

Given the following options detailing modifications to AutoTest, which of the following would you...;; Describe why you chose the option you did (r0\_AutoTest)

More feedback on what grade my submission received;;; I would like the auto grader to show more details because often it goes from straight developing to proficient. (r1\_AutoTest)

More detailed feedback;;; The AutoTest just showed a generic error and a suggestion on what to work on next but the feedback felt more from the point of view of the instructors than the students and it did not give any hints whatsoever about what was wrong in the code. Particularly in C3, we were stuck on a an AutoTest error for a long time and couldn't figure out what it meant even though things were working locally and it cost me a significant portion of my grade. (r2\_AutoTest)

More detailed feedback;;; it was very very annoying not to know what was failing. there were times that i could not identify at all what was wrong unless i could get in touch with a TA, and even they were unable to point what exactly was failing. it seemed to vague, which i understand is meant to push us, but the goal of testing should be to help identity where there are bugs, not that there are just a lot of bugs. (r3\_AutoTest)

More detailed feedback;;; For C1, I was unable to get to proficient and I kept adding tests and felt like I had so many but without additional feedback I didn't know where more tests needed to be added. For the last checkpoint autotest said I had passed all smoke tests but still wasn't at proficient and told me to work on get-datasets. The error ended up being in the addrooms portion of addDataset and so that feedback did not really help and I had talked to every TA, none of them knew what the issue was, and so it ended up taking me a full week to find the bug. (r5\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; it would've let me iterate on my work and improve it in real-time without waiting around. (r10\_AutoTest)

More feedback on how the grade was computed;;; We had a lot of difficulty figuring out why we were still in Developing for the last checkpoint even though we passed all of the smoke tests. Having more feedback about which test cases were passing would have helped us a ton, since the autobot just gives you a generic response to work on a certain function, and in our case that function wasn't really the issue but rather was showing issues from other functions. Having some sort of insight eg. a couple test cases (short descriptions of test cases (setup and test)) would be SUPER helpful and a lot more realistic of real life I think. In our case even the TAs and profs weren't able to figure out the issue which I don't think is fair. (r6\_AutoTest)

More feedback on what grade my submission received;;; It helps to gauge my progress better, as the grading bucket being arbitrary means that I have no idea how much work I need to do to reach the next tier. For example, if I'm in Developing, I need more information on what I'm missing to reach Proficient. (r7\_AutoTest)

More detailed feedback;;; For every checkpoint we struggled to pin point the errors we were encountering. Especially for C3, on the autobot it kept saying we needed to work on GET. From that my group automatically assumed that our call to listDataset was wrong, so we spent days looking through it. At the end it was because we casted the buffer into a buffer (redundant), that was causing the error. (r8\_AutoTest)

No changes;;; i don't want it as a crutch because that's not how software eng. works in real life (r9\_AutoTest)

More feedback on how the grade was computed;;; I feel as though that I had 0 clue at first what the smoke tests could possibly even be checking. If it gave an example dataset then it would be far more reasonable, however this would maybe make it too easy. (r11\_AutoTest)

More detailed feedback;;; I thought autobot was fine but sometimes I don't understand what the feedback means (eg. i got one that said work on valid edge cases, which could be more specific) (r12\_AutoTest)

More feedback on how the grade was computed;;; I believe it was c2 that had the most problems with AutoTest where the grading buckets (the checkmarks) were not actually correlated with what we needed to fix to get the highest bucket, so more feedback on what steps are needed would have been helpful there. (r13\_AutoTest)

No changes;;; my partner and I did our best to work incrementally and start early. We never wished or felt we needed to call the bot more than 4 times a day. We were confident in our tests too so once we saw we passed the smoke tests and were in the proficient bucket + all our tests passed - we were happy with the possible grade we would receive (r14\_AutoTest)

More detailed feedback;;; I wish I could choose all of them. But I would say more detailed feedback because it would not only help me know where to focus and improve my code, but in turn that would probably result in an increased grade and such, so the positive outcome from more detailed feedback I think would trickle down into the rest of the categories. (r15\_AutoTest)

More detailed feedback;;; It would have been nice to see which tests exactly were failing so we had a more clear understanding of what needs to be fixed. (r16\_AutoTest)

More detailed feedback;;; Hard to find bugs (r17\_AutoTest)

More detailed feedback;;; I found that the frequency of checking my grade was fine (apart from C0 which was a bit stressful). A lot of my issues with the autobot stemmed from the lack of description. For instance, in C1, all of our query tests were failing and all of the other dataset tests were passing. My partner and I spent over a day trying to find the root of the problem and it ended up being unrelated to the query implementation (it was an unresolved promise in the addDataset function). It would've been nice to know that the autobot tests failed due to a timeout or unresolved promise as that would've given us an idea of what to look for in our code. (r18\_AutoTest)

More feedback on what grade my submission received;;; Was good, sometimes frustrating with limited calls especially if you didn't know if your test suite was complete (r20\_AutoTest)

More detailed feedback;;; would have been nice to have a bit more context on why a given test case was failing (r21\_AutoTest)

More detailed feedback;;; Sometimes need to dig quite deep to check for a missing edge case. (r22\_AutoTest)

More detailed feedback;;; I don't need more frequent feedback, I would like more detailed feedback to give me more to work with. (r23\_AutoTest)

More detailed feedback;;; Honestly, our test suites was so thorough that we almost never got feedback beyond "proficient" but the one time we did, it was a bit vague about the direction. The fact that so many smoke test tests were interrelated meant that we were not sure if we had an issue in the smoke test error location or if it was because a different error meant that that test couldn't run. (r24\_AutoTest)

More detailed feedback;;; Some of the feedback was not clear. Something changing something not at all related to the feedback was what improved our score. Feedback should be more accurate and detailed. (r25\_AutoTest)

More detailed feedback;;; There were times where we weren't sure what specifically in the bucket we were failing, and additional feedback would have helped with more targeted solutions. (r26\_AutoTest)

More detailed feedback;;; Sometimes it was difficult to determine why specific tests were failing. Specifically for c2, our tests for MAX/MIN were failing and it was difficult to determine why. (r27\_AutoTest)

More detailed feedback;;; Most of the feedback was good, but sometimes it was an unhelpful failure notice. (r28\_AutoTest)

No changes;;; There is no need to change. (r29\_AutoTest)

More detailed feedback;;; The autograder can feel like a black box. For example on C2, we were failing the "MIN" or "MAX" test, but it wasn't because the "MIN" or "MAX" function was wrong, it was because of completely unrelated but larger issues in the implementation. (r30\_AutoTest)

More detailed feedback;;; It was terrible feedback that the autobot did. I wish it could give more details on the tests that were failing. Like in C2, it would mention work on Min/Max even though we can see with the table of checkmarks that it is X'ed out so it was giving useless feedback that we could see. Also for c3, it would show us 7/7 but then not give us proficient and say work on Get datasets. That gave my group and the TAs not enough info on where to go on. (r31\_AutoTest)

More detailed feedback;;; I still have no idea what happened in C2 for the MIN/MAX. Even if more feedback isn't given during the checkpoint, it should be given after so at least our code can be correct for future checkpoints. (r32\_AutoTest)

More detailed feedback;;; The feedback was not very helpful in determining where the issue was coming from, for example we ran into timeout issues in C2 with not very helpful feedback. (r33\_AutoTest)

More detailed feedback;;; Sometime the autograder tells you a certain part of your code is not functioning correctly, and that the said function did not passed the test. But later we realized the real problem lies in a different function which we have to fix first. So it would be helpful if the autograder can tell us what that specific test fail, and how it fails. (r41\_AutoTest)

More detailed feedback;;; Just asking us to work on one part (for example addDataset) was not that helpful feedback. (r37\_AutoTest)

No changes;;; I don't think there should be more detail because in the real world, no one is there with a perfect set of test cases. YOU, have to make them yourselves. I'm actually impressed with some of the edge cases I missed and how TAs and instructors accounted for those. But what would be the point if autotest told you exactly what was wrong and then you just fixed it. Because when you go out and build projects without having developed this critical thinking ability of what type of bugs might originate, you might think you have something perfect but then when you send it out to production, it will be riddled with new errors and you will have no idea why. (r34\_AutoTest)

More detailed feedback;;; A lot of the times the autograder just said that one test was failing, but there was no way for us to know why it was failing (it turned out to be an issue completely unrelated to that one test case), and we had to go to office hours for the TAs to tell us what error messages our code was throwing in order to figure it out. (r35\_AutoTest)

More feedback on what grade my submission received;;; Sometimes it was hard to figure out why something was failing (r36\_AutoTest)

More feedback on what grade my submission received;;; Sometimes I wouldn't know where to fix places like maybe addDataset passed but if it was not entirely correct it could effect perform query tests (r38\_AutoTest)

More detailed feedback;;; Like I said feedbacks are too vague. (r39\_AutoTest)

More detailed feedback;;; I value detailed feedback as it provides valuable insights into my implementations and helps me identify areas for improvement. Specifically, I am keen to receive more information regarding any logged errors in the code. (r40\_AutoTest)

More detailed feedback;;; more detailed feedback help me debug, track what did i do wrong (r52\_AutoTest)

More detailed feedback;;; I found that my partner and I were not reliant on the autobot for checking the functionality of our code, but when our code is working and passing our local tests, it was frustrating to not know the issue, especially when at the end of the day, our goal was to get the grades we wanted. (r42\_AutoTest)

More detailed feedback;;; Because sometimes letting me know hey this thing fail because I try a query and it did not worked really does not help to think about anything specific I can work on. (r43\_AutoTest)

More detailed feedback;;; A lot of problems pointed out by autotest were super opaque. A lot of the time, the feedback was work on this overall topic but we didn't know why it was wrong. (r44\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; Sometimes I want to know if the small change i did fixed it but i felt held back (r45\_AutoTest)

More detailed feedback;;; The bucket grading autobot was kind of useless. It only gave feedback like: We recommend you work on GET-datasets, but if students knew if something was wrong with listDatasets then they would have fixed it already. The autobot feedback was very vague and gave no clue on where to begin debugging. (r46\_AutoTest)

More detailed feedback;;; The messages were not able to help us or the TA helping us to help triangulate where a bug was occurring. (r48\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; Less stress of only having 2 calls to the autobot each day. (r49\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; More checking is always useful to fix silly mistakes. Makes it easier to get a good project grade (r50\_AutoTest)

More feedback on how the grade was computed;;; I think that giving a hint to one possible test case that is failing could be a good idea. It will lead to more time spent coding and possibly lead the students to discover more cases that they have not thought of. (r51\_AutoTest)

More detailed feedback;;; If a bucket was missing a point, there could be many reasons why. With more feedback, I could figure out where the problem was more quickly. (r53\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; PrairieLearn and other automated testing tools used at UBC frequently have a faster cooldown. I was able to debug most of my issues locally or figure out how to test the issues, so I wouldn't have needed more detailed feedback, but figuring out how to 'time' #check usages felt very irrelevant. (r54\_AutoTest)

More detailed feedback;;; I wish we could see exactly how many tests were failing and in what categories. Knowing if a change actually made a difference i.e. 14 tests failing vs 13 would be really great to know. Ideally, the detail in feedback could also include actual and expected results for some tests, maybe a select few. (r55\_AutoTest)

More detailed feedback;;; I remember specifically one instance where autotest determine that my implementation of min and max were failing but it did not say anything more than that, so we were left with very little clue as to what was going wrong, because locally the implementation all of our test. (r56\_AutoTest)

More detailed feedback;;; For some tests that are dependant on each other, it was hard to tell if the error it was reporting was actually caused by a certain area of code or if it was on some other dependant section, and since autotest had no further feedback, it was up to independent debugging to exhaust every possible place the error may have occurred. (r58\_AutoTest)

More detailed feedback;;; Hints toward what path forward to take. For example, if the code failed on one method, what specific tests actually failed. (r60\_AutoTest)

More feedback on how the grade was computed;;; Often times I had several of the smoke tests passing but was still in a lower than expected bucket. This being said I would have liked to have known where the issues were coming from. (r61\_AutoTest)

More detailed feedback;;; It was oftentimes very hard to pinpoint where our issue was, especially when it timed out. (r62\_AutoTest)

More detailed feedback;;; i would have appreciated more detailed feedback as to what specific smoke tests represented. i remember being confused as to what exactly query ebnf means, and what partitions i was not meeting. (r63\_AutoTest)

No changes;;; My view would likely not be popular, but I don't believe AutoTest should be overly helpful or more frequent than it currently is. I believe the purpose of AutoTest should be to let groups know when they have done enough work to stop and to give some general pointers on where to focus if the submission is deficient. Making AutoTest too helpful might disincentive groups from writing their own tests which would be a shame. In the real world there isn't really some higher order test written by higher ups to validate the work being done. Things largely sink or swim on the strength of the tests written by the implementors, not tests conveniently given to them from above. (r64\_AutoTest)

More feedback on how the grade was computed;;; Sometimes what was causing a problem was extremely unclear. Even the TAs at times were unable to help improve our score due to a lack of specific information from the grader. (r66\_AutoTest)

More detailed feedback;;; A lot of the times we didn't understand why we are not passing the autotest fully and the TAs couldn't help us (r67\_AutoTest)

More detailed feedback;;; more detailed feedback, as to exactly which tests we failed would be very useful. (r68\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; Lots of feedback often is a good way to keep track of the project progress. (r69\_AutoTest)

More detailed feedback;;; Some tests are vague and we don't know what is failing exactly (expected/actual values). At times, we were passing all local tests but failing the bot tests. This especially applies when the returns value of certain functions weren't as expected by the bot, ie. bot expects int, our tests printed the correct number as a String, and the bot test fails ambiguously. (r70\_AutoTest)

More detailed feedback;;; Sometimes, it was difficult to know what was going on when a test was failing - for instance, our group got stuck for a while trying to figure out that our implementation was converting a numerical field to a string, for some strange reason. (r71\_AutoTest)

More detailed feedback;;; more detailed feedback would be useful, as many times we had to go to OH to check why our code failed. The limited feedback made it challenging for us to figure out what to debug and fix. (r72\_AutoTest)

More detailed feedback;;; Sometimes, the autotest would give feedback such as "you should look into working on X" but the problem actually occurs in another section of the code and only appears on X. more detail could help identify these kinds of cases. (r74\_AutoTest)

More detailed feedback;;; During C2 we were not able to get the MAX/MIN buckets but we figured out during office hours that it was because of issues with addDataset. Having more detailed feedback would have helped us figure out the problem and debug more efficiently. (r75\_AutoTest)

More feedback on how the grade was computed;;; I think having guidance over the bucket style of grading would have helped make more informed decisions of where to spend our time. (r76\_AutoTest)

More feedback on what grade my submission received;;; sometimes, in the beginning phases, it was very discouraging to not know our grade and how far along we were, especially if there were other persisting issues. (r77\_AutoTest)

More detailed feedback;;; Sometimes when the tests are failing it can be hard to tell what caused it to fail. One time the test failed because of some other defect in the system that was not directly related to the part under testing, and that took us a long time to fix without any detail from AutoTest. (r78\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; Limiting ability to check grade felt like I had to space out the times when I worked on the project, which didn't feel great especially as a student taking a lot of courses. (r79\_AutoTest)

More detailed feedback;;; More detailed feedback for why my code was causing issues and what was limiting it from fulfilling the next bucket. (r80\_AutoTest)

More feedback on what grade my submission received;;; I think buckets were nice since they grouped my progress, but sometimes I had a minor issue that I might have had to spend a lot of time to debug before I was able to move on, which would have already fully met the test requirements. (r81\_AutoTest)

More detailed feedback;;; Some of AutoTest's feedback was misleading. For example, my groupmate wrote code that didn't correctly store the id for rooms datasets (separated using '-' instead of '\_') which resulted in AutoTest saying to "work on geolocation" despite geolocation working fine. I understand that it is extremely difficult to test code independently, so this may not be something that can be fixed. (r82\_AutoTest)

More detailed feedback;;; I had some very irritating errors where I could not understand where I was going wrong. With more specific feedback I would be able to solve the issues myself instead of going to the TA. (r88\_AutoTest)

More detailed feedback;;; In a few small cases, the specification did not clearly indicate how to handle some edge cases. It would have been very helpful to get back information about what kind of test was failing, or what the result vs. expected result of the test was, to help clarify those cases. In the real world, I imagine I would make those judgements myself, or work with my team to define those. (r83\_AutoTest)

More detailed feedback;;; I chose that option because there were many cases where we got feedback from the bot and had no idea what it was talking about and it's feed back wasn't very actionable. so I would like more detailed feedback like the post checkpoint feedback or maybe test outputs. something more detailed and actionable. (r84\_AutoTest)

More detailed feedback;;; So we know what kind of errors are being thrown at us (r85\_AutoTest)

More detailed feedback;;; A lot of the time we didn't know what aspect of a bucket was incorrect. Sometimes, the buckets were misleading as well as they could depend on other features that were broken. For example, one time the bucket indicated that there was an issue with our wildcard implementation but our fix for it did not involve fixing any of our wildcard code. (r87\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; It felt like it took a lot to figure out the optimal times to choose to grade and sometimes it felt like my thinking and reasoning for writing code or writing certain portions of the code revolved more around thinking about how it was going to get graded than the actual functionality or effects on the project as a whole because the grading amounts felt limited and sacred. (r89\_AutoTest)

More feedback on how the grade was computed;;; We need more details on what kind of tests are failing. Or the error messages (r90\_AutoTest)

More detailed feedback;;; When some smoke tests failed it was sometimes unclear what exactly is wrong (r102\_AutoTest)

More feedback on how the grade was computed;;; it was tough spending many hours discovering and fixing bugs only to see the same message on autograde. It would be nice to have a more detailed feedback on issues that I was able to resolve. (r91\_AutoTest)

No changes;;; I felt busy at times but the workload was reasonable, and I think I have a relatively busy schedule. If autotest were available more often, I could've just cheated it. and I found the amount of feedback quite helpful. (r92\_AutoTest)

More feedback on what grade my submission received;;; Sometimes, we had the prompt giving us a very basic response when it could have been more involved. (r93\_AutoTest)

More feedback on how the grade was computed;;; I wasted a lot of time trying to fix something not worth a lot because I would affect the other portions of the grading. (r94\_AutoTest)

More feedback on how the grade was computed;;; In some cases, we would pass the required number of tests (e.g. 2/2, 3/3 etc) tests for each function but still have Developing or Acquiring grade. Then we would be given advice to work on a particular function, but often this wasn't actually the function that had problems.\\It would be nice to see precisely how many tests for each function were passing, as well as some particular test scenarios, to be able to hone in on the problem. (r95\_AutoTest)

More detailed feedback;;; I think it is probably for the #cx; it's sometimes frustrating after writing a lot of tests not to understand why it is still not working. But still, I guess it's part of the learning and is the goal of the project, so while it's frustrating, it makes sense. (r96\_AutoTest)

More detailed feedback;;; It was frustrating getting a test failing but not knowing why exactly, especially with c0. Even a little bit more of a hint such as the name of one mutant that failed would have helped a lot. (r97\_AutoTest)

No changes;;; I found the feedback to be sufficient. (r103\_AutoTest)

More detailed feedback;;; Trying to mimic real software development just doesn't work, this is a class and I should get feedback on the tests I'm failing and why. (r98\_AutoTest)

More detailed feedback;;; Sometimes, telling us that we should work on X can be helpful, but I often didn't get it. Maybe more feedback (more like the post-checkpoint feedback from the autograder) would be better. (r99\_AutoTest)

More detailed feedback;;; When most tests say 1/1, but I'm still getting 80% or less, I get discouraged because I have no idea where my bugs could be because my local tests are passing. So I tend to give up. (r100\_AutoTest)

More detailed feedback;;; Sometimes we got less point, but in Proficient. And another time we got full marks but Developing....\\In c2, our SUM is incorrect, but we have no idea how cause rest of the keywords passed (AVG, COUNT, MIN, MAX) (r101\_AutoTest)

More detailed feedback;;; Helps a lot when debugging (r104\_AutoTest)

More feedback on how the grade was computed;;; Failing smoke tests with no more details are not helpful, and also not an accurate reflection of how test-driven development works in reality. Being able to work backwards from the description of a failing test to a solution for the bug is a much more useful skill than blind guesswork. (r105\_AutoTest)

More detailed feedback;;; Sometimes just saying something like "We recommend you work on AddDataset" wasn't very helpful, since that encompasses quite a lot of functionality and it was hard to know exactly what was going wrong. The bots that provided feedback after the deadlines were much more helpful and allowed us to fix small issues that we couldn't identify earlier, before they turned into a bigger problem. (r106\_AutoTest)

More detailed feedback;;; More details in the feedback is going to be more helpful for pinpointing our issues and pointing us in the right direction. (r107\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; I think there were some instances where I had slight issues in the code but since I did not have any grading attempts left I had to wait for the next day to do it again to make sure that issue was fixed first instead of focusing on other parts of the project. (r108\_AutoTest)

More detailed feedback;;; Specifically for timeouts. The biggest issue for my group during the project was during c2 was timeouts and the autotest would take 5-10 minutes only to give zero feedback even when things were running quickly locally. It would be nice if it were possible to receive more feedback upon getting a timeout on the smoke tests, as this feedback was only given occasionally when using check. This feedback should show up more consistently and should definitely appear when calling autotest for smoke test results, especially since it normally tells you which part to focus on normally (it gives no feedback on what to work on for timeouts) (r109\_AutoTest)

More detailed feedback;;; More detailed feedback. At times, the autotest feedback was actively unhelpful. For example, on C2, many people found their implementations of MAX/MIN for example were not passing. The problem for most however wasn't with their implementation of MAX and MIN, but of the additional tests these suites had. Getting more insight about why tests were failing could be helpful for these students. (r110\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; Having only 2 checks a day is pretty scary and makes you not want to submit your code until you have more things implemented. Also, It would make a lot of sense for there to be better tests for add Dataset specifically since 99% of the adding requires perform query to be working, so you honestly don't even know if its working properly as they expect. Granted though, it is honestly not that hard to test yourself. (r111\_AutoTest)

More detailed feedback;;; Sometimes the feedback isn't really detailed enough. There could be multiple things wrong with your submission but the autobot only tells you one of the things that is wrong (r112\_AutoTest)

More feedback on how the grade was computed;;; Sometimes we misunderstood the grade so I think if we know more about how the grade was computed, it will be better. (r113\_AutoTest)

More detailed feedback;;; It was sometimes vague (r114\_AutoTest)

More feedback on how the grade was computed;;; It often felt like tests were failing for no reason (r115\_AutoTest)

More detailed feedback;;; Sometimes it was hard to pinpoint what's wrong as it's only the presence or absence of a checkmark beside the grade. For instance, my partner and I could not figure out why we weren't getting the caching progress check. (r116\_AutoTest)

More feedback on what grade my submission received;;; I would want to know how my code behaves/runs into error against the bot test. The info is extremely limited and sometimes I just don't get why I get stuck at a level and what to do to advance it (r117\_AutoTest)

More feedback on how the grade was computed;;; Sometimes the autograder was confusing on the feedback it was providing. For example, in one checkpoint we had two clusters with 'X' but the recommendation area to work on was neither of those areas and when we worked on the recommendation area which already had a checkmark, we reached proficient bucket (r118\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; It was stressful having so few autobot checks (r119\_AutoTest)



No changes;;; Most students would've liked more feedback but I think it defeats the purpose of local testing / test-driven approach. The minimal feedback we are seeing right now is more closely aligned to what you might get in the real world. (r120\_AutoTest)

More detailed feedback;;; It was really confusing when there was just an indication of one part of the project being wrong (r121\_AutoTest)

More feedback on how the grade was computed;;; (r122\_AutoTest)

More detailed feedback;;; Would have been helpful to figure out exactly which tests our code was not passing to make it faster to debug (r123\_AutoTest)

More detailed feedback;;; It was sometimes vague with what tests were / weren't passing, and the wording on what exactly was wrong was tricky (r124\_AutoTest)

No changes;;; I think it is okay as it is (r125\_AutoTest)

More detailed feedback;;; More detailed feedback would make it easier to reproduce issues locally. (r126\_AutoTest)

More detailed feedback;;; Failing smoke tests were sometimes confusing and more details would've helped figure out why we were wrong. Also since the autotest is supposed to be a "client", you should be able to get clarification from a client. (r127\_AutoTest)

More detailed feedback;;; Each submission we did we thought would be right so often when there was limited feedback (like the autobot just saying there was a timeout) we didn't know where to start. The checks were much more useful at addressing these issues however, there were still some instances where there was limited information to know where to focus our efforts. (r128\_AutoTest)

More detailed feedback;;; More detailed feedback would help with debugging. (r129\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; 2 tests a day is a lot of added stress, and disincentivizes students from getting a lot of work done in any single day (r130\_AutoTest)

No changes;;; it gave me good feedback that I could work on (r131\_AutoTest)

More detailed feedback;;; Sometimes the autotest feedback seemed vague and confusing, especially when our local tests were all passing. It was sometimes hard to figure out where the issues were in our implementation. (r132\_AutoTest)

More detailed feedback;;; i have no clue what 1/2 for AddDataset means (r133\_AutoTest)

More feedback on how the grade was computed;;; There have been times where the smoke tests through us off and the feedback after the checkpoint would show contradicting results to our submission tests. (r134\_AutoTest)

More feedback on what grade my submission received;;; Just because you get proficient, doesn't mean you get 100%. It makes it hard to know when you should stop working on it for students who want to get a perfect score. (r135\_AutoTest)

No changes;;; I think the Auto-test is good enough, more feedback will make student less care about their own tests. (r136\_AutoTest)

More detailed feedback;;; As with my answer in the previous portion, detailed feedback like "Focus on Invalid Cases in Query" is much more useful than "Query smoke tests failed". The ability to see which of our tests failed is also extremely helpful to help us identify misunderstandings of the specs early on, especially in the cases where the specs are unclear and we need to "guess" what the desired behaviour is. (r137\_AutoTest)

More detailed feedback;;; While the autograder did a good job of describing what part of the code to work on for a better grade, occasionally it could be very difficult to determine what exactly was going wrong (r138\_AutoTest)

More detailed feedback;;; i spend a week, on finding a really trivial test case because i didn't expect that to be tested. extremely waste of time and useless effort. (r139\_AutoTest)

More detailed feedback;;; We felt that the autotest feedback sometimes did not provide correct feedback or feedback that was accurate to where the issue was. (r140\_AutoTest)

More feedback on what grade my submission received;;; (r141\_AutoTest)

More feedback on how the grade was computed;;; efaef (r142\_AutoTest)

More detailed feedback;;; I would like to see the detail about the test that failed after run #c3.

Sometimes all block is  but my grade is still not Proficient. (r143\_AutoTest)

More detailed feedback;;; Knowing what's wrong is more important (r144\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; I would feel more secure knowing that we can check our grades more times because we would often want to “save” the check for later, and we end up not using them enough for the fear of not having any check left when we actually need to check. (r145\_AutoTest)

More detailed feedback;;; A lot of times I had very vague feedback like ValidEdgeCases and I had no way to see if my smoke tests were improving at all since all the checkboxes are ticked but I'm just stuck in the same grade bracket. (r146\_AutoTest)

No changes;;; I feel like the auto test feedback was fair. (r147\_AutoTest)

More detailed feedback;;; no idea what went wrong (r148\_AutoTest)

More detailed feedback;;; Sometimes I spend hours trying to debug off of the hints given, but it sometimes is not what I should focus on. (r149\_AutoTest)

No changes;;; I think the autotest is fair in all aspects. (r150\_AutoTest)

More detailed feedback;;; One thing that was very frustrating, was in checkpoint two, me and my partner received checkmarks for all the sections in the bucket grading, yet we still were developing meaning we were missing tests. It would've been nice to know from where the tests were missing from, as it does not make sense for check marks to be checked on all parts yet not being proficient. (r151\_AutoTest)

More detailed feedback;;; I think if we saw the actual error we got, we could debug the small things much faster. (r152\_AutoTest)

More detailed feedback;;; I felt the number of times you could call the autograder was fair, but there were times when I wasn't sure which parts of my implementation I should be working on. (r153\_AutoTest)

More detailed feedback;;; smoke test feedback is often misleading. often times the bug that causes a category we would get 0/1 has absolutely nothing to do with the category itself. for example if you put the result > 5000 check before transformations, you would get 0 on sum and max which is very misleading and does not point you in the right direction. should tell me to go read the spec for result limits more carefully or something (r154\_AutoTest)

More detailed feedback;;; I wish there was more feedback provided for each test. For example, we would be given a check mark for addDataset on c1 but later on, we found out that it only achieved acquiring. (r155\_AutoTest)

More detailed feedback;;; feedback was often vague, especially the hint/tip part. sometimes it would point in a direction that (at least in my opinion and experience) was not the source of the issue, which led us astray for a few hours before finding the problem. (r156\_AutoTest)

More detailed feedback;;; I feel like feedback was not sufficient at all, as most of the time you get asked to focus on when you are only in acquiring and developing, but sometimes you might have worked on the code for a lot of the time and thought that you are at least in developing finally, and then you get beginning with no feedback to what might be wrong (r157\_AutoTest)

More feedback on how the grade was computed;;; I felt that the auto bot feedback was too vague for me to actually figure out where to start looking. (r158\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; I like being able to change my code in small ways and then resubmitting it (r159\_AutoTest)

More detailed feedback;;; I believe the numeric scores of the auto-test is fine but the additional feedback really doesn't provide much in terms of giving us a direction of where to improve or fix. For instance, if it said to work on 'add datasets' but we can't really tell what's wrong with our implementation and all our tests pass then we'd be stuck for hours not knowing what to do, only for it to end up being a somewhat unrelated issue. (r160\_AutoTest)

More detailed feedback;;; The feedback given by auto test was often vague and opaque. It basically just told you if you needed to keep debugging or if you were done. Those were the only two options: keep going and done. Additionally, what little feedback was given was difficult to interpret. Try working on this component is unhelpful if we don't know in what way it is failing. (r161\_AutoTest)

More detailed feedback;;; Sometimes the feedback was so vague, that I couldn't identify where the problem was, especially if all my tests were passing locally. (r162\_AutoTest)

More detailed feedback;;; I think a more detailed check would be very valuable, sometimes it felt that searching for the edge cases were really challenging because the feedback was not very clear. (r163\_AutoTest)

More detailed feedback;;; Having to reach 80% to get detailed feedback from your project made the implementation harder especially in the earlier stages. If we were able to know what was failing and why it was failing earlier it would have made the coding process easier. I just think that gaining more detailed feedback is better than not getting feedback at all. The frequency was good. (r164\_AutoTest)

More detailed feedback;;; There was a time that the auto test keeps failing, but I have no way to figure out what is causing the problem locally because no matter how I change / add the tests, all the local tests pass. Going to the TA office hour also doesn't help much because they don't have as much insights into how I coded the part. In this case a single issue is costing a huge amount of workload, and if there is a feedback given, this should be a really easy fix. Also in real projects, we should have those debug information, instead of coding like checks & guesses. (r165\_AutoTest)

More detailed feedback;;; Could really go in detail about the failing/missing part (r166\_AutoTest)

More feedback on what grade my submission received;;; Now the AutoTest only gives bucket testing, but I want to see more details and the estimated grade I will receive, which would help me to be more specific in determining my bugs and mistakes. (r167\_AutoTest)

More detailed feedback;;; The c3 could be affected by c1 which the auto test could not show. This made it super hard to get full mark. (r168\_AutoTest)

More detailed feedback;;; I often got stuck at a certain place with no idea what was going on. Knowing the output my code had on failing tests on the auto bot end would have been helpful. (r169\_AutoTest)

More detailed feedback;;; I chose this mainly regarding the feedback about flaky tests, which turned out to be extremely hard to spot if they were not reported before the test suites became large. These feedback messages do not reflect anything about the test suites subject to flaky errors (although the nature of such errors make this hard as well), and could be more informative. (r170\_AutoTest)

More detailed feedback;;; Need more feedback! (r171\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; (r172\_AutoTest)

More detailed feedback;;; (r173\_AutoTest)

More detailed feedback;;; For Q0, my grade section for EBNF was stuck on 4/6 and it just won't improve, despite trying many times. It would be helpful to give us more clues. (r174\_AutoTest)

More detailed feedback;;; Sometimes you don't get full marks for an item but it doesn't say why (r175\_AutoTest)

More detailed feedback;;; (r176\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; It is insane that we can only check it twice. (r177\_AutoTest)

More detailed feedback;;; Thinking of C2 where there was a common issue people were having with MIN/MAX/SUM. I think the error message could have been more descriptive. (r178\_AutoTest)

More feedback on how the grade was computed;;; I think having more feedback on my computed grade can help me better determine the areas of code that I need to focus on to be able to go to the next grading bucket. (r184\_AutoTest)

More detailed feedback;;; I think more feedback would be helpful - sometimes it's very difficult to interpret the current feedback. While we were able to see which general functionality that we needed to work more on, it would have been more helpful to have a note about what generally we may need to work on specific to that function (e.g. "work on edge cases", "work your basic functionality, specifically in parsing the query"). This could help students better focus their limited time to working on important improvements (rather than getting very stuck and relying on TAs to spend lots of time finding the holes in their logic), and then be able to dedicate more of their time to making those improvements quality code. (r179\_AutoTest)

More detailed feedback;;; During C2, we constantly had timeout issues once we started to implement query. However, there was no detailed feedback as to where the server had timeout issue. This meant we had to blindly guess where the issue might have risen, but our attempt at fixing the code was ultimately fruitless. A detailed feedback would have allowed us to easily and quickly check which part of our program that needs fixing. (r180\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; It can help me to fix my problems more efficiently (r181\_AutoTest)

More feedback on how the grade was computed;;; I think in general, when an explicit grade isn't given in the early stages, more feedback would be helpful as to know how to advance to later grading stages. (r182\_AutoTest)

More feedback on how the grade was computed;;; I want to choose all :((( (r183\_AutoTest)

More feedback on how the grade was computed;;; Passing the smoke tests (1/1) but still receiving developing was a little disheartening (r185\_AutoTest)

More detailed feedback;;; The smoke test info was very vague and it was difficult to know what exactly to focus on. I lost a lot of time looking through working code to try and find small issues that caused the smoke test to fail (r186\_AutoTest)

More feedback on how the grade was computed;;; The we get feed back on which part of our code passed the smoke tests but we didn't get any information on the overall grade we'd receive in percentage form for that commit. (r187\_AutoTest)

More detailed feedback;;; knowing where tests fail/time out when everything is ok locally would be helpful and more reflective of real debugging, not semi-blind guessing (r188\_AutoTest)

More detailed feedback;;; A very frustrating moment was where our autotest implied there was only one thing which didn't work, but when fixed it unveiled a new issue which we didn't have time to fix but could have done at the same time as working on the previous bug if we had known about it. (r189\_AutoTest)

More feedback on how the grade was computed;;; Sometime you get no hint on where to even start looking for the issues. (r190\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; Just to be able to check if a fix we did was correct (r191\_AutoTest)

More detailed feedback;;; The feedback is sometimes really vague, such as what I received once: I recommended you to work on Geolocation! while the smoke tests passed for Geolocation. I later found it was the test running the queries with geolocation as a filter, that failed. In other words, it was problem of performQuery instead of Geolocation. That feedback really trapped me for a long time. (r192\_AutoTest)

More detailed feedback;;; Knowing what specific test failed would have been very valuable information. Not necessarily so specific that it would give us the answer but for example for performQuery telling us it was a problem with WHERE or COLUMNS would have been great. (r193\_AutoTest)

More detailed feedback;;; Sometimes gave incorrect error messages, so we would waste time working on aspects that were not causing issues. (r194\_AutoTest)

More detailed feedback;;; There were times when the AutoBot showed that a portion of the tests didn't pass but was too vague in describing the problem. This made it difficult to pinpoint where to make changes in order to fix the issue. (r195\_AutoTest)

More detailed feedback;;; To be honest, I didn't use the autotest feature much because I knew how limited tries I had and my time management was bad; so I ended up grinding out all the code and then using it in the end. I would've loved, however, more detailed feedback on the parts of my implementation that are lacking, or going wrong. Because I was LOSING my mind trying to debug let me tell you. But I agree that that was my own fault. (r196\_AutoTest)

More detailed feedback;;; What was my code doing exactly that got me the grade. Don't make the percentage of my own tests passing a condition for me to get my feedback. (r197\_AutoTest)

More detailed feedback;;; Sometimes the feedback was really vague and didn't help me at all figure out what was wrong with the code (r198\_AutoTest)

More detailed feedback;;; '- A lot of time is wasted barking up the wrong tree, and it's difficult to tell whether we're going in the right direction. If there were more hints we'd still learn what our mistakes are, and then know where to go, which makes it more efficient.\- In some cases, failure in one of the key functions will cause a spillover failure on the other functions. But it's difficult to tell through just "we recommend you to focus on X" (r199\_AutoTest)

No changes;;; it works pretty well through out the term. (r200\_AutoTest)

More detailed feedback;;; feedback from autobot sometimes very vague and hard to figure out where is the problem and what is the problem in the code (r201\_AutoTest)

More detailed feedback;;; the feedback was so vague that we didn't know where to start fixing the issues (r202\_AutoTest)

More detailed feedback;;; In C2, we were stuck on very limited feedback for hours, and we couldn't understand where to look for our mistake. We re-read the spec multiple times and the feedback was very vague. (r203\_AutoTest)

More detailed feedback;;; It's very difficult to narrow down the source of the issue, especially when different aspects of the tests are dependent on other functionality (query tests failing due to a data parsing issue, etc.) (r204\_AutoTest)

More detailed feedback;;; Because it would help me identify the error faster (r205\_AutoTest)

More detailed feedback;;; Had an issue with test coverage and would have been great to know more about the error message that was being thrown from the remote servers (r206\_AutoTest)

More feedback on what grade my submission received;;; (r207\_AutoTest)

More feedback on how the grade was computed;;; There were times where I felt that my part of the project (parsing queries) was correct, but especially in C2, there were issues with the aggregation despite our tests passing. It would be nice to maybe get a bit more detail, such as the length of the result array, to better understand if the issue is with validating the queries or something else. (r208\_AutoTest)

More detailed feedback;;; The current feedback only includes the clusters of smoke tests that I failed, which didn't help in pinpointing my issue sometimes. (r209\_AutoTest)

More feedback on what grade my submission received;;; Because of the bucketed grading, it was hard to know what my grade was. (r210\_AutoTest)

More detailed feedback;;; Often I'd find that the autotest would tell me to look at one method for problems when the problem was not in that method (r211\_AutoTest)

No changes;;; Sometimes the feedback it gives right before proficient is not correct. (r212\_AutoTest)

More detailed feedback;;; Easier to narrow down the issue would be a huge help. We thought an issue was with one class, spent hours and hours on it, turns out it was completely different : ( Would be nice if feedback was detailed (r213\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; We all have a lot of different things going on and it feels very bad when we cannot get feedback instantaneously towards the end of an assignment. (r214\_AutoTest)

More detailed feedback;;; Ever since c1, I couldn't understand why my addDataset, removeDataset, and listDataset did not pass the bot. They appeared to work locally, but I couldn't figure out why it didn't pass even after asking TAs. I was never able to figure this part out. I feel I could've had a better chance if the AutoTest feedback was more detailed. (r215\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; I think we may need more chances to learn the feedback. (r216\_AutoTest)

More detailed feedback;;; Some of the auto test feedback was wrong (told me to fix something when it was not wrong) (r217\_AutoTest)

No changes;;; I am satisfied with Autotest's feedback quality. (r218\_AutoTest)

More feedback on how the grade was computed;;; sometime me and my partner just spend a lot of time on how to improve the code from developing to proficient, if the grades can be more clear it will save time (r219\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; for c0,c1 and c2 it almost felt like we needed to run lot more tests on our own before calling a grade. I am used to the idea that failing tests gives an insight into what to fix. Having restriction on how many times, grade could be called made it bit hard. (r220\_AutoTest)

More detailed feedback;;; I would have really appreciated more details as to when a test cluster was failing, maybe showing the actual test case or just a hint to point me in the write direction. Sometimes I was completely lost when the test cluster wouldn't pass. (r221\_AutoTest)

More detailed feedback;;; Preferably have a stack trace or error message to indicate what was wrong (r222\_AutoTest)

More feedback on what grade my submission received;;; i wanted to know why my local tests were passing but failing in a autobot (r223\_AutoTest)

More detailed feedback;;; I chose this because it felt like I would submit a request for grading, but not know why things were failing. I feel like it's hard to know what you've done wrong, and after a certain point of trying I was really just guessing that the cases I was missing which was very frustrating and time consuming. (r224\_AutoTest)

No changes;;; I think it's good. (r225\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; Sometimes I needed more grades in a day because otherwise I was restricted to wait for the next day to see if I improved or not (r226\_AutoTest)

More detailed feedback;;; Some of the feedback when you have a problem can be very specific, leaving out feedback for other areas. (r227\_AutoTest)

More detailed feedback;;; Auto test was so frustrating to work with because sometimes it really gives no indication of where an issue might be (r234\_AutoTest)

More feedback on how the grade was computed;;; Sometimes we didn't really understand how our grade was computed and hence we did not know where the problem was hence hard to improve. The #check showed us that our 22 local tests were failing during a submission and all the code parts were 0. We were not sure if it was a problem of tests or code, so we spent lots of time on both sides, which turns out was the bug with the function that reads zip files in the test file. (r228\_AutoTest)

More detailed feedback;;; The feedback is too general. I think what kind of tests I failed can be provided that will be very informative. (r229\_AutoTest)

More feedback on what grade my submission received;;; More feedback from the bucket will help a lot in the initial stage of each milestone. Since this term was my first time interacting with ts, more guidance for C0 will help a lot! (r230\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; Often times I would know that an issue was likely occurring because of 2-3 possible reasons for example, but then after checking 1 possibility I couldn't check another so I would waste time on trying both possibilities. (r231\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; I always do smaller changes and I want to see if those changes made the grade go up more frequently. (r232\_AutoTest)

More detailed feedback;;; more detailed feedback would be useful. Also being able to run a previous checkpoint at any time (i.e running c1 while on c3) would have been helpful as the current way of accessing previous checkpoint feedback with your current code does not show nearly as much feedback as when you ran #c2 while on checkpoint 2 (r233\_AutoTest)

More detailed feedback;;; Sometimes specific grading clusters got 0/1 despite me thinking I have done everything correctly. A slight hint at what could have gone wrong could help a lot. (r235\_AutoTest)

More detailed feedback;;; Timeout errors were very cryptic, Especially when not using any loops. (r236\_AutoTest)

More detailed feedback;;; There was no information on why my code was failing and even the TA's had no advice. (r237\_AutoTest)

More detailed feedback;;; The Min/Max in C2 is very confusing, and no unit test I wrote seemed to be able to find problems locally. Also in post-checkpoint feedbacks we realized we got proficient for those two and only the smoke tests are failing, but that makes collaboration really hard because ppl would falsely take blame. (r238\_AutoTest)

More detailed feedback;;; In a realistic working setting, it would be impossible to tell what the error in our code is since we wouldn't have an AutoGrader grading our work in the industry, however, just for the purposes of learning during the course, having more detailed feedback on AutoTest, such as what internal tests were failing, would have sped up the coding process for the project a lot. (r239\_AutoTest)

More detailed feedback;;; (r240\_AutoTest)

More detailed feedback;;; The main problem we ran into was that the Autograder would tell us which of the checkpoints we had failed, but not much else. It wasn't a huge problem, but it took us a while to piece together what the actual problem was from that result. (r241\_AutoTest)

More detailed feedback;;; Going through and looking at my entire implementation was annoying :( (r242\_AutoTest)

More feedback on how the grade was computed;;; sometimes the feedback would throw us in a completely wrong direction. For example it said to check the wildcard implementation even though the issue was in add dataset. The tests are coupled and very depended on other code that it wasn't testing so certain issues would only pop up in tests that weren't related at all. (r243\_AutoTest)

More detailed feedback;;; Sometimes, the feedback would be "Valid Edge Cases", which was incredibly vague. It would be helpful to even have a sub category to tell us where the edge case was, like in addDataset or performQuery. I ended up having to work on performQuery all night because my partner was convinced addDataset was implemented correctly and the edge case was in performQuery, so it was a bit frustrating that the only feedback we could get was "Valid Edge Cases". (r244\_AutoTest)

More detailed feedback;;; It will help better debug the code. (r245\_AutoTest)

More detailed feedback;;; For the earlier checkpoints, it is really hard to determine the exact issue. Sometimes the issue reported seemed too broad and not necessarily related to the actual issue I later found to be causing the bug. (r246\_AutoTest)

More detailed feedback;;; This would make the project less stressful overall. A majority of debugging time was dedicated to locating where potential issues could be, not learning about and fixing issues. Combined with increasingly long AutoTest wait times, the project was less enjoyable and educational than it could have been. More fine-grained feedback about how the code is deficient would be nice. (r247\_AutoTest)

No changes;;; I thought the project was of appropriate difficulty given the info we had (r248\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; i think it would be helpful to see more of the change in grade as we make progress in the code (r249\_AutoTest)

More detailed feedback;;; More feedback -> easier to pinpoint the exact nature of the problem and solve it. The blackbox test suite doesn't need to be \*that\* black, y'know? (r250\_AutoTest)

More detailed feedback;;; Sometimes it was hard to determine what was wrong. (r259\_AutoTest)



More feedback on how the grade was computed;;; Sometimes it's hard to figure out what went wrong that resulted in not getting full marks. Feedback like this would be very helpful (r251\_AutoTest)

More detailed feedback;;; Sometimes it is really difficult to understand what went wrong with our code from the feedback given, and we often force ourselves going back to the specifications to try really really hard to guess what is expected and read between the lines. (r252\_AutoTest)

More detailed feedback;;; more feedback would help me figure out what specifically isn't working with my code (r253\_AutoTest)

More detailed feedback;;; Only giving the score of a component (remove for example) did not help us to figure out what was incorrect with our solution. (r254\_AutoTest)

More detailed feedback;;; We had something weird happen where many of our tests were failing #check but we got almost perfect scores for the smoke test. We asked our TA for hints but he wasn't able to see more details about it either. I think this means either our tests were bad, or our implementation was wrong despite passing enough to consider full marks. So more detail on why this may happen could be helpful (although in our case we still got full marks despite failing #check) (r255\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; Keeping track of how much I could check feedback today, especially closer to deadlines, caused some stress. (r256\_AutoTest)

More detailed feedback;;; Smoke test gave almost no feedback and was hard to decipher what was wrong with the code. Many times I have been to the TA and no other feedback was given, most issues were left in the dark. (r257\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; I feel like due to my limited capacity in my laptop, I am dependent on the autographed for my code feedback. (r258\_AutoTest)

No changes;;; In my opinion the AutoTest delivers what it promises to deliver and does so effectively for students to successfully complete the project. (r260\_AutoTest)

More feedback on what grade my submission received;;; It would be better if the actual grade could given to us (r261\_AutoTest)

More detailed feedback;;; The detailed feedback is largely only available after the deadline has passed. Beforehand it provides vague reviews. (r262\_AutoTest)

More feedback on how the grade was computed;;; It was very difficult to understand where our server implementation was failing, because it would run on our local instance, but would receive 0/1 on the autotest, I wish we could've known more precisely where the failure was happening. (r263\_AutoTest)

No changes;;; I think it gave a reasonable amount of feedback. If there was a problem it pointed out where it would be which made it easier to find and fix. (r264\_AutoTest)

More detailed feedback;;; Often, autotests were failing and we had to spend hours figuring out why. It's hard to tell if a test case failing is because of that test case, or just some other bug encountered during the database parsing. Additionally, in C3, we ended up getting 100% even though 3 of the smoke tests were still failing... still not sure what was going wrong there. (r265\_AutoTest)

More detailed feedback;;; (r266\_AutoTest)

More detailed feedback;;; Sometimes, the AutoTest would tell us to work on a specific part of the code more, such as the AND parts of the query, but we would not know exactly what was wrong with it. (r267\_AutoTest)

More detailed feedback;;; We encountered timeout issues in c2 and could not determine what exactly caused these issues. (r268\_AutoTest)

More feedback on how the grade was computed;;; since the c1 and c2 doesn't give much detail about the grade we already cover (r269\_AutoTest)

More feedback on how the grade was computed;;; Autograder is buggy in telling me what's wrong for C2 and C3. They are pointing at wrong places where the mistake is happening elsewhere. Definitely need to write more concrete tests to increase observability and actionability. (r270\_AutoTest)

More detailed feedback;;; I think the feedback was quite vague, which was frustrating. Specifically when using #check, this is purely because we had an issue with our code being too slow apparently and both my prof and TA did not know what was causing the error because my code was working perfectly fine. (r271\_AutoTest)

More feedback on how the grade was computed;;; There were instances in which the feedback was really misleading like showing a problem with 'MAX' when everything is correct. Or even when everything is correct, just not reaching proficient. I believe something of the sort that if we pass most or all tests then we get access to more feedback should have been there. (r272\_AutoTest)

More detailed feedback;;; More details on what type of smoke test I am failing would definitely help me understand what parts of my code needs more work. (r273\_AutoTest)

More detailed feedback;;; Sometimes the score would be a max score but not be proficient, or proficient but would not have a max score. So I would just like to have more feedback what is missing from our functionality. (r274\_AutoTest)

No changes;;; i felt comfortable with the amount of feedback (r275\_AutoTest)

More detailed feedback;;; I felt like the autograder often lacked any useful feedback (r276\_AutoTest)

More detailed feedback;;; 4 checks for each person is good enough, but i thought that sometimes the feedback was not helpful at all. Gaining more info that the TAs have access to would be a lot more helpful to gain insight and be more productive. (r277\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; The limit just holds the momentum of working the project (r278\_AutoTest)

More detailed feedback;;; While the smoke tests were helpful on a broad sense, I often wished that I there was more feedback telling us WHY certain smoke tests were failing, rather than it just failing. I recognize it part of the process, however it has led to going down the wrong rabbit hole. Sometimes certain parts of the program can break another seemingly unrelated smoke test, and that can be a bit hard to diagnose without more details. (r279\_AutoTest)

More feedback on how the grade was computed;;; I liked to start and finish my portions of the project early, while my partner preferred to work on the checkpoints later. When I completed my implementation, the smoke tests would pass, but I wouldn't be sure if I had hit all the edge cases until my partner did his work and we reached the proficient bucket. It would have been nicer to know that my portion was complete so I could safely move onto other course work. (r280\_AutoTest)

More detailed feedback;;; The AutoTest feedback is a bit vague sometimes and I understand it's for us to write local tests. However, more detailed feedback would be great since sometimes the code passes locally but not on AutoTest, so that makes it a bit harder to identify the issue. (r281\_AutoTest)

More feedback on how the grade was computed;;; The feedback gave enough of a hint on what next to work on, but not enough info on how close we were to our goal. (r282\_AutoTest)

More detailed feedback;;; More detailed feedback could help me find the problem in my code more easily and quickly. (r283\_AutoTest)

More detailed feedback;;; So we can figure out what exactly is wrong with our code and make corrections easier (r284\_AutoTest)

More feedback on how the grade was computed;;; It was stressful having no idea what wasn't working in a test cluster, especially when I was having difficulty understanding what in my implementation wasn't working when my tests were passing locally (r285\_AutoTest)

More detailed feedback;;; For c2, we were stuck with getting the MIN and MAX incorrect. despite literally exhausting nearly all possibilities and collaborating with our TA, we still could not get it, even though it was working properly locally. The autograder just said we should work on these two, which provided no help in this particular scenario. (r286\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; Being restricted to how many times to be graded would leave me unsure of if I was going in the right direction with my code, as it would run on my local device but not work on the autobot and I would be trying to get the autobot to work, but not having enough times to try to see what is going wrong. (r287\_AutoTest)

More detailed feedback;;; (r288\_AutoTest)

More feedback on what grade my submission received;;; (r289\_AutoTest)

Ability to check the same amount of feedback but more frequently;;; I think the auto grader gave specific enough feedback, but I did not like we had only 2 attempts at getting feedback (r290\_AutoTest)

More detailed feedback;;; Pinpointing errors were difficult without detailed feedback (r291\_AutoTest)

No changes;;; I thought that the autotests were really well implemented not giving away too much information but also providing lots of hints to help. (r292\_AutoTest)

More feedback on what grade my submission received;;; I think some of the bugs would be better fixed for the next implementations (r293\_AutoTest)

No changes;;; I wish i had more time to put into the project, but nothing about it was bad in my view. (I'm not suggesting it take less time either, since I think that might take away from the knowledge-gain.) (r294\_AutoTest)

More detailed feedback;;; so i know which part to edit and fix (r295\_AutoTest)

More detailed feedback;;; I think sometimes exposing an autograder test that failed could help save so much time! (r296\_AutoTest)