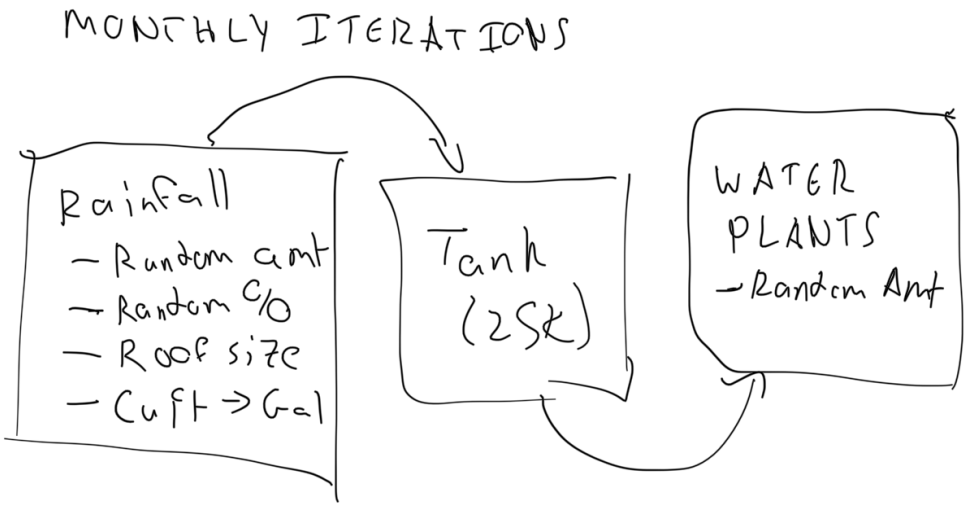
1. Construct a flowchart to represent the rainwater / harvest system

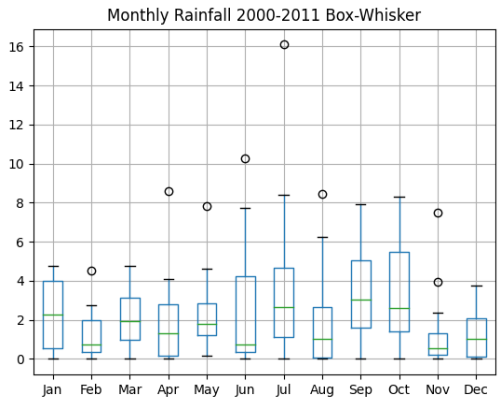
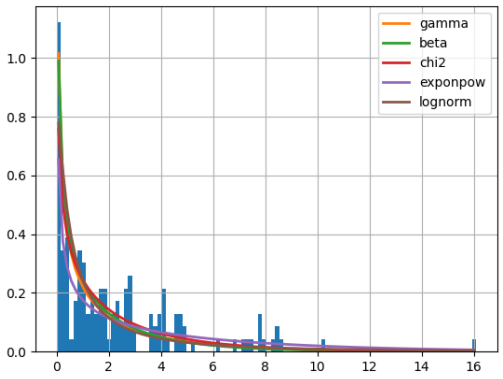


Using SimPy terminology, there are two agents - "Rainfall" and "Watering Plants" - and a water tank as a Container. Every month, the rainfall agent executes its process of pulling an expected volume of rain (see next question for details), identifying a capture efficacy, and converting this amount into a volume of water measured in gallons. The agent fills the water tank container up to the tank's maximum. Then, the "Water Plants" agent executes, identifying a volume of water needed that month, pulling that volume from the tank. If the tank runs out of water, the process ceases.

Note: the two agents could be collapsed into a single process where rainfall in and water out were netted, and if net rainfall was positive it got pushed to the tank and if negative it got pulled from the tank. This single-process approach generates a marginally higher end-of-month water level if the tank is at or near its full capacity. For the purpose of this simulation the two approaches would yield similar outcomes.

2. Generate a table of monthly rainfall and describe how you used that data in simulation.

I used the data as provided (available in the supporting documentation; not represented in this paper), and evaluated whether it made more sense to think about monthly rainfall as a per-calendar-month process or whether I would ignore calendar months. Seeing as we only had 12 observations per calendar month, creating a per-calendar-month distribution seemed unnecessary.

Monthly box-whisker	All observations histogram (w/ fitted lines)
	

I used a python package called Fitter to identify which of the common data distributions best fit the monthly rainfall data. Given the prevalence of "no rainfall" months, a key determination was to identify a distribution that would appropriately capture the heavy left tail of the data. Of the distributions tested, the gamma distribution had the lowest squared-error, and so was selected. To confirm its appropriateness, I compared the mean of the observed data (~2.3) versus the mean of a 10k-random sample draw from the fitted curve (~1.9). The difference is likely driven by the few large outliers in the observed data.

3. Build a simulation. How does the simulation answer the rancher's questions?

I built the simulation in SimPy, as described above in question 1. As the simulation runs, the month-end water level is captured in a data series. The simulation's 1,000 runs are captured in a master dataframe which can be interrogated for *post hoc* analysis.

4. Run 1k 30-year simulations starting with 10k gallons. What can you say about minimum water levels?

5. Consider possible effects of climate change. What will change in the simulation and what should the rancher do about it?

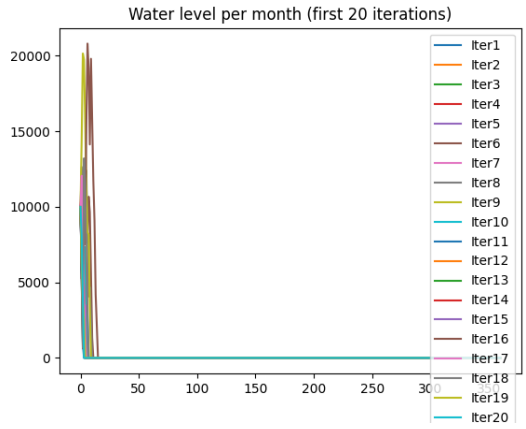
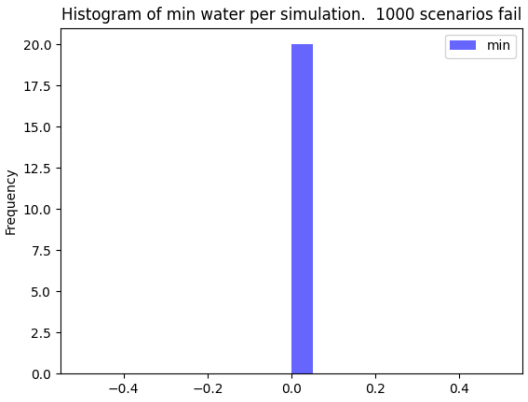
With the assumptions given (~2" rain / month, 3k sqft capture area = ~3800 gal/month), the rancher will run out of water 100% of the time. With an expected net outflow of 800 gal / month, tank size doesn't matter, the rancher will run out of water in a 30-year horizon.

If we increase the catchment area by a factor of 10 to 30k sqft, then the rancher will still run out of water ~15% of the time. If the catchment increased and the size of the tank is increased 10x then the rancher will run out of water <5% of the time.

In a world impacted by climate change, the most logical input to adjust is rainfall, which I represented as a scalar 20% haircut off of the estimated monthly rainfall. This change has a material impact on failure rate in the scenario with a normal-sized tank and larger catchment area, largely driven by the greater likelihood for a long-enough string of low-rainfall months. There was essentially no impact to the failure chance if there is also a very large tank.

		% scenarios fail (run out of water)	
Roof Size (000s sqft)	**Tank size (000s gal)**	**No Rain Haircut**	**20% Rain Reduction**
3	25	100%	
3	250	100%	
30	25	16%	25%
30	250	3%	3%

Scenario Outputs with 3k Sqft Roof + 25k Gallon Tank

Time Series of First 20 Runs	Histogram of Scenario Min Water Level
	

Scenario Outputs with 30k Sqft Roof + 25k Gallon Tank

Time Series of First 20 Runs	Histogram of Scenario Min Water Level