The Little QR Booklet

Don Blaheta
Department of Math and Computer Science
Longwood University

Quantitative Reasoning

What is "Quantitative Reasoning"? Broadly—and obviously—it is the process of reasoning through an answer to a question of interest using one or more quantitative modes of thought. Let's break that down a little:

Reasoning is what we do when we want to answer a question with more than just an opinion, and with more than just an intuition or a gut feeling. Some questions might not be answerable! But any reasoning process that does come up with an answer, which we'll refer to as *drawing a conclusion*, will be able to *justify* that conclusion by giving the reasons why it's true. Logical deduction and the scientific method are other well-known examples of reasoning processes. Sometimes a reasoning process can draw a conclusion with certainty; other times the best conclusion might be that a claim is only "probably true" or "probably false", or even that the claim can't be judged definitively without further evidence. Such conclusions are often still valuable!

Quantitative processes are those that deal with numbers ("quantities" or other measurements). As opposed to a qualitative analysis, where the focus is typically on deep details for a small handful of subjects, quantitative analyses will often focus on a few specific quantifiable measures over a very large data set: dozens, hundreds, maybe millions of data points. In other kinds of quantitative analyses, we have a smaller number of measurements but we use them to model growth patterns or decay rates with a greater degree of precision, so that we can predict future patterns.

So, quantitative reasoning is a process that lets you answer questions by making use of numbers. Importantly, the numbers themselves are not usually the answers to the questions, just a step towards the conclusion. You have certainly been taught many different computation tools in math classes, from adding two numbers to finding areas of polygons to solving for x in an equation, and depending on your specific background, maybe also taking a derivative or writing a `for` loop or calculating a z-score. All of those are important to the people that use them, but they are only a small part of the overall quantitative reasoning process (and not at all a focus of this booklet). Over the next few pages we'll be talking about how computation tools like those fit into the larger task of answering questions that you care about.

QR process overview

As we learn more about the quantitative reasoning process, always keep in mind the primary goal here: starting with an interesting research question, and doing stuff to answer it.

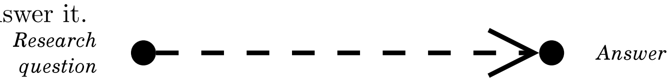

Fig 1: What we're trying to do

In this booklet, we're going to formalize the "stuff" of the QR process as three main steps:

1. Formulate a numbers question
2. Solve for numeric answer(s)
3. Interpret the numbers

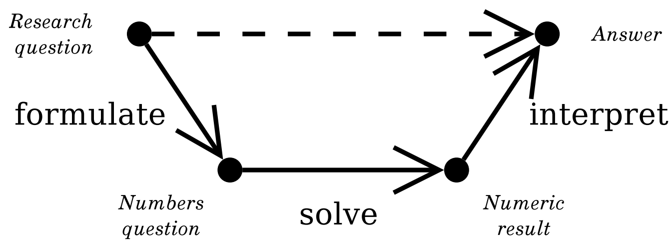

Fig 2: How we understand the process

plus a fourth overarching principle of "Reasonableness" that will guide you to reflect on your formulation, solution, and interpretation steps to give you confidence that your conclusion is valid.

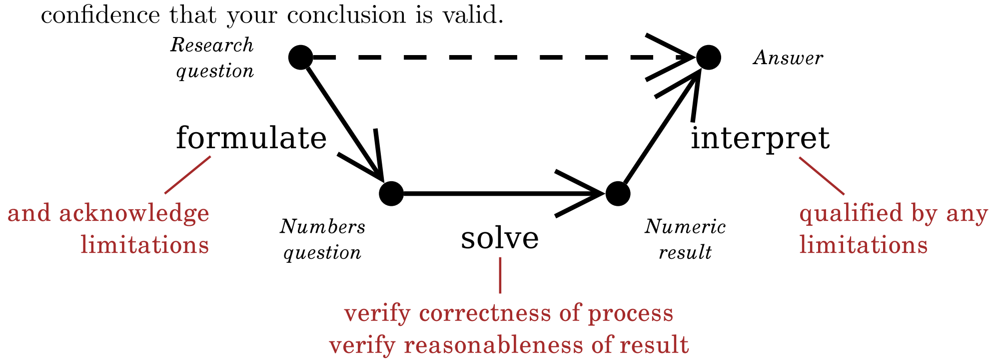

Fig 3: Building confidence in the process

Why make the process so formal? It probably seems like a lot, and it might be a bit extra if you're only investigating something for your personal curiosity—in that case, you might not need to explicitly lay all the steps out in such detail. The real power of the QR process is narrative. If you want to be able to convince *someone else* that your conclusion is valid, this multi-step framework provides a sort of checklist to help you lay out the reasons your conclusion is valid. As with logical deduction and the scientific method, sometimes the actual solving is a little messy—try something, back up, try something else—but when you want to explain your process and persuade others of your result, the reasoning framework is your narrative roadmap.

Prologue: What makes a good research question?

Before we dig into the QR process, we should dwell a little longer on the original dotted-line diagram:

Fig 4: What we're trying to do

What kinds of questions are we talking about here? Questions that can be answered through a quantitative process must have some sort of relationship with numbers, but (perhaps surprisingly) the core interesting question is not, itself, about the numbers. A question like, "I wonder how many runs our softball team scored this season" might interest you, but probably only if you already know some other numbers to compare it to, like how many they scored in a previous season, or how many a different team scored, or what the average number of season runs is. Otherwise, it's just a number.

On the other hand, "Was our softball team unusually strong this season?", or "Whose softball team was stronger this season, ours or our rival's?", or "Did our softball team score above the league average this season?" would all be plausible choices, all potentially addressed with hard numbers about the number of runs scored by the softball team(s).

In this booklet, we'll refer to these as *research questions*, that is, the core, underlying questions that you care about, and that you're using a QR process to answer.

One common format for a good research question is a yes/no question, often comparing two things and asking whether one is better or stronger or easier or more *something* than the other one:

> Is a V-T&T plan cheaper in the long run than a US-Mobile plan?
>
> Is chocolate ice cream more popular than vanilla?
>
> ...

Research questions of this type can also be equivalently stated in their hypothesis form as a true/false statement ("Chocolate ice cream is more popular than vanilla") instead of as a yes/no question; this is especially common when using statistical approaches.

Another good type of research question is to pose a comparison between two or three or a closed set of options:

> Of Wisconsin, Michigan, and Pennsylvania, which has the highest proportion of rural counties?
>
> Which North American country has the best-paid teachers?
>
> Which sport gets more viewers, diving or gymnastics?

The "closed set" doesn't actually have to be that small ("Which US state has the highest proportion of rural counties?") to be a valid research question, but the more options there are, the more numbers need to be reported. To keep it short, this booklet will stick to questions with 2–3 options, but the principles are not limited.

What all these have in common is that the answer to the research questions, as posed, would be categorical—yes or no, this thing or that thing—rather than a number. Less obviously, they also make clear from the start what the possible answers could be. One danger when devising a research question is that you might only identify an area of interest:

> I'd like to look at numbers about deforestation *topic statement*

or ask a question but state it too broadly:

> How does deforestation affect erosion? *question, but very vague*

That's a valid start, but as it stands it will be hard to formulate quantitatively, and hard to know what an answer would even look like. When you've gotten that far, try to further firm up the research question by identifying categories to compare:

> Do areas with high rates of deforestation have worse erosion than areas with low rates of deforestation? *much better*

This version of the question identifies the specific categories being compared and what scale ("worse erosion") they're being compared on. It doesn't indicate exactly what criterion is being used to evaluate that scale; that could be mentioned in the research question:

> Do areas with high rates of deforestation have higher volumes of soil lost to erosion than areas with low rates of deforestation? *specific criterion, also good*

but it's also ok to leave some of that detail for the quantitative formulation.

Don't feel like you have to get all the way there immediately. As shown above, devising a good research question can be a matter of starting generally and refining it from there.

Step 1: Formulate a numbers question

If we have devised (or been given) an interesting research question or hypothesis, the first step of the quantitative reasoning process is to formulate a question, or a series of related questions, whose answers will be the numbers we need in order to answer the question. Within this document, we'll use the term "numbers question" to refer to these, to distinguish them from the core research question or hypothesis.

Let's say I was interested in the question "which of the ballroom dances is most popular?" and had refined that to

> Which ballroom dance is more popular, waltz or swing?

(There are a total of nineteen or so ballroom dances, but focusing on just two makes this step a little easier to illustrate, for now.) This is a research question that can clearly be addressed quantitatively, so our first step is to formulate the relevant numbers questions. This involves identifying

a. some actual quantities that you can measure or compare

b. the specifics of those measured quantities

c. the data you need to let you compute those quantities

Possibly not in that order: having a particular data set sometimes drives the particular quantities that will be used. For instance, if I had a data set of registrations for one year of East Coast college ballroom competitions, that might lead me to formulate the numbers questions as

> Over the whole 2017-18 college ballroom team season on the East Coast, how many couples registered for waltz?

> Over the whole 2017-18 college ballroom team season on the East Coast, how many couples registered for swing?

Here we have two specific numbers (whose answers will be compared later in the process) that share most of their details, and we'll typically abbreviate these kinds of groups of numbers questions like this:

> Over the whole 2017-18 college ballroom team season on the East Coast, how many couples registered for (waltz/swing)?

for convenience.

A good numbers question should be:

- specific,
- justified, and
- answered with a single number.

To be *specific*, your question should clearly indicate the parameters of the data set you'll be using. This often includes time limits (e.g. 2017-18) and geographic scope (e.g. East Coast), and sometimes also demographics (e.g. college-age). An early draft might omit details:

> How many couples register for waltz in ballroom competitions? *not specific enough*

but this is still too open-ended. What kind of competitions? Over all of time and space? Even if you'd like to use a question like that, you probably don't have the data to answer it!

To be *justified*, numbers questions (with their constraints and parameters) should be linked to their research question well enough that you can expect to draw conclusions from them. Although the numbers questions above are restricted to certain regions and years (and to college students), we can expect that their answers will at least help us understand the research question. On the other hand, questions like

> Within the Ballroom Plus catalogue, what is the average track length for (waltz/swing) songs? *not very justified*

are sort of distantly related but wouldn't clearly supply an answer—why should we think that longer (or shorter?) track lengths indicate greater popularity? (We'll talk more about this below, under "Limitations".)

Finally, the answer should be *a number*, which might seem obvious (for "numbers questions") but bears repeating. If you get wrapped up in the other criteria, you might forget this last piece:

> Over the whole 2017-18 college ballroom team season on the East Coast, which ballroom dance was more popular, waltz or swing? *not a number!*

which is *almost* a numbers question but is still hiding, slightly, what actual numbers will be computed.

Which brings us back to the suggested numbers question for this data set:

> Over the whole 2017-18 college ballroom team season on the East Coast, how many couples registered for (waltz/swing)? *specific, justified, numbers*

None of this is to say that there is only one possible formulation from a particular research question. Much is driven by what kind of data you can find: with a different data set available, something like

> Among songs played by members of the Ballroom Dance DJ Association between 2016 and 2019, how many were associated with (waltz/swing)?

could be perfectly valid.

At the point in the process where you're trying to devise an interesting question, and the overall topic is one you know a lot about, you might end up with a question that seems interesting to you but is already numeric, like, "how many people signed up for each kind of ballroom dance?" This is sort of splitting the difference, but not in a satisfying way: it's number-oriented but not specific about the data source and details, and also, it's not that interesting, except from the implicit comparison you would be drawing among those numbers. If you have a question like this (that seems interesting to you) but the answer to the question would be one or two numbers or a list of numbers, try to work out what implicit comparisons you're making, and then, make them explicit. Most "interesting" questions, in this sense, will have an answer that is either yes/no, or the name of one of the data points, rather than a number.

As a rule of thumb, good numbers questions will do whichever of these are relevant to the problem:

- identify the specific population or population frame you're drawing from
- identify any assumptions you need to make about your data or populations
- identify any relevant thresholds

If all the specifics make it unwieldy, note that it doesn't need to be a single sentence!

Reasonableness checks in step 1

Identifying interesting research questions is most of the first step of the QR process (and narrative), but overlaid on that step is the first of our "reasonableness" checks: acknolwedging limitations, and setting expectations.

Limitations in the numbers questions

For virtually any numbers question, it's important to admit when it is an imperfect measure. It's ok if your quantitative formulation is not a perfect match for the interesting question you care about! You want it to be close enough to be persuasive, obviously, but part of the deal here is acknowledging weaknesses in your process and the extent to which your formulation is driven by the data you have access to even if it's not a perfect fit for your original question. (In a scientific reasoning context, these sorts of thing are often called "threats to validity".)

To continue the running example about ballroom dance popularity: for the research question

> Which ballroom dance is more popular, waltz or swing? *research question*

we proposed the formulation

> Over the whole 2017-18 college ballroom team season on the East Coast, how many couples registered for (waltz/swing)? *numbers question*

The answer to that question could be persuasive in answering the original research question, but there are caveats. College dancers may not be fully representative of the larger population. The "most popular" might have changed since 2018 (although this limitation would be a larger problem if the data were from 2008, or 1998). And it's possible that signups for particular dances reflect something more like available training, rather than popularity as such. Of these three limitations, the second and third are worth mentioning (though minor); the first is a bigger issue but not so bad that it's a dealbreaker. So we mention it as a limitation, and continue with the project.

Why do this? Recall that the motivation for the formal process is rhetorical, to persuade someone that your eventual conclusion is correct. If someone hears your formulation and immediately spots a limitation that you haven't thought of, it reduces their trust in your reasoning, and thus in their conclusion. By contrast, if you have already seen and noted all their objections, you seem like a thoughtful expert on the topic and they'll be more willing to hear our your argument.

Expectations

In addition to acknowledging limitations, it's good to go into this with some guesseseven if they'll turn out to be wrong!as to what the answers will approximately be. For instance (to continue the example), we've already thought a bit about our expectations when we chose waltz and swing for our two dances to consider—more difficult dances like viennese waltz and paso doble are likely to have substantially lower numbers. But it's a good idea to take this a little further, look at the data set, and assess what kinds of registration numbers to expect. Tens? Hundreds? Thousands? What kind of output, roughly, will we get when we run the numbers? Going through this now, early, will prepare the way for later steps when we evaluate how reasonable the results look.

What if you start with the numbers question?

In the introduction, we mentioned that problems solving using quantitative reasoning is sometimes a little messy. One way that's true is that sometimes our initial observation or inspiration comes from already having a data set, and thinking about the numbers we can pull from it, and *then* distilling a research question we can answer. How does the QR framework handle that?

Although we'll always aim to *narrate* the process by introducing the research question first, and then formulating a suite of numbers questions from it, there's no real contradiction in actually devising the numbers questions first. Or, starting with a research question, formulating a numbers question, refining the numbers question in light of the available data sets, and then going back and adjusting the research question to be answerable from the data.

For instance: the CORGIS data project has a data set about various cars and their fuel performance. If it catches our interest, we might start from the data set and see what we can learn from it. This particular one includes the following columns of data (among others):

- Make (the manufacturer, like say "Audi")
- Model year (e.g. 2009)
- Classification (manual vs automatic)
- Fuel type (gasoline, diesel, electric, or a couple other options)
- Highway mileage (in miles per gallon)
- City mileage (in miles per gallon)
- Power (in horsepower)

These can give rise to a variety of numbers questions, both in terms of categories to compare (different makes of car; manual vs automatic; different fuel types) and in terms of data to count or measurements to average or compare. Assuming we are interested in cars, we might start with something like

> What's the average gas mileage a car gets? *"who cares?"*

but the answer to that question is unlikely to be of interest to someone else, because they lack a sense of anything to compare it to and in the end it's just a number.

One easy way to work towards something interesting is to compare two categories within the data. For instance:

> What's the average gas mileage for cars with (manual/automatic) transmissions? *better...*

because while the numbers themselves might not be specifically interesting, they set us up for a comparison that might be.

From there, we can observe that there are many rows of the data set that are not quite relevant (because of other fuel types), and we can note the parameters of the data set and drop those details in, to make it a well-formulated numbers question:

> Of cars sold on the North American market from 2009–2012, what's the average highway mileage (MPG) for gasoline-based cars with (manual/automatic) transmissions? *pretty good NQ*

But if we *start* with the numbers questions, we do still need to work backwards to the research question. We *could* go with something hyperspecific, like

> Did gasoline-based North American cars with manual transmissions get better highway mileage than those with automatic transmissions, in 2009–2012? *ok, but kind of extra*

but with so many specifics baked directly into the research question, it might limit its interest to the general population. On the other hand, something like

> Do stick-shift cars get better mileage than automatics? *pretty good RQ*

is probably more interesting, and although there are some limitations to acknowledge, we would be well-justified in taking the answers to the numbers questions and drawing conclusions from them to answer this research question.

But remember, the *narrative* still starts from the research question!

Step 2: Solve for numbers

This section will be the thinnest in the booklet, because the techniques here are most specific to the particular subdiscipline being studied; the solution process for a programmer will be different from that for a statistician, and both will be different from someone performing an exponential growth analysis.

However, one aspect common to all of them is that the *process* to follow is not dependent on the specific numbers in the data. Once the numbers questions are clear, and the exact solving process is being selected, that same solving process could be followed with a different data set to derive a different (but analogous) number. Perhaps the data set in hand is one for the year 2013; we can answer our numbers questions relative to that data set, but if the corresponding data set from 2014 turned up, we should be able to follow the exact same process for that data set too.

It's certainly helpful, when learning a new process or algorithm or technique, to be able to work with specific, concrete numbers. But try to keep in the back of your mind: would this work if I tried it with different data?

Solving with a program

The overall quantitative reasoning process doesn't have a ton to say about the details of writing a program, but you have a number of decisions to make in building that program. For instance, you should ask: "do I need an accumulator, and what should it accumulate?", and "what kind of loop is the best way to loop through the data?", and "should I read it all into a list and *then* process it, or what?", and "are there any builtin Python functions that can save me some effort on this?". When you talk about your work on this step of the process, these are the sorts of design questions you can address.

Another set of design choices you can make revolve around how you're handling the suite of numbers questions. If you're following the patterns laid out in earlier sections, you probably have a matched set of numbers questions, with a one-word variation between them: waltz vs swing, for instance, or manual vs automatic. If so, you could decide to write a program that, in a single run, will produce both numbers; but it's also a reasonable choice, and possibly easier, to build a program that runs once for each version of the question, with a minor edit to the program in-between, or with user input to decide which version to answer.

However you design your program, though, remember that at the end of this step, after the dust settles, you'll have some numbers. These aren't your final answers yet! We will still need to interpret those numbers to answer the original question.

Testing: are you sure you did it right?

A big part of the "reasonableness" checking comes during the solution step. Someone reading or listening to your description critically will certainly wonder—did you get the math right? Did you get all the bugs out? Are your final numbers accurate? Indeed, *you yourself* should wonder those things. How can you convince yourself and others that your numbers are right?

Some kinds of obviously *unreasonable* answers are easy to notice. If you are counting how many examples match criteria, and your answer is a negative number, that's clearly wrong! If you are calculating any sort of proportion and the number suggests that more than 100% of the examples have some property, that too is clearly wrong. But not all cases are so clear-cut.

One approach to judging whether your output is reasonable—and thus plausibly correct—is to compare it with your expected results. If your expectation was that your numbers would be in the hundreds, and you get answer in the thousands, it *might* be right but is probably worth double-checking. If you expect hundreds and get millions, you should probably be very very skeptical of the result, and check your work extremely carefully.

Another approach is to persuade yourself and your audience that the process itself is correct, using hand-built *test cases*. Perhaps your ntended data set has thousands or millions of data points, and it's great to be able to use software to process it all, quickly. But if you can build a data set, artificial but realistic, that is in the same basic format but only has five or ten lines of data, you could reasonably process that data by hand, to know the *exact* correct answer for that data.

A *test case*, then, is a realistic input or data set paired with its exact expected numeric result—and preferably, one that is very easy to calculate!

When you have built a spreadsheet or program or software configuration that can process the simple data exactly correctly, passing the test case, you can be substantially more confident that it will work correctly on the larger, "real" data set as well.

Step 3: Interpret the numbers

With a lot of quantitative reasoning problems, it can be easy to feel "done" as soon as the numbers come out, at the end of Step 2. But while the numbers are the answers to the numbers questions, they're not the answer to the original research question. The final step of the process is to interpret those numbers to draw a conclusion and answer the interesting question.

To actually close the loop here, you need to take the numbers that your program prints out and interpret them in light of the thing you actually cared about in the first place. Continuing the running example: step 2 told me what the numbers were. If waltz had 2,243 registrations, and swing had 1,982, we can conclude that waltz is the most popular ballroom dance—at least within the context of the data I had and the particular way that I chose to quantify the question.

Basically, just don't forget that while the numbers are important for justifying your answer, the actual answer is the thing that anyone *else* is likely to care about.

Revisit limitations: confirm that results are reasonable

Coming back to some of the thinking that you did in Step 1b, the limitations of your data or method may help you qualify your interpretation and your answer. For instance, instead of "waltz is the most popular ballroom dance ever", you might write something like "at least among college team dancers, waltz draws the most registrations and seems to be the most popular of the competitive dances".

The other thing to do with your thinking from Step 1b is to see how well your predictions played out; if any of them were wildly off, you should consider that a bit of a yellow flag and investigate why you were wrong. *If* you were wrong. If you were wrong, that's bound to be interesting (since it means your answer is unexpected and therefore of particular interest); but investigating may reveal that you were right in the first place, and that your program was giving incorrect results. In which case you need to fix them and run it again!