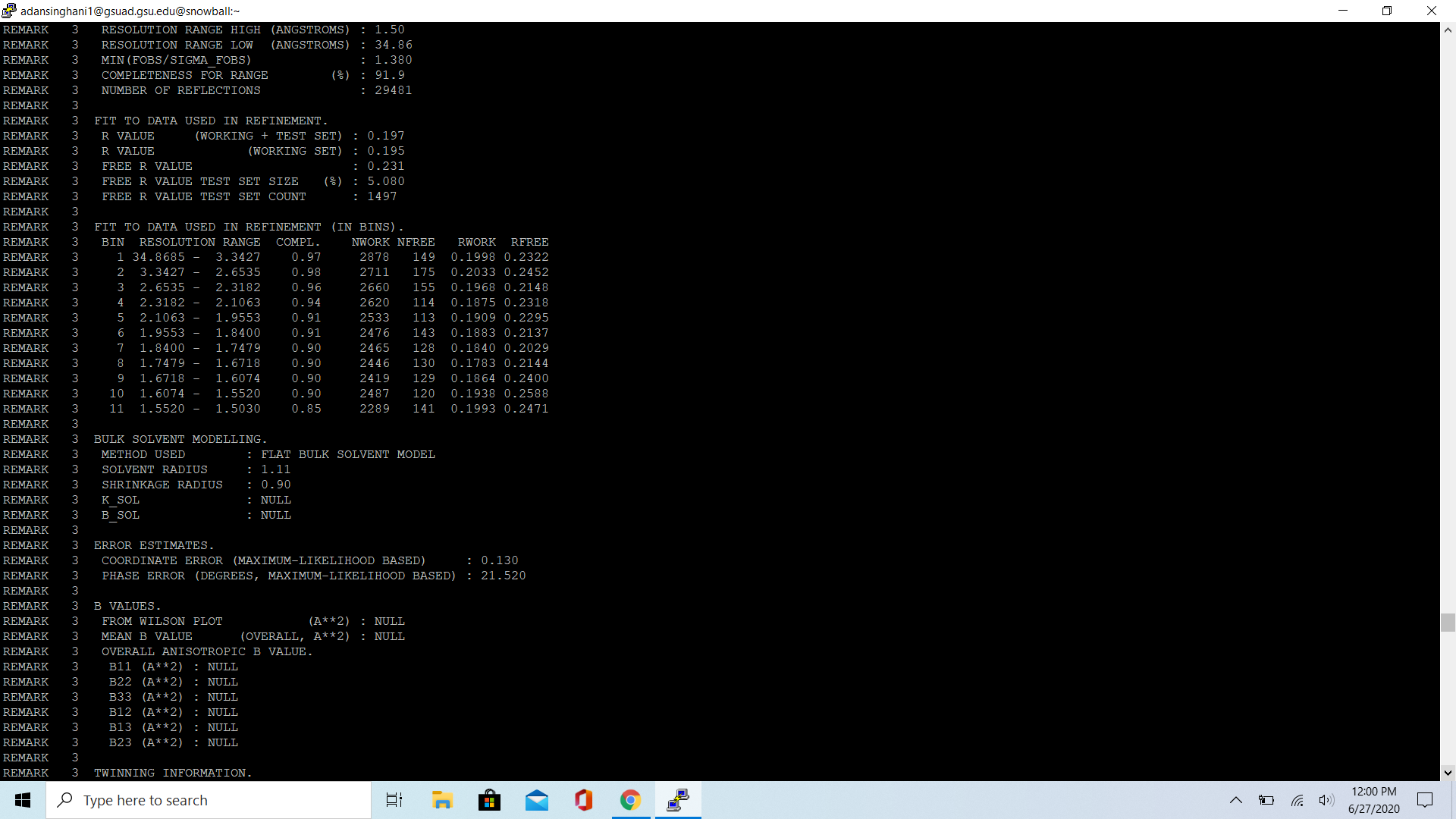
Lab 3

Akash Dansinghani

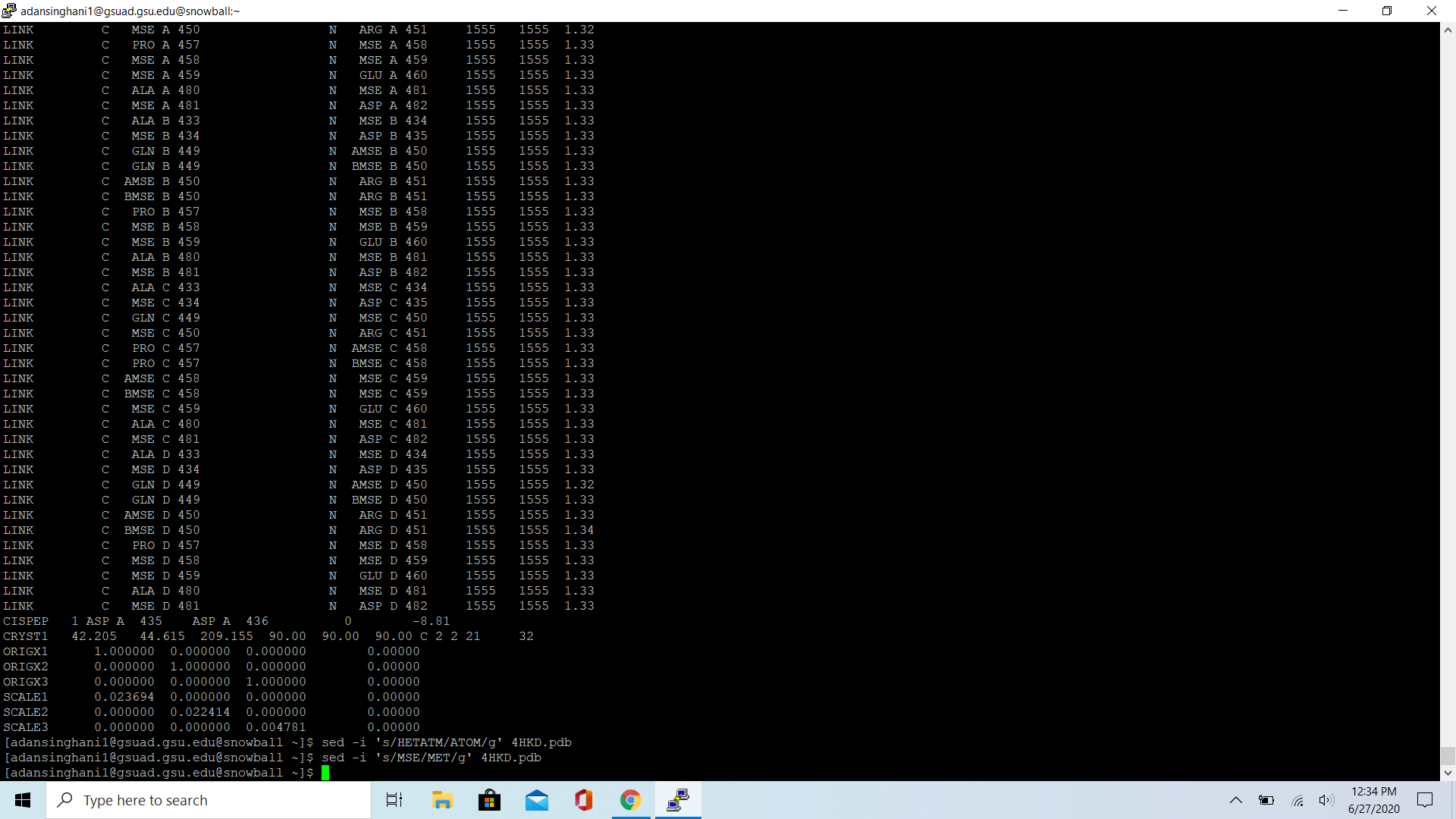
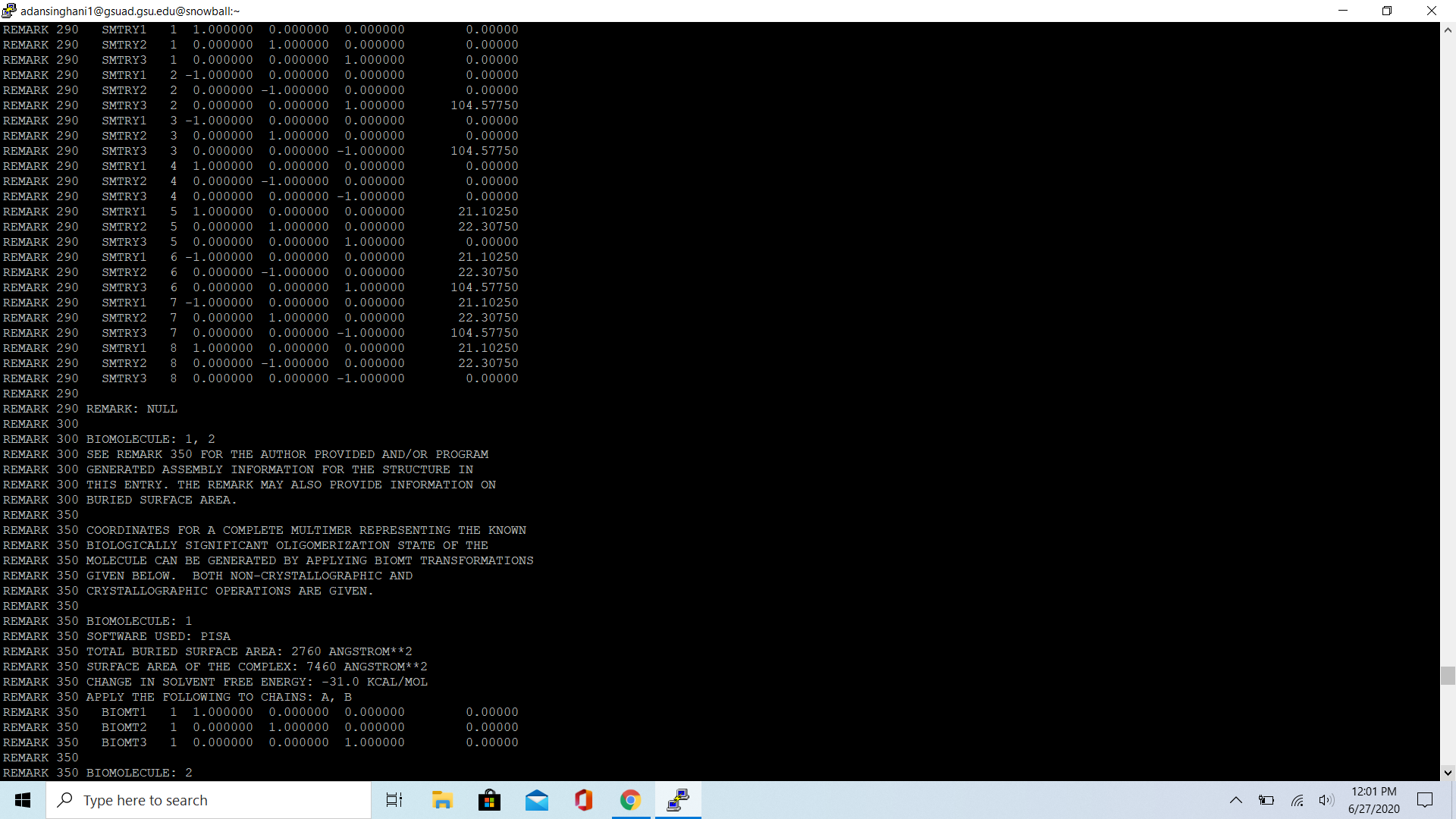
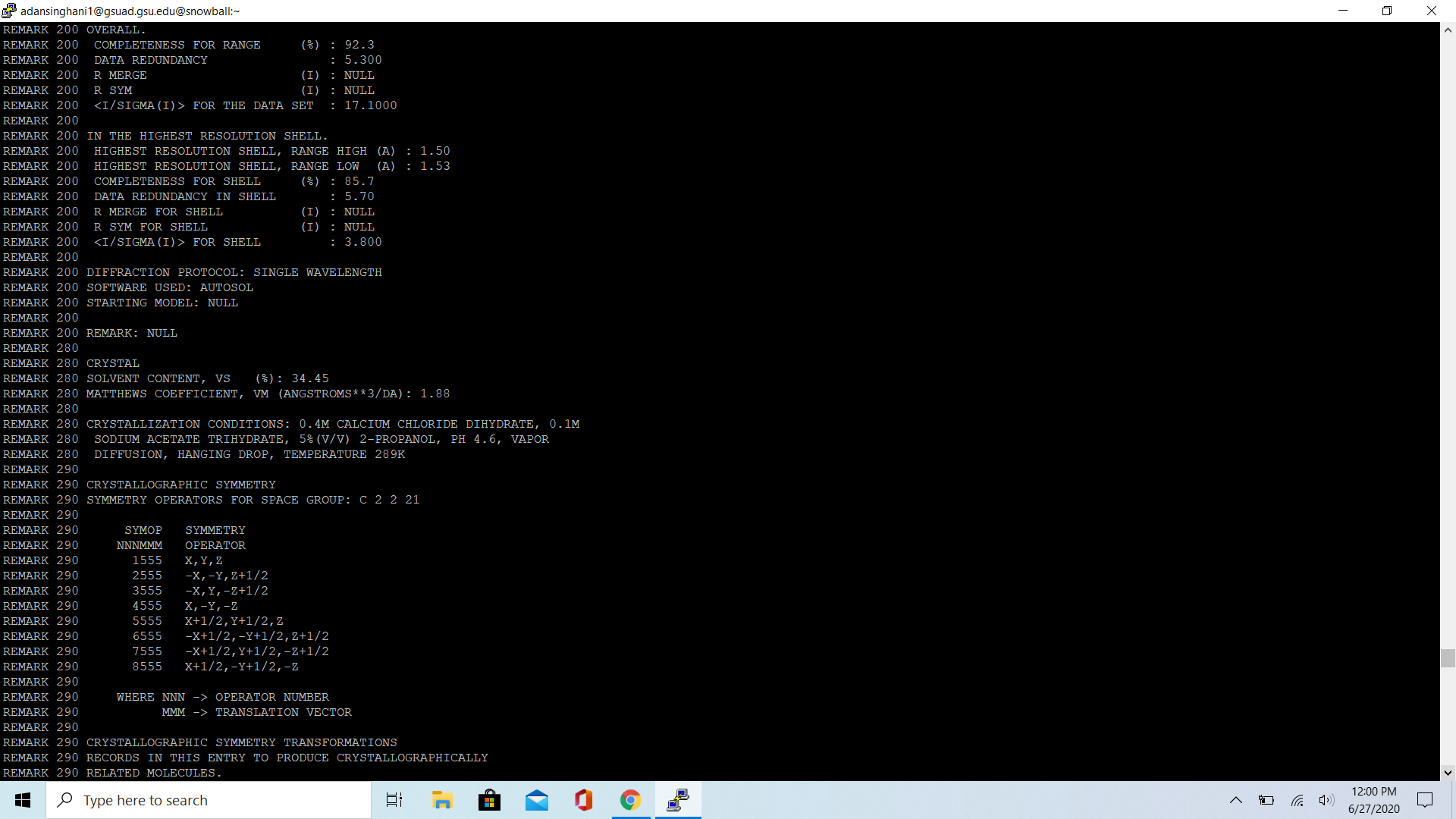
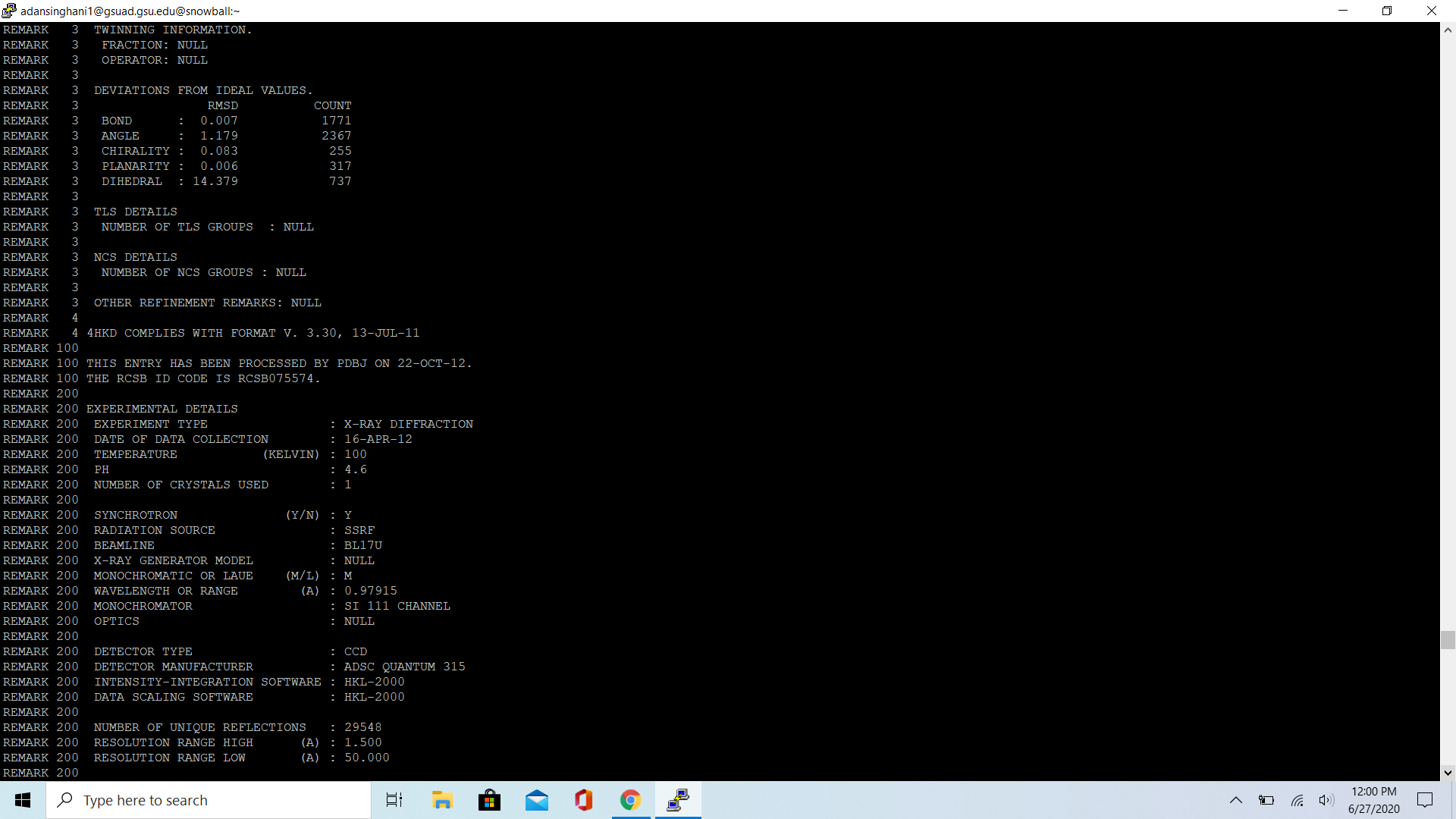
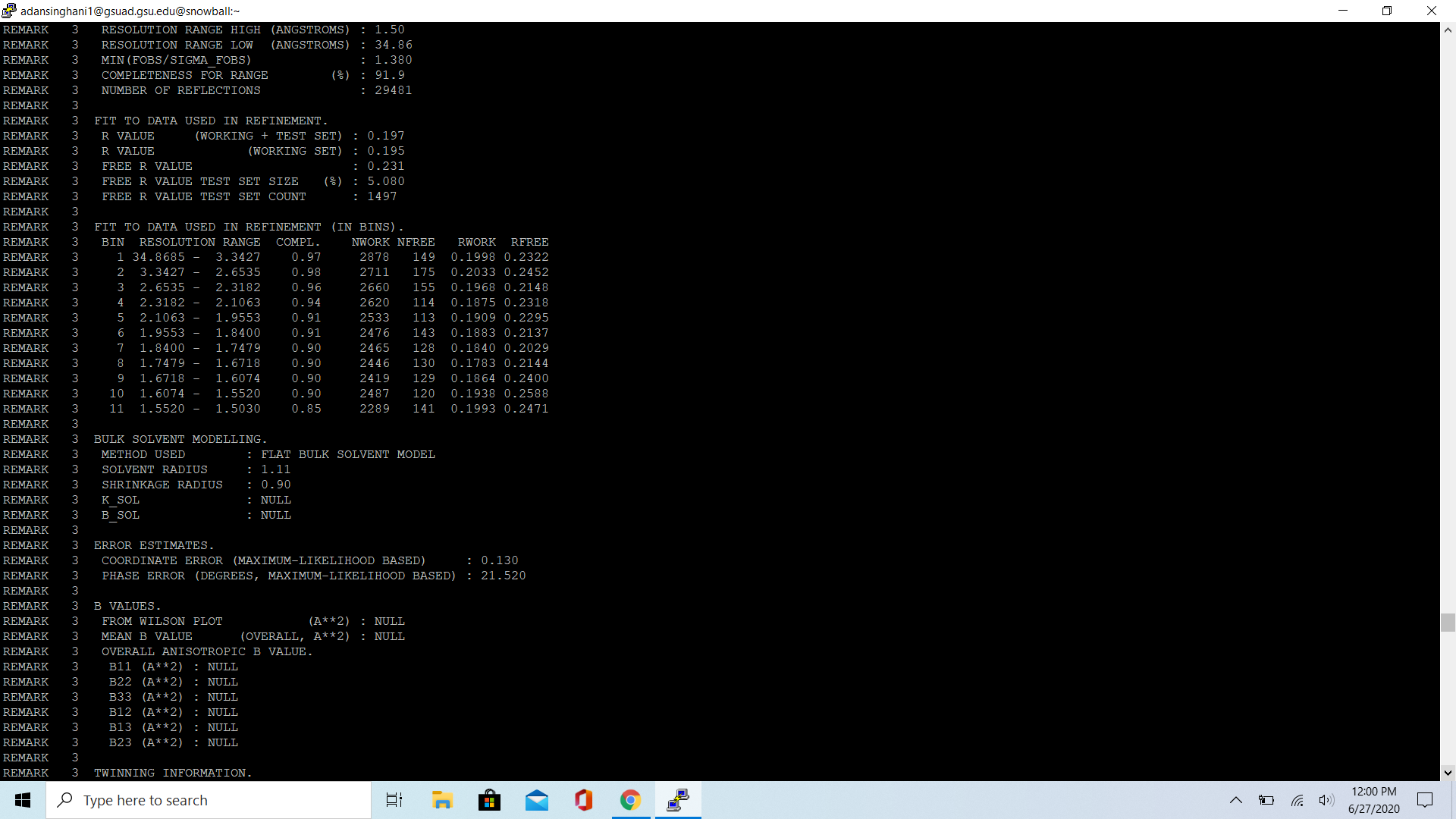
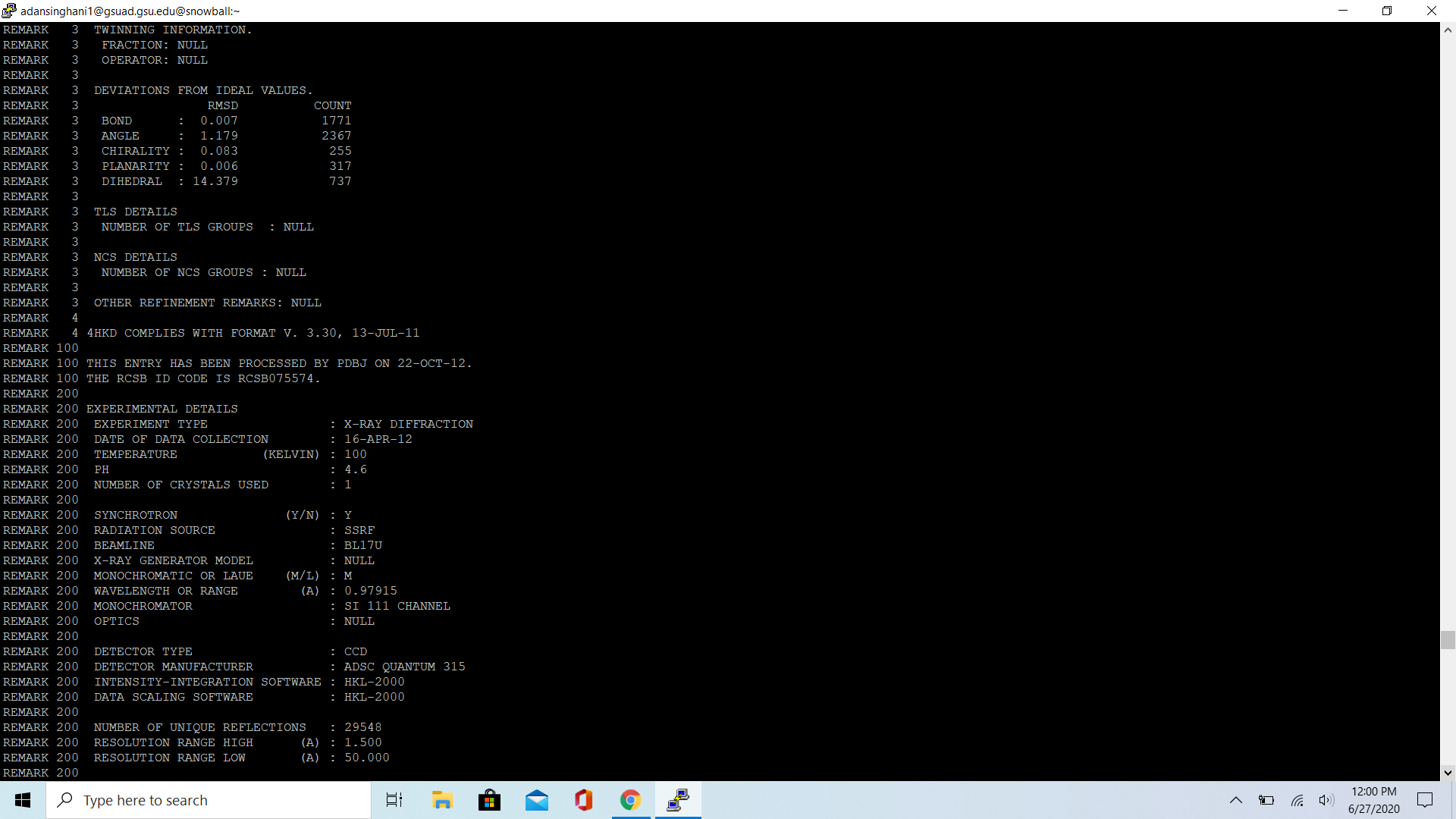
TA: Mahmood

Due: 06/30/20



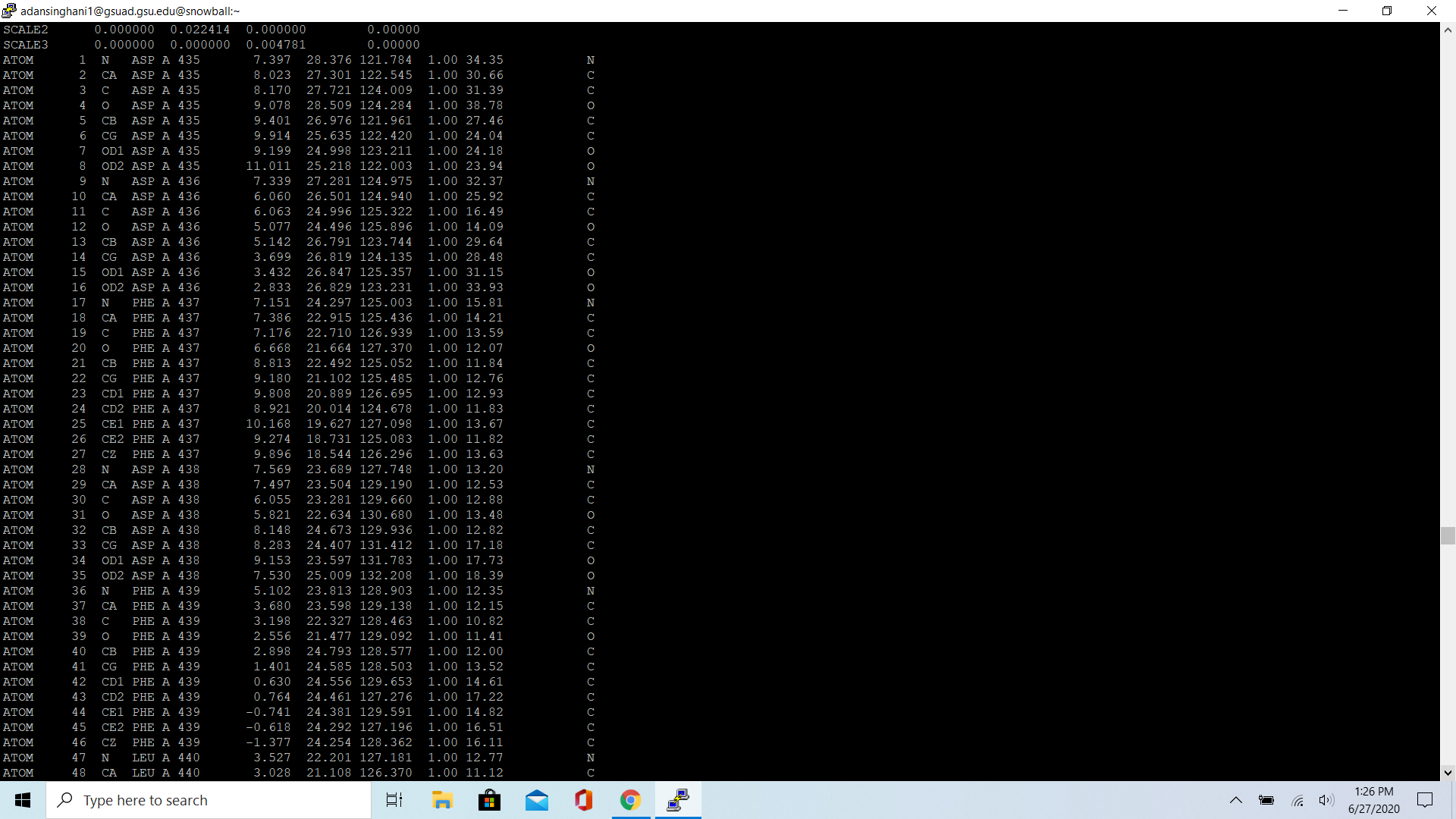
The code I used to get the records other than the following “ATOM”,”CONECT”, ”HETATM”, ”TER” and ”END” was: “*grep -v -e ‘ATOM’ -e ‘CONECT’ -e ’HETATM’ -e ‘TER’ -e ‘END’ 4HKD.pdb”*

Below shows the results.



For ”HETATM” and ”MSE” being changed to ”ATOM ” and ”MET” respectively, I used the following code: *sed -i ‘s/HETATM/ATOM/g’ 4HKD.pdb* and *sed -i ‘s/MSE/MET/g’ 4HKD.pdb*

The result is below.

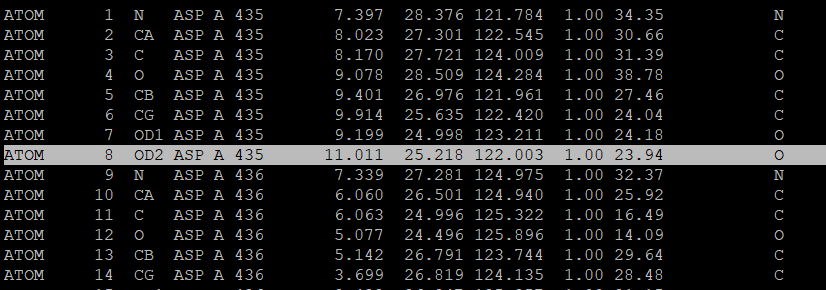


To find the max and min values, I used the command “sort -t, -nk3 4HKD.pdb”.

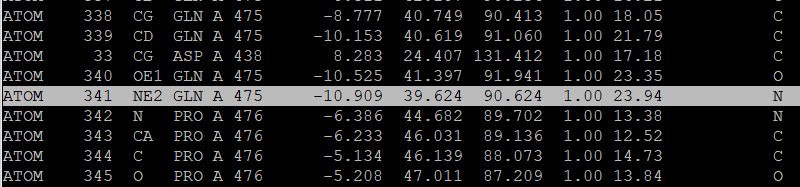
This did not work well with the negatives and with how long the lines were so I used

min=`awk 'BEGIN{min = 1000}{if ($column<0+min) min=$column} END{print min}' 4HKD.pdb`

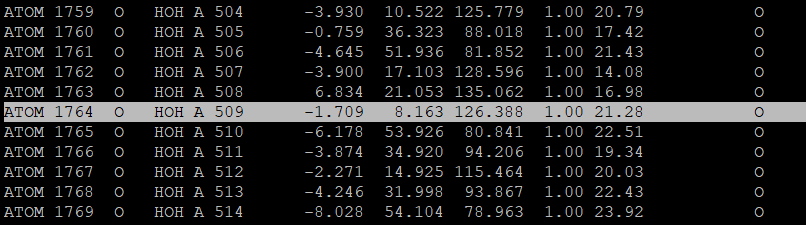
max=`awk 'BEGIN{max = 0}{if ($column>0+max) max=$column} END{print max}' 4HKD.pdb`



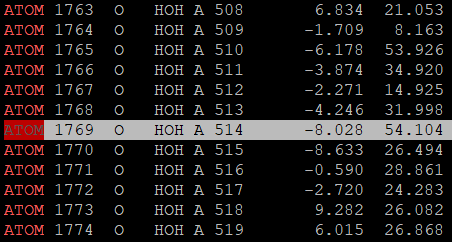
11.011 was the max for column 7 or x



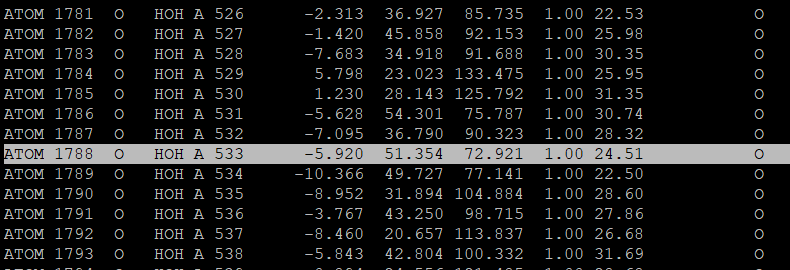
The min for column 7 is -10.909



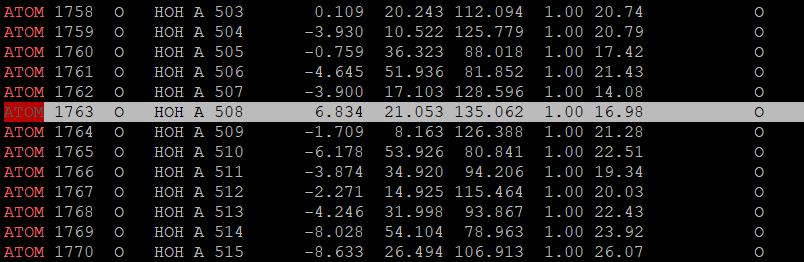
The min for column 8 or y is 8.163



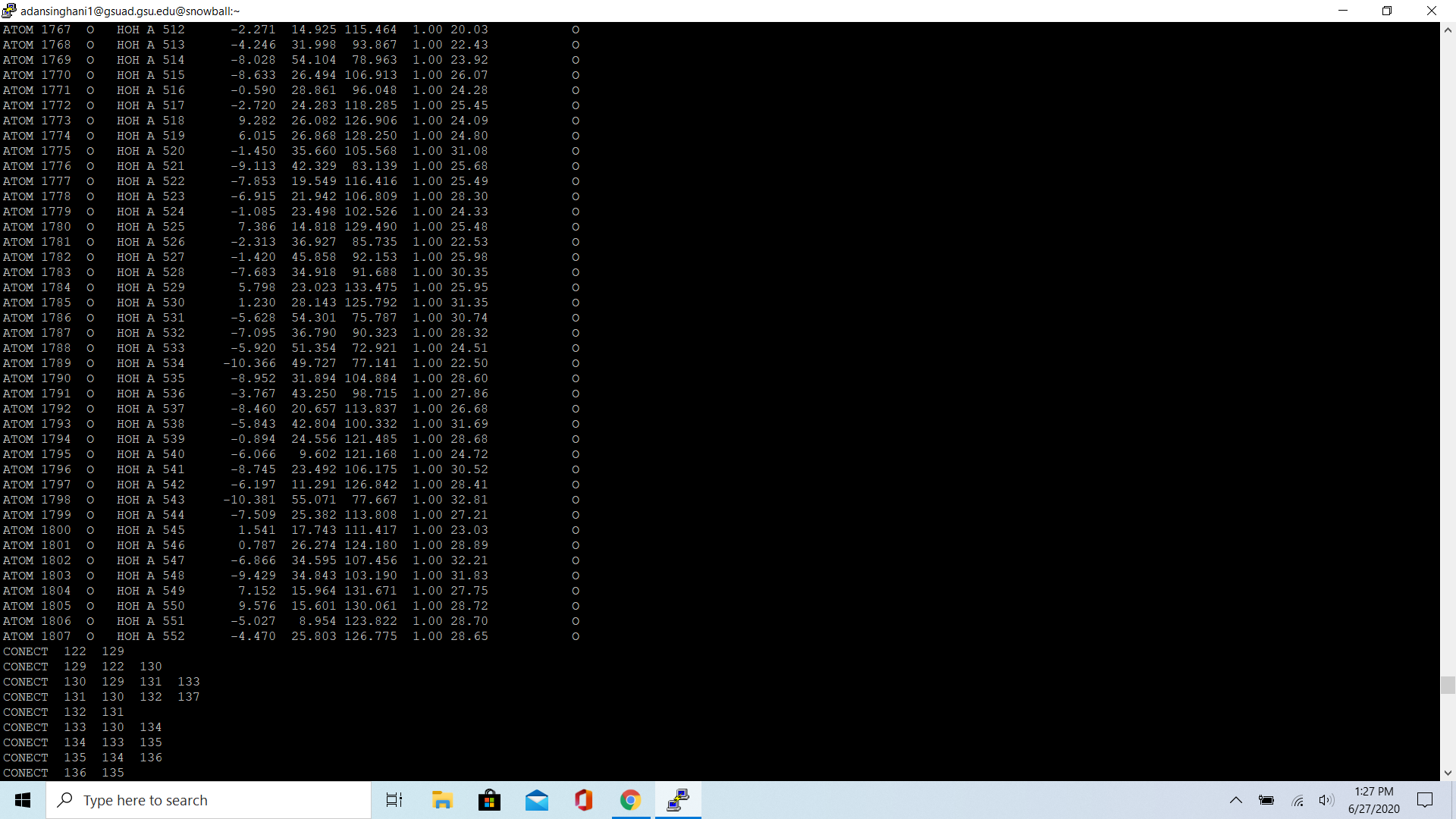
The max for column 8 or y is 54.104

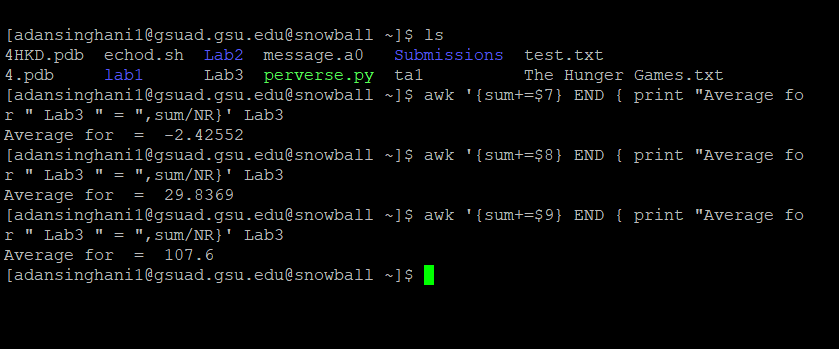


The min for column 9 or z is 72.921

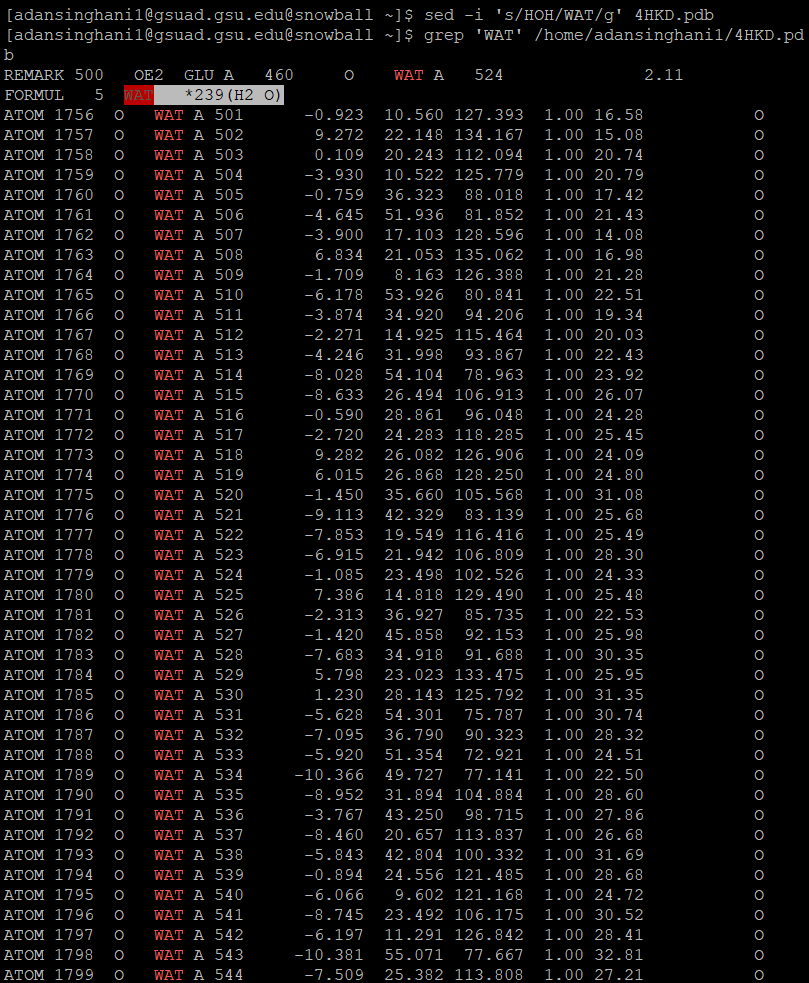
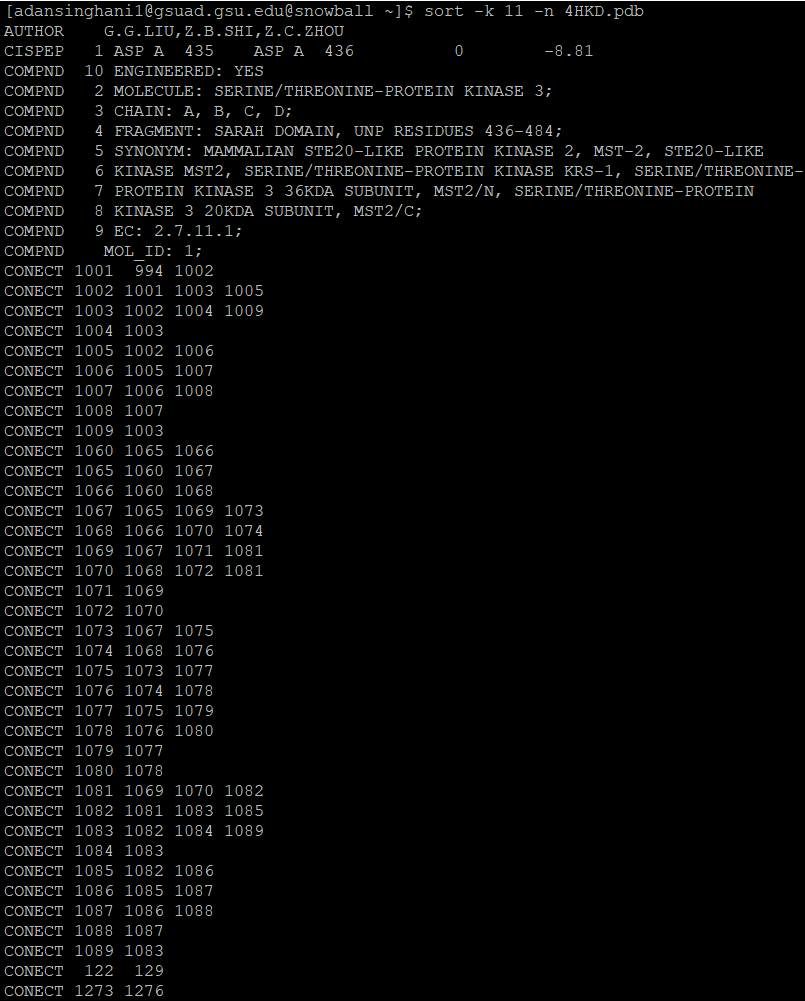


The max for column 9 or z is 135.062.

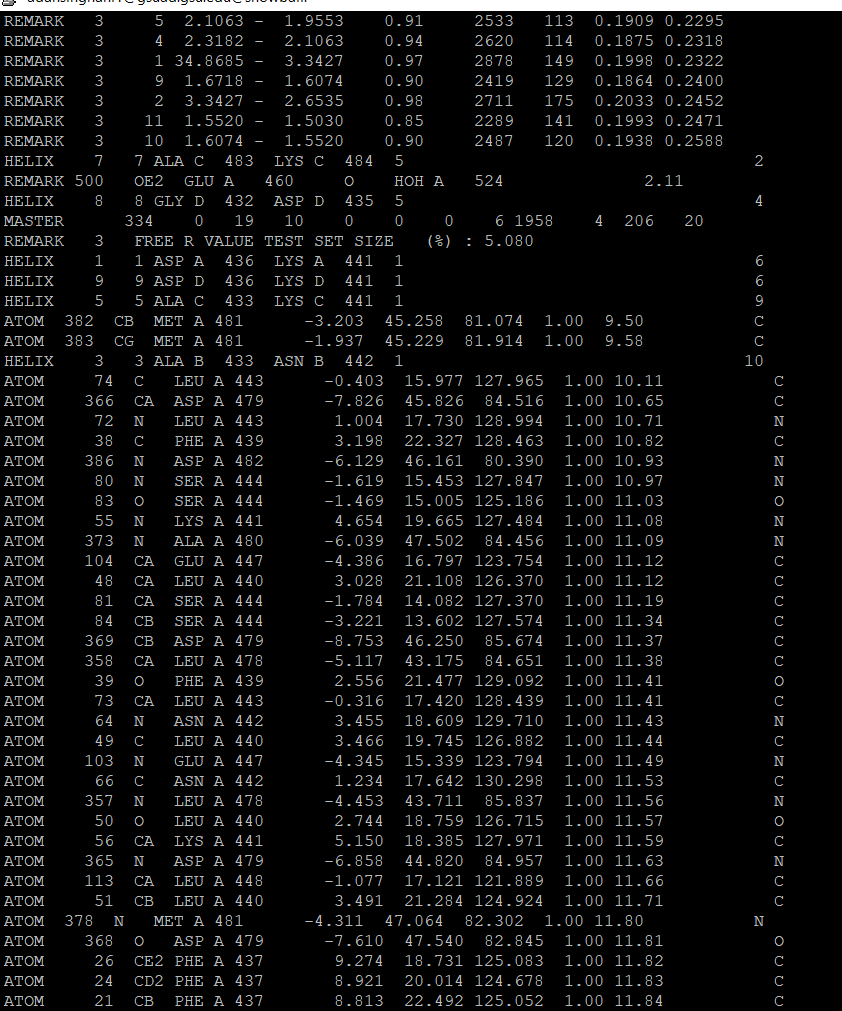




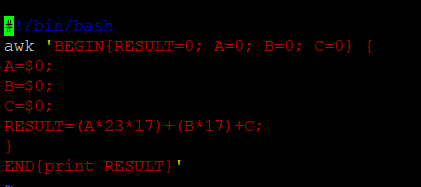
To find the average for columns x,y,z or 7,8,9, I used the command: *awk ‘{sum+=$column} END {print “Average for filename = “,sum / NR}’ filename.* As you can see above, the only factor I needed to change was the column number. A solution to a problem I had was creating a new file called “*Lab3”* that has the same contents as the 4HKD file. The only difference is the type since the original file was a .pdb extension. For some reason, I was not able to get a result with the code: *awk ' $1=="ATOM" {sum+=$7} END {print "Average for 4HKD.pdb = ", sum / NR}' 4HKD.pdb.* The new file had only contents of ATOM, so I did not need the $1==”ATOM”. I was able to just use the sum code and it worked.

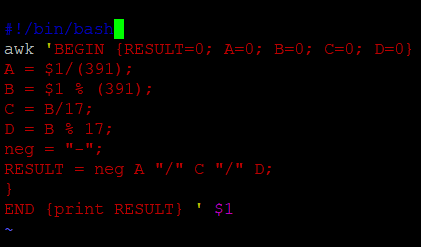
 To change HOH to WAT I used the command “*sed -i ‘s/HOH/WAT/g’ 4HKD.pdb*”

I used the command “*sort -k 11 -n 4HKD.pdb”* to sort column 11. In the screenshot below, we can see the headers named ATOM has a starting value of 9.50 and the values progress down the list the last value is 11.84 at the end of the screenshot

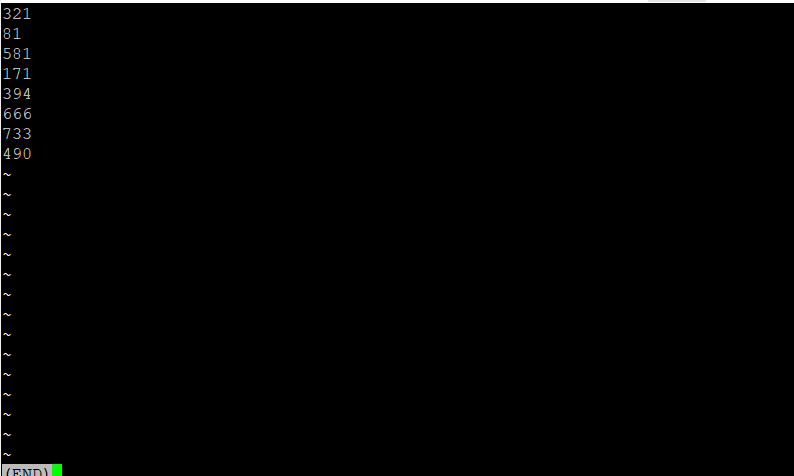


Lab 3 Part 2

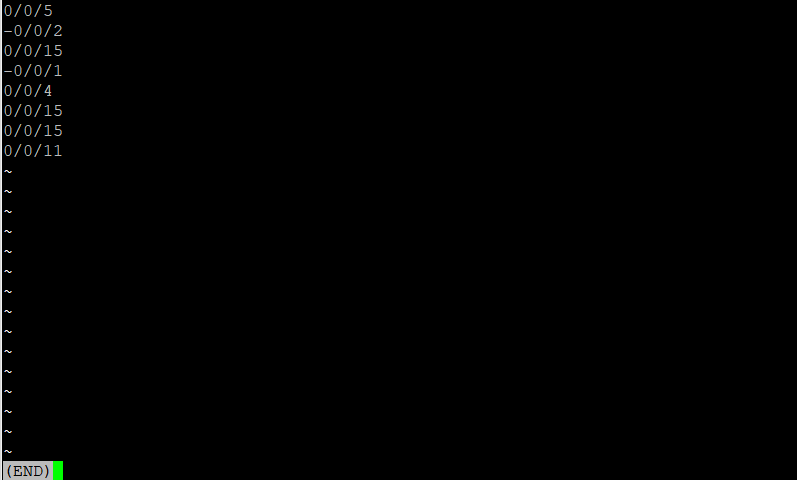
  
This is my first program called “*ScriptPt2Q1.sh”* which takes influence from string2knutz.sh program



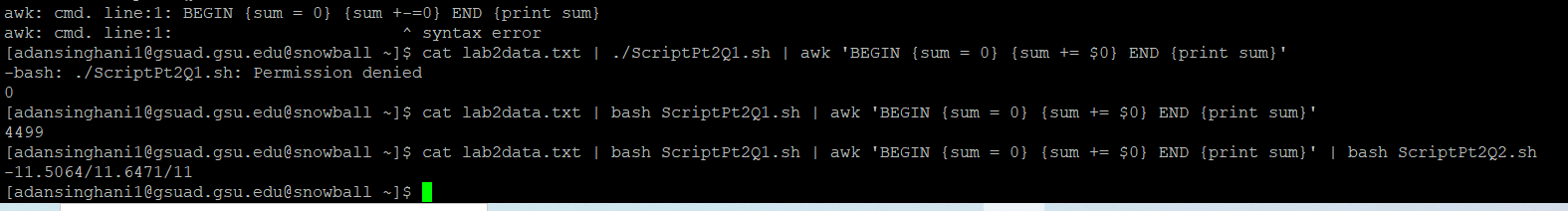
This is my second program called “*ScriptPt2Q1.sh”* which takes influence from knutz2string.sh program



This was my result from the first program by using the command: *“cat lab2data.txt | bash ScriptPt2Q1.sh”*



This was my result from the second program using the command:*“cat lab2data.txt | bash ScriptPt2Q2.sh”*



I used the command *“cat lab2data.txt | bash ScriptPt2Q1.sh | awk ‘BEGIN {sum = 0} {sum += $0} END {print sum} ‘* and *“cat lab2data.txt | bash ScriptPt2Q1.sh | awk ‘BEGIN {sum = 0} {sum += $0} END {print sum} ‘ | bash ScriptPt2Q2.sh*  to run both programs.

To generate the tmp file(s), I used “cat lab2data.txt | bash ScriptPt2Q1.sh >> tmp1.txt"

When doing the sum, I got a result of 4499 for the first program and -11.5064/11.6471/11 from the second program.