

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.

Problem Statement:

- Which customers need to get their power shutoff
 - How to prioritize shutoffs
 - How to shutoff the most efficiently with worker constraints

The data needed:

- Customer Data (Demographics, payment history, usage patterns)
- Geographic Data (Location of customers, distance from workers, travel time, customers per geographic location)
- Operational Data (Worker schedules, worker wages, lost revenue, time needed to shut off one provider)
- Payment Data (Prior data of customers who ended up paying after shutoffs)

The models possibly used:

- Predictive Models (K-nearest neighbors)
 - To determine which customers will pay their bill after they are shutoff.
- Optimization model (Linear or Integer Programming)

- Using worker schedule constraints to optimize the schedule of workers and their travel. Also optimize the amount of shutoffs that should be done first over others.
- Time Series Analysis (ARIMA)
 - For the customers that have not paid their bill, we can use ARIMA to take into account if they have cyclical patterns of payment like if they would rather pay closer to tax season or christmas.

The models output:

1. Customer List:
 - List of customers split by the category they fall into.
 - Paying customers
 - Late paying customers
 - Non-paying customers
2. Shutoff List:
 - With the help of the optimization models, we will be able to make a prioritization list that shows the areas of shutoff from highest priority to lowest priority.
 - The likelihood of customer payment will be taken into account along.
3. Shutoff Efficiency:
 - With the help of optimization models, we will be able to take into account the workers schedules and their location to better map out the order of shutoffs. This would mean minimizing travel time for workers.

Given customer demographic and payment data, I will **use** k-nearest neighbors **to** determine the category a customer falls into.

Given customers' previous payment history and their locations, I will **use** decision trees and time series analysis to determine which customers should be more prioritized for shutoff based on their payment history.

Given the customer and shutoff list, along with the workers location, schedule, and pay, I will **use** simulation (simPy) along with optimization models **to** find which worker schedule and shutoff list will make the shutoff process the most efficient.