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.

Given:

* Weather: Historical weather patterns during the seasons. In real life, there are regular power shutoffs where it gets too windy. The electric companies preemptively shutoff power to prevent fires.
* Cost: Costs associated with shutting off power (e.g., operational, logistical).
* Outstanding Payments: Amounts owed by customers.

Use:

* Time-Series Model:
  + Purpose: To understand historical shutoffs and predict future patterns.
  + Application: Analyze trends over time, such as seasonal variations or the impact of weather on shutoff rates.
* KNN Clustering:
  + Purpose: To identify “epicenters” of shutoffs.
  + Application: Group customers based on their geographical proximity and similar payment behaviors. This helps in understanding areas with high rates of shutoffs and optimizing travel for shutoff operations.
* Logistic Regression:
  + Purpose: To determine whether the outstanding payments have been paid or not (Y/N).
  + Application: Predict the likelihood of a customer paying their outstanding bill based on past behaviors.

To:

* Predict and prioritize which areas should have shutoffs based on weather (time-series), logistical efficiency (knn clusters), and financial standings (logistic regression)
* Assumption: Assuming power typically does not shut off due to political reasons, these models can help identify and prioritize shutoffs based on weather, geography, and financial updates.