Project statement

2.

3. Explanation of any APIs you are using

4. What sort of visualizations you have created and still plan on creating

5. Tell me a data story. This could be: (a) what you’ve found so far in the data. (b) what you were expecting the data to tell you.

A good scope statement includes the following things:

1. **Overall description of the work.**  This is where you state that the project is to “build a fence.”
2. **Deliverables.**  What [will be produced](http://www.projectengineer.net/25-example-project-deliverables/) by the project, and what are its key features?  Also, what client need is the project satisfying?
3. **Justification for the project.**  In order to provide a complete understanding of the scope, sometimes it is necessary to dive into the justification of why the project was [initiated](http://www.projectengineer.net/the-elements-of-a-project-charter/) in the first place. - **Review existing crime-**related data/statistics or reports related to a particular law enforcement jurisdiction; in our case study, it’s the city of Austin.
4. **Constraints.**  If the project faces certain physical boundaries, these can be a [source of risk](http://www.projectengineer.net/project-risk-checklist/)and thus should be defined further.
5. **Assumptions.**  All projects have assumed certain conditions as part of their existence. For example, the fence building project has assumed good weather, availability of tools, etc. What are those assumptions and what impact does their inaccuracy have on the project?
6. **Inclusions/Exclusions.**  Many projects have items that are uncertain because projects of that type/size sometimes do and sometimes don’t include those things.  They need to be explicitly included or excluded from the project.

**Why crimes should be analyzed?**

**Purpose of the study**

* The basic purpose of this study is to
  + Analyze data to identify trends in crime across the City of Austin over the years (from 2014 to 2016).
  + To determine a comparison of the crime rate with respect to the area of occurrence (districts) and time of the year (months)
  + To identify a particular crime type likeliest to occur at a certain season of the year.
  + And to determine the spread of particular crimes across the city with respect to the districts.

**Data sources**

* Austin\_Crime dataset was retrieved from website – csv format with 19121 kb of data and 159464 data entries – This had a lot of rows with missing data
* Austin Crime Dataset 2015 from Data.Gov website – csv format with 7702 kb of data and 38573 data entries.
* 2016\_Annula\_Crime\_Data dataset from Data.Gov website – csv format with 7432 kb of data and 37461 data entries.
* United States Census Bureau – to retrieve population estimate for the City of Austin for the years 2014, 2015 and 2016.

**Methodology for Data retrieval/gathering**

* For this project study, we collected and compiled the recent crime reporting numbers available for the City of Austin on five major crimes which are: Aggravated assault, auto theft, homicide, robbery and theft.
* We also made use of the US census Bureau to get the population estimate for the respective years.
* We tried to obtain 2017 reports, but we were unable to get the data.

**Analysis tools used**

* Python’s data evaluation and visualization packages, which includes:
  + Pandas (Data Analysis library) and
  + Matplotlib (Visualization library) were utilized extensively through out the project.

**Data Clean-up or Wrangling**

* The data wrangling process was initiated by reading into the retrieved CSV files and converting them into data frames.
* The basic cleanup process included dropping off unnecessary columns from the dataset, dropping rows with null values, reordering columns, changing the time format and aggregating similar crime types.

**Data Visualizations**

* Pie charts were plotted for all three years (2014-2016) individually to display the break up of each crime percentage in Austin.
* Line graph plotting of crime rates (per 100,000 inhabitants) for each month of the year(from 2014-2016) was performed.
* To make it easier to understand how the five crime types have changed over time in relation to each other, individual line plots for each offense type were created characterized by months of the year. These line graphs show the trending of specific crimes in the city of Austin from 2014 to 2016.
* The crime counts and rates represented on the graphs cover the period of 2014 to 2016.
* The crime rates are calculated per 100,000 residents of the City of Austin.

**Inferences**

* A comparison of crime data statistics for 2014, 2015 and 2016 shows that there has been a decrease in the crime rates. According to the crime data, theft has decreased by almost 15% while the rest of the crimes almost maintain a steadiness.
* The data for nearly all types of crimes show that, though the crime rate plummeted between September to April, seasonal spikes in summer remain.
* In our analysis of crime trends, the districts fell into 10 groups. Group A has an average of \_\_\_\_\_\_\_\_\_\_ crimes per 100,000 people in 2014. From 2014 to 2016, the group’s crime trend has dropped/increased to \_\_\_\_ percent.
* The burglary data is missing from the 2014 dataset. We are unsure if the theft data for 2014 has the burglary values included in them. ‘Theft’ has always been the highest crime occurred according to the graphs in all the years while the other crimes show pretty less occurrences but show a steadiness throughout the years considered.

**Reasons why the rise or increase in crime incidents in the summer months:**

* Some researchers reason
  + that increased temperatures—which drive many out of doors and to leave windows open in their homes
  + increased daylight hours—which can lengthen the amount of time people spend away from their homes—raise the amount of people in public and the amount of time that homes are left empty.
  + Others point to the effect of students on summer vacation, who are otherwise occupied with schooling during other seasons;
  + still others postulate that suffering heat-induced discomfort simply makes people more aggressive and likely to act out.

**Limitations of the Dataset**

https://www.thoughtco.com/why-does-crime-spike-in-summer-3026089

Crime represented by police crime data does not represent all crime occurring in society. Thus, we typically see reports that name this type of data, “reported crime” or “crimes known to the police.” This is an important distinction for anyone interpreting or trying to understand crime problems based on police data: that is, that we may not be aware of the entire problem and the police data may portray a biased picture. For example, we know that a very low number of rapes are reported to the police. Thus, if we report rape statistics rising or falling we must be very cautious and mention the fact that even though the police are seeing an increase or decrease, the actual number may be changing in a different way since we do not know the actual number of rapes being committed. This is particularly relevant for certain types of crimes (domestic violence, drug crimes, white collar crimes) and not as much for others (motor vehicle theft, arson, murder).

All crime reported to the police is counted based on the date it is reported, because it would be impossible to count it by when it occurred, since this is not always known. As noted above, crime data is dynamic and crimes can be reported days, weeks, months, even years after they occur. Counts would need to be constantly updated based on date of occurrence and the ranges would cause further issues in counting. Date of report is constant, it does not change, and thus counts of crime are based on this variable. However, it can affect the interpretation of crime statistics. For example, if a large number of burglaries occur over the Christmas/New Year’s holiday, victims may not report them until after the new year when they return home. Thus, a large number of crimes would be counted in the following month and even the following year when they may have occurred the previous month and year. Although there is nothing an analyst or police can do to correct this problem, it should be considered during the analysis process. Finally, the date of occurrence, though not adequate for counting crime, is important for analysis and is used for identifying patterns and series—when the crime occurred is much more important that when it was reported.

* The total incidence of reported and discovered crime is based on factors that change over time; therefore interpretation of these graphs should be done with caution. Some factors affecting crime incidence are:
* Actual amount of crime as interpreted through Uniform Crime Reporting rules
* Rate of reporting by crime victims
* Special efforts of police to reduce crime (which may increase the amount of crime known to police)
* Special efforts of victims to reduce crime, such as shoplifting prevention programs (which may increase the amount of crime known to police)
* Changes in the law or in criminal justice system policies and practices
* Figures from the 2013 Swedish Crime Survey (SCS) show that exposure to crime decreased from 2005 to 2013.[[3]](https://en.wikipedia.org/wiki/Crime_in_Sweden) Since 2014 there has been an increase in exposure to some categories of crimes, including fraud, some property crime and especially sexual offences (with a 70% increase since 2013) according to the 2016 SCS.[[4]](https://en.wikipedia.org/wiki/Crime_in_Sweden) Violence (both lethal and non-lethal) has been on a downward trend the last 25 years.[[5]](https://en.wikipedia.org/wiki/Crime_in_Sweden) The figures for fraud and property damage (excluding car theft) are in contrast with the numbers of reported crimes under such categories which have remained roughly constant over the period 2014-16.[[6]](https://en.wikipedia.org/wiki/Crime_in_Sweden) The number of reported sexual offences clearly reflect the figures in the 2016 SCS, and car related damages/theft are also somewhat reflected.[[7][8]](https://en.wikipedia.org/wiki/Crime_in_Sweden) The number of convictions up to 2013 has remained between 110,000 and 130,000 in the 2000s — a decrease since the 1970s, when they numbered around 300,000 — despite the population growth.[[9]](https://en.wikipedia.org/wiki/Crime_in_Sweden) Consistent with other [Western](https://en.wikipedia.org/wiki/Western_world) countries in the [postwar era](https://en.wikipedia.org/wiki/Post-war), the number of reported crimes has increased when measured from the 1950s; which can be explained by a number of factors, such as statistical and legislative changes and increased public willingness to report crime.[[10]](https://en.wikipedia.org/wiki/Crime_in_Sweden)
* For a longer-term perspective, we looked at the change in the violent crime trend between 2010 and 2015 in our 68-city sample. These 10 places had the largest increases and decreases over this period.
* This line graph shows how many tourists came to four countries in southern Europe from 1970 to 2006. Italy, Spain and Greece are shown from 1970, Portugal starts in 1980. The number of tourists increased in all countries. Most tourists went to Italy in 1970, but in in the beginning of the 90s most people travelled to Spain for their holiday. More than twice as many tourists as in 1970 went to Spain and Greece in 2006. Portugal increased its number of tourists five times. All lines show the highest numbers in 1998. Afterwards the numbers went down. If you compare the figures you see more and more tourists wanted to go to the four countries. This chart shows a steady growth of tourism in the four Mediterranean countries. The chart shows a trend in tourism. Many tourists choose the Mediterranean countries for their holidays. The main reason is their subtropical climate with a lot of sun and nearly no rain during the summer months. The chart does not say anything about the situation in Portugal before 1980. Portugal was quite poor and had to build hotels and streets before it was ready for tourism
* Take a look at this chart, and you’ll see how changing the city, the type of crime — murder, rape, robbery, or aggravated assault — or the span of time studied can affect your view of how crime is changing. You can scroll down for month-by-month trends and highlights.

Once we start looking at different types of crime, and how they have changed over time, it becomes possible to shed light on the possible causes of different crime trends. The figure below does this for four selected police recorded crimes – rape, burglary, vehicle and drug offences – setting them alongside the ‘overall’ crime rate.

. According to the crime statistics, murders had increased by 6.2% while rape, aggravated assault and robbery offenses had increased by 9.6%, 2.3% and 10.6% respectively. On the other hand, property crime statistics showed a decrease in property crimes in the first six months of 2015 compared to the same period in 2014. Larceny-theft cases had decreased by 3.2% while burglary cases decreased by 9.8%. However, there was an increase in motor vehicle theft cases during the same period by 6.7%.