**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.

**Answer:**

# When considering which power hookups to shut off there are a few parts to be considered. This problem can be broken down into three components:

# How much is an account costing the company?

# This aspect is likely the most straightforward among the challenges presented. Given:

# Historical data on customer power usage

# Seasonal variations

# Specific usage patterns of the customer under consideration

# Current outstanding debt from the customer

# We can utilize linear regression to anticipate the customer's power consumption for the upcoming month or months. By factoring in the existing debt and forecasting their upcoming usage, we can estimate the total debt each customer might accrue if payments are not made promptly. This approach offers a valuable means to assess and compare customers' financial situations in the upcoming phase of the problem.

# The seasonality aspect is crucial as power consumption tends to fluctuate based on weather conditions. Understanding the specific month under consideration, whether it's during summer or winter, plays a significant role in projecting power usage accurately.

# Which shutoffs should be done each month or rather, who is not going to pay, who cannot pay, and who will pay?

# Before deciding to shut off power to customers, we must figure out which ones should be considered for this action. We can do this by looking at past payment records and predicting what might happen with new customers.

# We'll look at things like how much power a customer has used before, how often they've paid on time, how many times they've missed payments, and how long it takes them to pay after getting a bill. We'll also consider other details like whether they rent or own their home, if they're married, how many people live there, and if they're a residential, business, or government customer.

# We'll use a simple method called logistic regression to predict if a customer will pay, won't pay, or might pay but chooses not to. This helps us guess what might happen based on what we've seen before. Then, we'll think about how much it costs the company if a customer doesn't pay, how much it costs to cut off their power, and how much they might use in the next month.

# If it seems like a customer probably won't pay and they owe a lot or will use a lot of power soon, we might consider cutting their power. But if they've only missed a few payments or don't owe much, we might not. We'll also think about different types of customers separately, like businesses or governments, because they might have different payment habits.

# Overall, we're trying to make a smart decision about who to consider for a power shutoff based on their payment history and how much they owe, while also considering how much it would cost the company.

# How do we optimize the limited manpower to perform the shutoffs?

# Given:

# Customers identified for shutoffs

# Predicted costs associated with each customer

# Location data for each customer

# Time required to perform each shutoff

# Estimated travel time between locations using mapping applications

# We can prioritize areas for shutoffs using communities in graphs, modularity, clustering, and optimization.

# By treating each customer as a node and travel times between nodes as connections, we can use algorithms like Louvain to create communities based on proximity. Setting a travel time threshold prevents considering every possible connection. Using methods like K-Means helps establish a manageable number of communities. We calculate the total predicted cost within each community and focus resources on the one with the highest expected cost. This model can be updated regularly based on changing conditions.

# We then optimize worker routes within these communities, aiming to minimize total cost considering travel and work times. We need to consider workers' total daily time constraints.

# In conclusion, combining various analytical methods like linear regression, logistic classification, clustering, and optimization provides a solid framework for recommendations. Access to customer data, location details, and public information like GPS and traffic data is essential. While improvements like queuing models and real-time updates would require significant investment, the proposed framework offers effective recommendations without such costs.