Comparison of incidents in San Francisco in 24 hours period

In this visualisation, I would like to deep down among the category of crime incident in San Francisco. For the first step, the data need to be categorise by crime category and summarise total number of incident in each hour period of the day. I was keen to use RDBMS advantages to summarise and retrieve the crime category. At the result of running by data aggregation function of RDBMS, the data can be shaped up as follow.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	hour	ARSON	NON-CRIMINAL	LARCENY/THEFT	DRUG/NARCOTIC	DRIVING UNDER THE INFLUENCE	OTHER OFFENSES	TRESPASS	VEHICLE THEFT	ASSAULT	FRAUD	SUSPICIOUS OCC	SECONDARY CODES	WEAP
2	0	5	144	354	56	9	262	8	66	169	25	74	30
3	1	2	72	230	30	17	106	7	51	115	10	32	7
4	2	1	59	143	31	8	89	12	36	126	10	26	10
5	3	5	36	88	24	5	54	9	24	68	2	17	14
6	4	4	21	65	3	3	23	5	19	39	0	13	5
7	5	3	28	66	8	1	45	5	19	52	2	11	5
8	6	2	42	91	9	2	59	6	33	47	4	17	10
9	7	0	75	153	25	1	81	12	33	70	7	31	5
10	8	5	135	235	27	0	135	16	71	90	9	41	14
11	9	0	139	293	62	1	134	16	66	103	6	65	11
12	10	2	183	380	47	2	129	15	52	114	9	63	22
13	11	4	166	417	79	0	150	8	44	117	11	55	24
14	12	4	217	508	68	0	267	18	89	154	33	80	33
15	13	0	189	440	94	0	170	11	61	112	15	79	19
16	14	0	150	477	82	0	161	16	60	125	11	84	24
17	15	0	180	513	99	0	212	19	85	135	13	77	19
18	16	0	217	544	105	6	189	16	100	155	11	90	25
19	17	5	192	654	117	3	263	18	136	182	13	73	30
20	18	2	151	818	97	2	207	8	177	138	16	69	22
21	19	3	149	719	62	7	183	14	144	130	10	81	25
22	20	2	122	681	85	5	150	9	172	164	5	62	18
23	21	0	134	593	26	7	146	6	130	153	7	60	26
24	22	9	115	562	51	10	165	8	161	178	4	54	32
25	23	5	107	462	58	11	178	19	137	146	9	46	12
26													

As a result, there was total 34 category of crime case in San Francisco.
For the next step, I would like to visualise all type of criminal case as below scenario :

By follow the aspect of Data Type and Visual Mapping, the data type, which will contain in the visualisation, are
1.   Crime Category (Nominal data type) which will present with various colour,
2. Time Stamp (Ordered data type) on X-axis
3. Total cases of crime category (Quantitive - Ratio) on Y-axis

In order to do visualisation, I choose to do with Stackable Bar Chart and Group Bar Chart for presentation.

The main purpose of laying out 2 presentation is for comparison between 2 types of visualisation and choose which is more understandable upon visualise. I use d3.js to do visualisation.

Figure 1 : Visualisation of San Francisco Crime Case with Stackable Bar chart
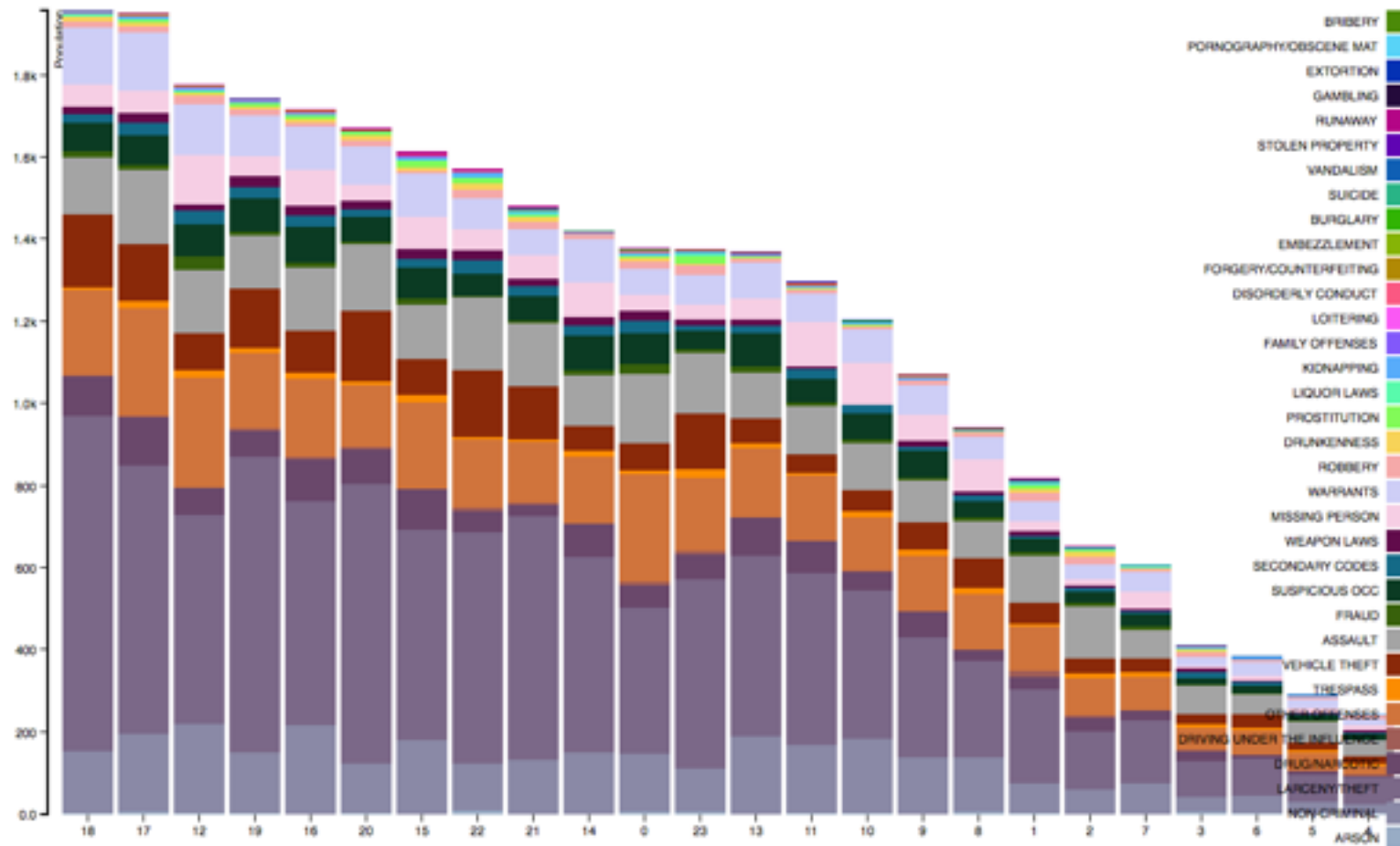
X - axis represent Hour in day, Y - axis represent total crime case per Different color on chart represent different type of crime category

Figure 2 : Visualisation of San Francisco Crime Case with Group Bar chart
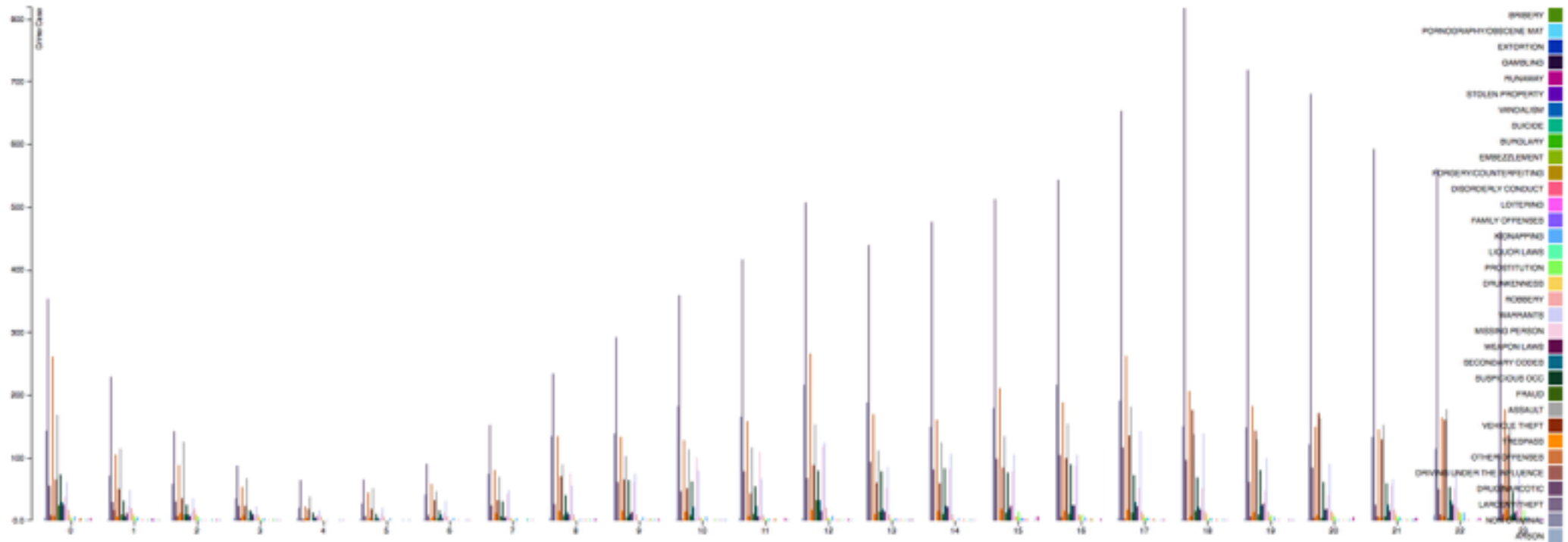
X - axis : Hour in day, Y - axis : total crime case per, Different color on chart represent different type of crime category

The original copy of enlarge graph can refer into main folder.
Plotting of 34 category over 24 hours is taking a huge space on X axis.

<div style="text-align: center">

What visualisation show?

</div>

Comparing the visualisation from above two graph, we can easily conclude that **stackable bar chart** can tell the complete story of which crime cases are most frequently on what period of time. As per stackable bar chart is showing, LARCENY/THEFT is the most criminal case in San Francisco and OTHER OFFENSES, NON-CRIMINAL and ASSAULT is the second most criminal case. And at the glance on chart, we can tell that criminal cases rarely happen on early morning and it happen most in evening.

Although the first graph is good enough to tell the story about criminal cases, I would like to visualise with another graph type to do more comparison. This time I choose line graph to visualise

Figure 3 : Visualisation of San Francisco Crime Case with line graph
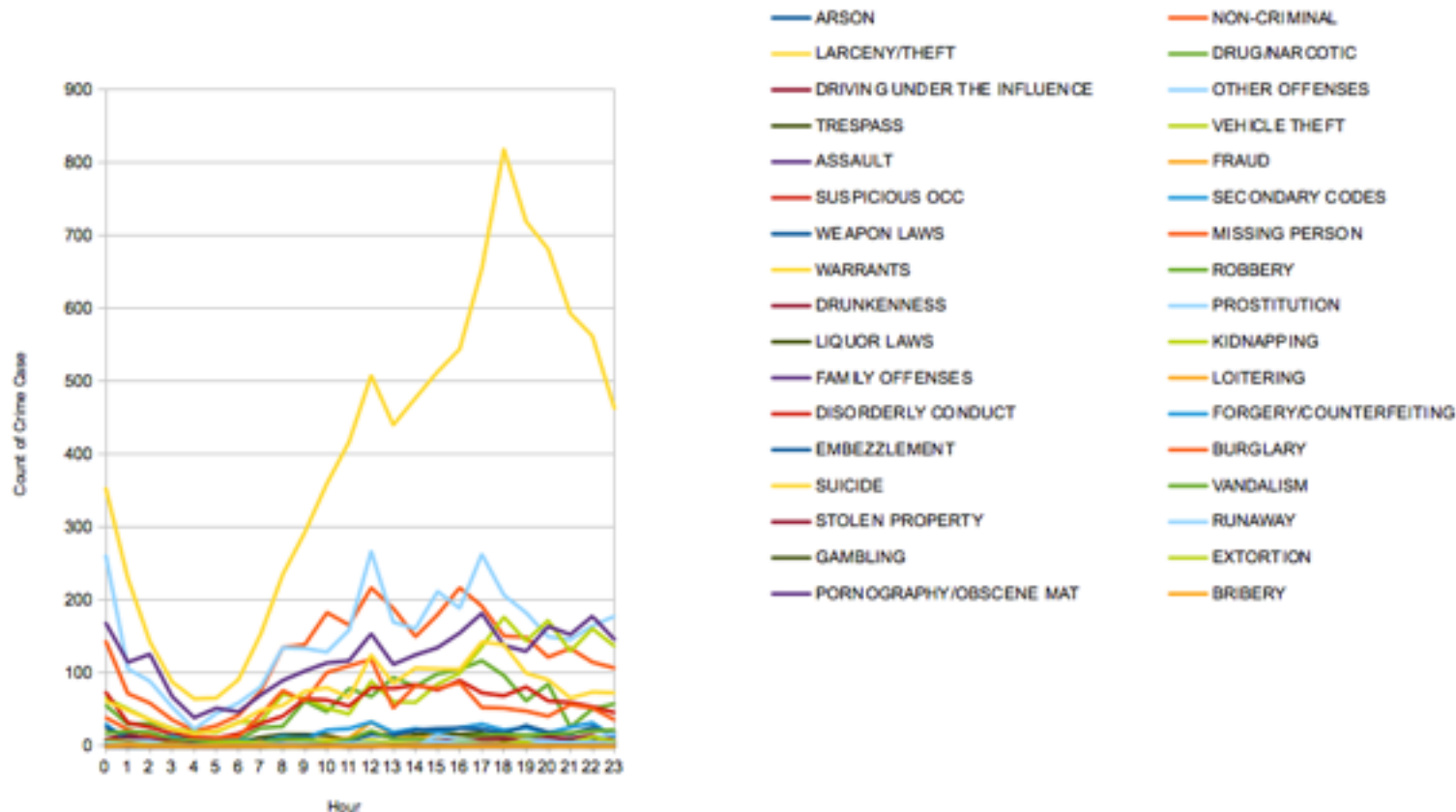
Visualising with line graph is also pretty good enough to know about the crime cases. I still prefer the visualisation with stackable bar chart.

Once I decided for the most preferable visualisation type, the next step to go in is to do comparison of 5 most occured criminal case.
Although I choose stackable bar chart as my preferable visualise type for this report, for this time round, I want to compare the stackable bar chart with line graph again. The following are the visualisation of 5 most occured criminal case.

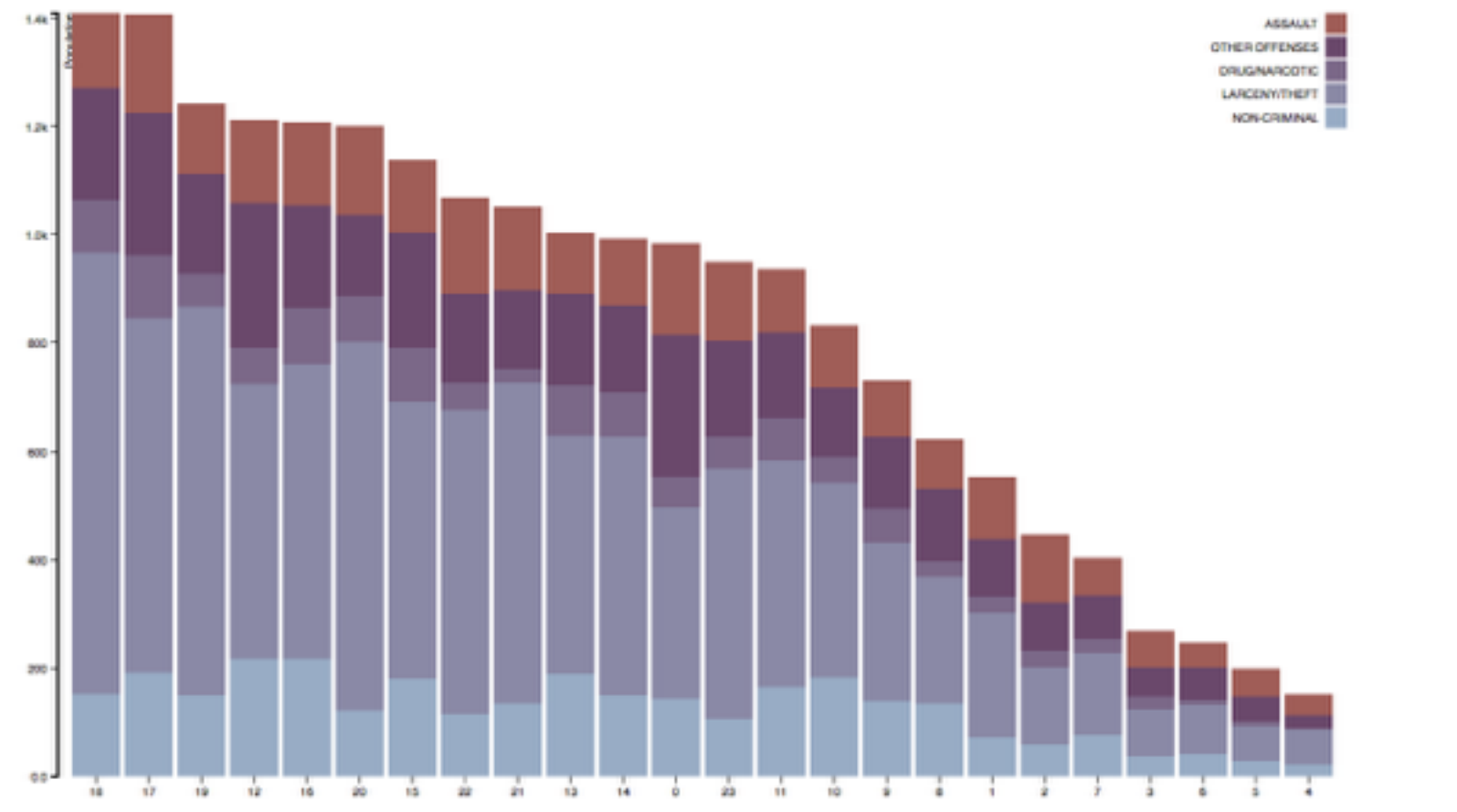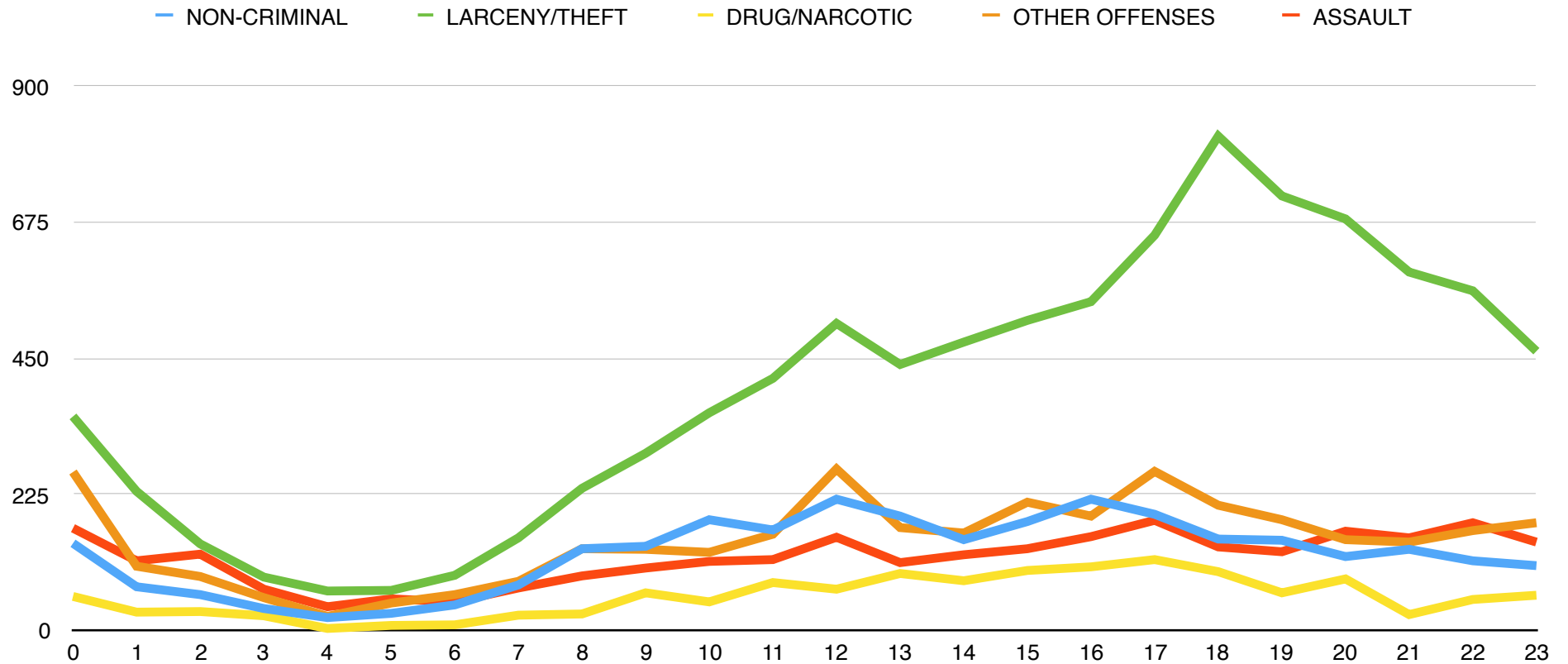**Figure 4 : Visualisation of top 5 criminal case with stackable bar chart**

Figure 5 : Visualisation of top 5 criminal case with line chart

From first analysing with large amount of category, I thought that stackable bar chart will be the best for visualisation. After I look into the visualisation of top 5 criminal case, I would like to change my thought which stick with stackable bar chart. The main reason why is with the stackable bar chart we cannot compare the dramatically data changes on first glance. According to the line chart visualisation, we can tell that on when LARCENY/THEFT crime case goes hight and when it go down. And in which period of time, all criminal cases cool down. etc.

In this research what we observed :
1. How do incidents vary by time of day , presented with 3 different graph chart

2. What type of graph is appropriate for presenting large amount of category
3. Top 5 Incidents variant from time to time
4. Again what type of graph is pleasing for presenting few category

As the research is more focus on visualisation result, I skip for the detail whole process of shaping up data in RDMS.

Overall aim of my visualisation is to decide what kind of visualisation perfect to tell the figure/story about the dataset rather on deep down on the detail case. As my aim was deviated from the instruction from the assignment, I hope that I could bring you well through my presentation.