

# **Investigating Sexual Assault Rates Using a Holistic Approach\***

Vyjayanti Vasudevan

Singapore American School

January 27, 2024

---

<sup>1</sup>\*Vyjayanti Vasudevan, Grade 11, Singapore American School, 40 Woodlands street 41, Singapore 738547.  
vasudevan773999@sas.edu.sg

Acknowledgements: Joel Ramirez Jr - Curriculum manager/instructor - InspiritAI  
Word count - 4162

## **Abstract**

The rates of sexual assault in the United States of America (USA) remain high with 282,043 incidents in females in 2021 alone. We wanted to determine the socio-economic and political factors associated with rates of rape.

We examined data from 1990 to 2021 and noted crucial political, socio-economic and other noteworthy events that occurred during this time period. Our dataset includes data from the Rape, Abuse & Incest National Network (RAINN), Federal Bureau of Investigation (FBI), Association of American Universities (AAU), and the National Crime Victimization Survey (NCVS).

A statistical analysis showed a link between unemployment, the year, the annual change in unemployment, crucial geopolitical events, and educational levels with the rates of rape in the United States. A significant link was found between education level, unemployment, and poverty on sexual assault rates; however, there was a lack of evidence for a link between the year and annual change in unemployment on sexual assault rates.

## **Introduction**

In 2021, the U.S. alone documented over 282,043 incidents of rape and sexual assault among females (1). According to the National Intimate Partner and Sexual Violence Survey (NISVS), though more non-Hispanic White women report sexual assault, Hispanic and non-Hispanic Black women have higher rates of sexual assault (2).

During the Civil Rights movement in the 1960s, student activism occurred through means such as non-violent protests. At the forefront of this movement were women of color who, unfortunately, were a big part of the population with a historically high assault rate. This also supports the data on how rape disproportionately impacts different ethnicities and races across the United States. Although anti-rape efforts have persisted since the 1960s, a lack of emphasis on investigation of these crimes lead to increase in rates of rape. A dataset from a specific demographic or race cannot be generalized to the entire population; moreover, it is essential to recognize that the relationship between unemployment, education level, and sexual assault may vary based on race.

During the 1960s, the link between levels of poverty, education, and unemployment with sexual assault was a large focus (3). The connection between the social issues of the 1960s and contemporary data (elaborated on below) highlights the complexity of this issue. While there is more access to education, we see drastic inequalities in educational services; while we see more services for the underprivileged, we also see a growing wealth gap, and while some may say that there's a decrease in violent hate crimes, we see an increase in targeted online hate speech. Even in the 2000s, these power imbalances manifest themselves, such as in educational environments. According to the RAINN, around 26.4% of female undergraduates experience sexual harassment compared to the national rate of 14.8% (4). Recognizing why sexual assault cases reported by undergraduates are much higher than the national rate is crucial to know how educational environments link to sexual assault.

My paper will focus on the following research questions:

1. How do policies or events affect sexual assault in the U.S.?
2. How does education level impact sexual assault rates in the U.S.?
3. How does unemployment impact sexual assault rates in the U.S.?
4. What are some factors governments can take into consideration when trying to mitigate sexual assault incidents?

These questions will help direct our investigation of the different socioeconomic factors and critical events in American history that might be linked with rates of rape. We hypothesize that the rate of rape is linked to unemployment, annual change in unemployment, poverty, year, and education level. This analysis provides an extensive review of the changes in sexual assault rates in the US and their contributing factors.

## **Methodology**

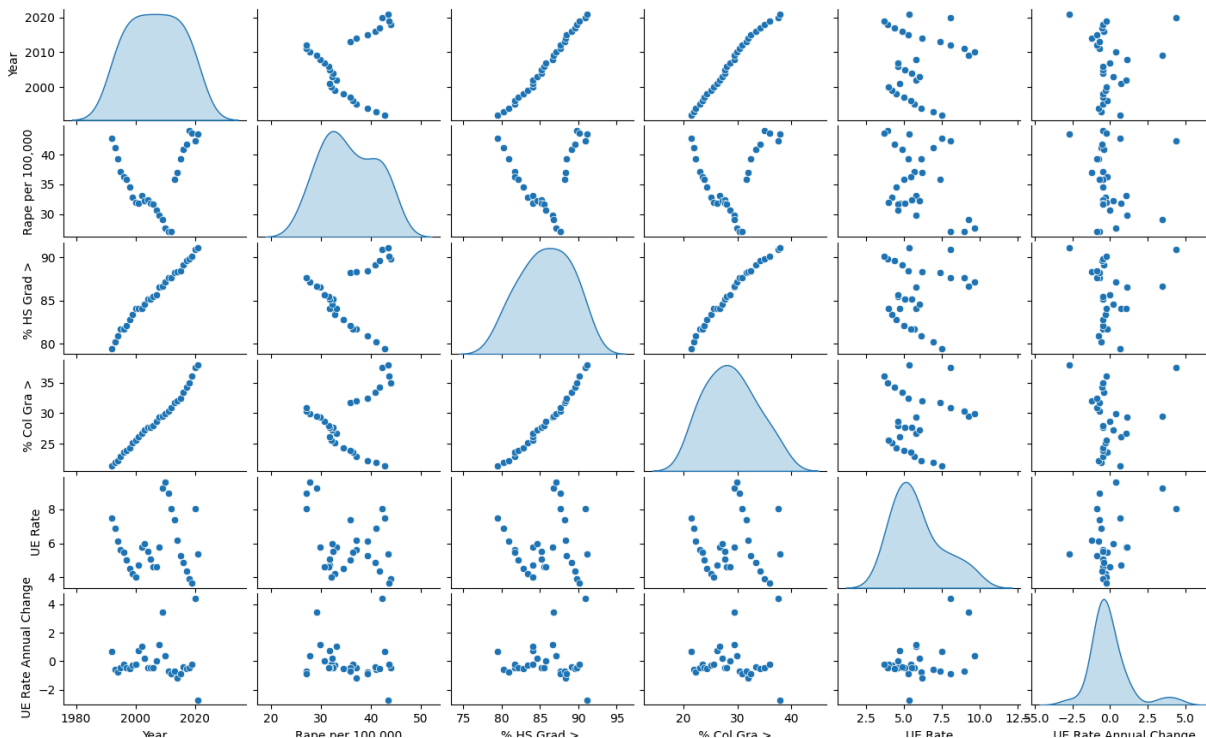
To conduct this analysis, we created a dataset including rates of rape, unemployment, education, and poverty in the US from 1990 to 2021, created primarily from US census data. A large part of our data sample was

collected from the Annual Social and Economic Supplement (ASEC), which significantly changed its sampling methodology starting in 2019 when it started weighting the responses of specific underrepresented households, mainly Hispanic households, to normalize the distribution of the responses (5). Weighing is assigning different levels of influence or importance to specific answers to increase the representativeness of the overall population. The unemployment data and annual changes (in %) from 1992 until 2021 were collected by the World Bank, which compiled data from the International Labour Organization’s estimations for unemployment (6). We also analyzed education level (in %) and poverty rates (in %) using the U.S. Census Bureau data. Next, we have our dependent variable, the rate of rape (per 100,000 people); the data was collected by the Federal Bureau of Investigation (FBI) and U.S. Bureau of Justice Statistics from 1990 to 2021 (7). The sudden increase in rates of rape in 2013 is partially due to the change in the definition of rape by the FBI. Before 2013, rape was defined as “carnal knowledge of a female forcibly and against her will” and then changed to encompass men who were raped, people who are unable to consent and to include any kind of penetration “no matter how slight” without “consent of the victim” (8). These cases, which previously would not be considered rape, now counted towards the annual rates of rape, increasing the reported rates of sexual assault. The destigmatization of rape victims in recent years could also contribute to some increases in rates of rape.

To create our database, we looked at the FBI, RAINN, AUU, and NCVS websites and the data they collected from annual surveys. The datasets we collected had to consist of annual data from 1992 to 2021. One important thing to note is that the range of years through which the data was collected for our variables differed: the data for education level range from 1960-2021, poverty data from 1990-2022, and unemployment data from 1991-2023. To manage some of this overlap, we eliminated the data from outside our range so that our database consisted of data from 1992 to 2021.

Our analysis used Spearman and Kendall correlation coefficients, distance correlations, and OLS regressions. In addition to these, we used Random Forests to help capture non-linear relationships. Random Forest is a collection of decision trees that help to cluster data. Lastly, we fitted the data into polynomial equations along our printed matrix of graphs. While the methodology mentioned above helped us modulate our data, it still does not account for the sampling biases and lack of general data on rape for us to ensure that our data is representative.

## Results



## Figure 1 - OLS regression and Distance correlation results for annual change, level of education, rates of rape per 100,000, unemployment (UE) rate and the year

The data for Highschool graduates and the rate of rape per 100,000; college graduates and the rate of rape per 100,000; year and rape per 100,000 have a strong negative linear relationship up until 2010, where we can see a significant gap in the data points and linear growth starting from that year again. The magnitude of this slope is the same as the slope before 2010. This is crucial to note since such a significant break in pattern could signify a major crisis, change in government policy, or updates in measurement standards for data, etc. To our surprise, we did not notice any apparent pattern between the annual change in unemployment, unemployment, the rate of rape, college graduates, high school graduates, and year.

As observed in Table 2, our R-squared value is 83.8%, which signifies that based on our factors, year, education level, and unemployment, our model accounts for 83.8% of the variation in our data for sexual assault rates. When observing our P-values, we can also see that for high school graduates, college graduates, and poverty, our P-values are extremely close to 0, with all values  $<0.00001$ , indicating great statistical significance.

One unexpected feature of Table 2, however, is that the year's p-value is 0.176, which goes against our predictions of the year having a strong link to sexual assault rates. Additionally, we only observed a slight difference in the R-squared percentage when removing the year's data from our OLS regression.

## Discussion

When looking at sexual assault rates per 100,000 and their correlation to years, we notice a sudden difference around the year 2010. We see lots of key events and policy changes throughout this period. Obama took office three years after the late 2000 recession, in 2008, during which we did not notice any significant changes in the annual change of unemployment, rape, and education level. This went against our predictions based on the significance we noted in unemployment and its link to rates of rape since the Bureau of Labor and Justice Statistics noted a doubling in unemployed persons from 2007-2013 (9). Then, in 2013, there was a change in the definition of rape which coincided with a drastic increase in the rates of rape per 100,000. This policy change expanded the definition of rape to include any unconsented penetration of the victim, explaining why rates of rape would increase so drastically.

To our surprise, we do not notice any apparent pattern in the correlations between unemployment and education level. Before this study, we hypothesized that education level, specifically attending college, would increase rates of rape since campus sexual assault rates are known to be higher than sexual assault rates off campus, with the RAINN estimating 26.4% of all female undergraduates having experienced sexual assault in 2020 (10). This could be because there was less emphasis on student mental health and physical well-being in universities until recently, despite the passing of Title IX in 1972. Additionally, many college students may have been afraid that the stigma surrounding sexual assault would affect their career opportunities. According to RAINN, 4 out of 5 college students, compared to 2 out of 3 non-college students, did not report sexual assault, with 20% stating that their "fear of reprisal" prevented them from doing so (4).

We also do not notice any significant pattern for the correlation between rates of rape and unemployment, with a seemingly random pattern. This is elaborated on in our *Limitations* section.

The reason we observe a strong link between poverty and sexual assault could be geographical location: their place of residence may not be located in a safe neighborhood. As documented by the Bureau of Justice Statistics, Americans with an annual income of less than \$7,500 are two times more likely to be victims of sexual assault (1996). The reason cited for this in the survey is that perpetrators of sexual assault often victimize those who are more vulnerable (aka. those in low-income neighborhoods, marginalized races, and those with disabilities).

This same strong link is observed between unemployment rates and sexual assault. It is difficult to point out a single reason as to why this might be the case; however, according to Harvard Business Review, as unemployment rates among men rise, they might feel a threat to their masculinity and, as a result, are more likely to be sexual assault perpetrators (11). They found that as rates of male unemployment were on the rise, so were rates of sexual harassment.

Another strong link can be noticed between education level (divided into two categories: high school graduate and college graduate) and sexual assault. This could be because, in low-income areas, as aforementioned, people are predisposed to conditions that leave them vulnerable to sexual assault. Another factor to consider would be higher dropout percentages to care for families. Rather than a direct correlation, educational attainment relates to unemployment and poverty in many ways. An article by BYU Education & Law Journal states that sexual education and rape are inversely proportional in their study examining the prevalence of sexual education and rates of rape in each state in the U.S. (12).

Moving onto the factors where we did not observe statistical significance, we have annual changes in unemployment and the year. The biggest question I have concerning this is why we saw a correlation between unemployment and sexual assault but not between the annual change in unemployment and sexual assault. A reason for this could be differentiating between long-term and short-term changes in unemployment and their impact on society. While prolonged periods of unemployment may have severe consequences on crime and society, the annual change in unemployment might not have an immediate impact. Additionally, the government often implements policies to mitigate the effects of economic recessions. For example, in the 2008 recession, the government provided broad-based stimulus by restoring the banking system and credit through bank bailouts that helped restore industry and industry hiring. Another factor that might make an annual change in unemployment less significant than unemployment in contributing to sexual assault rates could be the time lag in the effect of bailouts and monetary and fiscal policies. It can take time for communities to adapt or for economic hardships to influence sexual assault patterns.

Based on the results found from this study, we recommend that the government should take a more holistic approach to mitigate sexual assault cases. Simply educating more people or increasing incentives for victims to report such cases will not create drastic changes in the number of sexual assault cases if we do not address other factors, such as unemployment and poverty levels. Although redlining was banned in 1968 in the Fair Housing Act, the effects of redlining are still evident today and have had detrimental impacts on those communities. According to a study looking at modern consequences of redlining, they found that in communities historically redlined, they found “link[s] to increased risk of diabetes, hypertension, and early mortality due to heart disease with evidence suggesting [a systematic denying of services] impacts health through suppressing economic opportunity and human capital” (13). All of which, by our study, have been found to have strong links to sexual assault cases. In a National Community Reinvestment Coalition (NCRC) in 2018, they state, “Cleveland Neighborhoods ‘Redlined’ in the 1930s Are the Same Ones Dealing With Lead, Sexual Assault, Poverty and Poor Internet Issues Today” (14). Through these studies, it is clear that sexual assault has ties with so many factors that it cannot be addressed without looking at unemployment, poverty, systemic racism, gender, and even the internet. If the government wants to mitigate sexual assault effectively, it must address it through more than just education and support for victims. They must invest in infrastructure in historically neglected communities, expand Medicaid, and use subsidies to fund higher education in addition to increasing sexual education, support and encouragement for victims to report cases of sexual assault.

An analysis of sexual assault rates and levels of poverty, education, and unemployment can help identify misconceptions and potential safety nets for victims of sexual assault. Governments base their policies to prevent sexual assault on their understanding of its causes and effects; therefore, by knowing factors that link to the rise in sexual assault, they can make policies that mitigate sexual assault more effectively.

## **Limitations**

Some limitations in our study include the lack of specific and accurate data regarding sexual assault cases. Unfortunately, many victims of rape do not report the incident to authorities. Firstly, even with the increase in awareness surrounding rape, there are few resources for victims to turn to, specifically on school or university campuses. According to the National Crime Victimization Survey, around 13% of victims feel shame, which may lead to the underreporting of these incidents. Another major limitation in our study regards our analysis of unemployment rates. When looking at our data, there appears to be no evident association between unemployment rates and any other factors; however, looking closely at unemployment's definition, we realized that labor force participation may be a better indicator. Because the definition of unemployed includes just those who are unemployed and seeking employment, it confounds with poverty rates and educational attainment, which could impact the job opportunities one gets. As a result, analyzing the labor force participation rate, with binary categories

instead, could help us gain deeper insight into how having a job, not the hopes of attaining one, impacts rates of rape (16).

## Future Work

Based on our analysis of limitations, future work researching the effect of unemployment and education level should ask for these variables in the same survey that they use to find rates of rape. When we collect this data, we do not know whether the same people are filling out these surveys, which could make our data unreliable. Since people are more likely to over-report their educational level and under-report unemployment and rape since many feel shame in being unemployed or being a victim of sexual assault. Additionally, I would hope that future studies look into the correlation between labor force participation and rape rather than unemployment and rape because of its more accurate measure of how education levels impact job opportunities.

## Acknowledgments

I would like to thank my mentor, Joel Ramirez, and my parents for their continued support throughout the research process.

## References

1. <https://www.statista.com/statistics/642458/rape-and-sexual-assault-victims-in-the-us-by-gender/>
2. The National Intimate Partner and Sexual Violence Survey (NISVS) | Violence Prevention|Injury Center|CDC. (2021). Retrieved from <https://www.cdc.gov/violenceprevention/datasources/nisvs/index.html>
3. Crenshaw, Kimberle. "Mapping the margins: Intersectionality, Identity Politics, and Violence against Women of Color." *Stanford Law Review*, vol. 43, no. 6, 1991, pp. 1241-99. JSTOR, <https://doi.org/10.2307/1229039>.
4. Rape, Abuse, and Incest National Network Report On Campus Sexual Violence: Statistics. (n.d.). Retrieved from <https://www.rainn.org/statistics/campus-sexual-violence>
5. <https://www2.census.gov/programs-surveys/cps/techdocs/cpsmar22.pdf>
6. <https://data.worldbank.org/indicator/SL.UEM.TOTL.ZS?locations=OE>
7. FBI Crime in the U.S. 2018 Table 12. (2019). Retrieved from <https://ucr.fbi.gov/crime-in-the-u.s/2018/crime-in-the-u.s.-2018/tables/table-12>
8. <https://www.rainn.org/news/fbi-changes-its-definition-rape>
9. <https://www.bls.gov/opub/mlr/2019/article/tight-labor-market-continues-in-2018-as-the-unemployment-rate-falls-to-a-49-year-low.htm#:~:text=The%20unemployment%20rate%20declined%20to,the%20fourth%20quarter%20of%201969.>
10. Report on the AAU Campus Climate Survey on Sexual Assault and Misconduct (N.d.). Retrieved from [https://www.aau.edu/sites/default/files/AAU-Files/Key-Issues/Campus-Safety/Revised%20Aggregate%20Report%20%20and%20appendices%201-7\\_\(01-16-2020\\_FINAL\).pdf](https://www.aau.edu/sites/default/files/AAU-Files/Key-Issues/Campus-Safety/Revised%20Aggregate%20Report%20%20and%20appendices%201-7_(01-16-2020_FINAL).pdf)
11. When Male Unemployment Rates Rise, So Do Sexual Harassment Claims. (2021). Retrieved from <https://hbr.org/2017/08/when-male-unemployment-rates-rise-so-do-sexual-harassment-claims>
12. Sexual Education as a Form of Sexual Assault Prevention: A Survey of Sexual Education Among States with the Highest and Lowest Rates of Rape. (2020). Retrieved from [https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1006&context=byu\\_elj](https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1006&context=byu_elj)
13. Burgess KM, Campbell JA, Egede LE, Hawks LC, Linde S, Walker RJ. Modern Day Consequences of Historic Redlining: Finding a Path Forward. *J Gen Intern Med*. 2023 May;38(6):1534–1537. doi: 10.1007/s11606-023-08051-4. Epub 2023 Feb 6. PMID: 36746831; PMCID: PMC9901820.

14. NCRC. (2018). Cleveland Scene: Cleveland Neighborhoods “Redlined” in the 1930s Are the Same Ones Dealing With Lead, Sexual Assault, Poverty and Poor Internet Issues Today " NCRC. Retrieved from <https://ncrc.org/cleveland-scene-cleveland-neighborhoods-redlined-in-the-1930s-are-the-same-ones-dealing-with-lead-sexual-assault-poverty-and-poor-internet-issues-today/>
15. Lind, D. (2014). The FBI finally changed its narrow, outdated definition of rape. Retrieved from <https://www.vox.com/2014/11/14/7214149/the-fbis-finally-collecting-modern-rape-stats>
16. Hayes, Adam. “What Is Unemployment? Understanding Causes, Types, Measurement.” *Investopedia*, [www.investopedia.com/terms/u/unemployment.asp#:~:text=for%20work.%222-,The%20Bottom%20Line,rate%20represents%20a%20weak%20economy](http://www.investopedia.com/terms/u/unemployment.asp#:~:text=for%20work.%222-,The%20Bottom%20Line,rate%20represents%20a%20weak%20economy). Accessed 29 Oct. 2023.
17. Great Recession, great recovery? Trends from the Current Population Survey : Monthly Labor Review. (n.d.). Retrieved from <https://www.bls.gov/opub/mlr/2018/article/great-recession-great-recovery.htm>
18. United States. Bureau of the Census, & United States. Bureau of Labor Statistics. (2020). Retrieved from <https://www.icpsr.umich.edu/web/NADAC/studies/37652>

## Appendix

**Table 1: Rates of unemployment, annual change in unemployment, educational attainment, year, policies and events, and rates of rape per 100,000**



<b>Historical Event</b>	<b>Year</b>	<b>Rapes per 100,000</b>	<b>% High School Graduates &gt;</b>	<b>% College Graduates</b>	<b>Poverty Rate</b>	<b>Unemployment Rate</b>	<b>UR Annual Delta</b>
	1990	41.2	77.6	21.3	13.5		
	1991	42.3	78.4	21.4	14.2	6.8	
LA riots/policing ramped up in 93-94	1992	42.8	79.4	21.4	14.8	7.5	0.7
Bill Clinton elected into presidency/Marital Rape Laws	1993	41.1	80.2	21.9	15.1	6.9	-0.6
Bill Clinton passed Law Enforcement Act/Megan's Law/Violence Against Women Act	1994	39.3	80.9	22.2	14.5	6.12	-0.78
	1995	37.1	81.7	23	13.8	5.65	-0.47
	1996	36.3	81.7	23.6	13.7	5.45	-0.2
	1997	35.9	82.1	23.9	13.3	5	-0.45
US elections/invasion of Afghanistan/economic growth and Tax Relief Reconciliation Act of 2001	1998	34.5	82.8	24.4	12.7	4.51	-0.49
	1999	32.8	83.4	25.2	11.9	4.22	-0.29
	2000	32	84.1	25.6	11.3	3.99	-0.23
9/11	2001	31.8	84.1	26.2	11.7	4.73	0.74
<b>Historical Event</b>	<b>Year</b>	<b>Rapes per 100,000</b>	<b>% High School Graduates &gt;</b>	<b>% College Graduates</b>	<b>Poverty Rate</b>	<b>Unemployment Rate</b>	<b>UR Annual Delta</b>

	2002	33.1	84.1	26.7	12.1	5.78	1.05
	2003	32.3	84.6	27.2	12.5	5.99	0.21
Same-sex marriage legalized in Massachusetts (Goodridge v. Department of Public Health)	2004	32.4	85.2	27.7	12.7	5.53	-0.46
	2005	31.8	85.2	27.7	12.6	5.08	-0.45
	2006	31.6	85.5	28	12.3	4.62	-0.46
Late 2000s recession/first iPhone is released for sale	2007	30.6	85.7	28.7	12.5	4.62	0
	2008	29.8	86.6	29.4	13.2	5.78	1.16
Obama took office	2009	29.1	86.7	29.5	14.3	9.25	3.47
	2010	27.7	87.1	29.9	15.1	9.63	0.38
	2011	27	87.6	30.4	15	8.95	-0.68
	2012	27.1	87.6	30.9	15	8.07	-0.88
Change in definition of rape	2013	35.9	88.2	31.7	14.8	7.37	-0.7
	2014	37	88.3	32	14.8	6.17	-1.2
	2015	39.3	88.4	32.5	13.5	5.28	-0.89
	2016	40.9	89.1	33.4	12.7	4.87	-0.41
Trump took office	2017	41.7	89.6	34.2	12.3	4.36	-0.51
	2018	44	89.8	35	11.8	3.9	-0.46
	2019	43.6	90.1	36	10.5	3.67	-0.23
The first case of Covid in the US/bars was closed for social distancing reasons	2020	42.3	90.9	37.5	11.5	8.05	4.38
	2021	43.5	91.1	37.9	11.6	5.35	-2.7

**Table 2: Results from our statistical analysis**

OLS Regression Results			
Dep. Variable:	Rate of Rape per 100000 (counts)	R-squared:	0.824
Model:	OLS	Adj. R-squared:	0.788
Method:	Least Squares	F-statistic:	22.52
Date:	Fri, 05 Jan 2024	Prob (F-statistic):	2.41E-08
Time:	18:23:24	Log-Likelihood:	-66.195
No. Observations:	30	AIC:	144.4
Df Residuals:	24	BIC:	152.8
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std-err	t	P> t	[0.025	0.975]
const	569.718	61.932	9.199	<1X10 <sup>-4</sup>	441.897	697.539
High school graduate or more	-8.6945	0.958	-9.072	<1X10 <sup>-4</sup>	-10.673	-6.716
College graduate or more	6.6267	0.69	9.605	<1X10 <sup>-4</sup>	5.203	8.051
Unemployment Rate	-2.6252	0.651	-4.034	<1X10 <sup>-4</sup>	-3.968	-1.282
Annual Change	0.7185	0.518	1.388	0.178	-0.35	1.787
Poverty Rate	2.7284	0.79	3.452	0.002	1.097	4.359

	Year	Rate of Rape per 100000 (counts)	High school graduate or more	College graduate or more	Unemployment Rate	Annual Change	Poverty Rate
Year	1	0.570671	0.99588	0.994829	0.338443	0.344225	0.383893

Rate of Rape per 100000	0.570671	1	0.574897	0.611598	0.406715	0.320772	0.361915
High school graduate or more	0.99588	0.574897	1	0.993365	0.347196	0.34814	0.415181
College graduate or more	0.994829	0.611598	0.993365	1	0.334068	0.364189	0.427222
Unemployment Rate	0.338443	0.406715	0.347196	0.334068	1	0.484201	0.743473
Annual Change	0.344225	0.320772	0.34814	0.364189	0.484201	1	0.298397
Poverty Rate	0.383893	0.361915	0.415181	0.427222	0.743473	0.298397	1

Random Forest Feature Importances:	
Year	0.287873
High school graduate or more	0.221304
College graduate or more	0.223678
Unemployment Rate	0.141392
Annual Change	0.0323136
Poverty Rate	0.0934397