BYU Data and Modeling Case Study Competition 2024

Werewolves in Chicago

Purpose

The purpose of this case study competition is to provide an opportunity for students to use their skills in working with and modeling data in a more broad setting than is typically allowed in a classroom setting.

Instructions

Teams: Each team must consist of between 3 and 4 BYU Undergraduate students. Multiple teams can consult one with another in order to help guide, teach, or discuss, but the final product should be work created by each individual team.

Data and Methods: There will be data provided for you that will be sufficient to answer the question, but there are no limitations as to what data you can use. If you find other public data relevant to the problem you are trying to answer, you can include it in your analysis and your conclusions. You are not required to use any specific method. For the higher tier, some modeling is expected as long as the models used are done so correctly, are well justified, and effectively answer the problem you are given.

Presentation Presentations should be concise, with a maximum of 5 slides and an 8-minute presentation time, followed by a 2-minute Q&A session. One possible (but not required) way to organize the 5 slides is as follows:

- Slide 1 states the problem you have been given.
- Slide 2 describes the data you use and its appropriateness for answering the questions.
- Slides 3 and 4 describe the modeling and/or summarizing efforts that were taken, along with useful figures or relevant model results.
- Slide 5 includes a summary and conclusions that directly relate to the problem you have been given.

Each individual in the team must contribute to the presentation. Presentation skills will be evaluated as part of the competition.

Dates

March 6th: Details about the data and the problem statement will be released to registered teams.

March 12th: Registration deadline.

March 27th 6:00 PM to 8:00 PM: Data Contest Event. Dinner will be provided at 6:00 PM for presenters and judges. Guests are welcome to attend the presentations that will begin at 7:00 PM.

Presentation Rubric

The presentation will be evaluated based on the following criteria, with each section worth 10 points:

10 points: Slide Quality: Points are awarded for presentations with slides that are neat, clear, and concise.

10 points: Presentation Quality: Points are awarded for clear and effective descriptions of methods and results, and good general presentation skills.

10 points: Data Handling: Presenters who can clearly describe, display, or explain their data, including strengths and weaknesses in the context of the business problem, will receive more points.

10 points: Methodology: Points are awarded for methodologies that effectively address the business problem.

10 points: Conclusions: Presentations with conclusions that specifically address the business problem and are justified by the data and modeling work will receive more points.

Problem Statement

You belong to a vibrant young group of consultants called *Data Insight & Strategy Consultants (DISC)*, where "Our analytical skills light the path through the darkest mysteries" (a catchphrase your boss hoping will catch on). Today marks the beginning of an intriguing journey. Your new client, the Chicago Police Department (CPD), led by the esteemed Chief Landon, has presented us with a challenge that intertwines crime data analysis with elements of human behavior, meteorology, and perhaps, a touch of the mystical.

You have received the following correspondence from the chief's assistant:

Dear Team,

I hope this message finds you well. The Chicago Police Department is eager to explore innovative approaches to predict and, consequently, reduce the rate of violent crimes within our city. Your primary objective is to determine if and how various factors influence the rate of violent crimes in Chicago. Once we determine this, we can plan ahead with public outreach, community events, and other positive means of deterring violent crimes during particularly vulnerable periods. The Chief is particularly interested in understanding the potential impact of weather conditions, major holidays, and lunar cycles on criminal activity. While the notion that the full moon might affect crime rates is a topic of curiosity rather than conviction, we are open to exploring all possible correlations. Today, you find yourself seated across from Chief Landon in his office. The only light comes from the sprawling cityscape outside, casting a soft glow that dances across the Chief's earnest face. He's a man who has seen the city through its best and worst, known for his sharp insight and unexpected openness to explore beyond conventional wisdom.

Today, you find yourself seated across from Chief Landon in his office. The only light comes from the sprawling cityscape outside, casting a soft glow that dances across the Chief's earnest face. He's a man who has seen the city through its best and worst, known for his sharp insight and unexpected openness to explore beyond conventional wisdom.

"Thank you for coming," he begins. "We've been wrestling with the unpredictable nature of crime here in Chicago. It's a city of incredible resilience and complexity, but as you know, it's also facing a persistent struggle against crime."

Your colleague Bobby nods, understanding the significance. He's always been quick to dive into data, looking for patterns and answers in numbers. "We've been examining various datasets, Chief. There's an intriguing array of factors that could be influencing crime rates," Bobby says, eager to get into the details.

Chief Landon smiles slightly, appreciating Bobby's enthusiasm. "I've heard," he acknowledges, leaning back in his chair. "I suspect that weather and holidays might play a part. And did you see that I?m asking you to also consider the lunar cycle as a potential influencer of crime. I must say, I find the idea fascinating." His tone is thoughtful, inviting further discussion.

"I have to admit," Bobby jumps in, "the full moon theory caught my attention. Not that we're expecting to find werewolves roaming the streets of Chicago," he adds quickly, sharing a brief, light-hearted moment with the Chief.

"Werewolves in Chicago would be the day," Chief Landon chuckles. "But jokes aside, there's something peculiar about those full moon nights. Maybe it's the way the air feels, or perhaps it's all in our heads. Yet, I can't help but wonder if there's a pattern, something the data could reveal."

"Let's dive into the data with an open mind," Chief Landon concludes. "If there's a thread connecting the moon, the weather, or holidays to crime rates, it's our duty to find it and understand it. This could be a key piece in strategizing our efforts to make Chicago a safer place."

Bobby, always one step ahead in his role as the data maestro of DISC, had already begun the groundwork prior to your meeting with Chief Landon. He collected several datasets. These or other data you find could be used to answer the questions.

1. **Crimes.csv:** This dataset is a detailed record of every crime reported in District 11 in Chicago since 2001, including date, type, location, and outcome. We'll focus specifically on violent crimes. This dataset is quite large, with half a million records. It can be found at https://richardson.byu.edu/data_contest/2024/Crimes.csv.
2. **Weather.csv:** Daily weather data since 2010, capturing temperature, precipitation, wind speed, and other meteorological variables. This dataset can be found at https://richardson.byu.edu/data_contest/2024/weather.csv.
3. **Holidays.csv:** A dataset listing major public holidays since 2010. This dataset can be found at https://richardson.byu.edu/data_contest/2024/holidays.csv.
4. **Full_moon.csv:** Contains the dates of all full moons since 2005. This dataset can be found at https://richardson.byu.edu/data_contest/2024/full_moon.csv.

You are not limited to these datasets or even these factors. Feel free to use only what is given here or to go find additional features of your own. You will reconvene with Chief Landon and some CPD analysts on March 27th to present your findings on what, if anything, could be used to predict more or less incidents of violent crimes.