Code Challenge

Goal: Please complete within 24 hours and send it back to Flipt for review. You should use Python/R for data manipulation, and you can choose any data visualization tool (preferably data viz package within Python/R, or Excel/Tableau if you have a license) you prefer to do visualizations.

Questions/Clarifications: Contact Kimi Yang directly at kyang@fliptrx.com

Datasets Provided:

1. **Simulation Results Sample Data.xlsx**
   The dataset has 1000 simulation results. Each result is by one of the 9 combinations (see table below), and each result is cost breakdown by retail/mail & generic/brand/specialty. (You don't need to understand what these labels mean if you don't have PBM background, but instead just treat them as different categories)

	Low AWP Trend	Mid AWP Trend	High AWP Trend
Low Vol Trend	**Combination 1**	**Combination 4**	**Combination 7**
Mid Vol Trend	**Combination 2**	**Combination 5**	**Combination 8**
High Vol Trend	**Combination 3**	**Combination 6**	**Combination 9**

Questions:

1. Explore the various costs by low/mid/high volume trends, and by low/mid/high AWP trends, then create visualizations.
2. Explore the various costs by the 9 trend combinations, and create visualizations.
   - *For the above two, it's an open question, so you should be creative and think about your own way of visualizing the data and tell the story.*
3. Regardless of the trend combinations, what does the distribution look like for the 1000 results (total cost)? Is it normal distribution? If not, why do you think it's not normal?
4. Subset the data for the **Mid Vol Trend** and **Mid AWP Trend** combination. Plot the total cost distribution for the subset.
5. For the same subset above, randomly assign one of the three saving percentages (10%/15%/20%) to each row, and calculate the total cost after savings in a new column: *total cost after savings = total cost * (1 - savings %)*

- What does the distribution look like? Plot the distribution curve in the same graph as in #4. Hint: it will look similar to #4, but with lower average.

Submission:

1. HTML file from R (using RMD) or equivalent presentable file/HTML from Python.
2. Visualizations should be included in the file above.
3. Rerunnable codes with your comments.