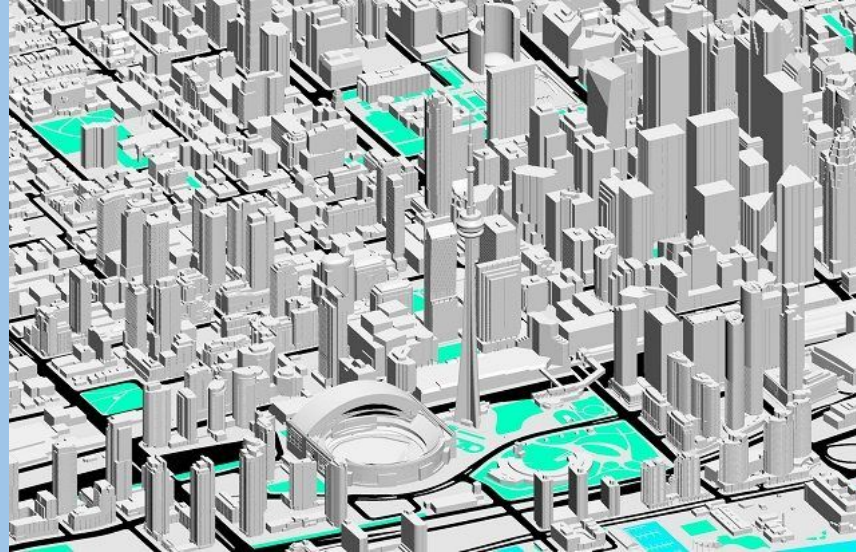


Exploring Toronto's Crimes



Neil Balazon
Mei Chang
Andrew Ivanov
Sam Lee
Skye Morgan

Objective

Our main objective is to study Toronto's crimes in hopes of learning how Neighborhoods are being affected

Motivation of Research Question

- Canada has a reputation of being a very safe country compared to the United States
- Toronto's crime rate has been rising steadily since 2015
 - 90% increase in people wounded by gunfire
 - 48% increase in shootings
- Studying the data would allow us to narrow down the neighborhoods with the most crime rates
- Government agencies could then use our findings in their initiatives to decrease crime in Toronto

Dataset Description



- Dataset shared by Toronto Police Service and posted on public safety portal website (<http://data.torontopolice.on.ca/datasets/mci-2014-to-2018>)
- MCI categories are Assault, Auto Theft, Break and Enter, Murder, Robbery, Sexual Violation and Theft Over
- 167,525 Records
- There are several variables of interest such as offence, neighbourhood, occurrence, hour and MCI (Major Crime Indicators)

Division Text	event_unique_id Text	Hood_ID Number	Index Number	Lat Number	Long Number	MCI Text	Neighbourhood Text	Objectid Unique ID
occurrence_date Date or Time	occurrence_day Number	occurrence_day_of_week Text	occurrence_day_of_year Number	occurrence_hour Number				
occurrence_month Text	occurrence_year Number	offence Text	premise_type Text	reported_date Date or Time	reported_day Number			
reported_day_of_week Text	reported_day_of_year Number	reported_hour Number	reported_month Text	reported_year Number	ucr_code Number	ucr_ext Number		

Variables of Interest	Description
Neighbourhood	The different neighborhoods (141) in the city of Toronto
Assault	Direct or indirect application of force to another person, or the attempt or threat to apply force to another person, without that person's consent
Auto Theft	Act of taking or another person's vehicle (not including attempts). Auto Theft figures represent the number of vehicles stolen.
Break and Enter	Act of entering a place with the intent to commit an indictable offence therein.
Robbery	Act of taking property from another person or business by the use of force or intimidation in the presence of the victim.
Theft Over	Act of stealing property in excess of \$5,000 (excluding auto theft)

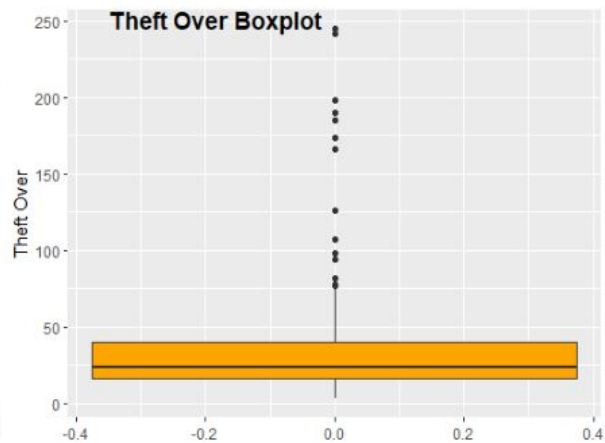
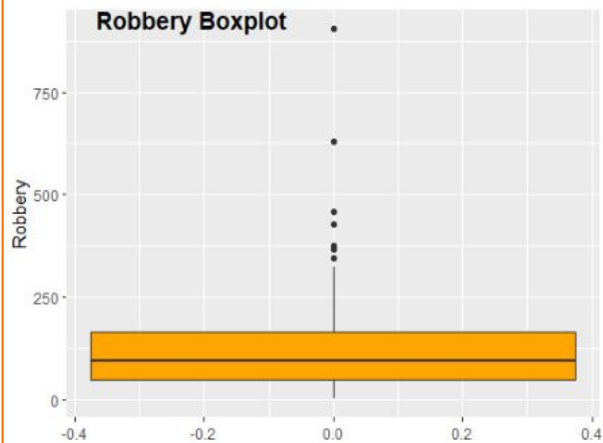
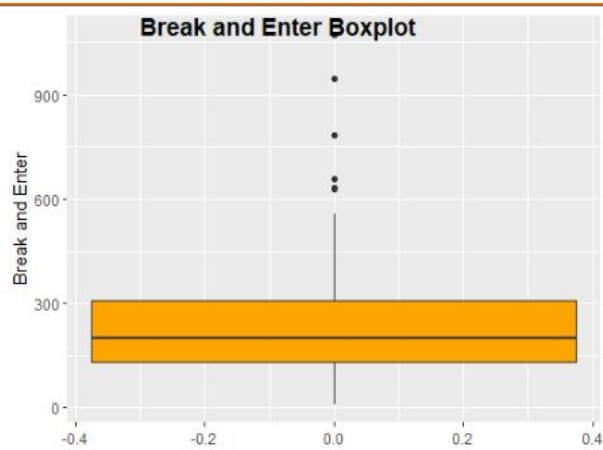
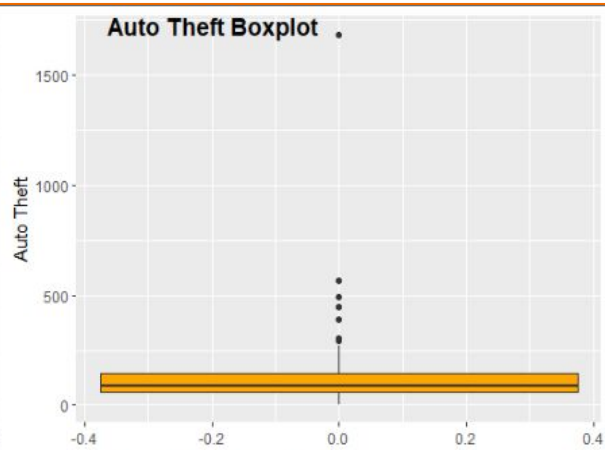
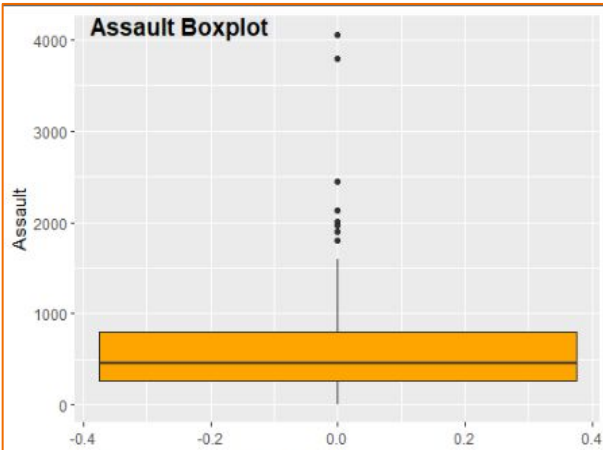
	Neighbourhood	Assault	Auto Theft	Break and Enter	Robbery	Theft Over
1	Agincourt North (129)	392	133	294	162	27
2	Agincourt South-Malvern West (128)	580	159	381	148	63
3	Alderwood (20)	182	84	124	36	34
4	Annex (95)	1523	115	782	314	174
5	Banbury-Don Mills (42)	398	84	341	70	45
6	Bathurst Manor (34)	417	174	240	69	51
7	Bay Street Corridor (76)	2446	109	538	286	185
8	Bayview Village (52)	415	88	217	38	40
9	Bayview Woods-Steeles (49)	180	61	157	19	11
10	Bedford Park-Nortown (39)	270	220	513	75	57
11	Beechborough-Greenbrook (112)	329	62	119	94	11
12	Bendale (127)	1246	257	357	347	70
13	Birchcliffe-Cliffside (122)	827	81	377	98	28
14	Black Creek (24)	1302	258	144	252	50
15	Blake-Jones (69)	261	22	83	56	3
16	Briar Hill-Belgravia (108)	613	114	197	167	26
17	Bridle Path-Sunnybrook-York Mills (41)	113	41	237	6	23
18	Broadview North (57)	180	20	70	24	5
19	Brookhaven-Amesbury (30)	422	145	140	93	29

Data Summary Statistics

```
> summary(toronto_wide)
```

	Neighbourhood	Assault	Auto Theft	Break and Enter	Robbery
Agincourt North (129)	: 1	Min. : 4.0	Min. : 4.0	Min. : 9.0	Min. : 3.0
Agincourt South-Malvern West (128)	: 1	1st Qu.: 264.0	1st Qu.: 60.0	1st Qu.: 129.0	1st Qu.: 50.0
Alderwood (20)	: 1	Median : 450.0	Median : 89.0	Median : 200.0	Median : 95.0
Annex (95)	: 1	Mean : 644.5	Mean : 128.9	Mean : 247.6	Mean : 128.6
Banbury-Don Mills (42)	: 1	3rd Qu.: 797.0	3rd Qu.: 145.0	3rd Qu.: 308.0	3rd Qu.: 165.0
Bathurst Manor (34)	: 1	Max. : 4061.0	Max. : 1684.0	Max. : 1074.0	Max. : 904.0
(Other)	: 135				

Theft Over
Min. : 3.00
1st Qu.: 16.00
Median : 24.00
Mean : 38.51
3rd Qu.: 40.00
Max. : 245.00



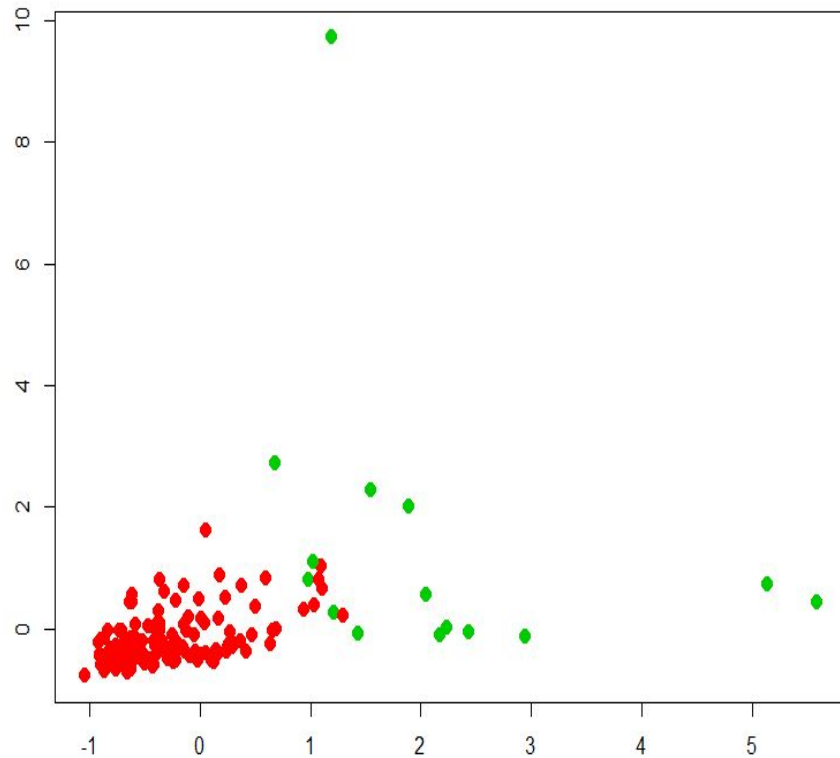
K-Means Clustering

- Setting out, we wanted to gain some insight into the nature of how crime was affecting the Neighbourhoods in Toronto
- That would involve looking into the individual crime statistics for each Neighbourhood
- We quickly recognized that Clustering would be a suitable Data Mining Method because it would provide a structure in our collection of data
- Additionally, k-means was selected because it proved to be relatively simple to implement and understand

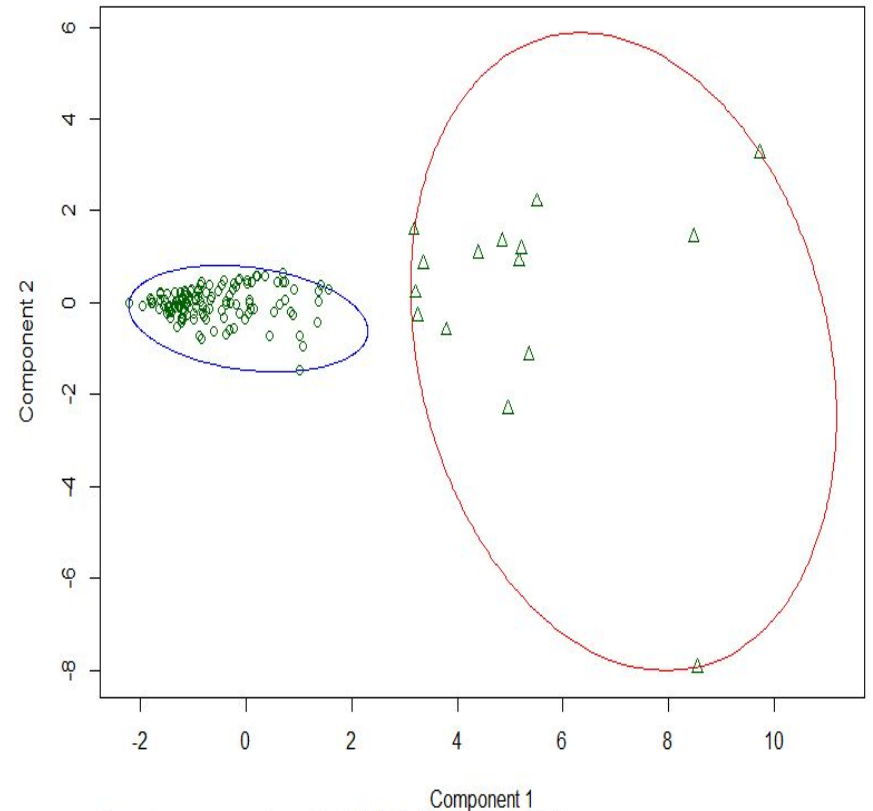
	row.names	Assault	Auto Theft	Break and Enter	Robbery	Theft Over
1	1	-0.4129285	0.02547666	0.2711717	0.2743428	-0.2640418
2	2	-0.105511	0.187907	0.7795721	0.159461	0.5617598
3	3	-0.7563204	-0.280642	-0.7222543	-0.7595932	-0.1034692
4	4	1.436482	-0.08697509	3.122889	1.521631	3.107981
5	5	-0.4031173	-0.280642	0.5458248	-0.4805946	0.148859
6	6	-0.3720485	0.2816168	-0.04438712	-0.4888005	0.2864926
7	7	2.945771	-0.124459	1.69703	1.291867	3.36031
8	8	-0.3753189	-0.2556527	-0.1787918	-0.7431815	0.03416437
9	9	-0.7595908	-0.4243303	-0.5294128	-0.8990925	-0.6310647
10	10	-0.6124228	0.5689935	1.550938	-0.4395654	0.4241262
11	11	-0.515946	-0.418083	-0.7514727	-0.2836544	-0.6310647
12	12	0.9835319	0.8001443	0.6393237	1.792424	0.7223324
13	13	0.2983833	-0.299384	0.7561974	-0.250831	-0.2411028
14	14	1.075103	0.8063916	-0.6053806	1.012869	0.2635537
15	15	-0.6271396	-0.6679758	-0.9618453	-0.5954764	-0.8145762
16	16	-0.05154938	-0.09322241	-0.2956655	0.315372	-0.2869807
17	17	-0.8691491	-0.5492767	-0.06191817	-1.005768	-0.3557975
18	18	-0.7595908	-0.6804705	-1.037813	-0.8580633	-0.7686983
19	19	-0.3638725	0.1004445	-0.6287554	-0.2918603	-0.2181639

Clustering with $k=2$

K-Means Clustering Results with K=2

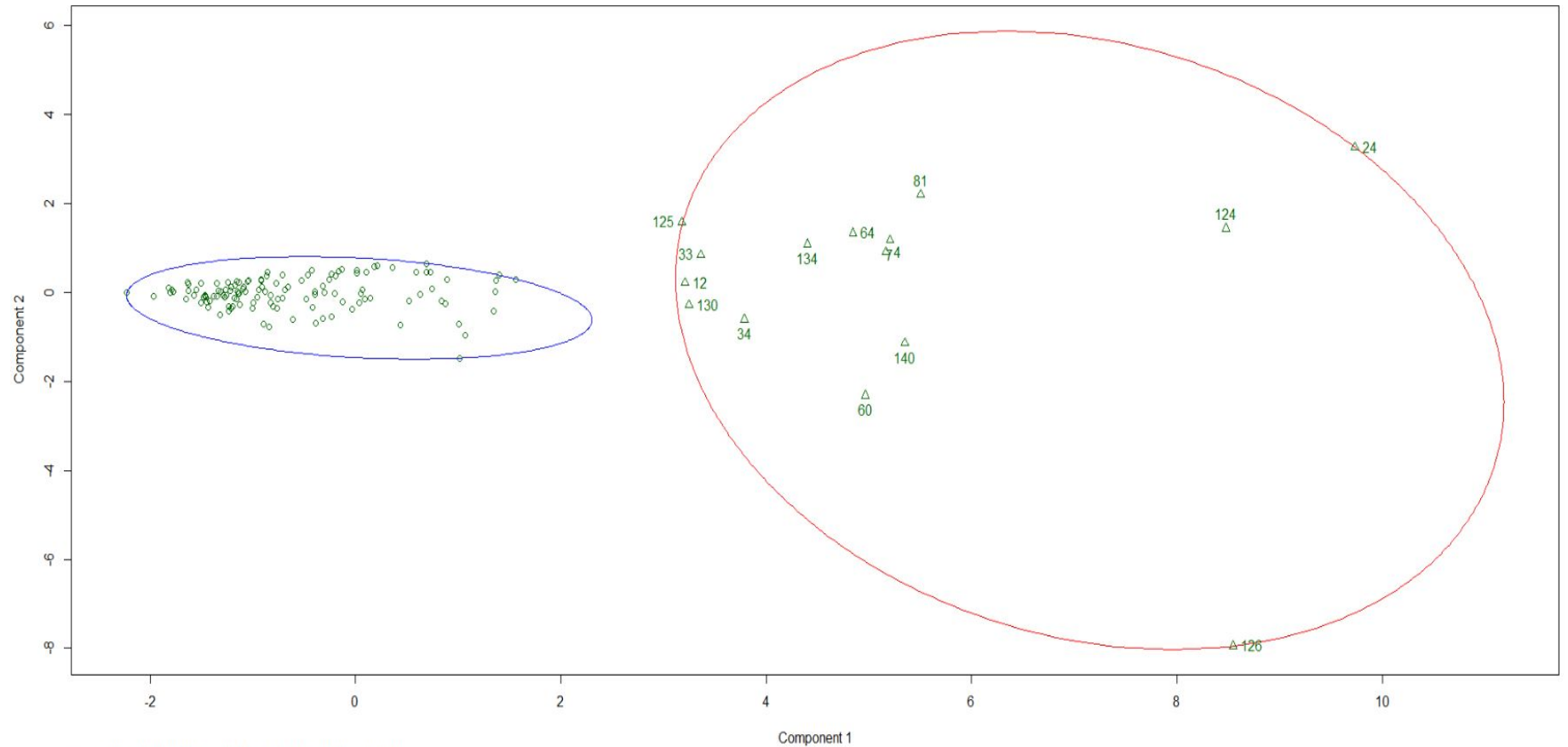


k-Means Cluster Analysis



These two components explain 86.13 % of the point variability.

k-Means Cluster Analysis



Results using K=2

- The first cluster has 126 Neighbourhoods while the second one has 15
- Cluster 1 has neighbourhoods with low assault, low auto theft, low break and enter, low robbery and low theft over
- Cluster 2 has neighbourhoods with high assault, high auto theft, high break and enter, high robbery and high theft over

```
> km.out
```

```
K-means clustering with 2 clusters of sizes 126, 15
```

```
Cluster means:
```

	Assault	Auto Theft	Break and Enter	Robbery	Theft Over
1	-0.2582983	-0.1608521	-0.245113	-0.2452954	-0.2607648
2	2.1697057	1.3511579	2.058949	2.0604810	2.1904241

```
Within cluster sum of squares by cluster:
```

```
[1] 167.3139 200.0593
```

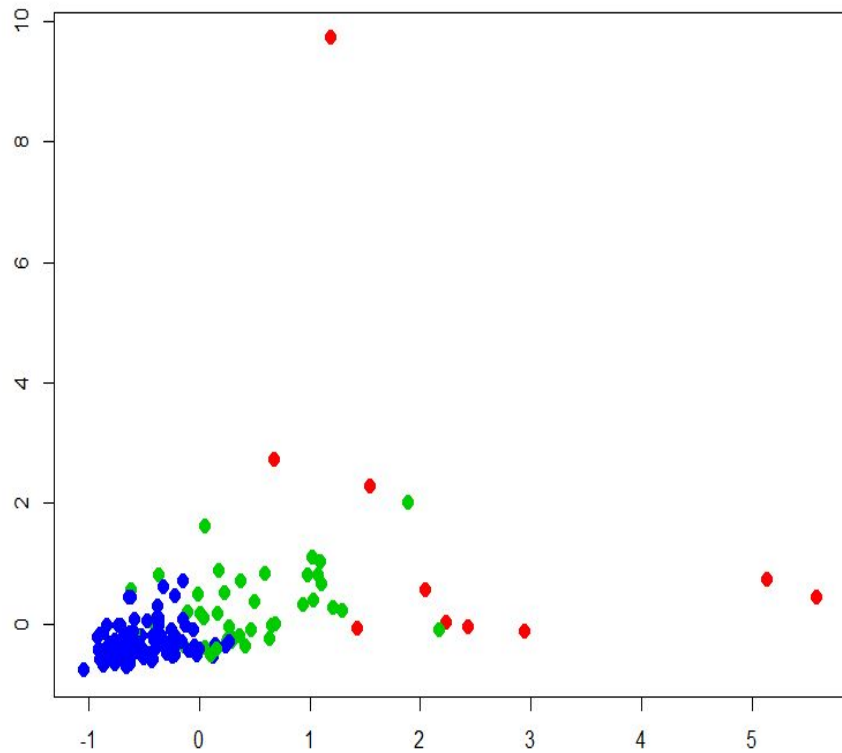
```
(between_SS / total_SS = 47.5 %)
```

```
> km.out$tot.withinss
```

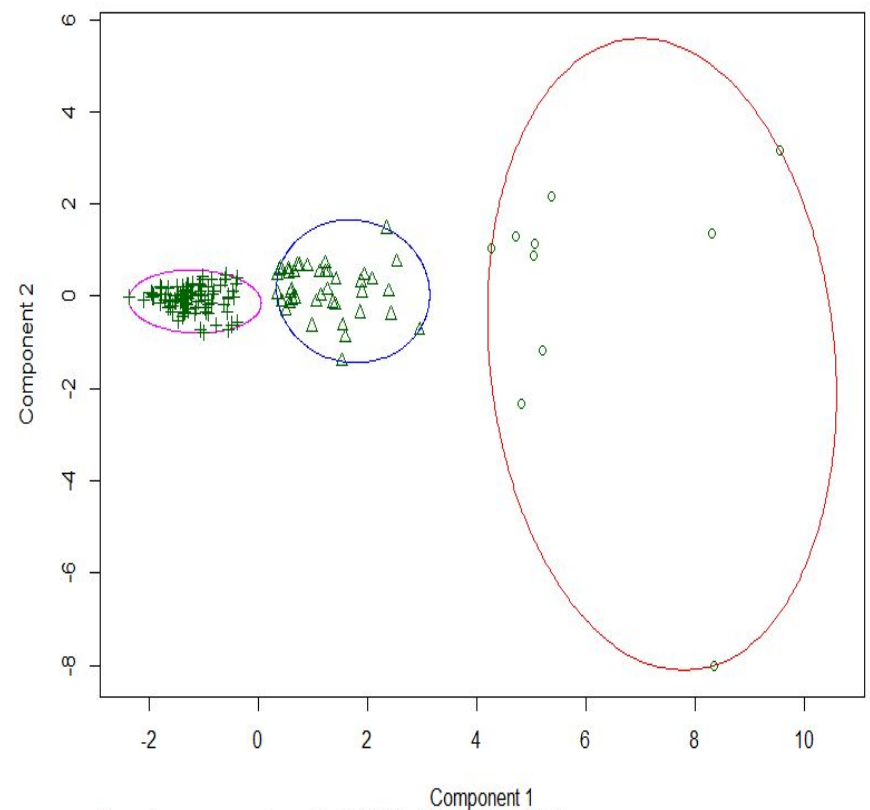
```
[1] 367.3733
```

Clustering with $k=3$

K-Means Clustering Results with K=3



k-Means Cluster Analysis



Component 1
These two components explain 88.01 % of the point variability.

Results using K=3

- The first cluster has 10 Neighbourhoods, the second one has 41 and the third cluster has 90.
- Cluster 1 has neighbourhoods with high assault, high auto theft, high break and enter, high robbery and high theft over
- Cluster 2 has neighbourhoods with medium assault, medium auto theft, medium break and enter, medium robbery and medium theft over
- Cluster 3 has neighbourhoods with low assault, low auto theft, low break and enter, low robbery and low theft over
- **Minimized tot.withinss**
- **Maximized Within cluster sum of squares**
- **k-means minimize the within group dispersion and maximize the between-group dispersion**

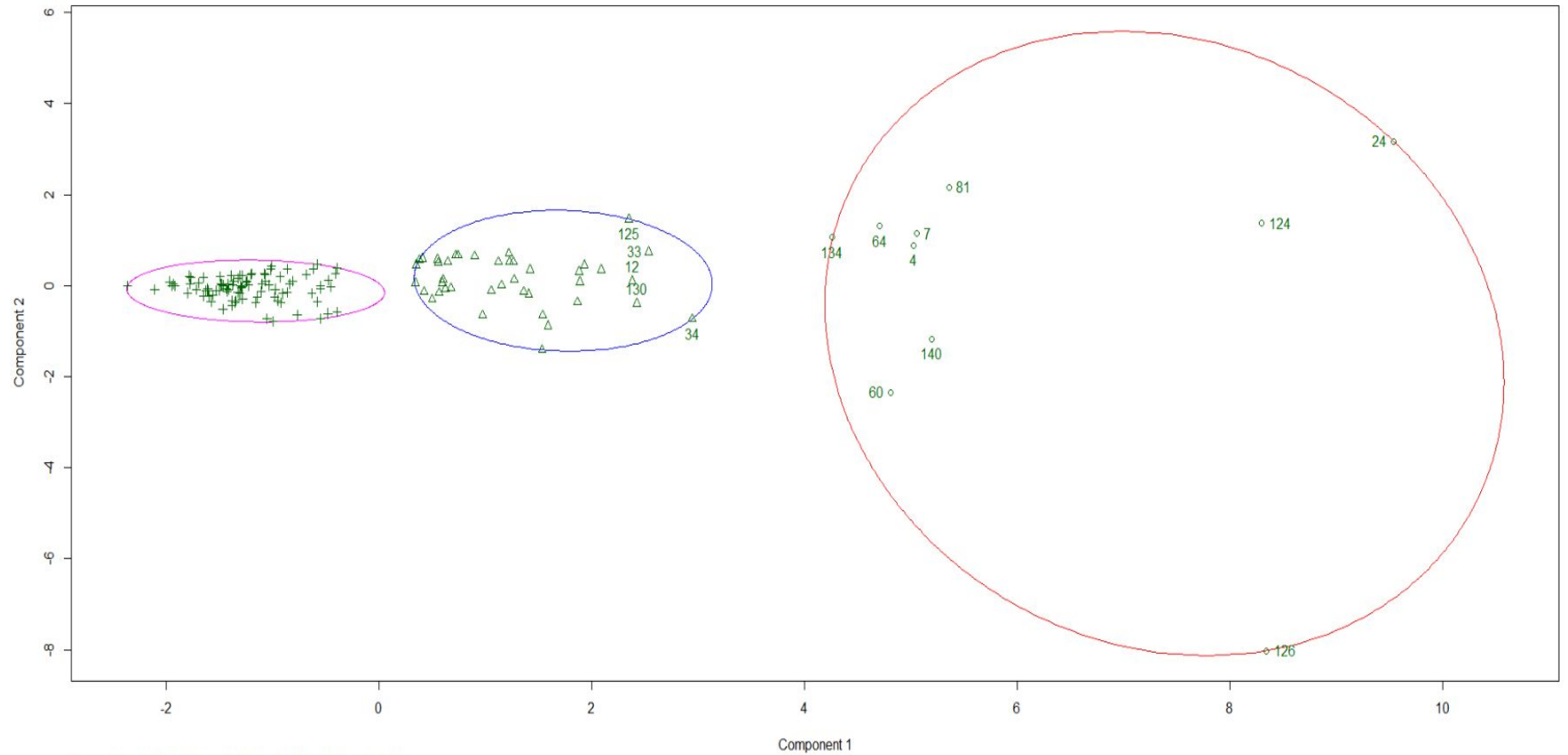
```
> km.out
K-means clustering with 3 clusters of sizes 10, 41, 90

Cluster means:
      Assault Auto Theft Break and Enter  Robbery Theft Over
1  2.5265062  1.6191679      2.5654014  2.3963735  3.0070501
2  0.4569977  0.2643985      0.5438294  0.4934988  0.2562804
3 -0.4889107 -0.3003558     -0.5327891 -0.4910799 -0.4508666
```

```
within cluster sum of squares by cluster:
[1] 157.01345  68.48119  39.38121
(between_SS / total_SS = 62.2 %)
```

```
> km.out$tot.withinss
[1] 264.8759
```

k-Means Cluster Analysis



Comparison of Clustered Neighbourhoods for k=3

Cluster 1 - High Crime

row.names	Neighbourhood	Assault	Auto Theft	Break and Enter	Robbery	Theft Over
24	Church-Yonge Corridor (75)	4061	200	946	904	190
81	Moss Park (73)	2134	119	632	630	94
64	Kensington-Chinatown (78)	2013	130	627	367	126
7	Bay Street Corridor (76)	2446	109	538	286	185

Cluster 2 - Medium Crime

row.names	Neighbourhood	Assault	Auto Theft	Break and Enter	Robbery	Theft Over
125	West Hill (136)	1971	112	416	323	35
33	Dovercourt-Wallace Emerson-Junction (93)	1386	173	512	318	61
12	Bendale (127)	1246	257	357	347	70
130	Wexford/Maryvale (119)	1270	306	494	202	78

Cluster 3 - Low Crime

row.names	Neighbourhood	Assault	Auto Theft	Break and Enter	Robbery	Theft Over
98	Pelmo Park-Humberlea (23)	254	197	128	91	22
133	Willowridge-Martingrove-Richview (7)	450	227	231	151	22
115	Stonegate-Queensway (16)	417	148	253	88	36
57	Humbermede (22)	510	203	129	97	29

Neighbourhoods with Higher than most Crime Rate

- Church-Yonge Corridor (75)
 - Moss Park (73)
 - Kensington-Chinatown (78)
 - Bay Street Corridor (76)
 - Waterfront Communities-The Island (77)
- Woburn (137)
 - Annex (95)
 - York University Heights (27)
 - Islington-City Centre West (14)
 - West Humber-Clairville (126)

Cluster 1 From K=3

Conclusion

- We were able to identify 10 Neighbourhoods where the amount of crime being committed is higher than anywhere else in Toronto
- Furthermore, we were able to identify 3 separate groups of Neighbourhoods where the occurrence of crime was generally low, medium or high
- Overall, most Neighbourhoods in Toronto are generally safe. However, there are a few exceptions. This is not surprising as most cities tend to have similar statistics. Nonetheless, this will remain useful to both law enforcement as well as residents/tourist to the area

Practical Implication

- Crime statistics can be useful in helping criminal justice professionals anticipate increased risk of crime in specific neighborhoods
 - Budget Formation
 - Resource Allocation
- Government agencies and other private entities such as real estate developers could also use statistics gathered within the field of urban planning and development in order to reduce crime
 - Publically Funded Initiatives
- Urban planners could develop plans and programs in order to reduce crime