hw13

**Question 18.1**  
Describe analytics models and data that could be used to make good recommendations to the power  
company.  
Here are some questions to consider:

• The bottom-line question is which shutoffs should be done each month, given the capacity  
constraints. One consideration is that some of the capacity – the workers’ time – is taken up by  
travel, so maybe the shutoffs can be scheduled in a way that increases the number of them that  
can be done.  
• Not every shutoff is equal. Some shutoffs shouldn’t be done at all, because if the power is left  
on, those people are likely to pay the bill eventually. How can you identify which shutoffs  
should or shouldn’t be done? And among the ones to shut off, how should they be prioritized?

Think about the problem and your approach. Then talk about it with other learners, and share and  
combine your ideas. And then, put your approaches up on the discussion forum, and give feedback and  
suggestions to each other.  
You can use the {given, use, to} format to guide the discussions: Given {data}, use {model} to {result}.  
Have fun! Taking a real problem, and thinking through the modeling and data process to build a good  
solution framework, is my favorite part of analytics.

**Power Case Problem:**

A power company is facing an issue where they have some customers that continually miss payments, which loses profit for the business. While the company doesn’t have an issue for customers who cannot afford basic power needs or customers who are behind on their power payments, the company does have an issue with customers who can pay their bills but chose not to pay. They want to turn off power for people who are not ever going to pay. Therefore, the 1st problem is to identify the customers that are not going to pay and turn off their power. Once the customers have been identified, the company needs to go to the physical locations to turn off the power. Unfortunately, there are more people each month whose power should be shut off than the company had capacity to handle. The 2nd problem seems to be route optimization problem in order to maximize the number of power shut offs in a single month.

**Data and Analytic Models:**

1. **Determining whose Power should be Shut-Off**

In order to determine whose power needs to go off, we should 1st determine if individuals could pay off their bills based on their earning and borrowing power. We can determine if individuals have access to capital by checking their credit history, the price of living in their zipcode, and their annual income. Some important factors such as annual income may be missing, so we could impute this data using regression + perturbation to help fill in the gaps. By at least using these factors, we can classify delinquent customers into groups: ones who do ***not*** have access to capital to pay their utility bills vs ones who do have access to capital to pay their utility bills.

For classification, we can use KNN (K Nearest Neighbors) to determine the customers in each group. We could also use SVM Hard (Support Vector Machine) to split the delinquent customers. Once we determine who can pay their bills but are deciding not to, we can reach out to those customers by giving a heads-up, and receiving feedback on why they are not able to pay their bills. This will help us collect data on why customers are missing their payments, but also serve notice that their power will be cut off at a future date (possibly a month?) unless action is taken. We can prioritize the shut-off per customer depending on their communication with us; if we can build a relationship with the customer, and understand them better, they should be more likely to pay off their bill. Those customers with negative responses or no responses should be the 1st ones to have their power shut off.

1. **Route Optimization for Power Shut-Off**

In order to optimize the number of shut-off we can do per day and per month, we need to determine the hours of shut-off operation, the number of vehicles out for dispatch, the locations of the shutoffs, the minimum, average, and maximum amount of time for a shut-off, and the location/s of our vehicle/s dispatches. For the sake of simplicity, we can assume that we have 1 vehicle and 1 start location. The company should investigate the optimal number of vehicles or dispatch locations once the initial route optimization is studied well enough. The vehicle would need to travel such that they can shut off the maximum number of locations within the hours of operation. We can determine the future locations for shut-offs as soon as we give the ~1 month or so notice to the customers that their power may be shut-off. Future routes can be pre-planned as locations to be visited are mostly predetermined.

Routes can be complex as we need to look at more than just the distance between locations; we need to factor in the average time to travel from one location to another location depending on the average traffic in that travel distance during that time as well. We also do not want to overbook the vehicle in case there is extra traffic in that a significant number of the shut-off locations scheduled to be visited that day are not able to be visited. This problem can be solved using a network optimization solution -- we can use data from Google Maps to help us with route optimization.

Depending on the company’s capital and operational jurisdiction, it also may be worthwhile to investigate 3rd party companies that have a fleet of vehicles and multiple locations that can perform the shutoffs. It may be more costly to shut off delinquent customer’s power, however it may save the cost of having many customers obtain power without paying for it. The power company only has limited time and resources, they could be spending their efforts in providing better services or obtaining more customers.