1. The platform that I have chosen to analyze is Twitter. Each rule that Twitter sets in place has a clear purpose in creating a safe community; however, some rules seemed to be of more importance to me such as hateful conduct, sensitive media, and private information as a whole. Hateful conduct, being anything aiming violence or negativity at an individual’s identity, group or ideas is prohibited by Twitter based on protected category. Twitter has always held up their end when determining what’s hateful and what’s purely a way of expressing opinion. Twitter allows the option for their users to quickly and easily report anything they feel promotes violence or hate against certain people or groups even if they’re not the one being specifically targeted. The same thing goes for Twitter deciphering what’s “appropriate” when it comes to sensitive media content. They define such sensitive content as consisting of adult content and graphic violence as a means to set guidelines and cover all bases, leaving no room for confusion when it comes to what’s sensitive and what’s not. If Twitter believes a video or post has the slightest potential to be sensitive to any party, a “sensitive content” warning label shows up on it. Whether the post be violent to raise awareness of a topic or is simply a bit on the more promiscuous side, this warning label gives viewers the ability to observe content at their own discretion. Twitter also gives the person who posts the media the option to list it as “sensitive content” which helps put more responsibility on the individual as well. Lastly, private information is one of the most significant rules that Twitter values. Giving out or posting any private information or pictures without permission like addresses, phone numbers or anything else that shouldn’t be made public are prohibited by Twitter. When it comes to privacy and protection, Twitter allows users the option to make their accounts and posts completely private. This platform allows people to pick who follows them, who sees their posted content, as well as what they like and share; if you aren’t a follower of theirs, you cannot see any their activity.

2. I compared Twitter’s rules to Snapchat’s rules and observed that both platforms enforce the same basic rules but have different ways of listing and defining such. Twitter and Snapchat prohibit and highlight the similar rules of harassment/bullying, violence/threats, terrorism, spam/impersonation, and sexually explicit content. These two platforms share these comparable rules as they’re the basic guidelines in making any social media safe and secure. On both applications, you can share and distribute images and videos with your friends/followers, but they have different purposes of use as well. The rules are similar due to the fact that users have the ability to post content that anyone can see. Snapchat is a simple platform that focuses on communicating and seeing the world through images and videos. I noticed that when comparing each platform’s community guidelines, Snapchat is more extensive with their list of rules, making them bullet-pointed and easy to read/understand. On the other hand, Twitter has more aspects to it that make it a more complex and opiniated platform. Not only does it allow the posting of photos and videos, but people can also write their own content, repost, like, and comment on other’s profiles. Twitter’s rules are more detailed and give thorough explanations of each one. For example, Twitter has bolded headings for each category that the rules fall under, and Snapchat just lists them with no organization. Twitter underlines and separates their rules for self-harm, violence, and hateful conduct as their own groups; contrastingly, Snapchat groups these rules together as threats, violence, and harm, giving less definition
to them as their own policies. Though the same basic rules are enforced by both platforms, Twitter having more usage than Snapchat for a wider variety of engaging activities makes it more apparent as to why the rules are more specific and organized into certain categories.

3. In studying Twitter over a period of a couple days, I saw a rather odd video of a man lighting his urine on fire and you could fully see his private parts with a funny caption over it. This was in violation of Twitter’s sensitive media policy highlighting adult content. This policy states that if the media produced and distributed includes close-ups of genitals, then it’s against the rules. It’s rare that I see content on Twitter in violation of the rules because I choose who I follow; however, sometimes certain posts randomly show up on my feed that may cross certain lines. This instance was the only post I saw during my analysis of Twitter over the past couple of days that is with no doubt against the what Twitter enforces. I never truly see anything on my Twitter feed that I’m offended by or feel targets certain individuals or groups, but that could just be my opinion and not how others feel if they were to see my feed.

4. A particular challenge of moderating content on Twitter would be to decipher what is truly violating and what is freedom of speech/expression. Every minute of every day, people share and post content all over Twitter. Due to all of this information, it’s hard at times to catch every single thing that’s in violation of the rules. By Twitter choosing what to allow and what not to, it can step on their billions of user’s toes. Twitter is intended to be used for public conversation and free expression, so it becomes difficult when deciding what to draw the line on. With public conversations and expression comes opinions meaning some people find things relatable, funny, useful, and newsworthy that others don’t. A post could be completely acceptable when it comes to Twitter’s post guidelines, but it could still offend some viewers. The main problem is that there’re ways to sneak around the rules and get away with certain posts due to the widespread user-base and consistently posted media.

5. One major thing I have learned from completing this exercise is that content moderation for companies is extremely difficult and they’re in charge of working to meet numerous standards at once. I never realized that every single post, thread of tweets, direct message, etc. can be offensive to just about anyone and that these platforms must decide if it’s fair or unfair to delete the post or even the account. Certain content can slip through the cracks Companies can have a number of people turn against them in an instant when it comes to certain content slipping through the cracks or disagreeing with some things that follows the guidelines enforced. It’s extremely hard for companies to create and enforce rules that meet every single user’s expectations; in fact, this notion more times than not can make the corporation look inconsistent or even biased. This knowledge I’ve gained will affect my interactions with social media moving forward. I know that I do have responsibility in knowing what is okay to post and what isn’t, also keeping in mind that I can report things I find offensive that aren’t seen by the company itself. This not only helps ease their investigation process of certain posts but further reminds me that I can have a say too.