Questions

Please run the Lab 8 QR Program and examine the code before answering the questions below. You can download the Python file from the Lab 8 QR Instructions page.

Type your answers in the box below each question. It will automatically expand to accommodate your answers.

1. State the Problem or Question

Run the program. The program generates and displays some statistics about the death penalty and about the rates of exonerations (people who were later found innocent) in states that have the highest number of death sentences. What question or questions can be answered with this program? What related questions might someone want to ask about the dataset, or about other data that might be available?

> Problem:
> We have data for 50 states regarding the death penalty, penalty. How do we display it?
>
> Relevant Questions:
> -What is the most important statistic to display? Is it the total number of executions, number of exonerations, difference, or expected rates?
>
> - Does the data / can the data display alongside majority party election results? Should that data be displayed with context from presidential or governor party election outcomes?
>
> - How does the data compare to the overall crime rate given these states varying population numbers?
>
> Answered questions:
> These are the states with the highest total number of executions. This is the total of exonerations per state, Expected exonerations, and difference in expectations.

2. Determine Information Needed

The data set is created in the `make_data()` function and is correct to the best of my knowledge. For at least one related question that someone might want to answer, do you have all the data you need to answer that question? If so, what data from the dataset would you need? If not, what other types of data would you need to answer the related question?

> The key factor missing here is the weighted percentage of the state's population comparative to the number of sentences handed down. While Texas and California have near-parity in the number of executions, Texas has a population 25% less than California. This drastically skews the data when accounting for proportionality and shows a much better impactful statistic.

Going off the proportionality, the data should be presented in a manner similar to:
California – Executions per 100,000 residents: 2.46
Texas – Executions per 100,000 residents: 3.23

This proportionality also works when we examine the exonerations compared to the number of residents in each state.
California – Exonerations per 100,000 residents: 0.033
Texas – Exonerations per 100,000 residents: 0.044

This shows that not only does Texas execute ~25% more people that California, Texas also has an exoneration percentage 33% higher than California.

I would argue that the exoneration percentage is a far more impactful data point to look at, and easier to explain to the general public. Having the capacity to say that the state the exoneration rate is a far more impactful visual to communicate.

3. Representation Part 1

Please describe the calculations which the program is making **in mathematical form**. You can include equations, graphs, diagrams, tables, etc. if they help you to communicate the processing that your program performs.

1: Take all data and place into list
2: Set index = 0
3: Select next object in list
4: Get objects total sentences
5: Sort under object with fewest higher sentences than current object
6: Repeat until all list indexes parsed
…
7: Pass organized list to calc_percent_exonerations
8: Set total_exonerations = 0
9: set total_sentences = 0
10: Set index = 0
11: Select next object in list
12: Get object total exonerations
13: Set total_exonerations = total_exonerations + objects total_exonerations
14: Set total_sentences = total_sentences + objects total_sentences
15: Return calculation (100 * total_exonorations / totalsentences)
…
16: Calculate the first five-states in the organized list' expected_exonorations
17: Set expected_exonorations = State percent_exonorated_sentences * State sentences
18: Set difference = expected_exonerations – State actual_exonorations

4. Representation Part 2

Please provide pseudocode for the `calc_percent_exonerations()` function. You only need to provide pseudocode for that one function.

```
START

SET total_exonerations = 0
SET total_sentences = 0

FOR state = 0 TO States.length
- - SET total_exonerations = total_exonerations + CALL state get_exonerations
- - - - RETURN self.__exonerations__
- - SET total_sentences = total_sentences + CALL state get_sentences
- - - - RETURN self.__sentences__

RETURN 100 * total_exonertaions / total_sentences
```

5. Interpretation

Please explain the mathematical information you provided in question 3, in English. What does it mean in simple terms? As stated above, there is a mathematical error in the calculations. When you run the program, something about the results doesn't make sense. Describe in English what doesn't make sense and then try to find and fix the bug.

When looking at the outputted data, the tables related to Expected Exoneration and Difference and showing abnormal numbers. While the data does indicate the correct variation, the extra 2 digits indicate a potential double multiplication issue.

Upon further examination, we need on line 94 percentage calculation where total_exonerations is divided by total_sentences before being multiplied by 100. A few steps later, we see a 2nd round of multiplication done to this figure.

This second multiplication is a more state specific calculation and the is the data we need whereas the multiplication earlier was at a broader scope level or the entire data set. Since we are calculating the states' individual statistics for the end result, we should remove the multiplication from the first division and just return the base value post division.

Line 94 is the problem, remove the multiplication and leave just the division.

6. Conclusions

Explain the results that the program produces once you've fixed the bug. There is a difference between the actual and expected number of exonerations. What **does** this difference imply? Do

you believe that the program is using the correct data in the correct way to support this implication? What other factors might be in play?

> Resolving the arithmetic error alleviates the abnormal numbers observed in the data output and shows a correct expected exonerations vs difference number that's inline with the sentences and exonerated data tables.
>
> The difference now displays correctly negative values where the actual number of exonerations was under the expected based on the data average. This indicates a lower likelihood of those states executing a prisoner who is later exonerated compared to other states.

7. Community and Environmental Implications

Describe how the output that the program produces could be used to change public policy, if the output is believed.

There was a bug in the original program. Consider the negative consequences that might result from policy changes based on the incorrect version of the program, or from any version that produces incorrect results, or doesn't consider the right factors. Describe those consequences.

Finally, consider ways that the calculations could be changed to produce different results (not necessarily correct ones), with different effects on policy. What is our ethical responsibility as computer programmers to ensure that the correct calculations are being performed?

> Accurate data communication isn't just an important aspect in a mathematical sense, but in a human one as well. Being able to discern data in a way that is communicative to parties not intimately familiar with it is the primary initiative needed for change, especially in policy.
>
> When we examine data based on just big number vs. small number, so much contextual importance is lost that undermines critical conversations around sensitive topics like the death penalty. That contextual information is powerful, and it's vital that not only we as computer scientist recognize that capacity, but also reflect on data we consume.
>
> I work in a complicated environment and having the ability to report to my supervisors or respective VP's business data both in a raw and contextual terms has been my most powerful way to invoke change even in lesser critical roles / issues. It is how I have been able to smartly manage my own personal assets and avoid predatory companies / services that do prey on those who can't as easily discern manipulative data they are being presented with omitted context.
>
> This sort of behavior, socio communicative power in computer science, is such a powerful force of change in positive and negative contexts.