Safecity India: Spatial Relations of Sexual Harassment in Delhi

Sasa Tang

5/2/2018

Abstract

This paper will use geospatial sexual harassment incident data from Safecity in India to examine the relationship between different types of sexual harassment in public space. This paper looks into the relationship between five main categories of sexual harassment: nondirect, direct verbal, physical, stalking, and others and asks how these behaviors relate to each other and to different spaces. Using ArcMap and R, I produce maps and tables to analyze these relationships.

Introduction

Gender-based harassment and sexual assaults are an important issue which affects people's safety and mobility throughout the world. Post-industrial democracies have many laws that codify harassment and sexual assault as illegal actions, and these laws tend to be generally well-enforced if victims have the agency to accuse their attackers. Most of these laws pertain to sexual harassment in the work place, with varience on laws that capture sexual harassment in public spaces. The governmentality of post-industrial democracies results in large amounts of data capture, and this facilitates spatial analysis of harassment and sexual assaults. It is less clear to what extend laws in developing states protect people from harassment and sexual assault, and even less clear whether authorities in those developing states enforce laws of harassment and sexual assault if victims have enough agency to accuse their attackers.

Sexual harassment laws in Indian can be divided up into two different categories: workplace laws and penal code laws for public spaces. The POSH Act (Sexual Harassment of Women at Workplace Prevention, Prohibition, and Redressal Act) of 2013 govern and protect against sexual harassment in the work place where the harasser is known and there is a structured relationship between harasser and victim. For laws governing sexual harassment violations in public spaces, there are four sections under the Indian Penal Code from 1860 that address sexual harassment:

1. Section 294: law addressing obscene acts or songs "in or near any public place"

2. Section 326: acid attacks that lead to physical damange.

3. Section 354: laws that address crimes that "outrage her modesty" including: physical contact of sexual nature, sexually colored remarks, showing pornography against someone's will, voyeurism, stalking, and sexual assault.

4. Section 509: law that punishes "word, gesture or act intended to insult the modesty of a woman," that intrudes upon the privacy of the person.

In both categories, there is not a strong enforcement of these law, coupled with the social stigma against reporting sexual harassment, and psychological barriers of self blame, all contribute to the continual prevelence of sexual harassment in both the work place and in public spaces. This paper focuses on sexual harassment in public spaces because gender based restraints on mobility affect female's participation in public life, their access to education, their economic empowerment, and even their political participation. So far in India, no good information exists which analyzes the spatial distribution of sexual assaults, and it seems that the government does not systematically catalogue reports of such incidents. To overcome this problem, Safecity

has created a novel way to collect data on such incidents. Safecity began as a website and has recently moved to an app for Android and iOS enabled devices, which allows harassment and sexual assault victims to report the exact location of their harassment and the characteristics of the incident. To date, it does not appear that a spatial analysis on public space using this Safecity data has been carried out—this is the research gap that I intend to address in this research project.

Literature

Sexual Harassment

There are three main frameworks to understand sexual harassment: social-structural, social control, and sociocultural. Sexual harassment in the social-structural framework regards this behavior as "markers of passage" for women because it communicates the hegemonic normative rules of power that negotiates behaviors of conduct between genders (Kissling 1991). It is used as a tool that demonstrates unequal power relations between males and females, "micro-aggressions" that intrudes into personal space without consent (Davis 2004). Furthermore, Kissling and Kramarae (1991), and Wise and Stanley (1987) regard sexual harassment on the street as a language of "sexual terrorism" and argue that one reason that sexual harassment is performed in public spaces is to oppress women through fear. Sexual harassment in the social control framework views this behavior as a tool to control populations. Ilahi (2009) studies the Arab society and says that Islamic doctrines believes that culturally men are rational and capable of self-control whereas females lack control particularly over sexual drives, and therefore is the male's job to maintain social hierarchy over women through the control of the female body. Similarly, Amar (2011) study of protestors in Tahrir Square, Egypt shows that the regime's security apparatus hired thugs to sexually assaults female activists to produce a "sexualized state terror" as a result of females taking agency over their own bodies to challenge the "manhood" of Egypt through political action. Following the tradition of body politics in feminist literature, the female body in Amar's view is symbolic of male control and a means to practice power dominance through sexual harassment. Lastly, the sociocultural framework understands sexual harassment as socially constructed gender roles and norms that are played out to promote violence against women (Fitzgerald 1993; Koss and Association 1994). Sexual harassment here is performed by men who need to demonstrate their masculinity. While these three frameworks differ in only minute ways, all three views sexual harassment a demonstrated behavior that indicates rules of behavior that is permissible to be performed in certain places. Rules that define behavior is one of the main definitions of the concept of logic of appropriateness.

March and Olsen (2011) say logic of appropriateness "are rules that prescribe, more or less precisely, what is appropriate action" (p. 843). Sometimes these rules are tied to the notion of morality, but most importantly, these rules are guided by conformity and stabilized expectations. These rules vary based on their location and ambiguity. Rules that written and enforced, such as behaviors of pickpocketing, are criminalized and variation of this behavior depends on levels of policing among other variables. Rules that are more ambiguous such as dress code, body behavior, access are more flexible and easier to be challenged. Chappell (2006) studied the logic of appropriateness in institutions and argued that the important thing about understanding logic is the process, not necessarily the rules. Using a neo-institutional framework, Chappell argued that the "gendering process" of institutions shape political opportunity structures for change. This logic of appropriateness can be applied to this research paper by connecting the behavior of sexual harassment and the location of public spaces. This leads to the question is there a spatial relationship between the type of sexual harassment and the place and locality it occurs? Are some areas more or less appropriate for certain behaviors of sexual harassment more than other areas? Are certain types of sexual harassment more prevalent in one type of space over another?

Space and Place

As previously stated in the concept of logics of appropriateness, there are both written and unwritten rule that determine behavior within certain places. How do these rules vary? Sack (1986) in his seminal piece

Human Territory talks about the theory of territoriality, conceptualized in terms social construction of human interaction. He says that territoriality is a tool that used "to affect, influence, or control actions, interaction, or access by asserting and attempting to enforce control over specific geographic area" (1983, p. 55). Territoriality establishes differential access to resources, dividing groups and thus asserting control. Sack highlights two ways which the dominant group can create this power: physical contact or domination over specific space. Applying this to the behavior of sexual harassment, Sack would see that physical touch or assault indicates to the victim that he/she is within the territory of the dominate group and such physical contact is within the privilege of the dominant group to exert. The second side of this coin is when the dominate group, in this case the hegemonic male group, restricts access for the other group through information dissemination. This paper argues that this second type of method has to do with logics of appropriateness within public space. Imagines or presence of male bodies, masculine behaviors that indicate entitlement and ownership to certain spaces product this domination over the public space in the form of sexual harassment.

Therefore, this paper asks, what types of sexual harassment behavior, or tools of territoriality play out in different spaces? Do these different tools have any relationship with each other? If a space already experiences the presense of one type of harassment, is it more likely to influence other types of sexual harassment? How are these acts influence each other? This paper will attempt to use spatial analysis and time series regression analysis to answer these questions.

Methodology

This paper uses data from Safecity, a platform that crowdsources personal stories of sexual harassment and abuse in public spaces. Safecity is a small nonprofit formed in Decemeber 2012. The data collected are anonymous and are mapped out to provide more awareness to the issue of sexual harassment in public spaces, as well as to inform intervention methods. This paper uses R to tidy and wrange the data and ArcGIS to map incidents.

In November 2017, Safecity launched an app for iOS and Android enabled platforms in hopes that the use of GPS data sourced from smartphones can increase the dataset's spatial accuracy as well as accessibility. The most recent observation for this paper ends at Feburary 22, 2018 so the effects of an app based platform is still undetermined. Safecity relies on victim reporting as the sole source of data collection for its database currently; victim reports are textual in nature and geocoded with GPS information, meaning that the victim submits a written description of a harassment incident with as much or as little information as they deem necessary. Safecity developers then code these written victim accounts and add them to the database. Therefore, due to the augmentation of data between the victim's report and the final dataset, the Safecity database is considered a secondary data source. It should be noted here that secondary sources should be treated as such– there is no way to verify that the programmer coding was appropriate for what the victim described. Furthermore, there is no way to determine the level of intercoder reliability between programmers tasked with coding victim reports.

There are a couple data considerations to address. First, while there are empirical limitations of crowd-sourced data since it not collected in a systematic manner, it is still a powerful source of information; this is especially true of a country with low levels of crime reporting and governmental statistical reporting. Second, it is impossible to determine whether the Safecity database suffers from collection bias. Bias could come be based on differential socio-economic access to smartphones, computers, internet, and the demographics of people reporting the incidents. The Caste system of India could preclude certain segments of the population from reporting sexual harassment incidents. Cultural norms could depress reporting of sexual harassment. Culture could also affect the concept of sexual harassment as conceived of in post-industrial Western democracies, and the concept may not have an exact analog within the totality of India. Education rates among different segments of the population could also play a role– education could teach females that sexual harassment is a crime that should not be tolerated. While there are 22 national languages in India, Safecity is avaliable only in three: English, Hindi, and Spanish. With these considerations in mind, the analysis of this data is still warranted as there is a lack of focus on addressing sexual harassment in India. This research paper is not intended to be definitive but, instead, is intended to shed light on any patterns in sexual harassment incidents

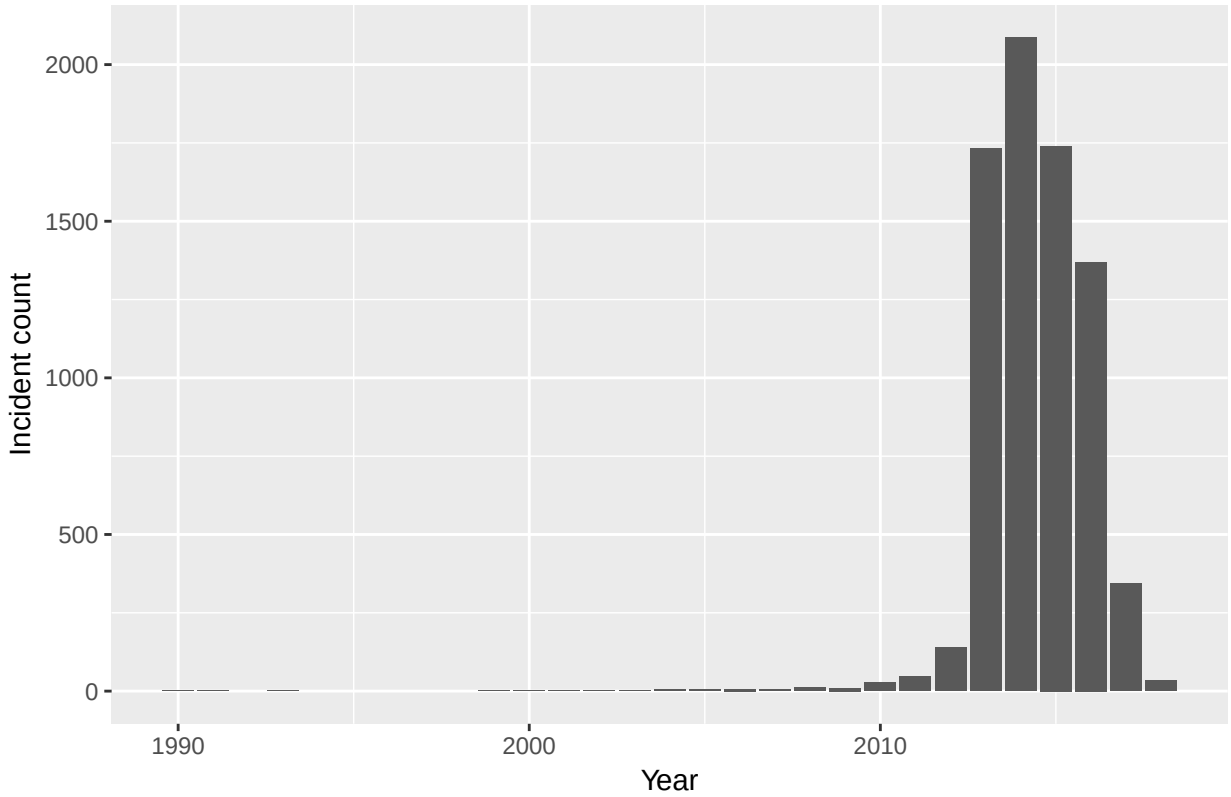
within India. This preliminary analysis will highlight areas within India that appear to have problematic, non-random, spatial distributions of incidents. These highlighted areas would then be primed for a systematic research project or focused intervention designed to lower the number of harassment incidents and reduce non-random patterns of harassment and sexual assault.

Tidying Data

Safecity neither maintains a codebook for their data nor does Safecity include metadata in their database, which made understanding and working with their dataset challenging. This dearth of information on the dataset made it necessary to communication with the developers of Safecity to clarify the meaning of the data before it could be used for this research. The full Safecity dataset contains 7,572 observations collected between March 5, 1990 to February 22, 2018. Safecity maintains data for the period between 1992-2012, which is prior to the NGO's founding in 2012; this pre-Safecity data was collected from secondary sources like old newspaper articles and it is unclear how reliable this data is. After removing all incidents that 1) were located outside of India, 2) did not have geospatial coordinates, and/or 3) occurred prior to 2013 (pre-2013 data was not crowd-sourced and source information is unknown), our modified dataset contains 5,360 observations collected between the 2013 and 2018. The reduction in observations was necessary for several reasons.

```
## Warning: Removed 505 rows containing non-finite values (stat_count).
```

Figure 1. Incident Report Frequency by Year



The first task completed was to acquire information from Safecity's developers in India regarding the dataset's coding logic to ensure that we fully understand the data. The second task was to remove data points that were invalid; many of these invalid data points turned out to be test points which Safecity added to the dataset but failed to remove. This task led to the exclusion of 41 data points which were either 1) located outside of India's boundaries, 2) plotted as tests by Safecity developers, or 3) reported as occurring in neighboring countries (Nepal specifically). The third task involved deciding what types of harassment should

be included to answer our research questions. The dataset contains 1907 observations that happened on a transit system (taxi, car, train, bus, metro, auto, rail, airport, travel.) Differentiation between non-transit and transit related sexual harassment incidents resulted in the exclusion of non-transit data because the geocoordinates of incidents which occurred in a closed transit system have a different spatial meaning than incidents which occurred on the street. Of eleven types of sexual harassment included within the dataset,I removed the category of *chain snatching* because it is not a form of sexual harassment as defined in this research paper. *Chain snatching*, as explained by the developers of Safecity were included in the dataset because it is a common petty crime that people were using the platform to report.

The fourth task was to determine the best way to analyze the types of sexual harassment in the Safecity database. The Safecity database contains eleven different sexual harassment types and offers a nuanced account of harassment in India. However, some of the harassment types in the database are very similar in character and would complicate the use of mapping techniques and spatial analysis. Inductive methods are commonly used in qualitative research and prove a powerful strategy for generating typologies and theories during the research process (George & Bennett, 2005). Inductive categorization was used to overcome the dataset's nuance problem. Then, the eleven sexual harassment types were effectively collapsed into larger and more manageable categories based on the spatial relation between the harasser and victim.

The inductive process was iterative and resulted in the generation of five broad harassment categories: 1) *non-interactive harassment*, 2) *direct-verbal harassment*, 3) *direct-physical harassment*, 4) *stalking*, and 5) *other*. These broad categories are explained below. It should be noted that the categories are not exclusive, meaning that one harassment incident could be coded as more than one sexual harassment type.

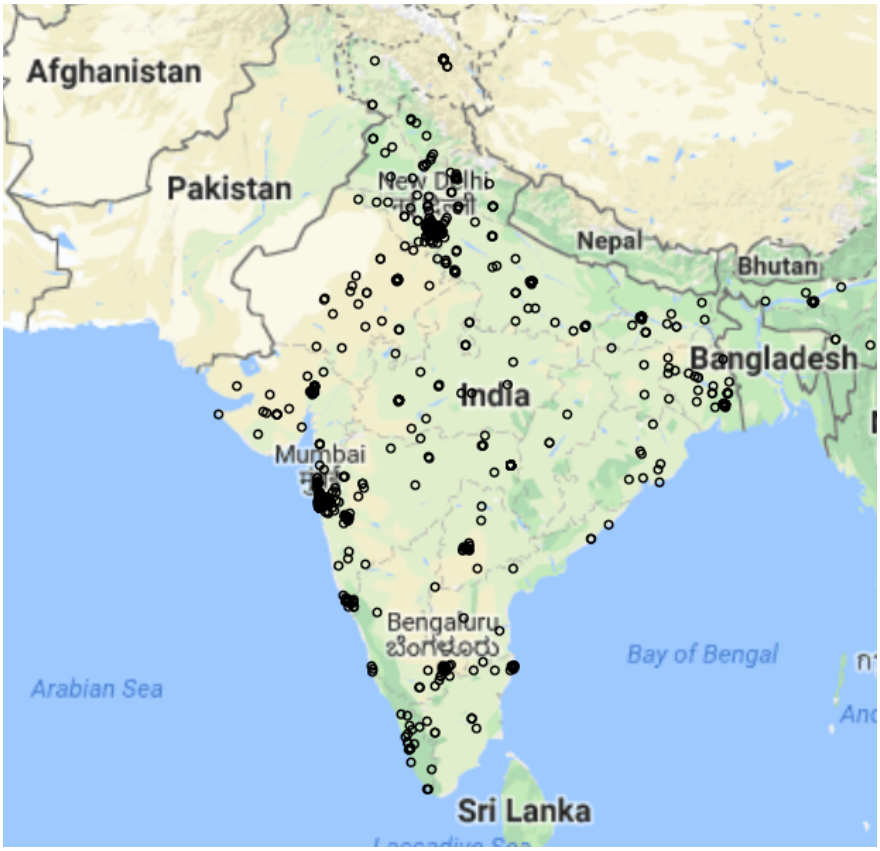
The *Non-Interactive* Harassment category includes 1) *indecent exposure*, 2) *staring*, and 3) *taking pictures*. This grouping was chosen due to the non-interactive nature of these types of sexual harassment between the harasser and the victim. These behaviors may not be directed at the specific victim but still violate people's body boundaries. The spatial aspect of this category of sexual harassment is considered distant since we assume that harassers do not have to be in close proximity to the victim.

The *Direct-Verbal* Harassment category includes 1) *catcalls*, 2) *comments*, and 3) *sexual invites*. This grouping was chosen due to the direct and purposive behavior of the harasser towards the victim. The spatial aspect of this category of sexual harassment is considered to be proximate since I assume that a victim must have been in close proximity to the harasser to hear their speech.

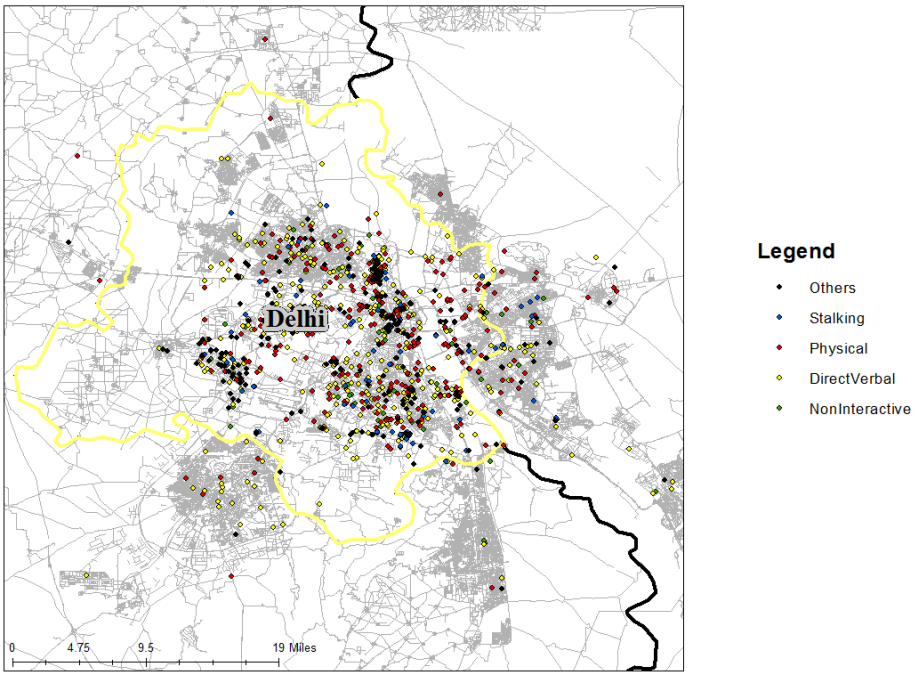
The *Direct-Physical* Harassment includes 1) *touch* and 2) *sexual assault*. This grouping was chosen due to the direct violation of body boundaries. These behaviors are direct, physical, and often violent in nature. The spatial aspect of this category is considered to be proximate because I assume a harasser must be close enough to physically violate a victim.

Stalking comprises its own category because this form of sexual harassment could be spatially proximate or distant depending on the specific occurrence. Furthermore, stalking could be a transient activity and that the reported location of stalking incidents could have a different meaning than the other incident types since the location where a stalking incident was reported could have been one of many locations where this episodic incident-type took place between one victim and harasser. In this case, the location recorded could be where the victim first become suspicious that they were being followed, where they realized they were being followed, or something else. It is impossible to tell since this information is absent from the Safecity dataset. According to India's penal codes, the law defining stalking states "any man who follows a woman and contacts, or attempts to contact such woman to foster personal interaction repeatedly despite a clear indication of disinterest by such woman." This law is punishable up to five years imprisonment and covers electronic communication. Finally, The *Other* category was left as is because of the indeterminacy of the incident type. There were 281 sexual harassment incidents that were reported and categorized as *other* and nothing else.

Analysis

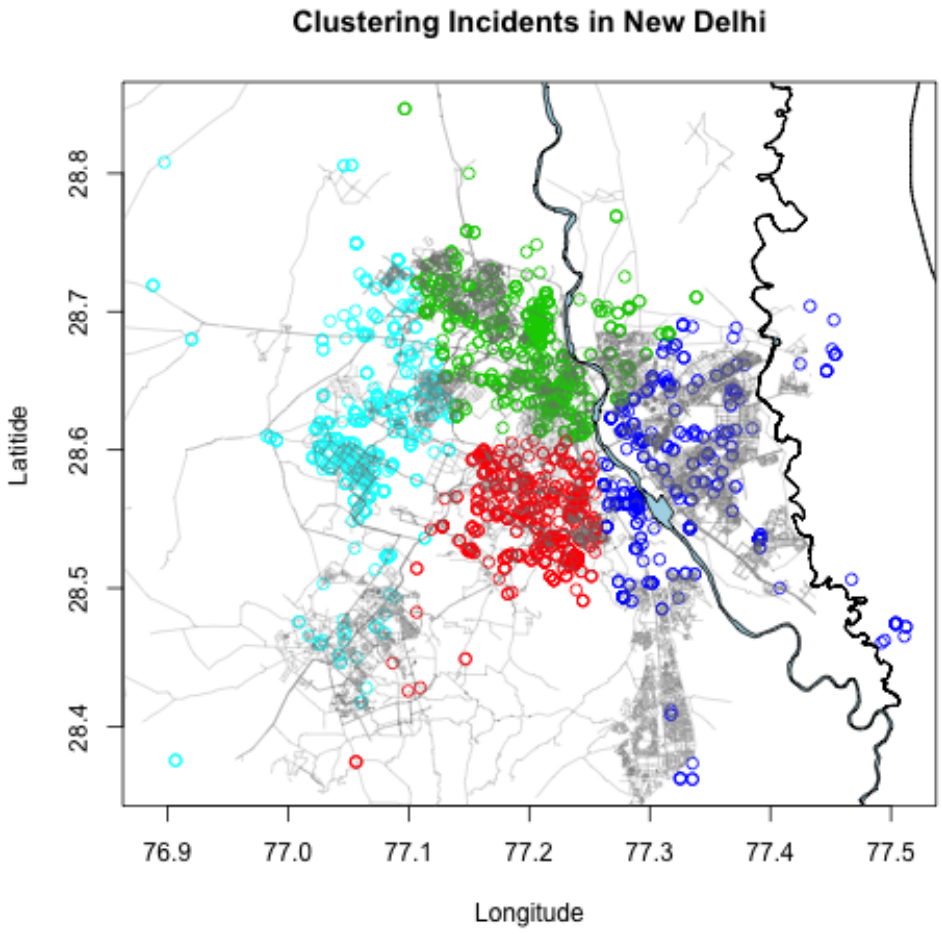


To focus in on the spatial relation between sexual harassment types, I have decided to zoom into focus only on New Delhi, where there are the most incident points, a total of 3165 observations.



Map 2 shows the spatial distribution of different types of sexual harassment layered on top of administrative boundaries and roads.

First, looking at Map 2, there seems to be some clusters. So using a kmeans cluster algorithm, I have plotted the same incident points grouped into centroids. This is a iterated process where the centroid of each cluster is calculated by calcluating the closest observation in terms of Euclidian distance to the within group mean until the cluster assignments no longer change (Imai 2017).

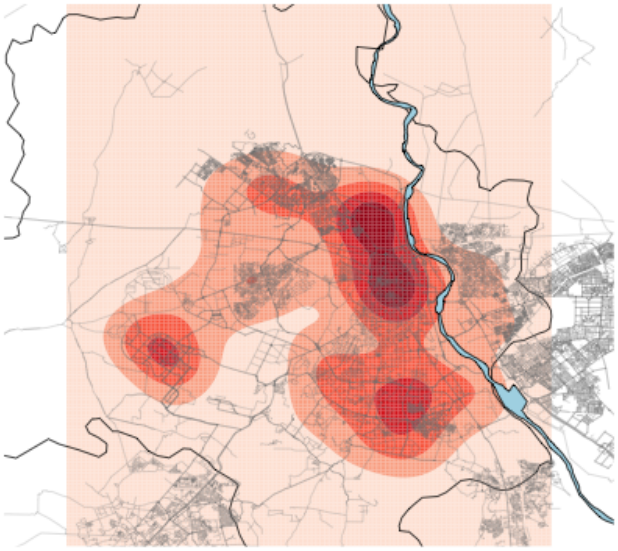



```
##                 cluster
## catagory          1   2   3   4
##   directverbal   413 789 252 498
##   noninteractive 274 389 176 303
##   others          56 140  31 145
##   physical       150 315 124 232
##   stalking        44  38  20  40
```

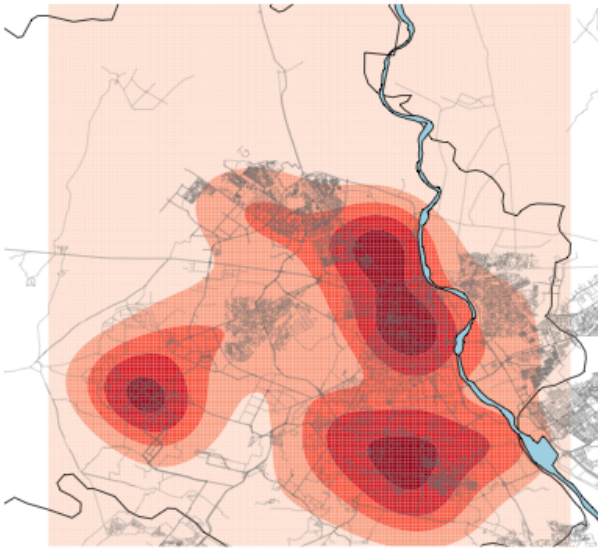
The output of this table shows the number of incident counts according to the cluster that they belong. It seems like cluster two has the lowest count incidents of all the categories but it is unclear where the clusters are.

To advance the clustering analysis further, I would like to see the probability density of each individual type of sexual harassment. I used the kernel density estimate which averages the probability distribution centered around each observation point (Brunsdon and Comber 2015).

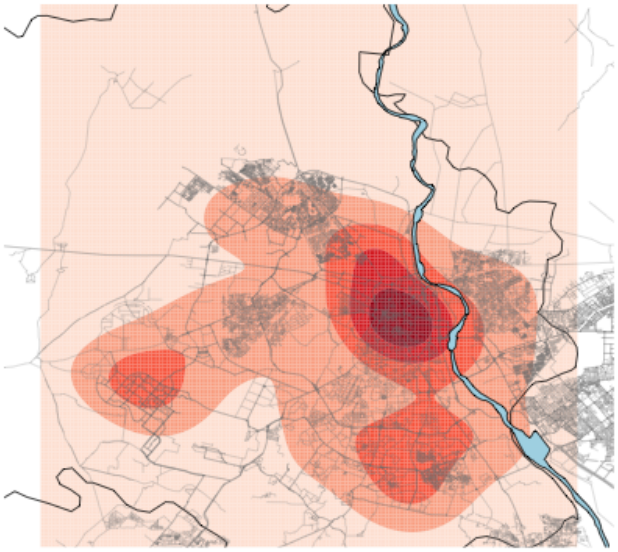
Direct Verbal Harassment



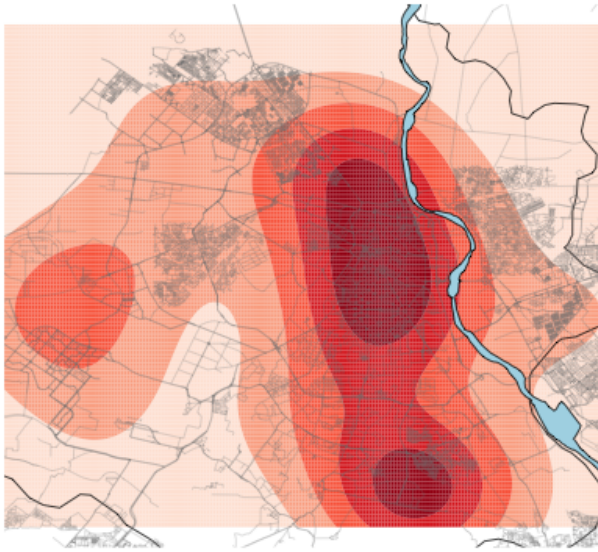
Non-Interactive Harassment

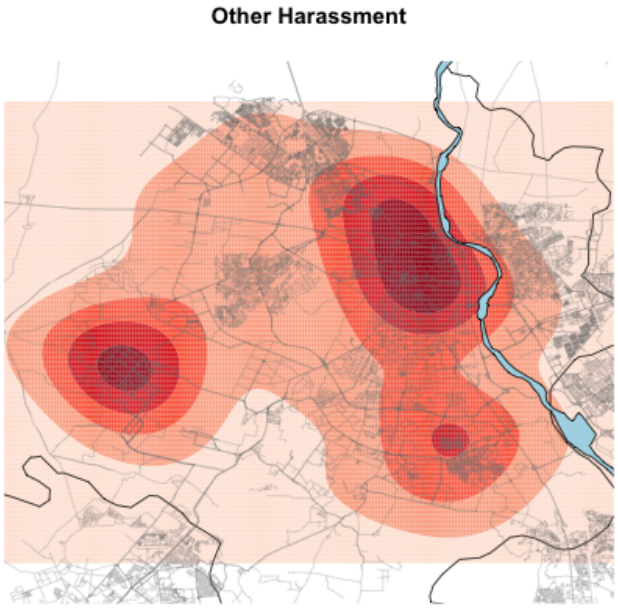


Physical Harassment



Stalking Harassment





Plotting the kernel density of the different sexual harassment categories is a comparison tool. In all the maps I have layered the roads, the administrative boundary of Delhi the subdistrict, of which New Delhi the captial is more spread out than the administrative boundary, and the Yamuna river. There is a pattern that in all sexual harassment follow a similar clustering layout in that they overlap the most condense roads. The kernel density of *Stalking* is most spread out, and this makes logical sense given the fact that stalking is a more transient behavior than the other types of harassment. There is visible variation between *Physical* and *Direct Verbal* and *Non-Interactive* in that *Physical* harassment is more concentrated in one area close to the Yamuna river. The *Direct Verbal* and *Non-Interactive* kernel density maps look very similiar to each other, with *Non-Interactive* being a bit more spread out than the *Direct Verbal*. This also makes more sense because as I have explained in the recategorization of the original categories of sexual harassment, that I have conceptualized types of sexual harassment to its spatial relation between the harasser and victim. Therefore, the prevalence of certain types of sexual harassment is more wide spreading and covers more area than others. Though, it is also interesting to note that there is not a lot of difference between *Physical* and *Direct Verbal* in their kernel density maps.

Times Series Regression

In order to explore how different types of sexual harassment interact with each other, and to answer the questions of the role of logic of appropriateness among different spaces, I will conduct a spatial time series regression analysis. First, I will create quadrants to overlay on top of the areas, and then join the incidents into the polygon quadrants to get a count data for each quadrant. The units of my analysis are each quadrant, with a count data for the different types of incidents within each quadrant. My model is the following:

$$DV_t = \beta_0 + \beta_1 X_{t-1} + \beta_2 DV_{t-1} + \mu_i$$

Where DV_t is one type of harassment, X_{t-1} is another type of harassment at a time before t, and DV_{t-1} is the dependent variable at a prior time.

I will run this equation with a few different iterations to see if there is a correlation among the different types of harassment.

Discussion

While these maps do demonstrate that there exists some sort of spatial relation among type of sexual harassment, there is not enough information on the physical enviornment of New Delhi. What type of spaces are these harassment incidents occuring in? Are these places shopping malls, tourist destinations, industrial zones, business districts, parks, or slum markets? How does the type of space affect the variation of sexual harassment? And do how these places are used also affect the variation of sexual harassment? As shown from the criminology literature, time also plays a role on the occurrence of crime. Therefore, how are these spaces occupied and used throughout the course of the day? Do these compositions change and therefore change the variation of type of sexual harassment? To answer these questions, I would need more information on different spaces.

Conclusion

I will spend more time looking at different ways to analyze spatial data and work on the time series regression analysis. I hope this will bring to light more understanding about how different types of sexual harassment interact with each other.

Moving forward, there are two main directions this project would like to explore. First, this research question would like to ask, is there a relationship between the build enviornment of a place and sexual harassment? Previous research in the field of criminology have looked at build factors such as building heights (Newman 1972; Rengert 1980; Rubenstein 1981), vegetation (Kuo and Sullivan 2001; Michael and Hull 1994; Shaffer and Anderson 1985), lighting (Painter, n.d.; Farrington and Welsh 2002), open spaces (Taylor 2002) and their relationship to crime rates and types of crime. My research would like to add to this field by looking at the different behaviors of sexual harassment and its build enviornment. The lack vigilance, as Jacobs (2011) points out can be a driver of crime. Taking this point further to sexual harassment, where this umbrella crime can emcompass close vicinity interactions (touch) or distant interactions (stare), I would like to research how the build enviorment contributes to the type of sexual harassment. Do more crowded spaces experience higher physical sexual harassment? Do bigger squares see higher prevalence of verbal and non-direct harassment because of the distance avaliable? Do areas of higher traffick lead to more sexual harassment? Does the type of market and the physical composition of the markets affect sexual harassment? To tackle these questions, I would need data on the build enviornment.

The second direction this project leads to is through on the ground quantitative data on people's perceptions. This paper has shown that the way space used by people may affect certain behaviors. Therefore, usage of space is influenced by perception of space and safety. This is why it would be beneficial to do on the ground mapping projects to better understand how people think about space, rights to space, ownership of space and of bodies, and what they would like to see change.

Work Cited

Amar, Paul. 2011. "Turning the Gendered Politics of the Security State Inside Out?" *International Feminist Journal of Politics* 13 (3): 299–328. doi:10.1080/14616742.2011.587364.

Brunsdon, Chris, and Lex Comber. 2015. *An Introduction to R for Spatial Analysis & Mapping.* Los Angeles: SAGE.

Chappell, Louise. 2006. "Comparing Political Institutions: Revealing the Gendered 'Logic of Appropriateness'."

Politics &Amp; Gender 2 (2): 223–35. doi:10.1017/S1743923X06221044.

Davis, Mike. 2004. "Planet of Slums." *New Left Review*, II, no. 26: 5–34.

Farrington, David P., and Brandon C. Welsh. 2002. "Improved Street Lighting and Crime Prevention." *Justice Quarterly* 19 (2): 313–42. doi:10.1080/07418820200095261.

Fitzgerald, Louise F. 1993. "Sexual Harassment: Violence Against Women in the Workplace." *American Psychologist* 48 (10): 1070–6. doi:10.1037//0003-066X.48.10.1070.

Ilahi, Nadia. 2009. "Gendered Contestation: An Analysis of Street Harassment in Cairo and Its Implications for Women's Access to Public Spaces." *Surfacing* 2 (1): 56–70.

Imai, Kosuke. 2017. *Quantitative Social Science: An Introduction*. Princeton: Princeton University Press.

Jacobs, Jane. 2011. *The Death and Life of Great American Cities*. 50th anniversary ed., 2011 Modern Library ed. New York: Modern Library.

Kissling, Elizabeth Arveda. 1991. "Street Harassment: The Language of Sexual Terrorism." *Discourse & Society* 2 (4): 451–60. http://www.jstor.org.proxyau.wrlc.org/stable/42888749.

Kissling, Elizabeth Arveda, and Cheris Kramarae. 1991. "Stranger Compliments: The Interpretation of Street Remarks." *Women's Studies in Communication* 14 (1): 75–93. doi:10.1080/07491409.1991.11089751.

Koss, Mary P., and American Psychological Association, eds. 1994. *No Safe Haven: Male Violence Against Women at Home, at Work, and in the Community*. Washington, D.C: American Psychological Association.

Kuo, Frances E., and William C. Sullivan. 2001. "Environment and Crime in the Inner City: Does Vegetation Reduce Crime?" *Environment and Behavior* 33 (3): 343–67. doi:10.1177/0013916501333002.

March, James G., and Johan P. Olsen. 2011. "The Logic of Appropriateness." In *Oxford Handbook of Political Science*. http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199604456.001.0001/oxfordhb-9780199604456-e-024.

Michael, S, and R Hull. 1994. *Effects of Vegetation on Crime in Urban Parks*. Blacksburg: Virginia Polytechnic Institute; State University.

Newman, Oscar. 1972. *Defensible Space: Crime Prevention Through Urban Design*. New York: Macmillan.

Painter, Kate. n.d. "The Influence of Street Lighting Improvements on Crime, Fear and Pedestrian Street Use, After Dark." *Landscape and Urban Planning* 35 (2-3): 193–201.

Rengert, G. 1980. "Spatial Aspects of Criminal Behavior." In *Crime: A Spatial Perspective*. New York: Columbia Press.

Rubenstein, H. 1981. "The Link Between Crime and the Built Environment." *US Department of Justice, National Institute of Justice*.

Sack, Robert David. 1986. *Human Territoriality: Its Theory and History*. Cambridge Studies in Historical Geography 7. Cambridge [Cambridgeshire] ; New York: Cambridge University Press.

Shaffer, Garnett, and L Anderson. 1985. "Perceptions of the Security and Attractiveness of Urban Parking Lots." *Journal of Environmental Psychology* 5 (4): 311–23.

Taylor, R. 2002. "Crime Prevention Through Environmental Design: Yes, No, Maybe, Unknowable, and All of the Above," In *Handbook of Environmental Psychology,* 413–26. New York: John Wiley; Sons.

Wise, Sue, and Liz Stanley. 1987. *Georgie Porgie: Sexual Harassment in Everyday Life*. Pandora Press. http://eprints.lancs.ac.uk/53762/.