TO 640 - Group 9- Final Project
Zewen Ai, Stephanie Chan, Reshma Erukulla, Eric Messinger, Samuel Perlmutter

Exploring the Potential Effects of Weather on Crime Statistics in Detroit Michigan

Description of Business Problem

It is often stated that there exists a relationship between weather and crime, specifically that crime rates increase during warmer summer months. In addition to investigating this relationship, we sought to determine whether other weather factors, such as moon visibility, daily sunlight hours, and humidity, have a relationship with changes in crime levels.

Hypothesis

Within Detroit, Michigan, we expect that there will be an increase in the number of 911 calls and/or the number of crimes committed when the temperature is hotter, day more humid, moon more visible, and sunlight more present. The team believes that these four factors - temperature, humidity, lunar illumination, and precipitation - play a role in determining the number of 911 calls placed and/or the number of crimes committed.

Methodology:

In order to pull and aggregate the pertinent information, the team took the following steps:

1. Use an API to connect to the World Weather Online API and pull weather-based information for the analysis (See Appendix for code)
2. Use Python code to aggregate and calculate weather data (ex. temperature, lunar illumination, humidity, precipitation, etc.) on a daily basis (See Appendix for code)
3. Use SQL to pull the crime data from the DPD database (See Appendix for code)
4. Export both weather and crime data into Tableau
5. Join both excel files together in Tableau in order to create the pertinent visualizations.

Data

The team focused on pulling data from the DPD (Detroit Police Department) database and utilizing an API to pull the pertinent weather data from the World Weather Online API. The team further honed in on focusing on data from 2017 for the analysis, which is a full calendar year available through both the API and database. From the DPD database, we pulled:

- callTime, incidentId, Totaltime, Zipcode, Call description; CrimeId, incidenttime, charge code, neighborhoodId

From the World Weather Online API, the following fields were pulled:

- date, max temp, min temp, total snowfall, sun hours, uv Index, moon illumination, moonrise, moonset, sunrise, sunset, dew point, feels like, heat index, wind chill, wind

gust, cloud cover, humidity, precipitation, pressure, temperature, visibility, wind direction, wind speed

Data Visualizations:

First, we visualized the aggregated number of reported crimes across a time series (months) to see if there was a general pattern of increased crime across seasons (Exhibit 1). From this analysis, there is a slight increase in the number of crimes in warmer months, such as May through August. Second, we plotted a square chart for the number of calls of each type of 911 call (Exhibit 2). We found the top 3 to be special attention, traffic stops and disturbances and charted this on a monthly basis to see if the number of calls made during certain months are higher (Exhibit 3). For example, the bar chart shows that most of the calls regarding disturbances happen during warmer months, especially in June and July.

Exhibit 1

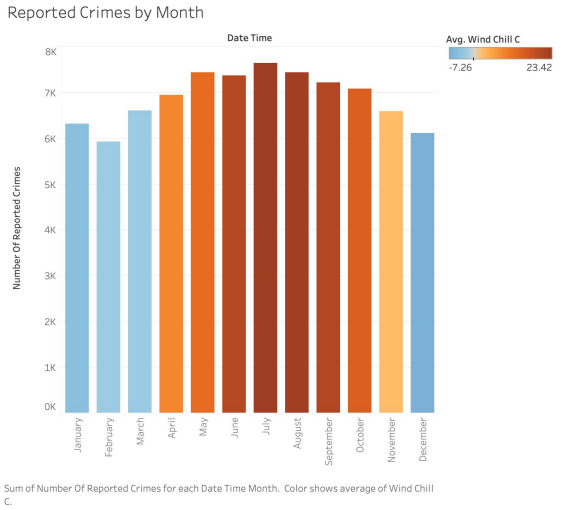


Exhibit 2

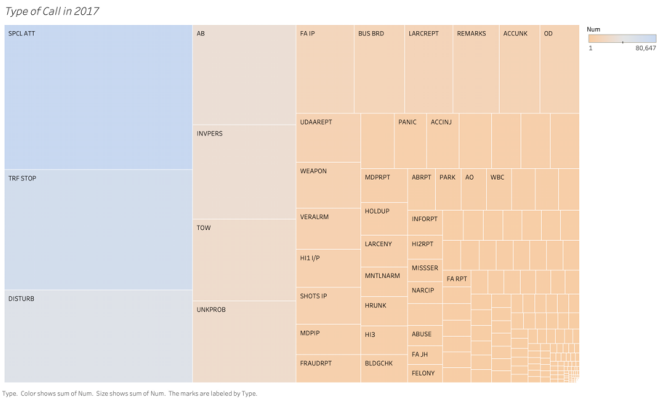
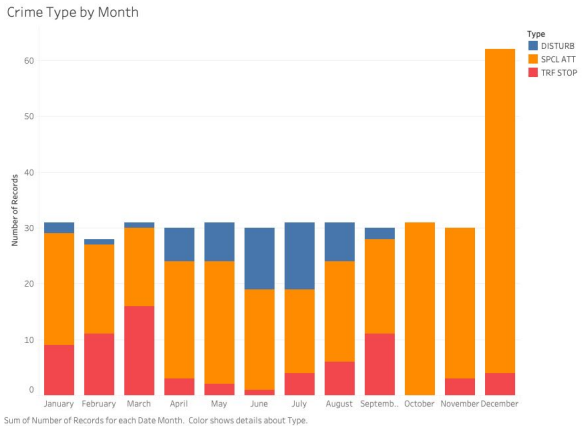
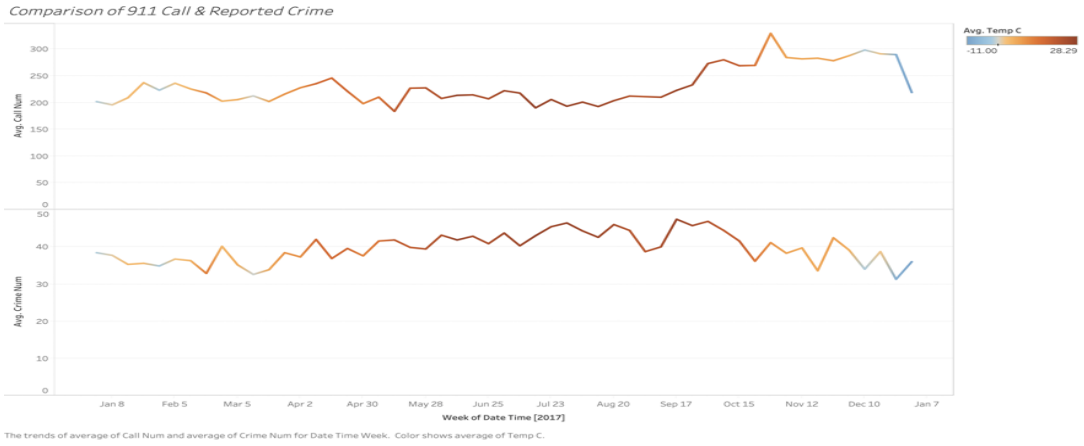


Exhibit 3



Next, we compared the average monthly 911 call levels against crime levels, transposing color for the average monthly temperature on top of each scatter plot line for 2017 (Exhibit 4). We found 911 calls started consistently increasing in the Fall months and then fell sharply back to normal rates through December when the temperature drops.

Exhibit 4



We then broke down the types of 911 calls and reported crimes by their volume across a temperature scale (Exhibit 5-6), with regression lines added for clarity. For 911 calls, disturbances increased with warmer weather, while special attention calls and traffic stops decreased. For crimes, fraud and assault increased with warmer weather, while motor-related crimes decreased. Most other crimes did not noticeably change volume with an increased temperature.

Exhibit 5

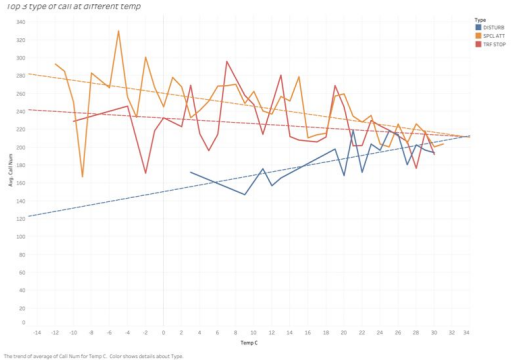
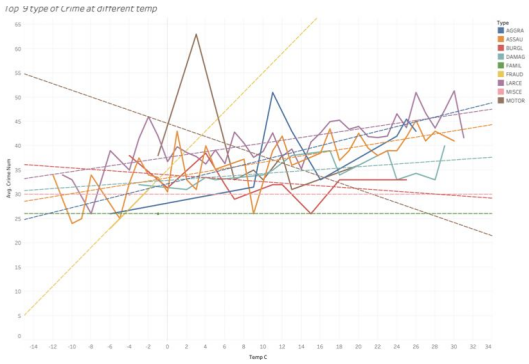


Exhibit 6



In addition, we calculated the crime ratio, which is the total number of reported crimes divided by the population in a zip code area, at different neighborhood locations in Detroit. Our analysis revealed that an even number of neighborhoods have a higher crime ratio in the summer (the blue side) than in the winter (the red side).

Finally, we attempted to understand if the number of crimes committed was affected by any other variables such as moon illumination and wind chill. Based on the number of crimes committed in a given day, bucketed into ranges of 10, we analyzed the accompanying average moon illumination. For example, of all the days with an average number of crimes reported

between 320-330, the average moon illumination was about 71. Ultimately, there is no discernable relationship between these variables, meaning no conclusion can be drawn.

Exhibit 7

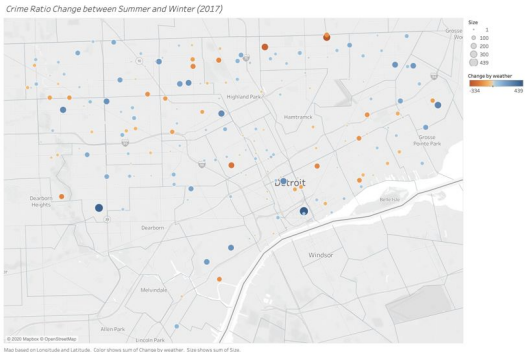
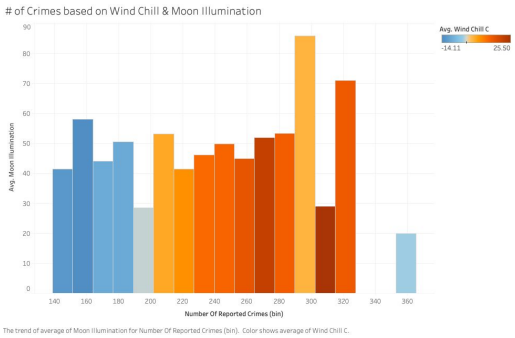


Exhibit 8



Key Takeaways

To summarize our Tableau visualization analysis, our key takeaways were:

- Overall crime rates do increase with warmer weather, as the number of crimes peak during the hottest month of the year (July) and are lowest during the coldest month of the year (February)
- The types of crime vary with the temperature
    - Disturbances, which are typically defined by problems occurring within the home, rise with the temperature (Exhibit 3). Other physically violent crimes and activities like aggravated assault, simple assault, and larceny increase at warmer temperatures (Exhibits 5-6). One potential cause is that warmer weather brings out more emotion in people, leading to arguments, fighting, and violence.
    - Auto-related crimes (Exhibits 6) tend to drop with warmer temperatures. One explanation for this is that there is less activity in colder months, giving thieves more opportunity to target idle vehicles.
- There does not exist any discernible relationship between moon illumination (how bright the moon is) and the number of crimes. This runs counter to our hypothesis that a brighter moon would lead to more criminal activity at night.

Research Shortcomings

We noticed a few shortcomings in our research and analysis, including:

- The team observed one full-year of data in order to examine a clear time period rather than multiple years of partial data. Unusual or unseasonal weather, or idiosyncratic criminal activity could therefore lead to misleading results.
- We didn't analyze correlation versus causation, thus weather factors may be indicators rather than precise drivers of behavior. As a result, certain takeaways and conclusions based on this data may be unsubstantiated.

Appendix

Python Code to Access Weather API:

```
#accessing Weather API to get historical daily accumulations for Detroit during 2017
#first step - loading wwo_hist in prompt
#importing packages
from wwo_hist import retrieve_hist_data
import os
#setting directory
os.chdir("/Users/sampe/.spyder-py3")
#setting parameters of data pull
frequency=24
start_date = '1-JAN-2017'
end_date = '31-DEC-2017'
api_key = '4be50bac89294352a8e221102201702'
location_list = ['detroit']
#defining how to talk to api and what to pull
hist_weather_data = retrieve_hist_data(api_key,
                                location_list,
                                start_date,
                                end_date,
                                frequency,
                                location_label = False,
                                #saving as detroit.csv
                                export_csv = True,
                                store_df = True)
```

SQL Code to Access DPD Crime Database:

1. Total number of 911 calls for each type of call in 2017

```
SELECT year(911call.callTime) AS Year, COUNT(*) AS CallAmount, 911callcode.category AS type
FROM 911call INNER JOIN 911callcode ON 911call.callCode = 911callcode.code
WHERE year(911call.callTime) = '2017'
GROUP BY type
ORDER BY CallAmount DESC
```

2. Total number of calls in each day of 2017

```
SELECT DATE(callTime) AS date_time, COUNT(incidentId) AS number_of_911calls FROM 911call
WHERE YEAR(callTime) = '2017'
GROUP BY date_time
ORDER BY date_time
```

3. Total number of crimes in each day of 2017

```
SELECT LEFT(crime.incidentTime,10) AS date, COUNT(DISTINCT charge.description) AS type FROM
crime INNER JOIN charge ON crime.chargeCode = charge.chargeCode
WHERE YEAR(crime.incidentTime) = '2017'
GROUP BY date
ORDER BY type DESC
```

4. Type of crime with the highest number in each day of 2017

```
SELECT DATE, TYPE, max(Num) AS CrimeNum
  FROM
     (SELECT LEFT(crime.incidentTime,10) as DATE,
          LEFT(charge.description, 5) as TYPE,
          COUNT(charge.chargeCode) as Num
       FROM crime INNER JOIN charge
       ON crime.chargeCode = charge.chargeCode
       GROUP BY LEFT(crime.incidentTime,10), TYPE
       ORDER BY Num DESC) AS Table2
  WHERE year(DATE) = '2017'
  group by DATE
  order by DATE
```

5. Type of 911 call with the highest number in each day of 2017

```
select DATE, TYPE, max(Num)
  FROM
     (SELECT left(911call.calltime,10) as DATE,
          911callcode.category as TYPE,
           count(911callcode.category) as Num
       from 911call inner join 911callcode
       on 911call.callcode = 911callcode.code
       group by left(911call.calltime,10),911callcode.category
       order by Num DESC) AS Table1
  group by DATE
  order by DATE
```

6. Crime Ratio at each season in each neighborhood
   Winter: 12-2, Summer: 6-8, Spring: 3-5, Fall: 9-11

```
SELECT COUNT(crimeId) AS Num_W, neighborhoodId,longitude, latitude FROM crime
WHERE LEFT(incidentTime,7) = '2017-12' OR LEFT(incidentTime,7) = '2017-01' OR
LEFT(incidentTime,7) = '2017-02'
GROUP BY neighborhoodId
ORDER BY neighborhoodId DESC
```

```
SELECT COUNT(crimeId) AS Num_S, neighborhoodId,longitude, latitude FROM crime
WHERE LEFT(incidentTime,7) = '2017-06' OR LEFT(incidentTime,7) = '2017-07' OR
LEFT(incidentTime,7) = '2017-08'
GROUP BY neighborhoodId
ORDER BY neighborhoodId DESC
```