Todd Faulkenberry – Stat 215a Lab 1 homework

Hypothesis Testing
See attached piece of paper

Questions from Freedman
1. The similar death rates from other causes. We know this number is legitimate because of the random sampling done at the beginning of the experiment
2. No because there a lot of confounding variables comparing those that accepted to everyone (i.e. wealth.) It would be better to measure everyone to whom you offered treatment vs everyone to whom you didn't.
3. Natural experiment because he didn't set up treatment and control groups or restrict anyone from drinking water from certain types. It matters that the Lambeth company moved their pump because it allowed Snowe the opportunity to see the impact of changing water supplies on the amount of cholera patients in the area, giving him a chance to compare and contrast the amount of cholera patients between two companies / sources of water.
4. Observational study because he didn't set up anything to randomize groups. He just recorded where people lived as well as a few other statistics.
5. If the number were negative, Yule would have had to conclude that more welfare earners outside the poor house actually decreased the numbers of paupers. If the number were positive, he would have had to conclude the opposite.

9. No, it does not because it does nothing to eliminate bias. The people keeping diaries may have a predisposal to believing God is helping him, or may have more resilience than non-religious people. I would explain the data as effective useless in any rigorous statistical sense, though the diary entries may be interesting.

Questions about John Snow
- If all I had were this map, I would make a simple table with three columns: "Block," "# of Cholera Patients," "Nearest Water Pump." If I had access to more robust information, I'd create tidy data that had one row for each observation, with values like "House," "Address," "Name," "Cholera Status," and "Nearest Water Pump."
- Quantifying would give you good leads on which water pumps were dispensing cholera-infested water because you would be able to see which blocks a) had the most cholera patients, and b) what their nearest pump was. At the same time, I think this map already provides that on some level. The visualization w/r/t the sheer amount of cholera patients near certain pumps like the Broad Street one is striking.