Chicago Crime Detection (Week 2)

Total points 19/20

Greetings from Consulting and Analytics Club, IIT Guwahati.

We hope you had an amazing learning experience till now. Reaching till Week 2 Assignment is a big feat! This is your assignment/mini project for Week 2.

You are hired as an Analyst by Chicago Police. They have been maintaining a database of "motor vehicle theft" (sometimes referred to as grand theft auto) since last few years. This is the act of stealing, or attempting to steal, a car.

Due to increasing crimes, they want to make use of data to strategize better to maintain peace and order in their State. As a part of their Analyst team, you have been posed with some questions given below. Use your gained Data Analysis skills to answer these questions and help of the Police Department!

For The Dataset refer to the drive link :
https://drive.google.com/drive/folders/1Wy4UppKe_guGnzuC-Xk2zriAECyt0_vt?usp=sharing

Data Columns details:
1. ID: a unique identifier for each observation.

2. Date: the date the crime occurred.
LocationDescription: the location where the crime occurred.

3. Arrest: whether or not an arrest was made for the crime (TRUE if an arrest was made, and FALSE if an arrest was not made).

4. Domestic: whether or not the crime was a domestic crime, meaning that it was committed against a family member (TRUE if it was domestic, and FALSE if it was not domestic).

5. Beat: the area, or "beat" in which the crime occurred. This is the smallest regional division defined by the Chicago police department.

6. District: the police district in which the crime occured. Each district is composed of many

beats, and are defined by the Chicago Police Department.

7. CommunityArea: the community area in which the crime occurred. Since the 1920s, Chicago has been divided into what are called "community areas", of which there are now 77. The community areas were devised in an attempt to create socially homogeneous regions.

8. Year: the year in which the crime occurred.
Latitude: the latitude of the location at which the crime occurred.

General Guidelines:
1) We recommend you to download the data set from the link above and perform analysis in your respective Jupyter Notebooks, to answer the questions given below.
2) All questions are compulsory and thus should be attempted.
3) Each Question has weightage and will contribute in the final grading of the course.
4) Please attempt this if you have completed all the 6 days of Week 2.
5) Violation of the honor code will lead to harsh actions being taken.

NOTE: PLEASE DO NOT COMMIT SPELLING MISTAKES IN WRITING THE DATES/MONTHS/WEEKDAYS.

Note: Some questions may require studying certain topics from the web , no support material will be provided for the same, it is up to the participants to study it on their own.

Helpful Material:
1) https://www.shanelynn.ie/summarising-aggregation-and-grouping-data-in-python-pandas/
2) Helpful function: df.value_counts()

19 of 20 points

✓ What is the Month and Year of the median date in our dataset? * 1/1

Enter your answer as "Month Year", without the quotes. (Ex: if the answer was 2008-03-28, you would give the answer "March 2008", without the quotes.)

May 2006 ✓

✓ In which month did the lowest motor vehicle thefts occur? * 1/1

(Ex: if the answer was March, you would give the answer "March", without the quotes.)

February ✓

✗ On which weekday did most motor vehicle thefts occur? * ···/1

(Ex: if the answer was Monday, you would give the answer "Monday", without the quotes.)

Sunday ✗

✓ Each observation in the dataset represents a motor vehicle theft, and the Arrest variable indicates whether an arrest was later made for this theft. Which month has the largest number of motor vehicle thefts for which an arrest was made? * 1/1

(Ex: if the answer was March, you would give the answer "March", without the quotes.)

January ✓

✓ Select the true statements * 2/2

- [x] In general, does it look like crime decreases from 2002 - 2012? ✓
- [ ] In general, does it look like crime increases from 2005 - 2008?
- [x] In general, does it look like crime increases from 2009 - 2011? ✓

✓ Crimes for which more arrests were made are in the first half of the time period or the second half of the time period? * 1/1

(Note that the time period is from 2001 to 2012, so the middle of the time period is the beginning of 2007.)

- (●) First Half ✓
- ( ) Second Half

✓ For what proportion of motor vehicle thefts in 2001 was an arrest made? * 2/2

Note: in this question and many others in the course, we are asking for an answer as a proportion. Therefore, your answer should take a value between 0 and 1

- ( ) 0.06
- ( ) 0.08
- (●) 0.10 ✓
- ( ) 0.12

✓ For what proportion of motor vehicle thefts in 2007 was an arrest made? (round your answer to 2 decimal places) * 2/2

0.08 ✓

✓ Which locations are the top five locations for motor vehicle thefts, excluding the "Other" category? * 3/3

- [ ] Bank
- [x] Gas Station ✓
- [ ] Hotel/Motel
- [x] Street ✓
- [ ] Car Wash
- [ ] Restaurant
- [x] Parking Lot/Garage (Non-Residential) ✓
- [x] Alley ✓
- [x] Driveway (Residential) ✓
- [ ] Vacant Lot/Land

✓ Create a subset of your data, only taking observations for which the theft happened in one of these five locations, and call this new data set "Top5". How many observations are in Top5? * 4/4

177510 ✓

✓ On which day of the week do the most motor vehicle thefts at gas stations happen? * 2/2

- ○ Monday
- ○ Tuesday
- ○ Wednesday
- ○ Thursday
- ○ Friday
- ◉ Saturday ✓
- ○ Sunday

NOTE FROM MENTORS

Though the questions above can only test a small part of what have been taught, it is highly suggested that you should explore the dataset further and create your own charts, visualizations and perform an even greater in depth Exploratory Data Analysis. A good analysis can even be put up on Github to showcase as your Data Analytics project.

BEFORE SUBMITTING

The next page will lead you to accept the honor code and submit the User ID and Password. Make sure you have these ready.

THE USER ID AND PASSWORD FOR ALL QUIZZES/ASSIGNMENTS IS SAME THROUGHOUT COURSE

General Instructions 0 of 0 points

Please fill up all the relevant details provided to you.

You are allowed only ONE submission per one slack account, hence click the submit button wisely only after re-checking your answers.

Honor Code

1) You can submit the form only once, using a single User ID and Pass. Usage of multiple accounts for submission of quizzes will lead to harsh actions being taken.

2) Your answers to the questions must be your own work. (Very Important)

3) You may not share your solutions with anyone else unless explicitly permitted by the mentor. This includes anything written by you, as well as any official solutions provided by the course .

4) You may not engage in any other activities that will dishonestly improve your results or dishonestly improve or damage the results of others.

You can report Honor Code violations by contacting any of the members of Consulting and Analytics Club, IIT Guwahati.

Please Enter you Name *

Ashwath Kumar Shetty R

Please enter your User ID correctly in the following text field (This will be needed for grading) *

U011ZAC18Q1

Please enter the password *

QC4U2HE5

Please type " I Accept the Honor Code and will not violate it in any possible way" in the following text field. *

I Accept the Honor Code and will not violate it in any possible way