from the Taliban can be seen in Table 2, whereas activist tweets can be seen in Table 1; it appears that the Taliban tweets do not promote women’s education. Activists are tweeting more than ever before on women’s empowerment and education. In Afghanistan, advocates of both sexes are asking the public to promote girls' education. A greater number of activist tweets focus on the importance of women continuing their education.

Table 3: Tweets From Ordinary People (Female )

|  |  |
| --- | --- |
| Unclean | Clean |
| RT @Malala: While banning girls from school, the Taliban are also arresting champions of education. @matiullahwesa founded @penpath1 to pro… | malala banning girl school taliban also arresting champions education matiullahwesa founded penpath pro |
| Since 555 days, #Afghan #girls #education centers closed in #Afghanistan. Quality Education has the power to transform societies in a single generation. Afghan girls &amp; women need quality education. They need support &amp; help For Formal education, #LetAfghanGirlsLearn https://t.co/2hdSbFAIaC | since days afghan girl education centers closed afghanistan quality education power transform societies single generation afghan girl women need quality education need support help formal education let afghan girl learn |
| @matiullahwesa is the voice of Afghan generation. He is peaceful education activist. He want to Education for all in Afghanistan. We are all also have obligation to rise voice for Afghan girls and women education in Afghanistan. #RelaeseMatiullahwesa #LetAfghanGirlsLearn https://t.co/Yv4BBSNrB4 | matiullahwesa voice afghan generation peaceful education activist want education afghanistan also obligation rise voice afghan girl women education afghanistan relaesematiullahwesa let afghan girl learn |
| Negotiating the ban on women education in Afghanistan is so silly. Explaining to anyone why Afghan women have the right to get an education is so silly. Explaining how and why Islam supports is so silly. It is all just waisting time. #Taliban wont allow so the wait is also silly. https://t.co/sk5VtLjAiD | negotiating ban women education afghanistan silly explaining anyone afghan women right get education silly explaining islam supports silly waisting time taliban wont allow wait also silly |
| The war in Afghanistan is to destroy the idea of ​​education. This war started fifty years ago. The world cannot even imagine how many teachers, students and education workers have been killed and how many schools have been burned in Afghanistan. https://t.co/IkUwrNtI3Z | war afghanistan destroy idea education war started fifty years ago world cannot even imagine many teachers students education workers killed many school burned afghanistan |

Table 3 contains tweets of Afghan female, their tweets addressing education and support for women. These tweets illustrate the solidarity displayed by Afghan women for their fellow women and their education. One such example is the story of "matiullahwesa," an activist who was recently jailed by the Taliban for his passion to women’s education and his attempts to arrange private lessons for females. Unfortunately, his situation remain unknown following his arrest. The tweets illustrate that people are fully aware of the relevance of education and consequently speak up for it. Moreover, other tweets suggest that the impediments to education extend beyond the Taliban, pointing out that the continuing fighting in Afghanistan also restricts educational prospects.

Table 4: Tweets From Ordinary People (Male )

Tweets from Afghan males are explicitly blaming the Taliban for their bans on women’s education. These tweets shows that Afghan male strongly support women’s education. They frequently blamed the Taliban and called them to allow women educate. Apart from that its very interested to highlights, most of Afghan people blaming the Taliban that working for the Pakistan, and any order related to women’s education comes from Pakistan. The tweets are strongly supportive of education, and standing with women’s education.

|  |  |
| --- | --- |
| Unclean | Clean |
| @tomas\_niklasson @gmukhopadhaya Excellencey! I really appreciate your efforts for Afghanistan and Afghans. Afghans love to see you and others like you working for peace, political stability and protection of human rights particularly women education and rights in Afghanistan. | tomas niklasson gmukhopadhaya excellencey really appreciate efforts afghanistan afghans afghans love see others like working peace political stability protection human rights particularly women education rights afghanistan |
| Have men in Afghanistan also staged rallies in support of women education? Women have been banned by Taliban for last 2 years but men inside Afghanistan are as silent as they have been sniffed by snake. Unfortunately many Afghans/Pashtun have religion dearer than their country | men afghanistan also staged rallies support women education women banned taliban last years men inside afghanistan silent sniffed snake unfortunately many afghanspashtun religion dearer country |
| Taliban reclusive Ameer Haibatulla said that his Emirate will not surrender to int pressure over women education and rights. In 2001 also TBs had such stubborn stance over #Osamabinladen who latter on was killed by US in Pak. Is Pak Army eying for an other project over Afg? | taliban reclusive ameer haibatulla said emirate surrender int pressure women education rights also tbs stubborn stance osamabinladen latter killed us pak pak army eying project afg |
| Afghanistan needs women leadership and women need education. Taliban r afraid of educated women. We stand with Afghan women and demand to let them get educated. Without women education no nation can prosper. Afghanistan needs prosperity and prosperity needs educated girls. | afghanistan needs women leadership women need education taliban r afraid educated women stand afghan women demand let get educated without women education nation prosper afghanistan needs prosperity prosperity needs educated girl |
| Why Taliban in Afghanistan r so much against women education and empowerment? Who r Taliban to decide about dress and education of women? Let us stand against terrorist Taliban and their masters in Pakistan to emancipate women and Afghanistan from them. | taliban afghanistan much women education empowerment taliban decide dress education women let us stand terrorist taliban masters pakistan emancipate women afghanistan |

## 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.

## Research Management

The study challenge at hand is investigating difference in the perspectives of activists, the Taliban and the ordinary Afghan people, and how gender affect male and female activist regarding 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 activists perceptions of women’s education access to education. The dependent variable in this study is women 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 activists perceptions and beliefs regarding education for women.

# CHAPTER IV: RESEARCH RESULTS AND DISCUSSIONS

# Research Results

## Women’s Education: Activist Perceptions

In this section the researcher focuses on the view of activists both male and female, both consists of 54 different activists. The total amount of tweets which achieved from male is 17,079, while female is (34,837). Female has more tweets than male, this shows that female is more active on Twitter. At first, we use sentiment measurements as a starting point for investigation to get insight about their tweets, this is essential as starting point.

Male activists, (5,180) positive, (8,307) neutral, and (3844) negatives, and the percentage of their positive is (29.8%). While, female activists, (13582) neutral, (13,531) positive, and negative (7,724). Female are mostly impartial and favorable about different issues, including women’s education, the analysis found that positive sentiment (38.8%). Figure 1 shows, the sentiment distribution among activists, and female are slightly higher in their positive tone, while male are tend to be more neutral.

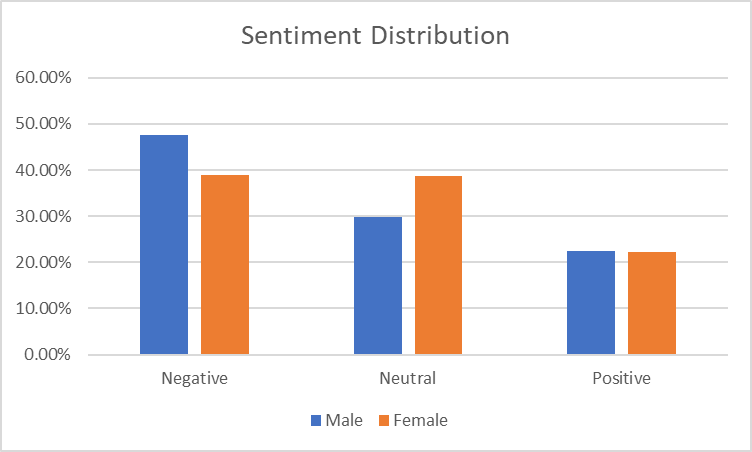


Figure 1:Sentiment Distribution (activists)

To narrow down the sentiment into five keywords which are “women,” “education,” “school,“ “university,” and “girl,” this helps us to extract the exact tones of their tweets. The sentiment analysis of tweets reveals interesting insights about how these topics are discussed. Starting with the keyword "women" male activists predominantly express a neutral sentiment, accounting for 53.76% of occurrences, see table 5 for more detail.

However, there is also a significant presence of positive sentiment at 29.11%. Negative sentiment, although relatively lower at 17.13%, still represents a notable portion of the discussions. When it comes to the keyword "education," male activists showcase a more positive outlook, with 43.27% of tweets expressing positivity. The sentiment distribution reveals that 41.62% of tweets maintain a neutral tone, while a smaller proportion of 15.11% leans towards a negative sentiment. Regarding the keyword "school," male activists tend to exhibit a relatively balanced sentiment distribution. Approximately 47.47% of tweets maintain a neutral tone, while positive and negative sentiments stand at 27.18% and 25.35%, respectively.

This suggests that opinions and feelings towards schools are varied among male activists. In the case of the keyword "university," male activists express a predominantly neutral sentiment, accounting for 56.84% of tweets. However, there is also a notable presence of positive sentiment at 27.24%. Negative sentiment represents a smaller proportion, accounting for 15.92% of occurrences. Finally, when it comes to the keyword "girl," male activists express a combination of sentiments.

The sentiment analysis reveals that 31.07% of tweets express a positive sentiment, while 47.15% maintain a neutral tone. Negative sentiment represents a significant proportion at 21.78%, indicating that discussions surrounding girls can evoke mixed reactions among male activists. Shifting the focus to female activists, the analysis indicates that they utilize these keywords more frequently than their male counterparts. Out of a total of 34,837 occurrences, women contributed 13,615 tweets related to these topics, suggesting their openness to discussing women’s education on Twitter. Analyzing the sentiment distribution among female activists, the keyword "women" elicited a mix of sentiments. While negative sentiment accounted for 19.22% of occurrences, a larger portion of 39.61% expressed a positive sentiment, indicating a more optimistic view.

Neutral sentiment constituted the majority at 41.17%. Regarding "education," women activists showcased a higher inclination towards positivity, with 48.55% of tweets expressing a positive sentiment. Negative sentiment accounted for 20.25% of occurrences, while neutral sentiment stood at 31.19%. Similar to their male counterparts, discussions around "school" among female activists exhibited a relatively balanced sentiment distribution. Negative sentiment represented 26.05% of tweets, followed by neutral sentiment at 39.98% and positive sentiment at 33.96%.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Male | | | | | | Female | | | |  |
| Sentiment | women | education | school | university | girl | women | education | school | university | girl |
| Neutral | 53.90% | 41.71% | 47.33% | 56.70% | 46.96% | 41.18% | 31.19% | 39.98% | 43.85% | 39.66% |
| Positive | 29.04% | 43.34% | 27.06% | 27.31% | 31.09% | 39.61% | 48.55% | 33.96% | 36.51% | 37.04% |
| Negative | 17.06% | 14.95% | 25.61% | 15.99% | 21.95% | 19.21% | 20.25% | 26.05% | 19.64% | 23.29% |

For the keyword "university," women activists leaned towards a neutral sentiment 43.85% as well, but positive sentiment 36.51% held a considerable presence. Negative sentiment accounted for 19.64% of occurrences. Lastly, the keyword "girl" evoked a mixture of sentiments among female activists. Negative sentiment represented 23.29% of tweets, while neutral sentiment accounted for 39.66%. Positive sentiment stood at 37.04%, highlighting a relatively optimistic outlook. We can say that men are more likely to have neutral tone, while women are more likely to express positive tone regarding the five keywords.

Table 5: Sentiment Analysis of Tweets Containing the Keyword (activists)

The sentiment analysis results reveal interesting patterns in the way both male and female participants expressed their opinions.

|  |  |  |  |
| --- | --- | --- | --- |
| Gender | Positive | Negative | Neutral |
| Male | 31.433426 | 19.524251 | 49.042323 |
| Female | 39.123090 | 20.593420 | 40.283490 |

Table 6: Calculate Sentiment for Males and Female on Five Keywords (activists)

Among male activists, a significant portion 31.43% of their sentiments was positive, while 19.52% leaned towards negativity. The majority 49.04% expressed a neutral sentiment, indicating a balanced perspective overall. In contrast, female activists exhibited a higher level of positivity, with 39.12% of their sentiments being positive. Similar to males, around 20.59% of their sentiments are negative. However, a slightly lower proportion 40.28% expressed a neutral sentiment. These findings suggest that there are differences in how males and females express their sentiments regarding the analyzed keywords. To make our result clearer, we provide the result in the format of graph, look at figure 2, the distribution for five keywords, based on the graph, female stay less negative and more positive.

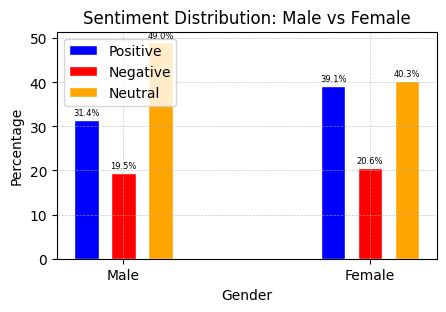


Figure 2: Sentiment Distribution Male vs Female (activists)

To completely narrow down our data, and specifically look for women’s education, we have used keyword of “women education”, this allows us to break down the data and look for those tweets which has only these two words.

|  |  |  |
| --- | --- | --- |
|  | Male | Female |
| Sentiment | Women Education | Women Education |
| Neutral | 59.09% | 46.32% |
| Positive | 31.82% | 40.00% |
| Negative | 9.09% | 13.68% |

Table 7: Sentiment Distribution on Women Education (activists)

The table 7 reveals that, among males, the sentiment appears to be predominantly neutral, with a significant percentage of 59.09%, indicating a balanced outlook or perhaps a lack of strong emotional inclination towards women’s education. Positive sentiment follows closely behind at 31.82%, suggesting a favorable perception and support for empowering women through education. However, negative sentiment among males remains present, though comparatively lower at 9.09%, highlighting some reservations or concerns regarding this topic.

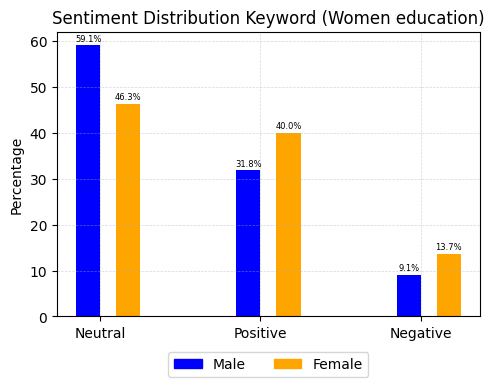


Figure 3: Sentiment Distribution on Women Education (activists)

Conversely, females exhibit a slightly different sentiment distribution. While the majority still falls under the neutral category at 46.32%, it appears that a significant proportion of females, accounting for 40.00%, express a positive sentiment towards women’s education, emphasizing their belief in the importance and benefits of educating women.

However, it is noteworthy that a slightly higher percentage of females, at 13.68%, indicate a negative sentiment, implying that some women may have reservations or face challenges regarding women’s education. These findings underscore the complexity of opinions and perspectives surrounding women’s education, with variations between genders. The table’s insights shed light on the diverse attitudes and sentiments towards this crucial aspect of education, emphasizing the need for continued efforts to promote and address the barriers faced by women in their pursuit of education.

Based on the analysis of sentiment distributions in male and female tweets, we observe the following results:

|  |  |  |
| --- | --- | --- |
| Gender | T-Statistic | P-Value |
| Male | 13.599393% | 6.610973e-42% |
| Female | 42.011182% | 0.000000e+00% |

Table 8: Calculate Sentiment Scores for Entire Dataset (activists)

Table 8 shows the sentiment score for the entire dataset, which we conducted between male and female. Even through this for the entire dataset, but it still compares both male and female. The T-statistic measures the magnitude of the difference between the means of the two groups, while the P-value represents the probability of observing such a difference by chance. In the case of male tweets, the T-statistic of 13.599393 suggests a substantial difference in sentiment compared to a hypothetical mean sentiment of 0.

The very low P-value 6.610973e-42 indicates that the probability of obtaining such a difference by random chance is extremely unlikely, providing strong evidence of a significant difference in sentiment. For female tweets, the analysis reveals an even more pronounced difference. The T-statistic of 42.011182 implies a substantial deviation from the hypothetical mean sentiment of 0. Moreover, the P-value of 0.000000e+00 (effectively zero) reinforces the notion that the observed difference is highly significant, leaving virtually no room for chance.

Further we have implemented the same method for our five keywords, which we have earlier described it. Table 6 show the analysis of sentiment in tweets related to education reveals interesting findings when considering the gender of the users. By focusing on men’s education tweets, the statistical comparison indicates a substantial difference in sentiment from the neutral standpoint. The computed t-statistic of 12.23 suggests a significant deviation from a neutral sentiment score of 0. This indicates that male’s tweets concerning education, encompassing keywords like “women,” “education,” “school,” “university,” and “girl,” tend to exhibit a distinct sentiment, whether positive or negative. Additionally, the associated p-value of 4.973316e-34 further reinforces the significance of this difference. This implies that male’s opinions and attitudes towards education, particularly when related to women, display a pronounced sentiment that differs significantly from neutrality.

|  |  |  |
| --- | --- | --- |
| Gender | T-Statistic | P-Value |
| Male | 12.231949% | 4.973316e-34% |
| Female | 28.321007% | 1.696955e-171% |

Table 9: Calculate Sentiment Scores for Five Keywords (activists)

When we refer to women’s tweets, we get a significant departure from neutrality, with a t-statistic of 28.321007 and an incredibly small p-value of 1.696955e-171. This indicates that women’s sentiments about education, as expressed in their tweets, differ significantly from neutrality. These findings highlight the distinct perspectives and attitudes of male and female when discussing education.

Understanding these divergent sentiments is crucial for addressing gender-specific challenges and promoting inclusivity in educational contexts. In short, the analysis provides strong evidence of gender-related differences in sentiment expression and emphasizes the need for considering gender perspectives in educational discourse.

To get solid result regarding women’s education, it’s time to only test “women education” keyword, both count as a single keyword.

|  |  |  |
| --- | --- | --- |
| Gender | T-Statistic | P-Value |
| Male | 2.595307% | 0.011091% |
| Female | 3.432136% | 0.000892% |

Table 10: Calculate Sentiment Scores for Women Education (activists)

Uncovering distinct sentiment patterns among male and female. Male’s tweets displayed a discernible shift from neutral sentiment, as evidenced by a t-statistic of 2.595307 and a p-value of 0.011091. Meanwhile, female’s tweets exhibited an even more pronounced deviation, indicated by a higher t-statistic of 3.432136 and a p-value of 0.000892.

These findings underscore the substantial differences in sentiment expressed by male and female when discussing women’s education. Recognizing and addressing these divergences can guide targeted approaches and foster inclusivity within educational contexts. Gender serves as a crucial factor in shaping perspectives and emotions surrounding this important topic.

The chi-square test of independence was performed separately for male and female genders. For the male gender, the test yielded a chi-square statistic of 53.38 and an extremely small p-value of 2.57e-12. Similarly, for the female gender, the chi-square statistic was 35.16, with a highly significant p-value of 2.31e-08. These results indicate a strong association between the gender of the users and the presence of tweets related to the five specified keywords. The cross-tabulation of gender and label shows that among females, 21,222 tweets were not related to the keywords, while 13,616 tweets were related. Among males, 10,934 tweets were not related to the keywords, and 6,474 tweets were related.

The significant chi-square test results suggest that the occurrence of tweets related to the specified keywords differs significantly between genders. The findings imply that there are notable differences in the engagement and discussion of the specified topics on social media platforms based on gender. Such insights can be valuable for understanding gender-specific perspectives and behaviors related to the given keywords and can aid researchers, policymakers, and educational institutions in tailoring strategies and initiatives accordingly.

## Female Activists Top Frequent Words

To discuss the female activists, the most commonly used words reflect the pressing issues and concerns prevailing in Afghanistan. The term "Taliban" takes the lead with a frequency of (13,079), underscoring the significant impact of the Taliban’s rule. "Women" follows closely with (11,054) mentions, emphasizing the focus on gender equality and women’s rights. Other frequently mentioned words include "Afghanistan" (8,665), "girl" (5,830), "Afghan" (5,339), and "people" (4,002), highlighting the broader context of societal discussions. Additionally, "rights," "school," "Kabul," and "education" feature prominently, shedding light on the key concerns related to human rights and access to education.

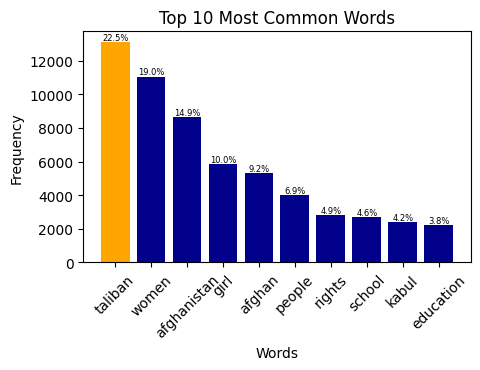


Figure 4: Female Activists Top Frequent Words

In Table 11, we can observe the frequencies of the top 10 words mentioned by female activists. Notably, the word "Taliban" holds the highest count, indicating its significance in their discussions. Following that, we find words like "women," "girl," "rights," and "schools," which align with the focus of our study. It is reassuring to see these keywords present in the dataset, as they reflect the concerns and topics that female activists actively engage with.

|  |  |
| --- | --- |
| Female Activists | Frequency Words |
| Taliban | 13,079 |
| Women | 11,054 |
| Afghanistan | 8,665 |
| Girl | 5,830 |
| Afghan | 5,339 |
| People | 4,002 |
| Rights | 2,843 |
| School | 2,694 |
| Kabul | 2,417 |
| Education | 2,239 |

Table 11: Frequent Words (female activists)

The word cloud visualization from female activists is displayed in Figure 5. The size of each word represents its frequency of occurrence in their discussions. As depicted, the word "Taliban" appears prominently, indicating its significant presence in their conversations. Additionally, we can see those words such as "women," "girl," "rights," and "schools" hold considerable importance, as they are larger in size, reflecting their frequent usage. This word cloud visualizes the key topics and concerns expressed by female activists in their discussions related to our study.

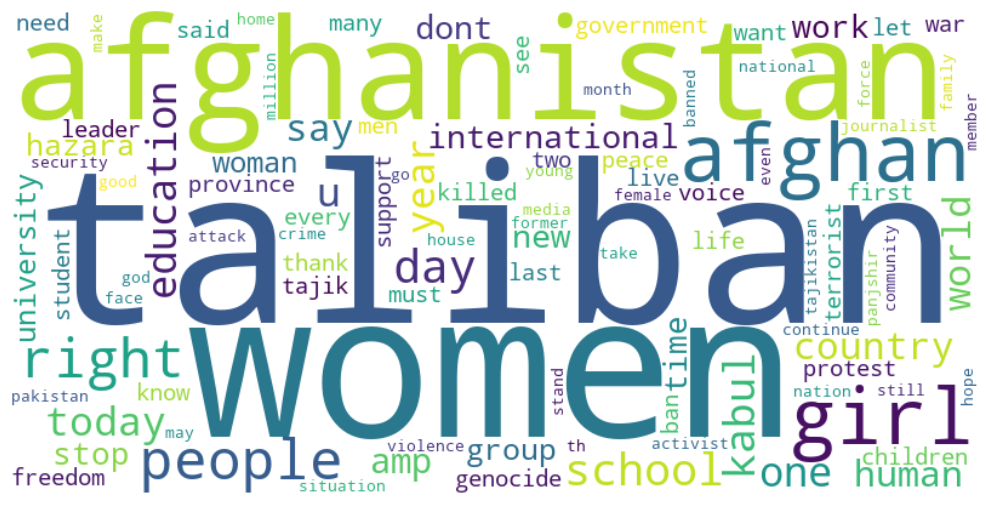


Figure 5: Female Activists Word-Cloud

To discuss the male activists, several notable terms in examining the word frequencies among male activists. "Taliban" emerges as the most frequent word same as the female activists, but with higher percentages, reflecting its significant presence in their discussions. Alongside that, we find mentions of "Afghanistan," "women," "girl," "Afghan," "education," "school," "Kabul," "Talibans," and "people." These terms collectively indicate the key topics and concerns that male activists actively engage with. It suggests that their conversations revolve around issues related to the Taliban, gender equality, education, and the broader context of Afghanistan. The prominence of these words sheds light on the focal points within the discourse of male activists.

## Male Activists Top Frequent Words

When we delve into the male activists, it becomes evident that the word "Taliban" takes center stage with a staggering frequency of 8099. This sheds light on a significant focus on addressing the actions and impact of the Taliban. Alongside this, other notable terms such as "Afghanistan" 3287, "women" 3218, "girl" 2422, and "education" 1495 emerge, revealing a strong commitment to championing gender equality and ensuring access to education. Additionally, the presence of words like "Afghan," "school," "Kabul," "Talibans," and "people" indicates the diverse range of topics that male activists engage with and advocate for.

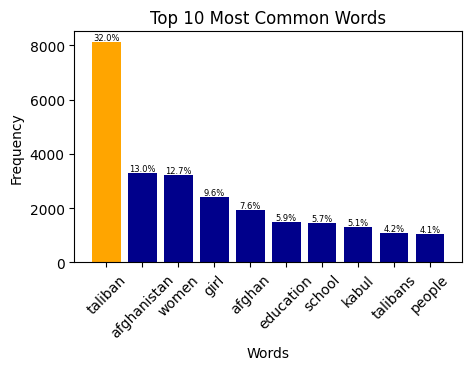


Figure 6: Male Activists Top Frequent Words

|  |  |
| --- | --- |
| Male Activists | Frequency Words |
| Taliban | 8099 |
| Afghanistan | 3287 |
| Women | 3218 |
| Girl | 2422 |
| Afghan | 1915 |
| Education | 1495 |
| School | 1444 |
| Kabul | 1291 |
| Talibans | 1072 |
| People | 1047 |

Table 12: Frequent Words (Male activists)

Also, the word cloud visualization from male activists shows a significant focus on “Taliban” followed by “Women,” “Afghanistan,” “school,” and “university,” This reflects a concern for gender equality and access to education.

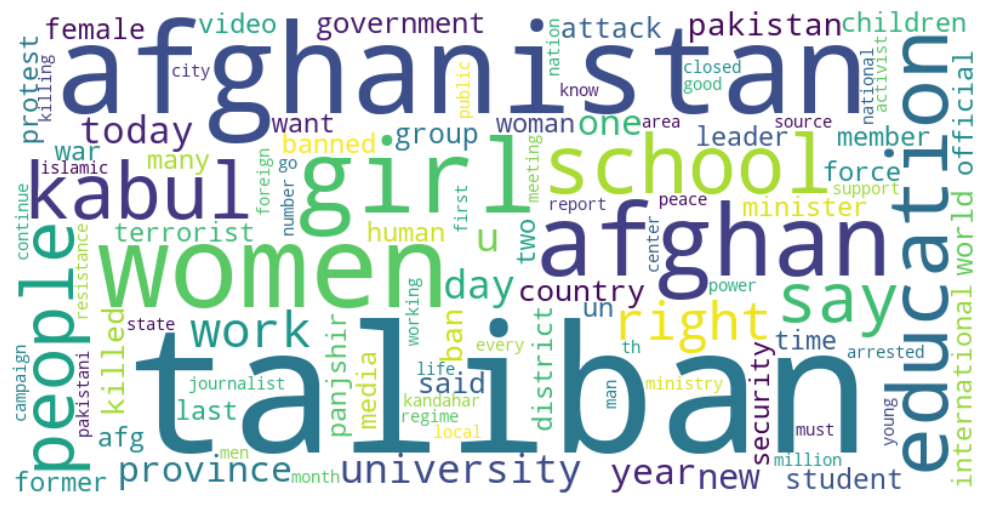


Figure 7: Male Activists Word-Cloud

In our study, we used logistic regression to examine the relationship between gender and support for women’s education. We treated gender as our independent variable (IV), where we assigned the value 1 to represent male and 0 to represent female. We began by merging the data of male and female activists using automated methods, ensuring a purely computational approach throughout. The logistic regression analysis yielded interesting findings regarding the association between gender and women’s education. Notably, the coefficient assigned to the “sex” variable indicated that females demonstrate a stronger inclination towards supporting women’s education compared to males. This discovery highlights the vital role played by female activists in advocating for and advancing women’s education. It emphasizes the importance of addressing gender disparities in educational access and underscores the need for empowering women through educational opportunities. The findings from this study shed light on the valuable contributions of female activists in promoting women’s education and serve as a call to action for fostering equal educational opportunities for all individuals.

Optimization terminated successfully.

Current function value: 0.666075

Iterations 4

Logit Regression Results

========================================================================

Dep. Variable: label No. Observations: 51916

Model: Logit Df Residuals: 51914

Method: MLE Df Model: 1

Date: Fri, 19 May 2023 Pseudo R-squ.: 0.0002553

Time: 00:26:46 Log-Likelihood: -34580.

converged: True LL-Null: -34589.

Covariance Type: nonrobust LLR p-value: 2.638e-05

========================================================================

coef std err z P>|z| [0.025 0.975]

---------------------------------------------------------------------------------------------------------------------------

const -0.4439 0.011 -40.423 0.000 -0.465 -0.422

sex -0.0809 0.019 -4.198 0.000 -0.119 -0.043

## Women’s Education: From an Activist, Taliban and Ordinary People Perceptions

This section consists of activists both male and female, Taliban officials is only male, there is no female in the cabinet of the Taliban, and also, ordinary people, which are Afghan people does not involve in politics. Except the Taliban the other two groups both consists of male and female, but our focus is not on the gender, instead we analyze them in a group, based on their background, like “activist, people, and Taliban”. At first, we would have a look at their tweets in group, then we move forward and analyze them in dept, using keywords, “education”, “school”, “university”, “women”, “girl”, and “women education”. These keywords narrow down the dataset and analyze only those tweets which related to these certain words.

This section is solely for activists, ordinary people and Taliban, in the rest of the section instead of (ordinary people) we use only people which is equivalent to (ordinary people) in this study. Taliban has a total of tweets 75,971 followed the people which is 57,689, and activists 52,247, Taliban has the largest total of tweets which shows how the Taliban active on Twitter platform, even though we have the same sample for each group. Further, figure 8 shows, a sentiment distribution among their tweets, according to the figure, Taliban has the dominant positive tweets and smallest neutral as well, the smallest negative tweets.

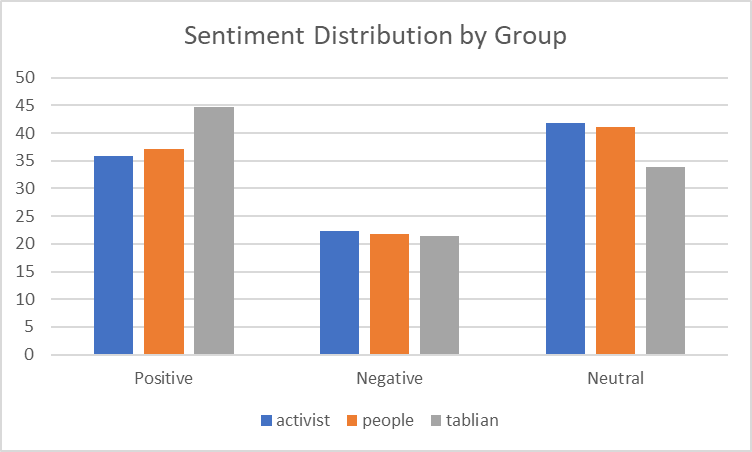


Figure 8: Sentiment Distribution of (activists, taliban, people)

Table 13 shows, the exact percentage of each group, activists and people has the same value on neutral side, while Taliban’s are seeming to be more positive. Table 13, and figure 8 are the same exact result but different approaches, one in chart format, while other in a table.

|  |  |  |  |
| --- | --- | --- | --- |
| Group | Positive | Negative | Neutral |
| Activist | 35.894522% | 22.282148% | 41.823330% |
| People | 37.163793% | 21.721719% | 41.114488% |
| Tablian | 44.766355% | 21.329472% | 33.904173% |

Table 13: Sentiment Distribution of (activists, taliban, people)

In our analysis of tweets related to the keywords "women," "education," "school," "university," and "girl," we aimed to gain insights into the level of support for women’s education in Afghanistan. However, it is important to note that the sentiment distribution for the Taliban was unavailable in the provided data, as all values were NaN (0). Meanwhile, our analysis primarily focuses on the sentiments of activists and the general public, including the Taliban, but for Taliban received NaN (0).

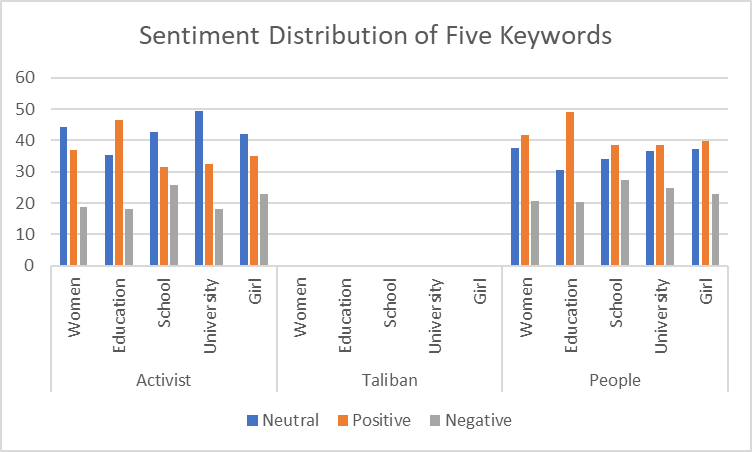


Figure 9: Sentiment Distribution of five Keywords (activists, taliban, people)

When examining the sentiment distribution among activists, we observed varying opinions on the issue of women’s education. For the keyword "women," the sentiment distribution showed diverse views. Approximately 44.42% of the tweets expressed a neutral sentiment, indicating a lack of clear stance or mixed opinions. On the positive side, 36.91% of the tweets supported women’s issues, emphasizing the importance of education and empowerment. However, it is worth noting that around 18.68% of the tweets reflected a negative sentiment and expressed reservations or opposition towards women’s education.

Moving on to the keyword "education," we found that among activists, 35.43% of the tweets had a positive sentiment, highlighting the value and significance of education for women. Conversely, 18.16% of the tweets expressed a negative sentiment, suggesting concerns or objections regarding women’s access to education. Notably, 46.41% of the tweets remained neutral, indicating a lack of strong sentiment or a balanced perspective among activists. Regarding the keyword "school," the sentiment distribution showed a similar pattern. Activists expressed a mix of sentiments, with 42.67% of the tweets being neutral, 31.53% positive, and 25.80% negative. This suggests that while a significant proportion of activists support the idea of women attending school, there are also concerns or criticisms expressed by a notable portion of the community.

For the keyword "university," the sentiment distribution among activists revealed that 49.43% of the tweets had a neutral sentiment, indicating a lack of a clear stance. However, 32.52% of the tweets were positive, demonstrating support for women pursuing education. Around 18.04% of the tweets displayed a negative sentiment, reflecting reservations or objections towards women’s participation in university education. Lastly, focusing on the keyword "girl," the sentiment distribution among activists showcased a similar trend to the previous keywords. Approximately 42.06% of the tweets had a neutral sentiment, while 35.13% expressed a positive sentiment, emphasizing the significance of supporting girls’ education. Around 22.81% of the tweets displayed a negative sentiment, indicating concerns or criticisms about girls’ education access.

However, it is crucial to acknowledge the absence of sentiment data for the Taliban in our analysis, which limits our ability to assess their stance on women’s education. It would be valuable to obtain sentiment data from the Taliban’s perspective to understand the different groups’ viewpoints comprehensively. In the sentiment distribution among people, we find that for the keywords of women, education, school, university, and girl, the sentiment percentages vary. When it comes to women, approximately 37.72% of the people express a positive sentiment, while around 20.70% hold a negative sentiment, and the remaining 41.59% maintain a neutral stance. Regarding education, about 30.45% of people express positivity, 20.39% hold a negative sentiment, and 49.16% maintain a neutral viewpoint. Similarly, sentiments related to school, university, and girl show variations in the percentages of positive, negative, and neutral sentiments among the people.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Activist | | | | | | Taliban | | | | | People | | | | |
| Sentiment | Women | Education | School | University | Girl | Women | Education | School | University | Girl | Women | Education | School | University | Girl |
| Neutral | 44.42 | 35.43 | 42.67 | 49.43 | 42.06 | Nan | Nan | Nan | Nan | Nan | 37.72 | 30.45 | 34.17 | 36.69 | 37.20 |
| Positive | 36.91 | 46.41 | 31.53 | 32.52 | 35.13 | Nan | Nan | Nan | Nan | Nan | 41.59 | 49.16 | 38.45 | 38.59 | 39.75 |
| Negative | 18.68 | 18.16 | 25.80 | 18.04 | 22.81 | Nan | Nan | Nan | Nan | Nan | 20.70 | 20.39 | 27.37 | 24.72 | 23.05 |

Table 14: Sentiment Analysis of Tweets Containing the Keyword (activists, taliban, people)

In Table 14, we can observe the percentage breakdown of each keyword, highlighting an unfortunate absence of data for the Taliban group. There are a couple of potential explanations for this absence. It could be that the Taliban either refrains from tweeting about these specific keywords or their involvement in discussions related to women, education, school, university, and girl is minimal, resulting in negligible representation in the collected data. This underscores the significance of the data obtained from the other two groups, providing valuable insights into their perspectives.

|  |
| --- |
| Tweets Related to Keywords |
| jobs allocated women educational sector today announced ministry education jobs teachers administrative staff male female states herat farah ghor badghis nangarhar laghman kunar nuristan |
| many years uzbekistan hijab allowed women educational institutions |
| public private universities throughout country open men women education process extended percent remaining parts country thousands women education higher education public health id passport offices airports police media |

Table 15: Tweets Related to five Keywords (activists, taliban, people)

Among the tweets we analyzed from the activists and people groups, we have curated a collection of the most suitable tweets corresponding to our chosen keywords. It is worth mentioning that a significant proportion of their tweets predominantly discuss subjects like Islam and the official visits of their representatives to foreign countries, which may not directly relate to our specific investigation on women’s education. Nonetheless, we have diligently selected tweets that closely align with our research topic, ensuring their relevance and significance.

To narrow down our data and focus on keyword “women education” only, we found a number of noteworthy patterns emerge. We first put our attention to the Taliban group, unfortunately, the sentiment distribution data is not available NaN (0). This lack of information hinders our understanding of the Taliban’s stance on women’s education specifically. It could imply that they do not engage in discussions or tweets about women’s education, or their stance is not reflected in the dataset. Again, you can take table 14, the best tweet related to education on their dataset.

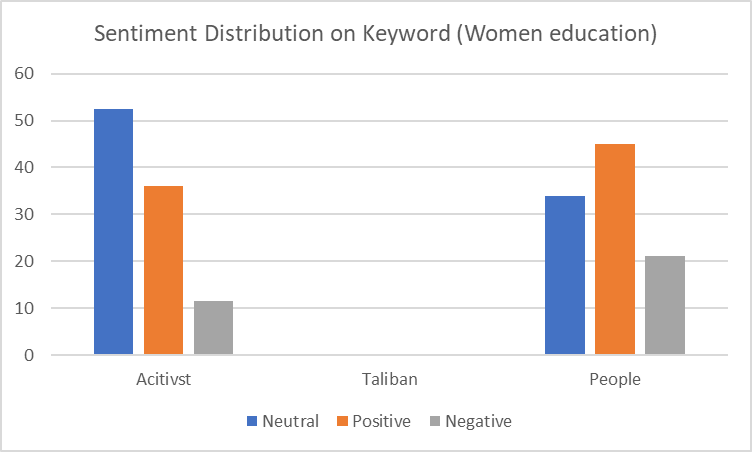


Figure 10: Sentiment Distribution on Women Education (activists, taliban, people)

Considering the sentiment distribution among people, it is evident that a significant proportion of tweets, 33.94%, express a neutral sentiment. This suggests diverse opinions and perspectives among the general public regarding women’s education. On the positive side, 44.95% demonstrate a supportive sentiment towards women’s education, indicating a widespread belief in the importance of empowering women through educational opportunities. Conversely, 21.1% of the tweets are negative, highlighting concerns or reservations people may have regarding various aspects of women’s education.

|  |  |  |  |
| --- | --- | --- | --- |
| Women education | Activist | Taliban | People |
| Neutral | 52.46% | NaN | 33.94% |
| Positive | 36.07% | NaN | 44.95% |
| Negative | 11.48% | NaN | 21.10% |

Table 16: Sentiment Distribution on Women Education (activists, taliban, people)

Based on the analysis of the sentiment distribution on the specific keyword "women education," activists show a mixed sentiment with a notable emphasis on support and advocacy. However, the lack of available data on the Taliban group limits our understanding of their stance. The general population’s sentiment distribution reveals diverse opinions, with a significant portion expressing support for women’s education while acknowledging some negative sentiments. These findings shed light on the complexities and varying perspectives surrounding women’s education within different groups, underscoring the need for further exploration and dialogue on this crucial topic.

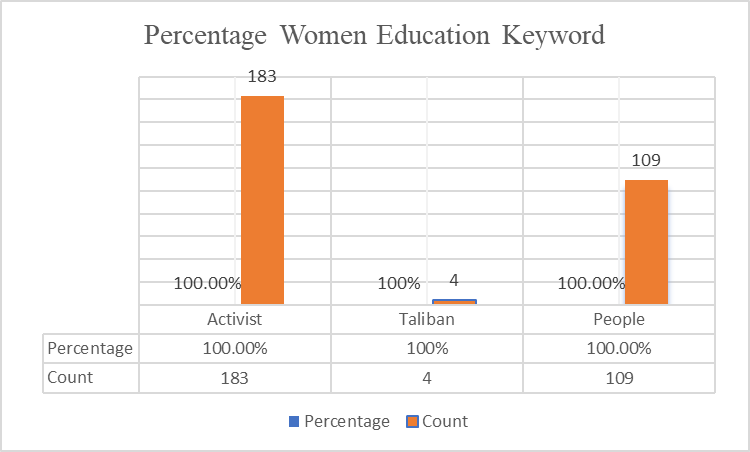
**

Figure 11: Percentage Women Education Keyword (activists, taliban, people)

In addition to recover Table 16, we conducted a more detailed investigation focusing on the keyword "women education." Our findings revealed that only four tweets mentioned "women education." The percentage for this keyword does not appear in Table 15 because either these tweets did not express a positive sentiment or the sample size was too small to calculate a meaningful percentage. These findings are based on our analysis compared to the other two groups.

|  |  |  |
| --- | --- | --- |
| Group | Percentage | Count |
| Activist | 100.00% | 183 |
| Taliban | 100.00% | 4 |
| People | 100.00% | 109 |

Table 17: Percentage and Count on Women Education (activists, taliban, people)

Based on Table 17, we found that people generally have a more positive sentiment toward "women’s education." However, activists tend to express their support for women’s education more passionately, resulting in a lower positive sentiment but a higher number of related tweets compared to the general population, table 15 shows, activists has tweeted (183) related to “women education,” followed by people (109) and Taliban which is (4).

## Analyzing the Frequency of Keywords Regarding Women’s Education

The figure 12 provides valuable insights regarding our keywords, namely "women," "education," "school," "university," and "girl." It illustrates the frequency of these keywords’ usage by different groups over the course of each year. Based on the figure, we can observe that activists have consistently used these keywords more frequently in their tweets, particularly between October 2022 and January 2023. These findings suggest that the activist group has shown higher engagement and emphasis on these topics during those specific time frames.

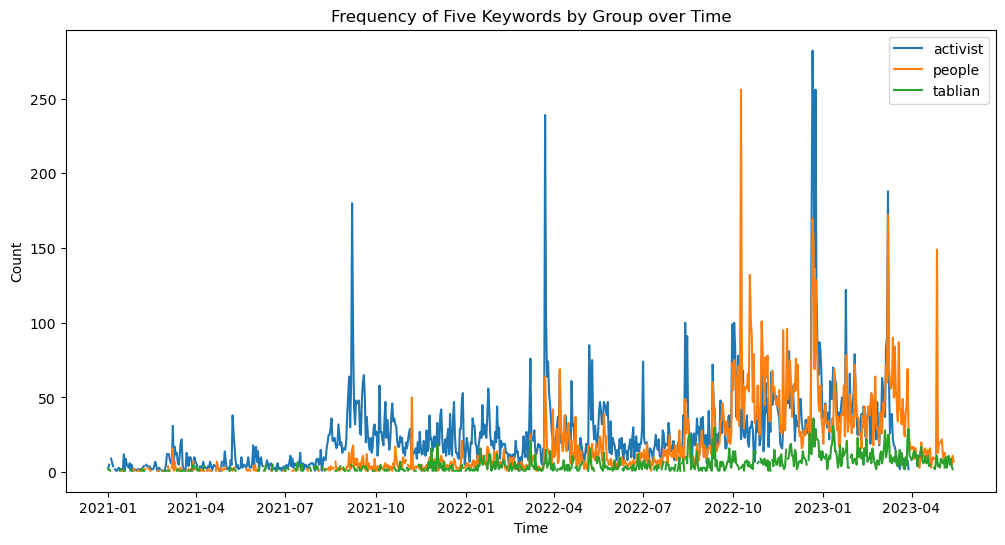


Figure 12: Frequency of Five Keywords by Group over Time (activists, taliban, people)

Additionally, the analysis reveals that the people has noticeably utilized the mentioned keywords. This indicates that people are actively engaged in discussions related to education and express support for women’s rights and equality in education. On the other hand, the Taliban has exhibited a considerably lower frequency of tweets containing these keywords. This suggests that the Taliban group may not prioritize or demonstrate concern for women’s education based on their limited usage of these keywords.

Last but not least, the results of the statistical test reveal a significant relationship between the groups and the label values. The chi-square statistic, which measures the discrepancy between the observed and expected frequencies, has a high value of (26655.35). This indicates a substantial difference between the groups regarding their support or opposition towards women’s education. The p-value, which assesses the statistical significance, is extremely low at (0.0). This suggests that the likelihood of observing such a large discrepancy between the groups by chance alone is negligible. Therefore, we can conclude that there is a significant association between the groups and their stance on women’s education, indicating varying levels of support or opposition across the different groups.

## Taliban Top Frequent Words

The Taliban’s most frequently used terms reveal their point of view. Words like "Afghanistan," "people," and "Islamic" reveal their connection with the region, their religion, and their political leanings, respectively. They appear to be leading towards an Islamic state, as terms like "minister," "country," and "emirate" all point that way. The words "ministry," "god," "said," and "foreign" all allude to their diplomatic and theological concerns. These reoccurring expressions shed light on the Taliban’s overarching goals and ideas.

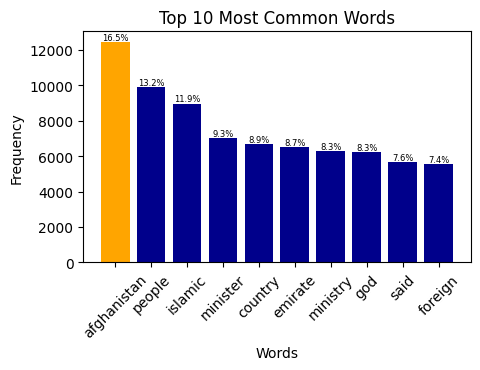


Figure 13: Taliban Top Frequent Words

The word cloud for the Taliban reveals key themes and priorities. It prominently includes words such as "Afghanistan," "people," and "Islamic," highlighting their close association with the country and their religious ideology. Terms like "minister," "country," and "emirate" suggest their involvement in governance and their pursuit of an Islamic state. This word cloud offers a concise glimpse into the Taliban’s central focus and core values.

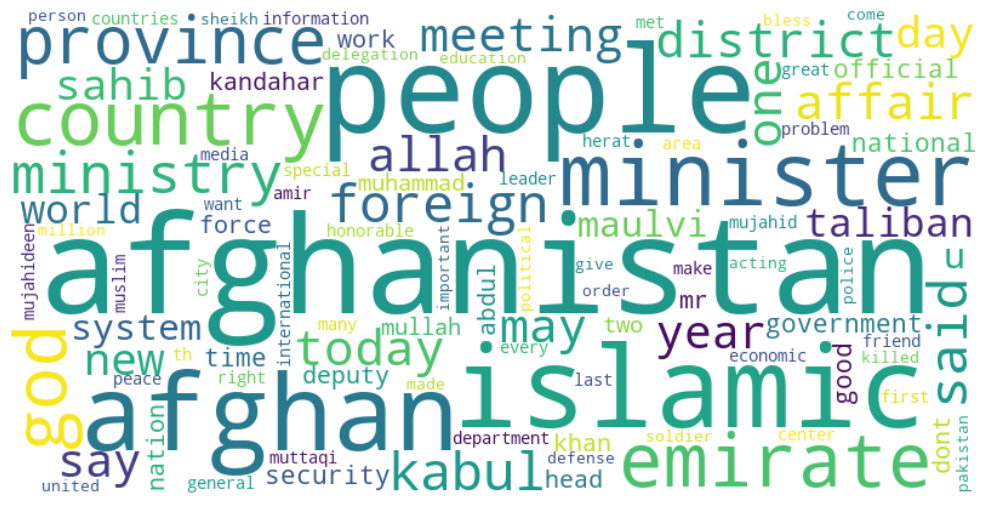


Figure 14: Taliban Word-Cloud

## Female Ordinary People Top Frequent Words

Ordinary women’s top 10 words represent their worries and activism. "Hazara," "genocide," and "stop" show their awareness and activism. "Afghanistan," "women," and "people" demonstrate their national identity. "Taliban" shows their resistance to the gang, while "girl" underlines their empowerment of young women. These remarks reveal the social and political priorities of regular women.

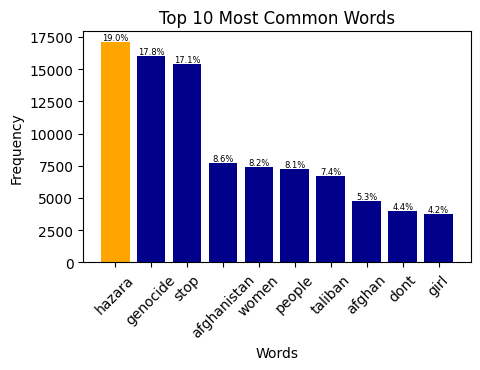


Figure 15: Female Ordinary People Top Frequent Words



Figure 16: Female Ordinary Word-Cloud

Male ordinary people, word cloud expresses their concerns, about Afghanistan and the Taliban. "Afghanistan," "Taliban," and "people" show their focus on the country’s predicament and its impact on Afghans. "Hazara," "genocide," and "Pakistan" indicate their grasp of regional affairs. "Stop," "country," and "one" emphasize their patriotism, action, and individuality. The word cloud the most pressing issues and areas of interest, including Afghanistan, the Taliban, and the Hazara people. This shows that they care about ending genocide, Pakistan, and violence. The cloud symbolizes the people’s love of their homeland along with their dedication to maintaining tranquility there.



Figure 17: Male Ordinary Word-Cloud

# CHAPTER V: GENDER AND WOMEN'S EDUCATION: A CLOSER LOOK AT THE ACTIVISTS

The struggle for women access to education has been there for quite long time, with activists playing an important role in supporting women’s eduction and push society for gender equality. The study have analyzed the differences in the perspectives of activists, Taliban and ordinary Afghan people on women access to education, alos, how gender affect activists both male and female perceptions about women’s education.

It is important to recognize the differences in the experiences of male and female activists when fighting for equal educational opportunities for women. Women’s rights advocates have first-hand experience with the discrimination, societal standards, and other factors that make it difficult for women to further their education. The challenge of overcoming such obstacles strengthens their resolve to work toward educational equity. They are highly motivated to take action because of their compassion for and identification with women’s difficulties.

However, male activists take a different tack when discussing the topic of women’s education. They are aware of the inequalities that exist between the sexes and value gender equality highly. They recognize that empowering women through education improves communities and countries as a whole. Male activists who support women’s education seek to eliminate discrimination and promote equality for all.

# VI Discussion

The discourse highlights a conspicuous inequality between the perspectives of activists, Taliban, and ordinary Afghan people concerning the education of women, as indicated by the research findings. Advocates demonstrate keen backing and promotion of women’s education, underscoring the significance of equitable access to educational opportunities.

The unwavering endeavors of the individual in question indicate their dedication toward the progression of gender equality in the realm of education, as well as their advocacy for societal transformation. Conversely, the limited involvement of the Taliban with regard to the education of women is a cause for concern. The absence of significant attention given to this crucial matter implies a disregard for the promotion of women’s empowerment and rights through educational means.

The conspicuous contrast in perspectives underscores the divergent ideologies and priorities held by these two factions. Moreover, the substantial backing exhibited by the general public towards the education of women reflects a more extensive societal preoccupation and aspiration for equality. The results emphasize the significance of enhancing the advocacy efforts of activists and implementing focused measures to tackle the inequalities in women’s education. It is imperative for policymakers, educators, and organizations to engage in collaborative efforts toward the development of inclusive policies, community engagement, and the establishment of strong support systems.

In addition to the discussion surrounding activists and the Taliban, it is crucial to consider the perspectives of ordinary people regarding women’s education. The findings highlight their substantial support for women’s education, indicating a broader societal concern and recognition of the importance of equal access to education for women. This support signifies a collective desire for positive change and highlights the potential for widespread social transformation.

Understanding the perspectives of ordinary people is vital for policymakers, as it provides insight into the societal landscape and helps identify areas where interventions and policies can be implemented to address barriers and promote gender equality in education. By harnessing this support and actively involving ordinary people in initiatives, it becomes possible to create a more inclusive and supportive environment for women’s education. The study’s findings underscore the need for collaboration among policymakers, educators, activists, and ordinary people to collectively work towards dismantling barriers and ensuring that every woman has the opportunity to access quality education, empowering them to contribute to their communities and achieve their full potential.

# HAPTER V: CONCLUSION AND SUGGESTIONS

# Conclusion

The study carries several important reasons; first and foremost, the study focused on women’s education, and we clarified the perspectives of various organizations, including activists, ordinary people, and the Taliban. Knowing the extent of support and public interest in women’s educational opportunities requires an appreciation of these viewpoints. Second, we emphasize the significance of activism in fighting for women’s access to education. When it’s about gender equality and promoting education activist play an important role and raising awareness, and advocating for policy reforms. This study analyzes the involvement and backing of activists to highlight the need to amplify their voices and support their activities.

Third, the research sheds light on how various communities prioritize and interpret the world. It shows how campaigners and the Taliban view women’s education very differently. This insight is critical for politicians, educators, and groups working to advance women’s access to education because it illuminates the obstacles that must be overcome to realize gender equality in the society. The importance of ordinary people’s backing for women’s education is also emphasized. The results indicate a cultural concern and desire for women’s educational equality.

This highlights the need for supportive legislation, active communities, and social networks to facilitate women’s access to education. The relevance and connections found are also supported by statistical tests used in the study, such as the chi-square test, cross-table along other techniques that enable us to analyze tones of tweets. The analysis is strengthened, and the links between the variables are better understood, as a result of this. This research is significant because it adds to the existing body of knowledge on women’s education, sheds light on the roles different groups plays in developing attitudes and support, and offers guidance to those working to ensure that women and men have equal access to educational opportunities. This study can inform targeted actions, policies, and advocacy activities to improve women’s education and empower women via knowledge and skills by analyzing the current situation and identifying gaps.

Throughout the analysis of the data and results, it is evident that different groups, namely activists, ordinary people, and the Taliban, hold distinct attitudes and sentiments towards women’s education. Activists stand out with their notable support for women’s education, as reflected in their frequent use of keywords related to women, education, school, university, and girl. This suggests their active involvement in advocating and promoting women’s access to education. The chi-square test confirms the strong association between activists and their support for women’s education, underscoring their consistent engagement and vocal commitment to advancing gender equality in education. On the other hand, the ordinary people group has shown a significant level of support for women’s education as well. Although their frequency of tweets related to women’s education is slightly lower compared to activists, it is evident that they share a genuine interest and concern. The word-cloud analysis reveals important themes such as the Hazara community, genocide, and the call to stop violence, indicating their awareness of the obstacles women face in education and their strong desire for positive transformations. The chi-square test results further solidify the association between ordinary people and their support for women’s education, confirming their active involvement and commitment to bringing about change.

Surprisingly, the Taliban group showed very limited engagement and concern when it comes to women’s education. Their frequency of tweets related to the specified keywords was significantly lower compared to activists and ordinary people. This observation highlights a clear inequality in priorities and ideologies between the Taliban and the other groups. It indicates that the Taliban does not prioritize or emphasize women’s education in their messaging, reflecting a disregard for the significant role education plays in empowering women. The word-cloud analysis for the Taliban further reinforces this finding, as their focus primarily revolves around Afghanistan, people, and Islamic values, with minimal mention of women’s education. This underscores the need for targeted efforts to address this gap and promote gender equality in education within the Taliban context. The analysis reveals a stark contrast between activists and the Taliban regarding their attitudes towards women’s education. Activists actively support and advocate for women’s education, while the Taliban’s engagement is lacking. This difference underscores opposing ideologies. Strong support from ordinary people further emphasizes societal concern and the desire for equality. Policymakers, educators, and organizations should address these findings to improve women’s education and promote gender equality. Supporting and amplifying activists’ voices is crucial. The data highlights the need for targeted interventions, policies, and continued advocacy to ensure equal access to education for all women.

# 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.

Allan, A., & Tinkler, P. (2015). ‘Seeing’ into the past and ‘looking’ forward to the future: visual methods and gender and education research. *Gender and Education*.

Allen, B. J. (2017). Standpoint Theory. *American Journal of Political Science*, 1.

Anderson, C. L., Reynolds, T. W., Biscaye, P., Patwardhan, V., & Schmidt, C. (2021). Economic Benefits of Empowering Women in Agriculture: Assumptions and Evidence. *The Journal of Development Studies 57*, 200.

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.

Butler, J. (2006). *Gender Trouble: Feminism and the Subversion of Identity.* 270 Madison Avenue, New York, NY 10016: Routledge.

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

Cin, F. M. (2017). *Gender Justice, Education and Equality.* Istanbul , Turkey: Palgrave Macmillan.

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

Domagoj Margan, A. M. (2015). LaNCoA: A Python Toolkit for Language Networks Construction and Analysis. *In 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics*.

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.

El Barachi, M., AlKhatib, M., Mathew, S., & Oroumchian, F. (2021). A novel sentiment analysis framework for monitoring the evolving public opinion in real-time: Case study on climate change. *Journal of Cleaner Production*, 1.

Evans, H. K., & Clark, J. (2015). “You Tweet Like a Girl!”: How Female Candidates Campaign on Twitter. *American Politics Research*, 10.

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.

Freedman, E. (1979). Separatism as Strategy: Female Institution Building and American Feminism, 1870-1930. *Feminist Studies*.

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.

Harding, S. (2004). A Socially Relevant Philosophy of Science? Resources from Standpoint Theory's Controversiality. *Hypatia 19, no. 1 (2004): 25-47*.

Hasse-Biber, S. N. (2013). *Feminist Research Practice: A Primer.* United States of America: SAGE Publications, Inc.

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

Hurwitz, H. M., & Taylor, V. (2012). Women’s Cultures and Social Movements in Global Contexts. *American Journal of Political Science*.

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.

Karamouzas, D., Mademlis, I., & Pitas, I. (2022). Public opinion monitoring through collective semantic analysis of tweets. *Social Network Analysis and Mining*, 1.

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

Kihlstrom, J. F., & Cantor, N. (2000). Social Intelligence. *Cambridge, England: Cambridge University Press*.

Kitzie, V., & Ghosh, D. (2016). #Criming and #Alive: Network and content analysis of two sides of a story on twitter. *American Journal of Political Science*, 2.

Lewis, D., & Burke, C. J. (1949). Categorical Data and Chi-Square Tests. *American Journal of Political Science*.

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*.

Parson, L. (2020). Researching Gender and Higher Education. *American Journal of Political Science*, 515.

Pedregosa, F. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 2826.

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

Plutzer, L. A. (1993). Contextual Determinants of Feminist Attitudes: National and Subnational Influences in Western Europe. *American Political Science Association*, 149.

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

Roby, J. L., Lambert, M. J., & Lambert, J. (2009). Barriers to girls’ education in Mozambique at household and community levels: an exploratory study. *American Journal of Political Science*.

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

Ropers-Huilman, R., & Winters, K. T. (2011). Feminist Research in Higher Education. *The Journal of Higher Education 82, no. 6 (2011): 667-690*, 671.

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.

Scott, R., & Kosslyn. (2015). Content Analysis. *American Journal of Political Science*, 9.

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.

Taylor, V., & Rupp, L. J. (1993). Women's Culture and Lesbian Feminist Activism: A Reconsideration of Cultural Feminism. *University of Chicago Press*.

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

UNESCO. (2013). Summary Report Of The 2013 UIS Cultural Employment Metadata Survey. *UNESCO Institute for Statistics*.

UNICEF. (2006). The State of The World's children 2007. *Division of Communication, UNICEF NY (3 UN Plaza, NY, NY 10017) USA*.

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 18. 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()

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 19. 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 20. 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

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 21. Top 10 Trends

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 22. Remove Hashtags

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 23. Expand Contractions

The table below, shows that our contractions has changed to complete words. This is important in data cleansing as well.

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 24.Remove ASCII Characters

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 25. Remove Whitespace

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 26. 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 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 27. Remove Punctuations

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)](https://corporatefinanceinstitute.com/resources/knowledge/other/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 28. Contingency Code

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

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)

Table 29. Transforming Textual to numeric

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 30. Evaluate Performance

1. def label\_tweets(text):

2.     keywords = ['education', 'school', 'university', 'women', 'girl']

3.     for keyword in keywords:

4.         if keyword in text.lower():

5.             return 1

6.     return 0

7. df = df.copy()

8. df['label'] = df['text'].apply(label\_tweets

Table 31: Labeling Data

1. x = df['text']

2. sentiment\_labels = []

3. sentiment\_scores = []

4. for tweet in x:

5.     # Create a TextBlob object

6.     blob = TextBlob(tweet)

7.     # Get the sentiment polarity (-1 to 1)

8.     polarity = blob.sentiment.polarity

9.     if polarity > 0:

10.         sentiment = 'positive'

11.     elif polarity < 0:

12.         sentiment = 'negative'

13.     else:

14.         sentiment = 'neutral'

15.

16.     sentiment\_labels.append(sentiment)

17.     sentiment\_scores.append(polarity)

18. # Add the sentiment labels and scores as new columns in the DataFrame

19. df['sentiment'] = sentiment\_labels

20. df['sentiment\_score'] = sentiment\_scores

21.

Table 32: Sentiment Analysis

1. # Define custom colors for sentiment categories

2. colors = {'negative': 'red', 'neutral': 'blue', 'positive': 'orange'}

3.

4. # Calculate sentiment counts

5. sentiment\_counts = df['sentiment'].value\_counts()

6.

7. # Plot the sentiment counts with custom colors

8. plt.figure(figsize=(8, 6))

9. sentiment\_counts.plot(kind='bar', color=[colors.get(x, 'gray') for x in sentiment\_counts.index])

10. plt.xlabel('Sentiment Category')

11. plt.ylabel('Count')

12. plt.xticks(rotation=0)

13. plt.title('Sentiment Distribution')

14. plt.show()

15.

16. # Calculate total counts

17. total\_counts = sentiment\_counts['positive'] + sentiment\_counts['negative']

18. print("Sentiment Distribution:", total\_counts)

Table 33: Sentiment Distribution

1. women\_positive\_count = df[(df['gender'] == 'male') & (df['sentiment'] == 'positive')].shape[0]

2. women\_total\_count = df[df['gender'] == 'male'].shape[0]

3. women\_positive\_percentage = (women\_positive\_count / women\_total\_count) \* 100

4. # Print the results

5. print("Positive tweets from taliban:", women\_positive\_count, "(", women\_positive\_percentage, "%)")

Table 34: Calculate Positive Tweets

1. coding\_scheme = {

2. 'women': ['women', 'woman', 'female', 'feminine', 'girl', 'lady'],

3. 'education': ['education', 'educational', 'learning', 'academics', 'knowledge'],

4. 'school': ['school', 'classroom', 'teacher', 'students', 'curriculum'],

5. 'university': ['university', 'college', 'higher education', 'academia', 'campus'],

6. 'access': ['access', 'opportunity', 'enrollment', 'attendance', 'entry'],

7. 'quality': ['quality', 'standard', 'curriculum', 'teaching', 'learning'],

8. 'benefits': ['benefit', 'advantage', 'improvement', 'progress'],

9. 'barriers': ['barrier', 'restriction', 'obstacle', 'challenge', 'difficulty'],

10. 'empowerment': ['empower', 'empowering', 'confidence', 'self-esteem', 'agency'],

11. 'gender': ['gender', 'discriminatioan', 'inequality', 'feminism', 'patriarchy']

12. }

13.

14. # Define a function to apply the coding scheme to each tweet

15. def apply\_coding\_scheme(text):

16.     codes = []

17.     for code, keywords in coding\_scheme.items():

18.         for keyword in keywords:

19.             if keyword in text.lower():

20.                 codes.append(code)

21.     return codes

22.

23. # Apply the coding scheme to the text column and create a new column for the codes

24. df\_edu['codes'] = df\_edu['text'].apply(apply\_coding\_scheme)

25.

Table 35: Code Scheme

1. from nltk.sentiment import SentimentIntensityAnalyzer

2. import nltk

3. # Initialize the sentiment analyzer

4. sid = SentimentIntensityAnalyzer()

5. # Loop through each keyword and calculate the sentiment of tweets containing that keyword

6. for keyword in ['women', 'education', 'school', 'university', 'girl']:

7.     keyword\_tweets = df\_edu[df\_edu['text'].str.contains(keyword)].copy()

8.     keyword\_tweets.loc[:, 'sentiment\_score'] = keyword\_tweets['text'].apply(lambda x: sid.polarity\_scores(x)['compound'])

9.     print(f"Sentiment analysis of tweets containing the keyword '{keyword}':")

10.     print(keyword\_tweets.groupby('sentiment')['text'].count())

Table 36: Anlyzing Five Keywords

1. from nltk.sentiment import SentimentIntensityAnalyzer

2. import nltk

3. import pandas as pd

4.

5. # Initialize the sentiment analyzer

6. sid = SentimentIntensityAnalyzer()

7.

8. # Create a DataFrame to store the sentiment percentages for each keyword

9. keyword\_sentiments = pd.DataFrame(columns=['Positive', 'Negative', 'Neutral'])

10.

11. # Loop through each keyword and calculate the sentiment of tweets containing that keyword

12. for keyword in ['women', 'education', 'school', 'university', 'girl']:

13.     keyword\_tweets = df\_edu[df\_edu['text'].str.contains(keyword)].copy()

14.     keyword\_tweets['sentiment\_score'] = keyword\_tweets['text'].apply(lambda x: sid.polarity\_scores(x)['compound'])

15.

16.     # Calculate the sentiment distribution for the keyword

17.     sentiment\_distribution = keyword\_tweets.groupby('sentiment')['text'].count()

18.     total\_tweets = sentiment\_distribution.sum()

19.

20.     # Calculate the percentage of each sentiment category

21.     sentiment\_percentage = sentiment\_distribution / total\_tweets \* 100

22.

23.     # Add the sentiment percentages to the DataFrame

24.     keyword\_sentiments.loc[keyword] = sentiment\_percentage.values

25.

26. # Calculate the average sentiment percentages across all keywords

27. average\_sentiments = keyword\_sentiments.mean()

28.

29. # Print the average sentiment percentages

30. print("Average sentiment percentages:")

31. for sentiment, percentage in average\_sentiments.items():

32.     print(f"{sentiment}: {percentage:.2f}%")

Table 37: Average Sentiment Percentages

1. sentiment\_scores = [TextBlob(tweet).sentiment.polarity for tweet in df['text']]

2. # Compare the sentiment distributions using a t-test

3. t\_stat, p\_value = stats.ttest\_1samp(sentiment\_scores, 0)

4. print("Comparison of Sentiment whole tweets:")

5. print("T-statistic:", t\_stat)

6. print("P-value:", p\_value)

Table 38: Comparison of Sentiment Whole Tweets

1. import scipy.stats as stats

2. from textblob import TextBlob

3. keywords = ['women', 'education', 'school', 'university', 'girl']

4. # Filter tweets containing the keywords

5. men\_education\_tweets = df[df['text'].str.contains('|'.join(keywords), case=False)]

6. # Calculate sentiment scores using TextBlob

7. sentiment\_scores = [TextBlob(tweet).sentiment.polarity for tweet in men\_education\_tweets['text']]

8. # Perform one-sample t-test

9. t\_stat, p\_value = stats.ttest\_1samp(sentiment\_scores, 0)

10. # Print the comparison results

11. print("Comparison of Sentiment for five keywords:")

12. print("T-statistic:", t\_stat)

13. print("P-value:", p\_value)

Table 39: Comparison of Sentiment Five Keywords

1. taliban\_education\_tweet = df[df['text'].str.contains('women education', case=False)]

2. sentiment\_scores = [TextBlob(tweet).sentiment.polarity for tweet in taliban\_education\_tweet['text']]

3. t\_stat, p\_value = stats.ttest\_1samp(sentiment\_scores, 0)

4. print("Comparison of Sentiment for Women Education Tweets:")

5. print("T-statistic:", t\_stat)

6. print("P-value:", p\_value)

Table 40: Comparison of Sentiment on Keyword Women Education

1. from scipy.stats import chi2\_contingency

2. contingency\_table = pd.crosstab(df['gender'], df['label'])

3. chi2, p\_value, \_, \_ = chi2\_contingency(contingency\_table)

4. print("Chi-square Test of Independence:")

5. print("Chi-square:", chi2)

6. print("P-value:", p\_value)

Table 41: Chi-square Test of Independence

1. cross\_tab = pd.crosstab(df['sentiment'], df['gender'])

2. # Print the cross-tabulation

3. print("Cross-Tabulation of Sentiment and Gender:")

4. print(cross\_tab)

Table 42: Cross-Tabulation of Sentiment and Gender

1. import matplotlib.pyplot as plt

2. from collections import Counter

3. # Combine all the text into a single string

4. all\_text = ' '.join(df['text'].tolist())

5. # Split the text into individual words

6. all\_words = all\_text.split()

7. # Count the occurrences of each word

8. word\_counts = Counter(all\_words)

9. # Retrieve the most common words

10. most\_common\_words = word\_counts.most\_common(10)

11. # Extract the words and their frequencies

12. words = [word for word, count in most\_common\_words]

13. frequencies = [count for word, count in most\_common\_words]

14. # Calculate the total frequency of all words

15. total\_frequency = sum(frequencies)

16. # Calculate the percentages

17. percentages = [count / total\_frequency \* 100 for count in frequencies]

18. # Set up the figure and axes

19. fig, ax = plt.subplots(figsize=(10, 6))

20. # Create a bar chart for the most common words

21. bars = ax.bar(words, frequencies, color=['orange', 'darkblue', 'darkblue', 'darkblue', 'darkblue', 'darkblue', 'darkblue', 'darkblue', 'darkblue', 'darkblue'])

22. # Add labels and title

23. ax.set\_xlabel('Words')

24. ax.set\_ylabel('Frequency')

25. ax.set\_title('Top 10 Most Common Words')

26. # Rotate x-axis labels for better readability

27. plt.xticks(rotation=45)

28. # Add percentage labels above each bar

29. for bar, percentage in zip(bars, percentages):

30.     height = bar.get\_height()

31.     ax.text(bar.get\_x() + bar.get\_width() / 2, height, f'{percentage:.1f}%', ha='center', va='bottom')

32. # Show the chart

33. plt.show()

Table 43:Frequent Words

1. sentiment\_counts = df.groupby(['created\_at', 'label']).size().unstack()

2. # Extract the time periods and convert them to datetime objects

3. time\_periods = pd.to\_datetime(sentiment\_counts.index)

4. # Set up the figure and axes

5. fig, ax = plt.subplots(figsize=(12, 6))

6. # Plot the stacked area plot for sentiment categories

7. ax.stackplot(time\_periods, sentiment\_counts.T, labels=sentiment\_counts.columns)

8. # Add labels and title

9. ax.set\_xlabel('Time')

10. ax.set\_ylabel('Count')

11. ax.set\_title('Tweet Frequency Time Series')

12. # Add legend

13. ax.legend()

14. # Show the chart

15. plt.show()

16.

Table 44:Time Series

1. from wordcloud import WordCloud

2. text = ' '.join(df['text'])

3. # Create a WordCloud object with customized options

4. wordcloud = WordCloud(width=800, height=400, background\_color='white', colormap='viridis',

5.                       max\_words=100, contour\_width=3, contour\_color='steelblue',

6.                       prefer\_horizontal=0.8, relative\_scaling=0.5,

7.                       collocations=False).generate(text)

8. # Add a central word

9. wordcloud.generate\_from\_text("Afghanistan")

10. # Manipulate font sizes for the central word and other words

11. wordcloud = WordCloud(width=800, height=400, background\_color='white', colormap='viridis',

12.                       max\_words=100, contour\_width=3, contour\_color='steelblue',

13.                       prefer\_horizontal=0.8, relative\_scaling=0.5,

14.                       collocations=False).generate(text)

15. wordcloud.words\_["Afghanistan"] = 80  # Set font size for the central word

16. # Display the word cloud using matplotlib

17. plt.figure(figsize=(10, 6))

18. plt.imshow(wordcloud, interpolation='bilinear')

19. plt.axis('off')

20. plt.tight\_layout()

21. plt.show()

Table 45: Word-Cloud