MIDNAPORE COLLEGE

(AUTONOMOUS)

DEPARTMERNT OF STATISTIC

End Semester Examination of Semester-VI

PROJECT WORK

Submitted By

Roll – MC/UG/S-VI/20

No -1612

Reg. No – 2017-331 of 2017-2018

PROJECT WORK ON

Statistical Analysis of Crime in West Bengal

Table of Contents

Introduction

"Crimes increase as education, opportunity, and property decrease. Whatever spreads ignorance, poverty and, discontent causes crime. Criminals have their own responsibility, their own share of guilt, but they are merely the hand…."

Rutherford Birchard Hayes. (Hayes,1922:629).

[The Oxford Dictionary of Sociology defines crime in a more complex way: 'an offence which goes beyond the personal and into the public sphere, breaking prohibitory rules or laws, to which legitimate punishments or sanctions are attached, and which requires the intervention of a public authority.' In ordinary language, a crime is an unlawful act punishable by a state or other authority. The term crime does not, in modern criminal law, have any simple and universally accepted definition, though statutory definitions have been provided for certain purposes.

If we want to understand what crime is, we have to discuss about history of crime. Crime is a very old concept and it is transmitted to the society from generation to generation. Crime produces law and order situation. It is a social evil. Crime as an integral part of a society existed from the very beginning of the civilization. In course of search for historical definition of crime this Study followed what the colonial rules and regulations perceived as crime. It is difficult to define what the historical definition of crime in colonial India was as the rulers who

codified the criminality of Indians had all kinds of prejudices against the people whom they ruled.

Before coming to my subject crimes of west Bengal, it would be useful to have a overall idea about types of crime occur. Many types of crime exist. Criminologists commonly group crimes into several major categories: (1) violent crime; (2) property crime; (3) white-collar crime; (4) organized crime; and (5) consensual or victimless crime. Within each category, many more specific crimes exist. For example, violent crime includes homicide, aggravated and simple assault, rape and sexual assault, and robbery, while property crime includes burglary, larceny, motor vehicle theft, and arson.

Now coming to my subject 'Crime in West Bengal in the period of 2000-2013'. In comparison to the western world, however, the research in social history of crime in Bengal is still in infant stage. Especially, micro level study of crime and criminality of colonial Bengal has, so far, drawn little attention from researchers and scholars. One historian aptly said, "the historians of Indian history once almost knew nothing about the overall patterns of crime –its incidence and forms at national, regional or local level." The period I have chosen to study is a period regarded as growth and development phase of Bengal. Along with economic growth and development and population growth the crime rate also increased. According NCRB, total violent crimes, which include murder, attempt to murder, rape, dacoity, dowry deaths, has always seen an increasing trend in

West Bengal during this phase 2000 -2013. There was a huge increase during 2008 and 2009. The year 2006 and 2011 had the lowest increase compared to the other years. Incidentally, the state assembly. The Director-General of Police is incorrect in saying that rate of violent crimes has come down in the state; while he is correct while saying rape, cases have decreased. The number of rape cases went down in 2012 but in the last decade it has mostly been increasing.

Objectives

1. Descriptive analysis
2. Is there any association between year and types of crimes?
3. Is there any association between districts and types of crimes?

Data Description

Data Source: https://www.kaggle.com/rajanand/crime-in-india

About Data: From the above link we pick only two files containing district and year wise data and then collect data only for West Bengal.

For each districts of West Bengal, 13 Years (2001-2013) data have number of crimes on following 24 variables:

1)MURDER, 2)ATTEMPT TO MURDER, 3)CULPABLE HOMICIDE NOT AMOUNTING TO MURDER, 4)RAPE, 5)KIDNAPPING, 6)KIDNAPPING & ABDUCTION, 7)KIDNAPPING AND ABDUCTION OF WOMEN AND GIRLS, 8)KIDNAPPING AND ABDUCTION OF OTHERS, 9)DACOITY, 10)PREPARATION AND ASSEMBLY FOR DACOITY , 11)ROBBERY, 12)BURGLARY, 13)THEFT, 14)AUTO THEFT, 15)OTHER THEFT, 16)RIOTS, 17)DOWRY DEATHS, 18)ASSAULT ON WOMEN WITH INTENT TO OUTRAGE HER MODESTY, 19)INSULT TO MODESTY OF WOMEN, 20)CRUELTY BY HUSBAND OR HIS RELATIVES, 21)IMPORTATION OF GIRLS FROM FOREIGN COUNTRIES, 22)CAUSING DEATH BY NEGLIGENCE , 23)OTHER IPC CRIMES AND 24)TOTAL IPC CRIMES

Methodology

Chi-Square Test for independence:

Suppose that n observations in a random sample from a population are classified into k mutually exclusive classes with respective observed numbers x_i (for $i = 1, 2, \ldots, k$), and a null hypothesis gives the probability p_i that an observation falls into the ith class. So, we have the expected numbers $m_i = np_i$ for all i, where

$$\sum_{i=1}^{k} p_i = 1; \qquad \sum_{i=1}^{k} m_i = n\sum_{i=1}^{k} p_i = \sum_{i=1}^{k} x_i$$

Pearson proposed that, under the circumstance of the null hypothesis being correct, as $n \to \infty$ the limiting distribution of the quantity given below is the χ^2 distribution.

$$X^2 = \sum_{i=1}^{k} \frac{(x_i - m_i)^2}{m_i} = \sum_{i=1}^{k} \frac{{x_i}^2}{m_i} - n$$

Pearson dealt first with the case in which the expected numbers m_i are large enough known numbers in all cells assuming every x_i may be taken as normally distributed, and reached the result that, in the limit as n becomes large, X^2 follows the χ^2 distribution with $k - 1$ degrees of freedom.

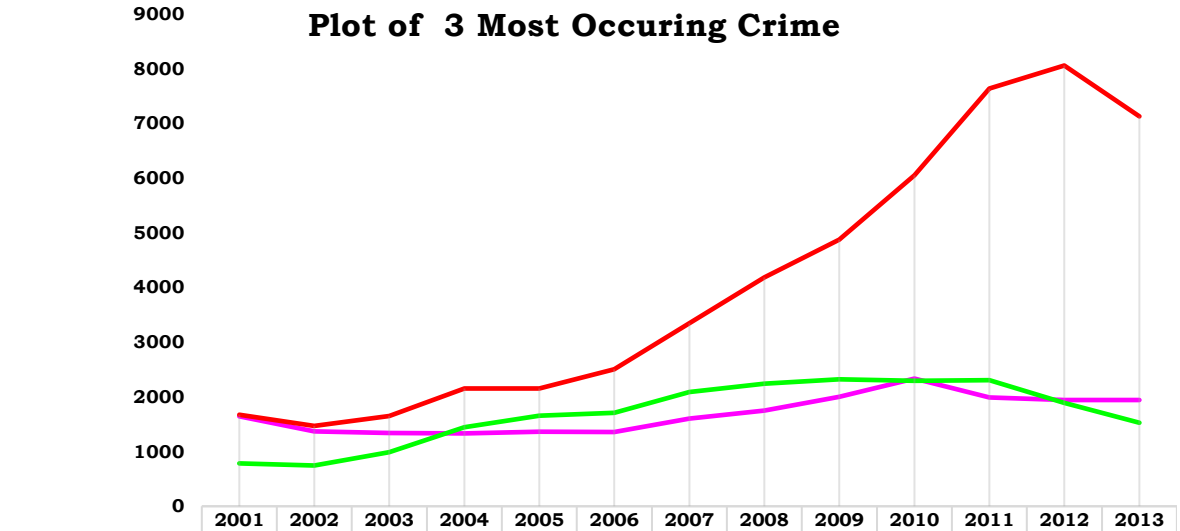
However, Pearson next considered the case in which the expected numbers depended on the parameters that had to be estimated from the sample, and suggested that, with the notation of m_i being the true expected numbers and m'_i being the estimated expected numbers, the difference

$$X^2 - X'^2 = \sum_{i=1}^{k} \frac{{x_i}^2}{m_i} - \sum_{i=1}^{k} \frac{{x_i}^2}{{m_i}'}$$

will usually be positive and small enough to be omitted. In a conclusion, Pearson argued that if we regarded X'^2 as also distributed as χ^2 distribution with $k - 1$ degrees of freedom, the error in this approximation would not affect practical decisions.

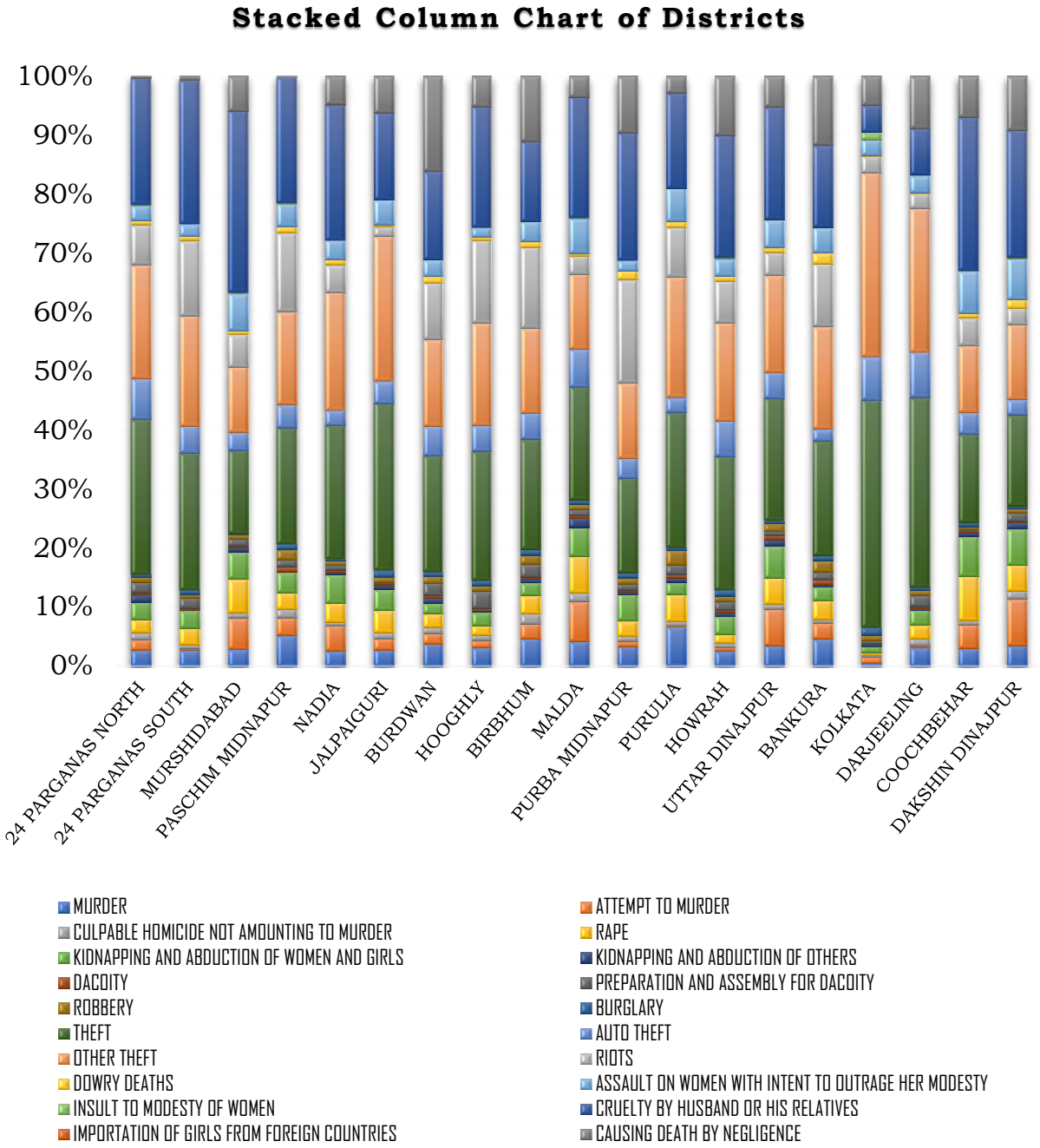
Data Analysis

❖ Descriptive Analysis



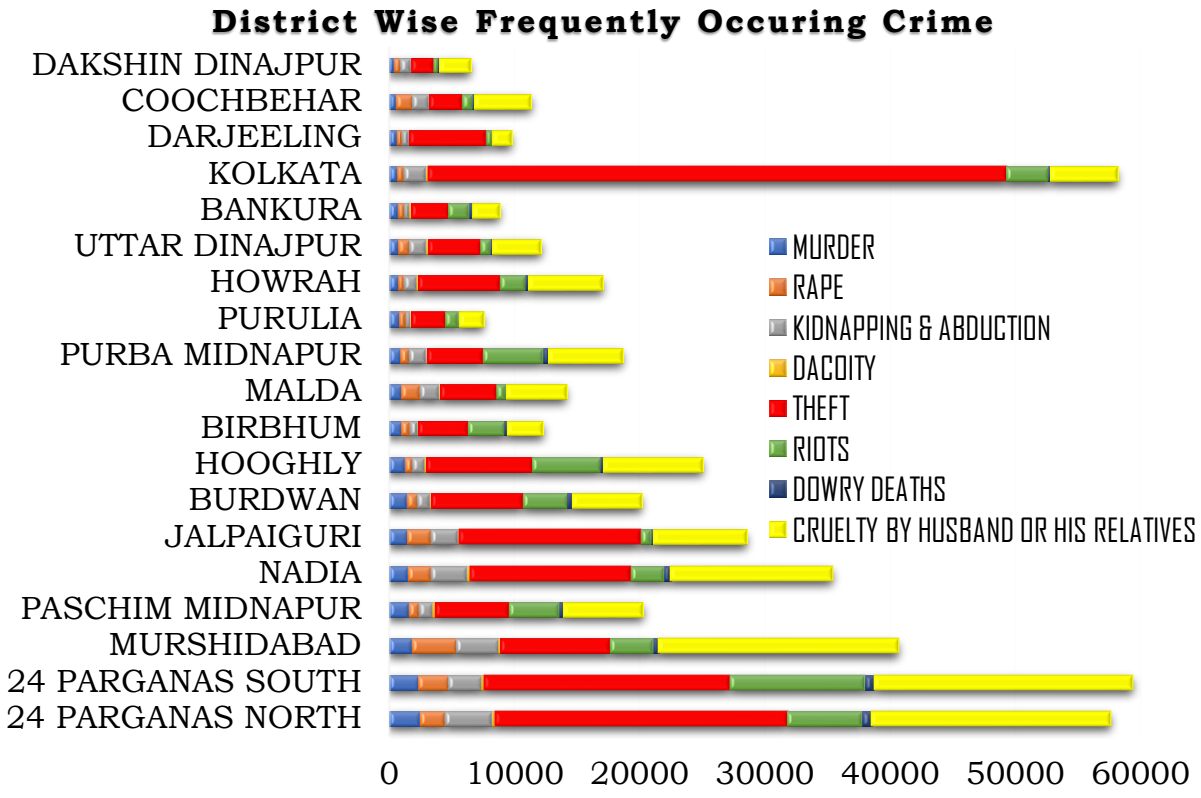
	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
MURDER	1641	1371	1338	1332	1364	1357	1604	1747	2000	2334	1987	1942	1945
RAPE	784	745	988	1445	1653	1711	2086	2240	2321	2291	2306	1887	1524
KIDNAPPING	1671	1469	1650	2151	2155	2504	3344	4183	4876	6049	7636	8057	7125

➢ Increasing rate of kidnapping is very high.



Most Frequent Crimes are:

1) Cruelty by husband or his relatives
2) Riots
3) Other theft and theft
4) Kidnapping
5) Rape
6) Murder and Attempt to murder



- Theft is mostly occurred among all the frequent crimes.
- In 2001-13, 25% theft occur in Kolkata.
- Cruelty by husband or her relatives also a main crime of West Bengal.

- Here we are interested in checking any association between year and types of crimes. So, our hypotheses are

  H_0: Time (year) and types of crimes are independent vs

  H_1: Time (year) and types of crimes are not independent.

```
dw= read.csv("C:/Users/SUSMITA/Desktop/sum_district_wise.csv")
row.names(dw) = dw$DISTRICT
dw = dw [, -c (1,2)]
r=chisq.test(dw)
r

Pearson's Chi-squared test
data:  dw
X-squared = 186413, df = 396, p-value < 2.2e-16
```

As the p-value is less than 0.05, we reject the null hypothesis and conclude that they are not independent i.e. there are some association between year and types of crimes.

- Now we are going to check, is there any association between districts and types of crimes. So, our hypotheses are

  H_0: Districts and types of crimes are independent vs

  H_1: Districts and types of crimes are not independent.

```
dw = read.csv ("C:/Users/SUSMITA/Desktop/sum_year_wise.csv ")
row.names(dw) = dw$YEAR
dw = dw [, -c (1,2)]
r=chisq.test(dw)
r

Pearson's Chi-squared test

data:  dw
X-squared = 42216, df = 264, p-value < 2.2e-16
```

As the p-value is less than 0.05, we reject the null hypothesis and conclude that they are not independent i.e. there are some association between districts and types of crimes.

Conclusion

- From the chi-square test, year and types of crimes are not independent.
- Also districts and types of crimes are not independent.
- Theft and Cruelty by husband or her relatives also a main crime of West Bengal.

Reference

1. Fundamental of Statistics- Volume I & II, Goon; Gupta; Dasgupta
2. An Outline of Statistical Theory- Volume I & II, Goon; Gupta; Dasgupta
3. Mathematical Statistics

Appendix

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1		YEAR	MURDER	ATTEMPT	CULPABLE	RAPE	KIDNAPPI	KIDNAPPI	KIDNAPPI	KIDNAPPI	DACOITY	PREPARAT	ROBBERY	BURGLARY	THEFT	AUTO THE	OTHER TH	RIOTS	DOWRY DI	ASSAULT (
2	1	2001	1641	448	607	784	1671	917	754	163	277	350	680	459	13381	2502	10879	3322	317	992
3	2	2002	1371	396	524	745	1469	797	672	125	192	472	492	345	11352	2337	9015	2386	266	949
4	3	2003	1338	383	464	988	1650	871	779	92	200	665	540	332	11567	2479	9088	2505	319	1148
5	4	2004	1332	371	449	1445	2151	1155	996	159	190	804	479	523	13201	2492	10709	2572	382	1497
6	5	2005	1364	418	388	1653	2155	1152	1003	149	201	1005	381	716	10805	2164	8641	2485	432	1530
7	6	2006	1357	650	379	1711	2504	1327	1177	150	166	1186	393	591	11357	2677	8680	2323	427	1813
8	7	2007	1604	863	385	2086	3344	1772	1572	200	135	1071	408	398	12853	2376	10477	3694	439	2250
9	8	2008	1747	1675	474	2240	4183	2299	1884	415	178	1007	599	457	16246	2473	13773	5375	442	2367
10	9	2009	2000	2094	916	2321	4876	2714	2162	552	206	670	714	323	16123	3569	12554	6617	485	1916
11	10	2010	2334	2084	614	2291	6049	3307	2742	565	278	594	771	383	18005	3946	14059	6741	486	2436
12	11	2011	1987	2224	459	2306	7636	4089	3547	542	220	809	684	412	18465	4105	14360	5827	498	2310
13	12	2012	1942	2694	463	1887	8057	4420	3637	783	226	638	576	398	16668	3528	13140	6078	522	2877
14	13	2013	1945	3338	368	1524	7125	3848	3277	571	170	701	485	387	16745	3367	13378	5649	423	4028

	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	DACOITY	PREPARAT	ROBBERY	BURGLARY	THEFT	AUTO THE	OTHER TH	RIOTS	DOWRY DI	ASSAULT (	INSULT TO	CRUELTY B	IMPORTA	CAUSING I	OTHER IPC	CRIMES
2	268	1649	785	398	23336	6209	17127	5995	718	2226	55	19162	8	188	64593	
3	212	1499	433	690	19652	3900	15752	10807	691	1650	21	20650	3	482	67134	
4	158	799	413	120	8766	1896	6870	3452	419	3894	40	19168	0	3587	19970	
5	245	377	497	300	5881	1166	4715	3993	338	1126	28	6374	0	43	21392	
6	254	468	420	179	12811	1466	11345	2618	544	1823	45	12962	72	2649	27086	
7	135	118	512	523	14480	1920	12560	818	135	2192	7	7524	0	3104	27484	
8	154	800	436	250	7279	1831	5448	3557	395	1047	14	5538	1	5904	23595	
9	138	1140	373	344	8464	1737	6727	5463	207	600	21	7962	0	1945	19235	
10	89	448	323	188	3973	941	3032	2923	188	725	2	2874	0	2316	7662	
11	72	251	220	156	4494	1510	2984	718	100	1382	19	4818	0	799	17057	
12	114	286	267	215	4425	944	3481	4795	427	471	6	5933	0	2602	20979	
13	68	200	300	51	2752	294	2458	1014	114	679	0	1936	0	330	8192	
14	158	411	271	288	6509	1732	4777	2024	236	826	65	5947	0	2850	22761	
15	141	151	262	86	4220	898	3322	791	155	946	12	3888	1	1039	12913	
16	151	193	306	127	2994	315	2679	1639	281	660	9	2136	0	1793	6826	
17	155	610	1049	1516	46227	8846	37381	3355	133	3279	1519	5367	36	5783	67803	
18	61	365	157	103	6099	1463	4636	479	33	540	29	1504	1	1650	6737	
19	49	32	99	127	2627	645	1982	811	154	1245	13	4556	0	1207	13592	
20	17	175	79	63	1779	302	1477	322	170	802	13	2494	0	1054	7864	

	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	ATTEMPT TO MURDER	CULPABLE HOMICIDE NOT AMOUNTING TO MURDER	RAPE	KIDNAPPING & ABDUCTION	KIDNAPPING AND ABDUCTION OF WOMEN AND GIRLS	KIDNAPPING AND ABDUCTION OF OTHERS	DACOITY	PREPARATION AND ASSEMBLY FOR DACOITY	ROBBERY	BURGLARY	THEFT	AUTO THEFT	OTHER THEFT	RIOTS	DOWRY DEATH	ASSAULT ON WOMEN WITH INTENT TO OUTRAGE HER MODESTY	INSULT TO MODESTY OF WOM	CRUELTY BY HUSBAND OR HIS RELATIVES	IMPORTATION OF GIRLS FROM FOREIGN COUNTRIES	CAUSING DEATH BY NEGLIGENCE	OTHER IPC CRIM	TOTAL IPC CRIM
2	23	61	53	81	63	18	14	0	56	46	1679	459	1220	285	23	63	0	447	0	0	3036	6167
26	55	47	88	91	63	28	10	0	60	21	1842	467	1375	248	29	84	1	630	0	0	3360	6894
51	30	50	92	119	103	16	29	0	73	29	1988	615	1373	243	24	105	3	672	0	0	3571	7351
76	43	50	176	180	124	56	32	72	72	60	2196	701	1495	208	55	154	4	810	0	3	4228	8827
101	70	38	217	211	153	58	18	196	39	78	1564	493	1071	203	71	151	5	886	0	7	3707	7958
126	82	39	174	186	140	46	15	397	29	37	1440	431	1009	180	74	175	5	1049	0	3	3200	7776
150	6	46	231	236	236	0	13	238	18	14	1183	269	914	545	53	187	2	1696	0	113	3917	9201
174	168	47	192	317	215	102	15	329	43	19	1452	257	1195	579	67	209	5	1946	0	12	4973	11247
198	238	241	166	368	249	119	21	99	88	22	1959	491	1468	778	54	203	3	2225	0	4	5549	13046
222	224	179	173	438	292	146	15	108	83	28	2312	576	1736	815	71	174	8	2378	5	11	7123	15151
246	215	126	203	587	414	173	32	150	117	27	2918	811	2107	720	88	218	10	2847	0	17	7918	17412
271	108	80	118	497	338	159	34	40	61	6	1453	381	1072	680	74	196	1	1860	3	10	6062	11765
300	280	52	115	428	258	170	20	20	46	11	1350	258	1092	511	35	307	8	1716	0	8	7949	13409