**Descriptive Analysis**

**Student:** *Srivishnu Maganti*

**Date Submitted:** *November-04-2021*

**Task Selected:** *Response time*

Now that we have provided our initial project plan, and conducted a check on our data sources to assess potential sources of bias or error that could be introduced, we are now ready to begin the pathway to modeling. Before we can actually build a predictive model, we need to first get a feel for our data sets, the relationships between different variables, and existing patterns that could be relevant to our ultimate predictive model.

For this assignment, you will be constructing a series of descriptive analyses on both your primary data source (the EMS calls for service data) and any secondary data sources you are providing (whether they are weather data, Census data, etc.). Unlike our previous assignment, you will not need to do a separate analysis for each secondary dataset that would be introduced. In short, if you use the Community District dataset provided by the instructor, you can just run one set of descriptive analyses for the Community District dataset and one for the EMS calls for service. If you add weather data to your analysis, you will need to run a separate set of statistics for those data.

**WARNING**: This document will become very long (20 or so pages), primarily because you will be posting tables, charts, and then your own initial written assessments of what patterns you are seeing in the data. Please make sure you complete this assignment fully, as it will dramatically help your predictive modeling approach.

**Part 1 – EMS Calls for Service Data**

In this section, you will first analyze your calls for service data. This section is **required for everyone**, regardless of whether you are doing the demand modeling or the response time task. You will be providing descriptive tables and charts of the **individual calls for service data**, so no transformation is needed yet. We will move to that in Part 2a for those doing demand modeling, while those doing the response time will be working on Part 2b.

For this section, you will need to complete the following tables, charts, and assessments. There is a space on each page for these items, and you will need to provide not only the formatted items, but also a short (2-3 sentences) description of what patterns you are seeing. Please conduct the requested analyses, insert the formatted items, and complete the descriptions. Make sure the formatting stays the same, so that the presentation is consistent across pages.

All the relevant code to create tables and charts you can find in the training\_data\_load.r file on the Blackboard page.

***# of calls by year***

[INSERT **BAR** GRAPH]

1. Distribution of calls by length of incident\_response\_seconds\_qy, for each Borough (i.e. five separate distributions)

Chart, bar chart

Description automatically generated

What patterns are you seeing?: *[Please note anything unusual or unexpected, and describe the general trend or pattern of the data.]*

This graph explains about number of calls/incidents each year, so if we see the pattern every year number of calls were increasing, so minimum number of calls were in 2008 after that every year number of calls were increasing. So, this might be a good interpretation for our analysis.

***# of calls by day of week***

[INSERT **BAR** GRAPH]

Chart, bar chart

Description automatically generated

What patterns are you seeing? *[Please note anything unusual or unexpected and describe the general trend or pattern of the data.]*

Minimum number of accidents were on weekends, although there is no big difference in the numbers but they vary. Here, we can conclude that working days has more number of calls, when compared with weekends.

***# of calls by month (i.e. Jan-Dec)***

[INSERT **BAR** GRAPH]

Chart, bar chart

Description automatically generated

[INSERT **TABLE**]

| **data\_month** | | **Freq** | |
| --- | --- | --- | --- |
|  |  | |  |
| **1** | 01 | | 976155 |
| **2** | 02 | | 881258 |
| **3** | 03 | | 976679 |
| **4** | 04 | | 953244 |
| **5** | 05 | | 1041125 |
| **6** | 06 | | 1029786 |
| **7** | 07 | | 1056384 |
| **8** | 08 | | 1023985 |
| **9** | 09 | | 986370 |
| **10** | 10 | | 999089 |
| **11** | 11 | | 938495 |
| **12** | 12 | | 1001189 |

What patterns are you seeing?: *[Please note anything unusual or unexpected, and describe the general trend or pattern of the data.]*

If we carefully observe the trend the number of calls in winter seasons typically from 10(oct) to 3(march) (according to NYC weather), were less in number than that of summer season. Because in summer months number of calls were the highest, for instance if you consider July it has nearly a number of 1056384 noted calls.

Whereas, number of calls in February were least 881258(winter), so here we can tell in summer seasons number of calls were increasing. So, while doing analyses we can mostly concentrate on the summer season.

***# of calls by hour of day (i.e. 12am-12am)***

[INSERT **BAR** GRAPH]

Chart, histogram

Description automatically generated

What patterns are you seeing?: *[Please note anything unusual or unexpected, and describe the general trend or pattern of the data.]*

If you see the details of the graph, we can say that number of calls increases in the mid-day hour, also called as busiest hours in a day. In afternoon from 12:00pm to 8:00pm number of call records were more. In the same way, number of calls that were reported in mid night were less in number. The graph start increasing rapidly from 10:00am, so when the traffic start in the city number of calls received were increasing, and number constantly reduces from 8:00pm to 6:00pm (city silent period).

***# of calls by month and year (i.e. Jan 2008 – Dec 2016)***

[INSERT **LINE** GRAPH]

A picture containing diagram

Description automatically generated

What patterns are you seeing? *[Please note anything unusual or unexpected and describe the general trend or pattern of the data.]*

Like we discussed before from 2008 to 2016 every year number of cases were increasing. But, this plot helps us to understand clearly which parts of the year exactly the numbers were increased compared to other year.

So, after clearly understanding the stats in this graph we can conclude that every year at February period numbers of calls were reduced. And, as we discussed in months graph here also, in summer seasons the numbers were increasing and every plot has maintained same pattern with respect to months, but the numbers were increasing according to year.

***# of calls by date (i.e. 1/1/2008, 1/2/2008, etc.)***

[INSERT **LINE** GRAPH]

A screenshot of a social media post

Description automatically generated

What patterns are you seeing?: *[Please note anything unusual or unexpected, and describe the general trend or pattern of the data.]*

By this graph, we cannot conclude anything specific related to our analyses, because this is not like a generalized version, this is like a continuous graph which explains very briefly, where finding patterns is very difficult.

But we can tell that there are certain days in year where number of calls were more that average calls in the highest year(2016- because it has more number of calls compared with other years)

***# of calls by Borough***

[INSERT **BAR** GRAPH]

Chart, bar chart

Description automatically generated

[INSERT **TABLE**]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Var1** | **Freq** |  |  |
|  |  |  |  |  |
|  | **1** | BRONX | 2733876 |  |
|  | **2** | BROOKLYN | 3441280 |  |
|  | **3** | MANHATTAN | 2909565 |  |
|  | **4** | QUEENS | 2288528 |  |
|  | **5** | RICHMOND / STATEN ISLAND | 490386 |  |
|  | **6** | UNKNOWN | 124 |  |

Compared with other locations, here we can tell that Brooklyn and Manhattan has more calls than any other borough, According to new York population we can tell that Brooklyn and Manhattan as more number of people compared with other locations, so this might be also a reason for that, Apart from this even Bronx has nearly the same number with no major difference. But, out of all Richmond/Staten Island has a very less number compared with others. Borough is a important tool for analyses and while computing correlation.

***# of calls by Community District***

[INSERT **TABLE**]

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 226 | 267 |  | 405 | 132217 |  | 206 | 215062 |
| 484 | 330 |  | 315 | 132537 |  | 401 | 216216 |
| 480 | 700 |  | 502 | 134469 |  | 107 | 217582 |
| 227 | 719 |  | 101 | 135536 |  | 212 | 219268 |
| 595 | 988 |  | 409 | 139112 |  | 203 | 237184 |
| 482 | 1204 |  | 309 | 142772 |  | 501 | 239086 |
| 228 | 1502 |  | 403 | 143267 |  | 302 | 241314 |
| 356 | 1613 |  | 311 | 143670 |  | 207 | 243956 |
| 164 | 2388 |  | 102 | 148308 |  | 316 | 250877 |
| 481 | 2512 |  | 202 | 149134 |  | 112 | 265852 |
| 355 | 4508 |  | 210 | 154082 |  | 110 | 268268 |
| 483 | 11531 |  | 414 | 164066 |  | 104 | 268995 |
| 411 | 68625 |  | 109 | 168924 |  | 201 | 281184 |
| 406 | 86781 |  | 211 | 171154 |  | 209 | 285014 |
| 402 | 107387 |  | 308 | 175755 |  | 205 | 290914 |
| 503 | 107787 |  | 304 | 181609 |  | 103 | 297128 |
| 310 | 112695 |  | 314 | 182561 |  | 204 | 312674 |
| 312 | 112757 |  | 407 | 183693 |  | 303 | 320418 |
| 208 | 112917 |  | 318 | 188851 |  | 105 | 324173 |
| 307 | 113626 |  | 106 | 190582 |  | 111 | 330581 |
| 306 | 113724 |  | 313 | 193533 |  | 412 | 361438 |
| 410 | 119779 |  | 413 | 194163 |  | 305 | 378032 |
| 408 | 124772 |  | 108 | 195272 |  |  |  |
| 404 | 131341 |  | 301 | 197397 |  |  |  |
|  |  |  | 317 | 202899 |  |  |  |

I have arranged them in a ascending order and as we can see 226 community district have least number of calls and 305 district has highest number of calls. This table also helps us when we are analyzing with weather data because, if we have weather data for each community district and identify if calls were based on weather. Apart from that, for each community district we can evaluate response time and compute a detailed correlation between response time with number of calls for each community district.

***# of calls by Initial Call Type***

[INSERT **TABLE** sorted in descending order from MOST to LEAST]

|  |  |
| --- | --- |
| **initial\_call\_type** | **Freq** |
| Sick | 2049109 |
| Non-Critical Injury | 1840886 |
| Difficult Breather | 1051494 |
| Psychiatric Patient | 843135 |
| Hx Drug Or Alcohol Abuse | 684123 |
| Unconscious Patient | 657792 |
| Abdominal Pain | 560576 |
| Caller Has No Pt Medical Info | 460021 |
| Cardiac Condition | 452863 |
| Auto Acc W-Injuries | 412764 |
| Unknown Condition | 264470 |
| Respiratory Distress | 220775 |
| Pedestrian Struck | 215551 |
| Asthma Attack | 195751 |
| Status Epilepticus | 183711 |
| Altered Mental Status | 161274 |
| Seizures | 158732 |
| Cardiac Arrest | 147686 |
| Major Injury | 123515 |
| Sick Pediatric, <5 Year Old | 109519 |
| Internal Bleeding | 98303 |
| Minor Injury | 94479 |
| CVA (Stroke) | 91116 |
| Female In Labor | 87489 |
| Anaphylaxis | 58733 |
| Stabbing | 54241 |
| Multiple Trauma Patient | 52761 |
| Stroke | 51084 |
| Obstetric Complications | 47349 |
| Request For Stand-By | 42534 |
| Gun Shot Wound | 37397 |
| Miscarriage | 31271 |
| Minor Illness | 29860 |
| Gyn Bleeding-Pt Not Pregnant | 27929 |
| Major Obstetrical Complaint | 26642 |
| Hypertension | 22773 |
| Reaction To Medication | 21772 |
| Choking | 20485 |
| Inhalation Of Smoke | 17544 |
| Police 10-13, Unconfirmed | 15865 |
| One Alarm Fire| | 15165 |
| Major Burns 18% Adlt 10% Child | 13443 |
| Sick - Cough & Fever | 10614 |
| Minor Burns <18% Adlt Or <10% | 10376 |
| Rape | 10191 |
| Injury Lower Ext In Elderly | 9096 |
| Heat Exhaustion | 7005 |
| Gyn-Severe Pain-Bleeding | 6798 |
| Jumper Up | 5698 |
| Auto Accident, No Confirmd Inj | 5608 |
| Hypothermia | 5594 |
| Jumper Down | 5572 |
| Drowning | 4802 |
| Fire75 Working Fire | 4701 |
| Sick Ped<5 Yrs-Fever & Cough | 4084 |
| Child Abuse | 3644 |
| Amputation, Fingers Or Toes | 3340 |
| Baby Out Or Imminent Birth | 2407 |
| Diff Breathing - Fever&Cough | 2269 |
| Electrocution | 1593 |
| Special Event | 1297 |
| Resp Distress - Fever&Cough | 968 |
| Abdominal Pain-Fever & Cough| | 865 |
| Asthma Attack - Fever&Cough | 705 |
| Sick - Rash And Fever | 705 |
| One Alarm Fire | 598 |
| Stat Transfer Request | 472 |
| Venom (Snake Bites) | 421 |
| Amputation, Arm, Hand,Leg,Foot | 266 |
| Occupied High-Rise Building| | 259 |
| Sick Ped<5 Yrs-Rash & Fever | 245 |
| Evac | 221 |
| Cardiac Condition-Fever&Cough | 194 |
| Internal Bleeding-Fever&Cough | 117 |
| Report Of Explosives| | 89 |
| Medevac, T-C Authority Only | 82 |
| ACC | 78 |
| Alt Mental Status-Fever&Cough | 70 |
| Police 10-13, Confirmed | 55 |
| Sick Patient Fever-Travel | 55 |
| Seizures - Fever & Cough | 52 |
| Difficult Breather Rf | 51 |
| Death Confirm By Medical Auth| | 47 |
| Two Alarm Fire| | 46 |
| Unc Patient - Fever & Cough | 36 |
| Fire77 High Rise Residential | 28 |
| Stroke - Fever & Cough | 27 |
| Abdominal Pain-Fever & Cough | 26 |
| Mult Or Prolong Seizur-Fev&Cou | 25 |
| Reaction To Med - Fever&Cough | 24 |
| Active Shooter | 22 |
| Anaphylactic Shock-Fever&Cough | 20 |
| Hazardous Materials Incident| | 18 |
| BBP | 17 |
| Hostage Situation - Barricaded| | 15 |
| MOSINJ | 14 |
| Fire76 High Rise Commercial | 13 |
| Ground Transport Incident| | 11 |
| Death Confirm By Medical Auth | 10 |
| Stroke Critical - Fever&Cough | 9 |
| Structural Collaspe [Specify] | 9 |
| Hx Drug Or Alchl Abuse-Fev&Cou | 7 |
| ADM | 6 |
| Occupied High-Rise Building | 6 |
| STUCK | 6 |
| Five Alarm Fire Or Greater | 5 |
| Report Of Explosives | 5 |
| All Other MCIs| | 5 |
| Two Alarm Fire | 4 |
| Four Alarm Fire | 4 |
| Aircraft Incident - Crash | 4 |
| Choking Fever&Cough | 3 |
| Three Alarm Fire | 3 |
| Rapid Transit-Rail Incident | 3 |
| Power Failure - Blackout | 3 |
| MOSILL | 3 |
| Unconscious Fever-Travel Patient | 3 |
| Unconscious Patient-Rash&Fever | 3 |
| Card Or Resp Arrest-Fevercough | 2 |
| Difficult Breathing Fever-Travel | 2 |
| MCI25 | 2 |
| Construction-Demolition Incid| | 2 |
| MECHE | 2 |
| Respiratory Distress Fever-Travel | 2 |
| Test Kdt-Modat | 2 |
| Abdominal Pain Fever-Travel | 1 |
| Asthma Patient Fever-Travel | 1 |
| DRILL | 1 |
| Criminal Detection Facil Incid | 1 |
| Explosion | 1 |
| Ground Transport Incident | 1 |
| Confined Space Incident | 1 |
| MCI37 | 1 |
| Civil Distrubance | 1 |
| Hostage Situation - Barricaded | 1 |
| Active Shooter| | 1 |
| Active Shooter|| | 1 |
| MECHV | 1 |
| NOVEH | 1 |
| RADIO | 1 |
| Status Epilepticus Fever-Travel | 1 |
|  |  |

What patterns are you seeing? *[Please note anything unusual or unexpected, and describe the general trend or pattern of the data.]*

If we carefully observe the trend in this data sick is the greatest number of initial calls are sick, non-critical injury.

And least was fever-travel. Mostly, sick, difficult-breather and non-critical injury cases are more in number. This data is not useful for analyzing response time because, for any call type response time is same even if the person who is calling is suffering from fever or accident he gets the same response from EMS, so I think this table might helps to extract additional information, rather than the insights which are required. This helps in summarizing the type of call that is received.

***# of calls by Final Call Type***

[INSERT **TABLE** sorted in descending order from MOST to LEAST]

|  |  |
| --- | --- |
| **final\_call\_type** | **Freq** |
| Sick | 1925873 |
| Non-Critical Injury | 1799152 |
| Difficult Breather | 1100885 |
| Psychiatric Patient | 808263 |
| Hx Drug Or Alcohol Abuse | 713540 |
| Unconscious Patient | 689235 |
| Abdominal Pain | 557298 |
| Cardiac Condition | 525473 |
| Auto Acc W-Injuries | 399053 |
| Caller Has No Pt Medical Info | 390175 |
| Unknown Condition | 221032 |
| Altered Mental Status | 217580 |
| Pedestrian Struck | 215638 |
| Cardiac Arrest | 213173 |
| Respiratory Distress | 207017 |
| Status Epilepticus | 202481 |
| Asthma Attack | 187808 |
| Seizures | 144941 |
| Major Injury | 138544 |
| Sick Pediatric, <5 Year Old | 108968 |
| Internal Bleeding | 105697 |
| Minor Injury | 92491 |
| CVA (Stroke) | 87309 |
| Female In Labor | 84578 |
| Anaphylaxis | 62877 |
| Multiple Trauma Patient | 59788 |
| Stabbing | 56913 |
| Obstetric Complications | 50497 |
| Stroke | 47994 |
| Request For Stand-By | 38488 |
| Gun Shot Wound | 37156 |
| Miscarriage | 31362 |
| Minor Illness | 29226 |
| Major Obstetrical Complaint | 28582 |
| Gyn Bleeding-Pt Not Pregnant | 28043 |
| Choking | 22145 |
| Hypertension | 21685 |
| Reaction To Medication | 21140 |
| One Alarm Fire| | 20318 |
| Inhalation Of Smoke | 18137 |
| Major Burns 18% Adlt 10% Child | 14933 |
| Police 10-13, Unconfirmed | 14479 |
| Sick - Cough & Fever | 10830 |
| Rape | 10584 |
| Injury Lower Ext In Elderly | 10373 |
| Minor Burns <18% Adlt Or <10% | 10044 |
| Gyn-Severe Pain-Bleeding | 7505 |
| Heat Exhaustion | 6975 |
| Jumper Up | 6462 |
| Jumper Down | 6183 |
| Hypothermia | 5951 |
| Auto Accident, No Confirmd Inj | 5189 |
| Drowning | 4883 |
| Sick Ped<5 Yrs-Fever & Cough | 4439 |
| Baby Out Or Imminent Birth | 4021 |
| Child Abuse | 3882 |
| Amputation, Fingers Or Toes | 3771 |
| Diff Breathing - Fever&Cough | 3049 |
| Ground Transport Incident| | 2563 |
| Hostage Situation - Barricaded| | 2060 |
| Electrocution | 1725 |
| Special Event | 1324 |
| Two Alarm Fire| | 1316 |
| Resp Distress - Fever&Cough | 1173 |
| Abdominal Pain-Fever & Cough| | 1023 |
| Asthma Attack - Fever&Cough | 806 |
| Report Of Explosives| | 769 |
| Sick - Rash And Fever | 746 |
| Hazardous Materials Incident| | 708 |
| Occupied High-Rise Building| | 691 |
| Stat Transfer Request | 504 |
| Venom (Snake Bites) | 456 |
| Amputation, Arm, Hand,Leg,Foot | 371 |
| Sick Ped<5 Yrs-Rash & Fever | 276 |
| Three Alarm Fire| | 261 |
| One Alarm Fire | 260 |
| Police 10-13, Confirmed | 243 |
| Cardiac Condition-Fever&Cough | 230 |
| Evac | 224 |
| All Other MCIs| | 208 |
| Internal Bleeding-Fever&Cough | 151 |
| Structural Collaspe [Specify]| | 137 |
| Alt Mental Status-Fever&Cough | 113 |
| Confined Space Incident | 112 |
| Medevac, T-C Authority Only | 92 |
| Four Alarm Fire| | 84 |
| Construction-Demolition Incid| | 84 |
| Unc Patient - Fever & Cough | 83 |
| Death Confirm By Medical Auth | 72 |
| Difficult Breather Rf | 69 |
| Seizures - Fever & Cough | 67 |
| Five Alarm Fire Or Greater | 58 |
| Mult Or Prolong Seizur-Fev&Cou | 51 |
| Marine - Harbor Incident| | 48 |
| Stroke - Fever & Cough | 44 |
| Fire75 Working Fire | 39 |
| Sick Patient Fever-Travel | 39 |
| Rapid Transit-Rail Incident| | 29 |
| Reaction To Med - Fever&Cough | 29 |
| Civil Distrubance | 28 |
| Abdominal Pain-Fever & Cough | 27 |
| Explosion | 26 |
| Anaphylactic Shock-Fever&Cough | 24 |
| Criminal Detection Facil Incid| | 24 |
| Aircraft Incident - Crash | 20 |
| Active Shooter | 18 |
| Stroke Critical - Fever&Cough | 16 |
| Card Or Resp Arrest-Fevercough | 14 |
| Hostage Situation - Barricaded | 14 |
| Hazardous Materials Incident | 14 |
| Hx Drug Or Alchl Abuse-Fev&Cou | 13 |
| Medical Facility Evacuation | 11 |
| Power Failure - Blackout| | 10 |
| Ground Transport Incident | 9 |
| All Other MCIs | 6 |
| Report Of Explosives | 5 |
| Unconscious Fever-Travel Patient | 5 |
| Choking Fever&Cough | 4 |
| Abdominal Pain Fever-Travel | 3 |
| Difficult Breathing Fever-Travel | 3 |
| Two Alarm Fire | 3 |
| Occupied High-Rise Building | 3 |
| Active Shooter|| | 2 |
| Unconscious Patient-Rash&Fever | 2 |
| ALMNFC | 1 |
| ARSTFC | 1 |
| Fire77 High Rise Residential | 1 |
| MCI27 | 1 |
| Structural Collaspe [Specify] | 1 |
| MCI37 | 1 |
| MCI42 | 1 |
| Respiratory Distress Fever-Travel | 1 |
| Status Epilepticus Fever-Travel | 1 |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

What patterns are you seeing?: *[Please note anything unusual or unexpected, and describe the general trend or pattern of the data.]*

Practically, if we consider the scenario regarding EMS calls, more than initial call types final call types will be more useful. Because, initial call types will not be the actual reason for the response, so for analyzing response time we can consider final call type as a factor(variable). So here in this table we can see sick, Non-Critical Injury, Difficult Breather are more in number.

|  |
| --- |
|  |
|  |

**# of calls, by Initial Severity Level and Final Severity Level**

[INSERT **CROSS-TABULATION**, with Initial Severity Level as rows and Final Severity Level as Columns]

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 1 | 164262 | 1418 | 735 | 1131 | 317 | 734 | 153 | 7 |
| 2 | 49054 | 2278633 | 8850 | 6458 | 2174 | 7619 | 1123 | 259 |
| 3 | 8628 | 54845 | 1424569 | 6032 | 5630 | 8294 | 2902 | 104 |
| 4 | 6200 | 71274 | 46332 | 2317895 | 15232 | 12876 | 3308 | 189 |
| 5 | 2587 | 27066 | 52970 | 37179 | 2141574 | 2152 | 3696 | 99 |
| 6 | 3870 | 42888 | 81912 | 44584 | 27471 | 1833836 | 8510 | 682 |
| 7 | 745 | 6515 | 9306 | 27108 | 11514 | 6709 | 938364 | 83 |
| 8 | 48 | 3819 | 371 | 454 | 767 | 232 | 211 | 39025 |
| 9 | 10 | 25 | 15 | 59 | 24 | 12 | 3 | 26 |
|  |  |  |  |  |  |  |  |  |

What patterns are you seeing?: *[Please note anything unusual or unexpected, and describe the general trend or pattern of the data.]*

So, this is a correlation matrix between initial and final call type, here we will get confused by numbers, but the numbers are nothing but index in those respective tables. So here for instance

1<- sick, 2<-non-critical etc

So by this correlation we can understand similarity between initial and final call type.

**Part 2 – Task Analysis**

Now that we have completed some initial descriptive analyses of the individual calls for service data, it is time to move onto our specific sections related to the task that we have chosen. In this section, you will complete either Part 2a (for those doing the demand modeling) or Part 2b (for those doing the response time analysis).

For the part you are not doing, please delete that part from your final submission.

Please complete all sections, and ensure that the analyses provided are formatted, well-described, and clear.

**Part 2b – Response Time Analysis**

For those who are doing the response time analysis, please complete the following descriptive statistic tables and/or charts. After each group of charts or tables, I want you to include the following section:

What patterns are you seeing?: *[Please note anything unusual or unexpected, and describe the general trend or pattern of the data.]*

Here are the charts to complete:

1. Distribution of calls by length of incident\_response\_seconds\_qy

**A picture containing graphical user interface

Description automatically generated**

Here by the chart we can clearly understand that most of the calls response time under 1500 seconds, mostly between 1100-1200, there are only few cases where the response time is very more, there are few cases response time is 5000seconds, which is nearly 1hr 40min. shortest one is less than 1100 seconds and longest is 5000sec. Nearly 1500 calls have response time of <1200.

1. Distribution of calls by length of dispatch\_response\_seconds\_qy

**Graphical user interface, application

Description automatically generated**

This explains that nearly 250 calls have dispatch response of 1000sec and most of the calls are under 2000sec , mostly they lie between 1000 and 1500. So, here averagely nearly 150-200 calls have dispatch response time of 1000-2000, there are few calls less than 200 have dispatch response of 5000. Here, 1000 is the most frequent dispatch time.

1. Distribution of calls by length of incident\_travel\_tm\_seconds\_qy

**Graphical user interface

Description automatically generated**

Here the most frequent travel time is 1300seconds with nearly 500 observations, and most of the observations have travel time less than 1800 seconds.

1. Distribution of calls by length of incident\_response\_seconds\_qy, for each Borough (i.e. five separate

distributions)

**A picture containing table

Description automatically generated**

**Graph explains incident response time for each borough and we already observed incidents for each borough in previous section and concluded that Manhattan, Brooklyn have more number of calls so here working with incident response indicates that most of the most number of cases have incident response nearly 1000 in each borough.**

1. Distribution of calls by length of incident\_response\_seconds\_qy, for day of the week (i.e. seven separate distributions)

**Diagram

Description automatically generated**

**This plot explains response time for each days of week where we can say that every day in a week follows the same trend and response time is not particularly varied upon days of week.**

1. Distribution of calls by length of incident\_response\_seconds\_qy, for each month (i.e. 12 separate distributions)

**Calendar

Description automatically generated**

**Even the same follows with month for each and every month there is the same pattern that is being followed where most of the response time is nearly 1000, and some are more than that. But, depending on the month number of calls were decreasing in this plot, that we understood from the above graph( months vs calls) where winters have less number of calls whereas summers have more in number.**

1. Distribution of calls by length of incident\_travel\_tm\_seconds\_qy, for each month (i.e. 12 separate distributions)

**Calendar

Description automatically generated**

This graph explains the travel time for each month, Although, each month has same pattern in travel time, but the number of calls vary. Mostly travel time is less than 1000, so we can conclude that the greatest number of cases travel time is less than or equal to 1000sec.

**Part 3 – Additional Data Analysis**

Now that you have completed your analysis of the overall data and your specific task requirements, it is time to focus on the additional data you are planning on using for the modeling. This could be weather data, this could be community district data, or other data sources.

In the following section, you are to conduct **1 additional table or chart (your choice) for each variable of interest** you plan on using from these alternative datasets. So, if you plan on using weather data (like precipitation), you will need to create a chart that shows precipitation by month of year, and/or precipitation by year. If you are using the Community District data, you need to do a chart or table for each variable you use.

You should have several charts/tables in this section, as everyone should be using additional data in their modeling approach.

Finally, after including all the relevant charts and tables, I want you to spend 1-2 paragraphs describing the additional patterns that you see in these data, and how they might be related to your dependent variable of interest (# of calls or response time).

[Please insert additional tables/charts here]

#Total snow fall over years

Chart

Description automatically generated

Here this figure explains total snowfall over years, and this tell that total snow fall is more in December every month. In 2008 November, December and January had same amount of snow fall. In 2010, snow fall is the highest. In 2011 starting the amount is higher but gradually in the end of 2011 and starting of 2012 amount was reduced. Again it was increased in 2014 which followed the same trend in 2016.

#Total precepitation over years

Chart, histogram

Description automatically generated

This explains the total rain fall from 2008 to 2016 which explains the 2011 to 2012 has most rain fall compared with other years and 2013 has the least amount of rain fall.

#mean of precipitation over years

Chart, line chart

Description automatically generated

As mean tells us the average amount of snowfall, here we have computed mean of snow fall over years where 2011 has the highest and 2016 has the lease average.

#mean of precipitation over months

Chart, line chart

Description automatically generated

This explains the average precipitation per month, where June has the highest precipitation among all other months, whereas January the least.

#mean of snowfall over years

Chart, line chart

Description automatically generated

This explains the mean of snow fall over years where this explains that 2010 has more snow fall average compared to other years and 2012 has the least.

#mean of snowfall over months

Chart, line chart

Description automatically generated

As we can clearly observe here snow fall is highest in January and feb among all the data from 2008 to 2016 and gradually decreases after that until november.

What patterns are you seeing in this additional data you are including in the model?: *[Please note anything unusual or unexpected, and describe the general trend or pattern of the data.]*

How might these additional factors be related to your modeling task (either # of calls or response time)? [Please answer here]

There are many factors that might effect the response time from weather data, we can first analyze with precipitation data each and every month, secondly, we can plot the response time with precipitation for each month and see that is precipitation affecting response time or not. Same analyses we can carry out with snow fall for both month and year.

In the similar manner, we can start analyzing travel time with snow fall and precipitation for respective month and year, because due to weather some times travel time might be reduced such as snow fall might effect the traffic.

Apart from this, we can check the final call type for each and every month with amount of snow fall, because cold weather might effect few cases such as sick, problems in breathing etc, so final call type can be varied with amount to snow fall and precipitation.

There are few more effect that might effect we can analyze when we compute correlation between them by cross tabulation. All these help us to understand what factors we need to consider for modelling.