**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.

ideas.for.hw8

how to identify paying vs not paying customers

* logistic regression - probability of x paying or not paying
* k means clustering could also help (id a definite bad customer and see how other customers cluster close to them based on demographics)
* how about using exponential smoothing (cyclical, or trend, or estimate vs real)
* heuristic / bayes model (get input from the customer as well?)

how to identify optimum routes

* k means clustering to identify the right clusters of target (dud-)customers,
* identify the optimum clusters, then for each clusters, identify the maximum cost cluster

How to identify minimum cost:

* this is an optimization problem.

Advanced topics:

* What about the cost when we have to turn service on to customers who paid?

What not to use:

* Weibull or geometric. Both suggest the customer (or cluster of customers) end up paying in the end, and we are after the customers who NEVER pay
* Cusum to track change detection for a customer going from paying , or potential paying, to never paying based on earlier piazza discussions
* Personal data is costly to use and can be risky. Minimal data is necessary for identifying dud-customers.