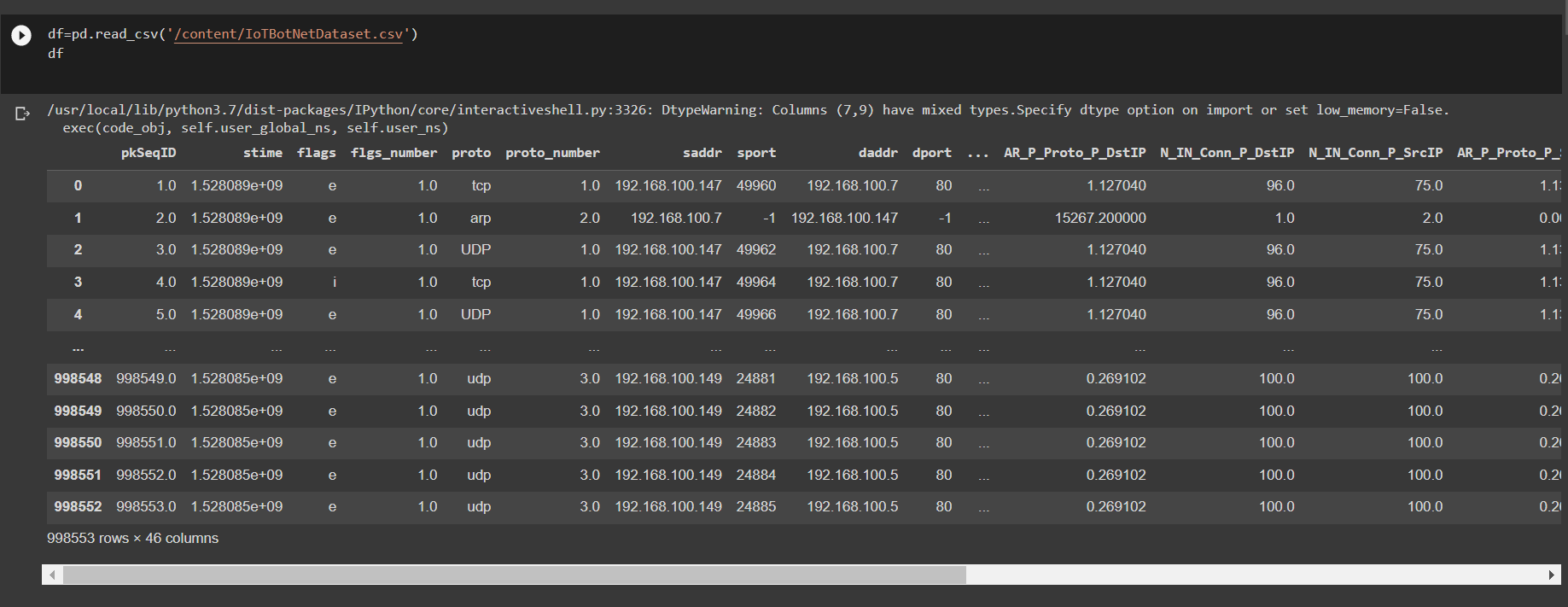
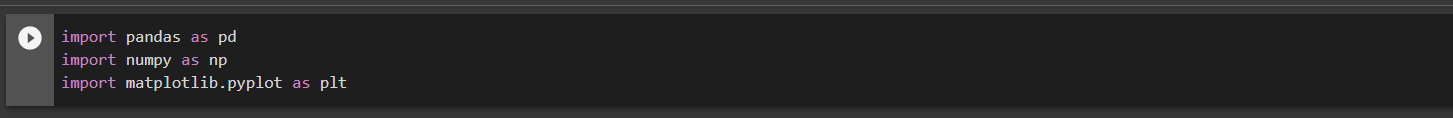
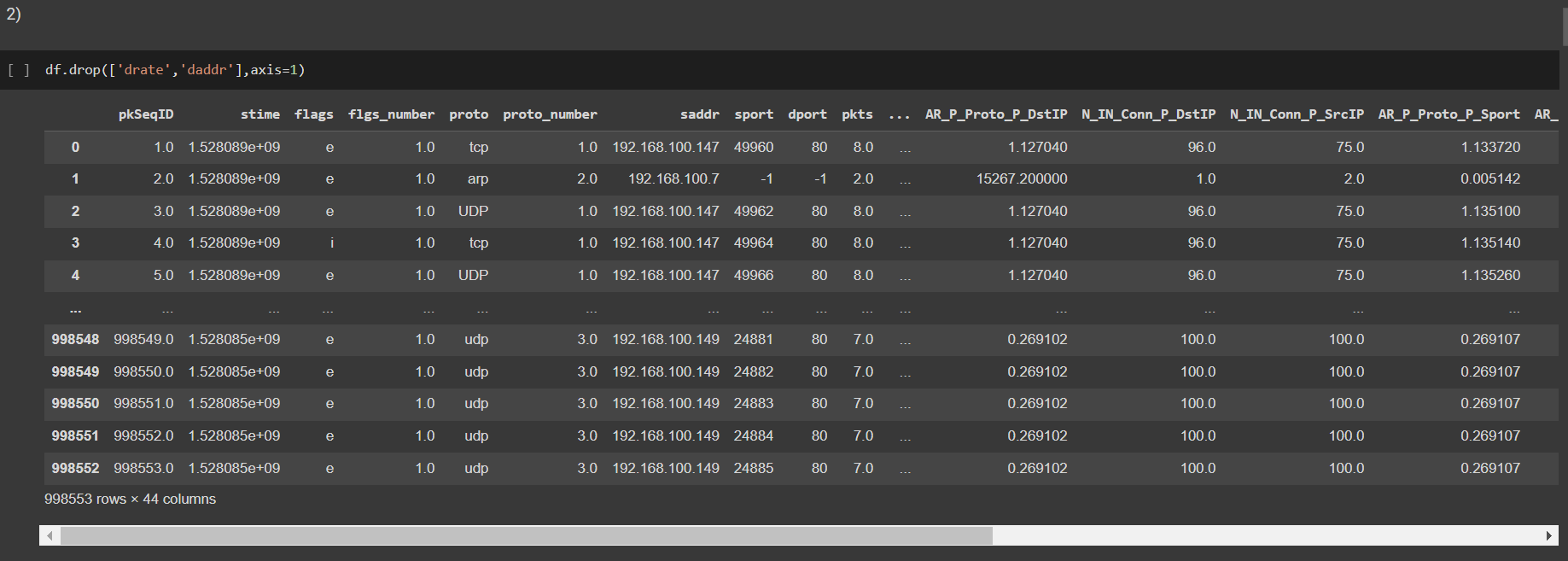
LABEVAL-1 CSE304-FODS

ROLLNO: CB.EN.U4CSE20226

1) Import data and save it in your disk space (1)

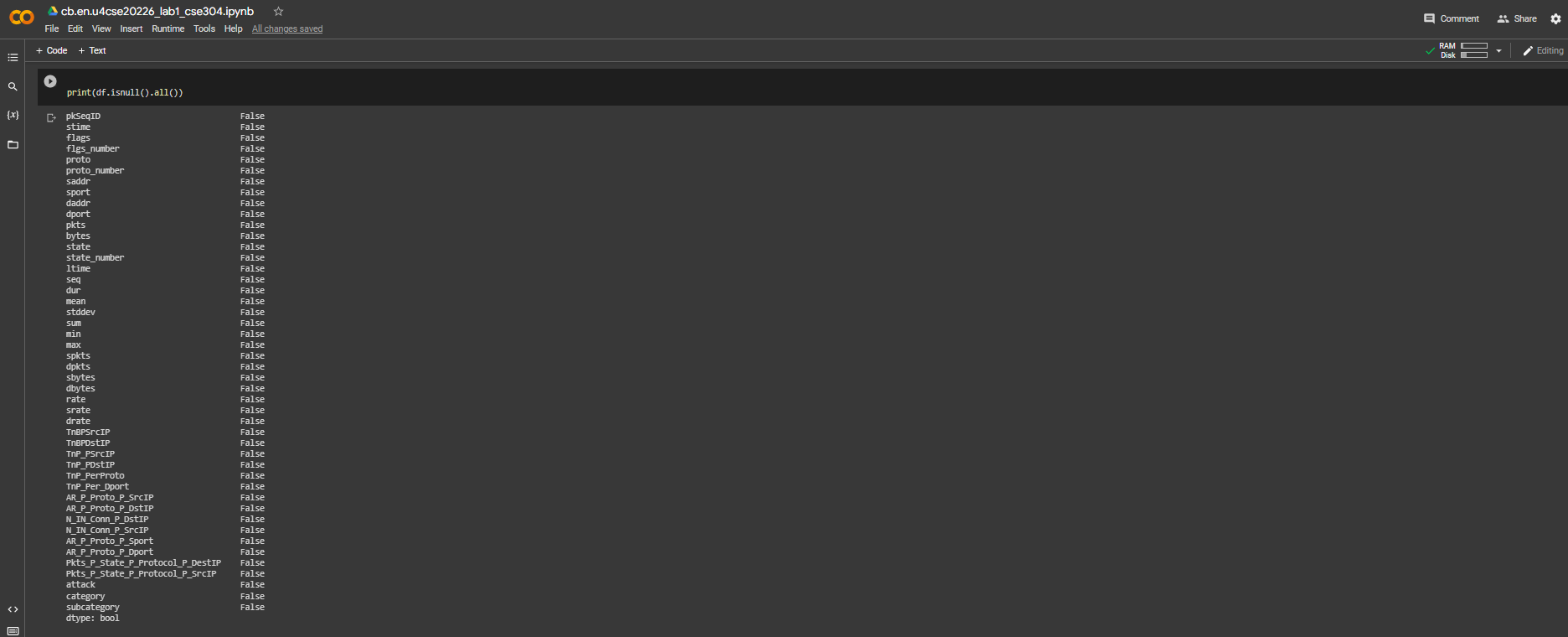


2) Drop the columns not required for your analysis. (1)

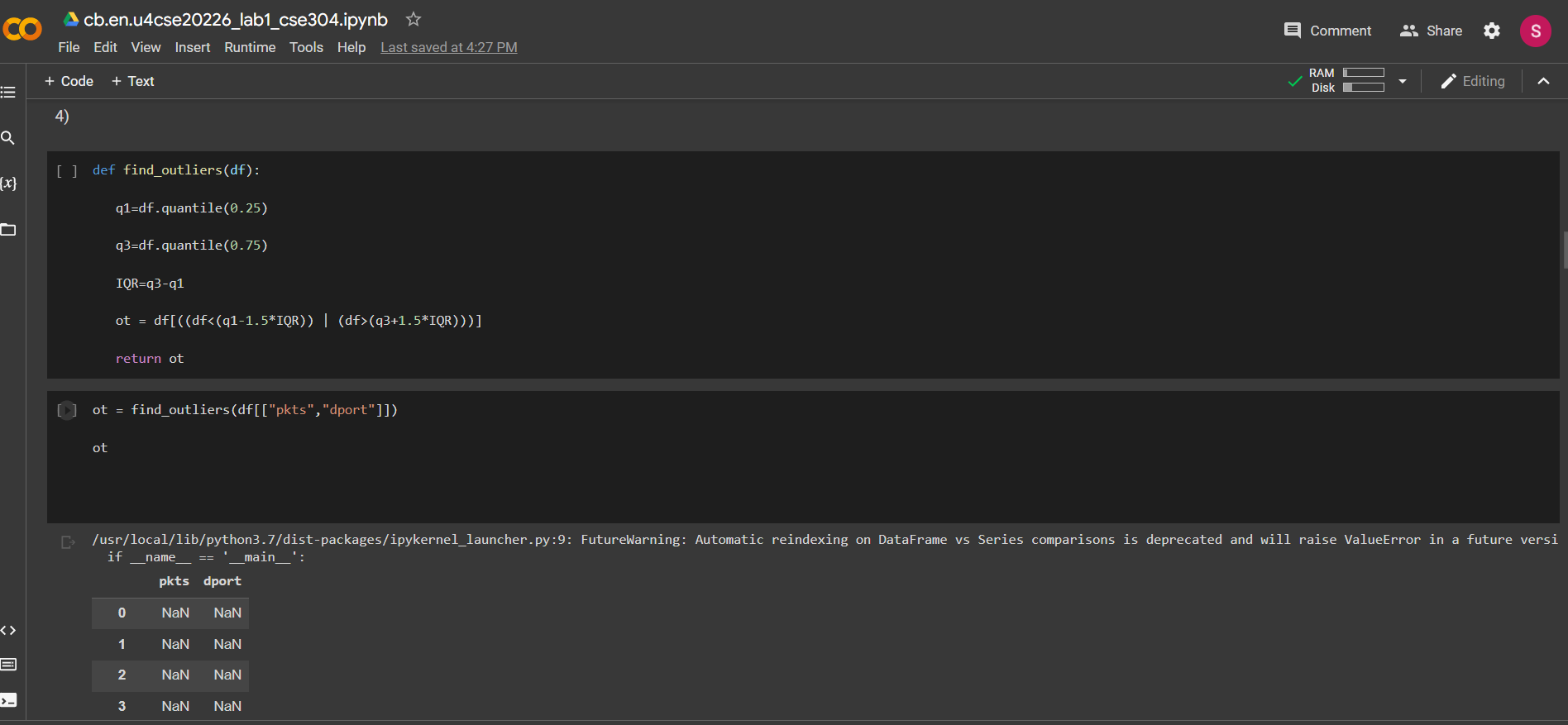


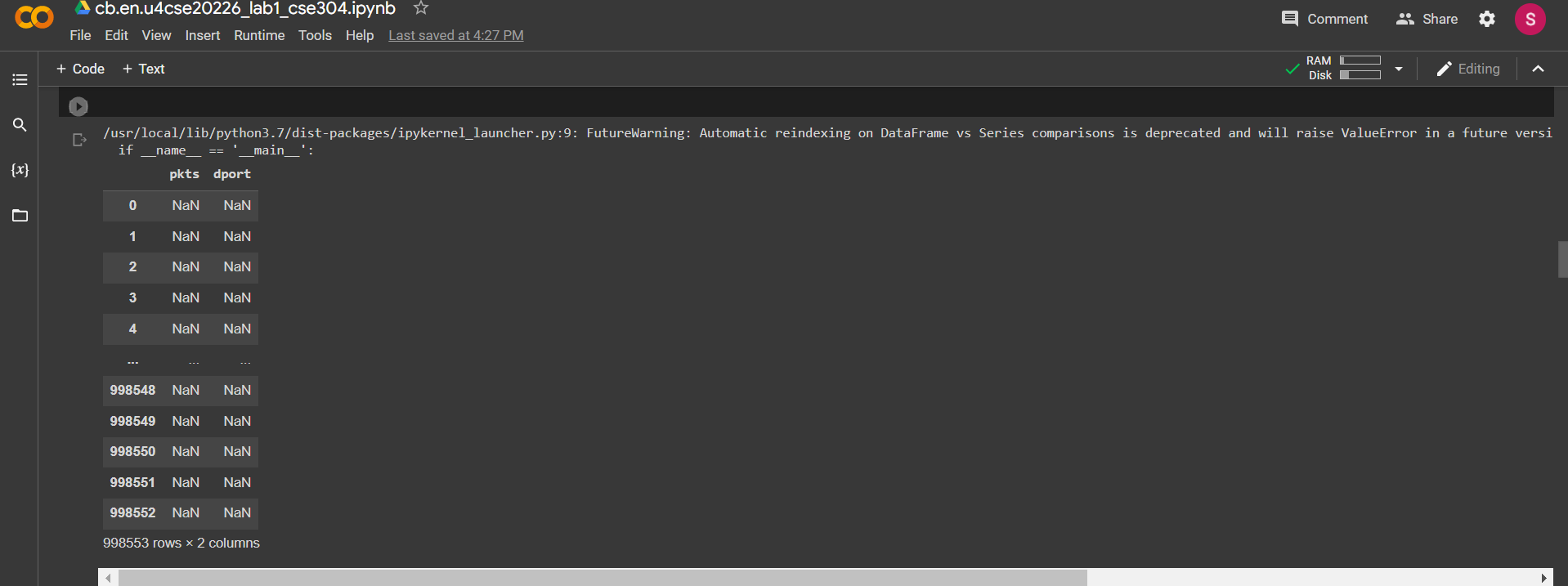
Here we have 44 columns so 2 columns we dropped

3) Do a missing value analysis and perform appropriate steps to correct them (2)

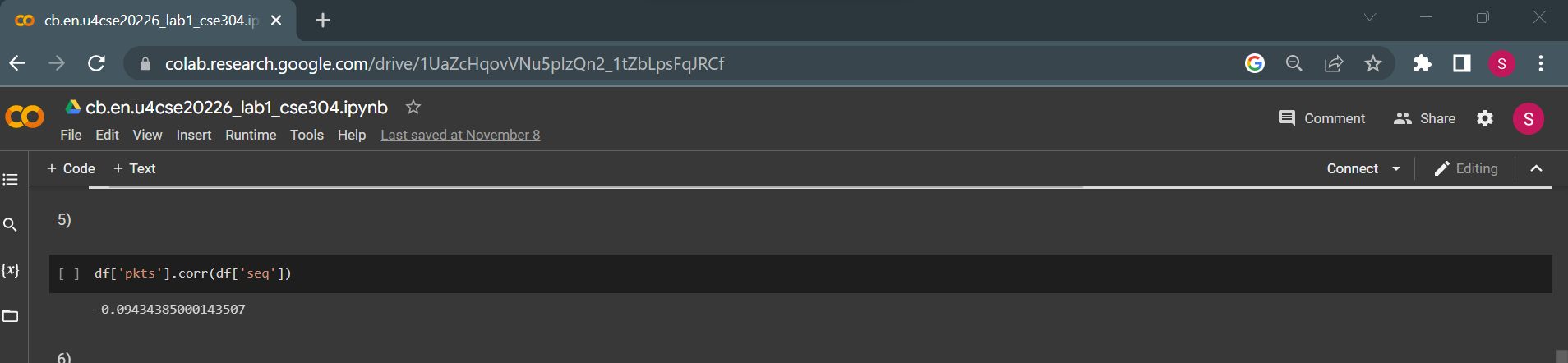


4) Perform an outlier analysis for any two columns (2)

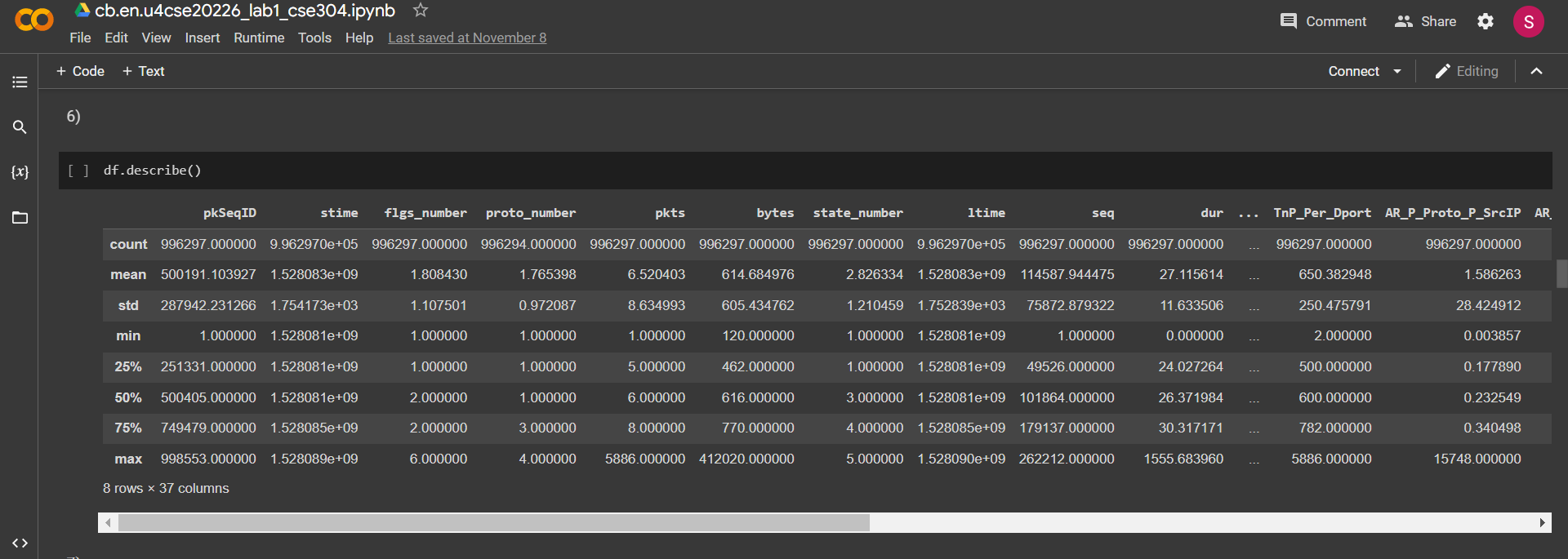




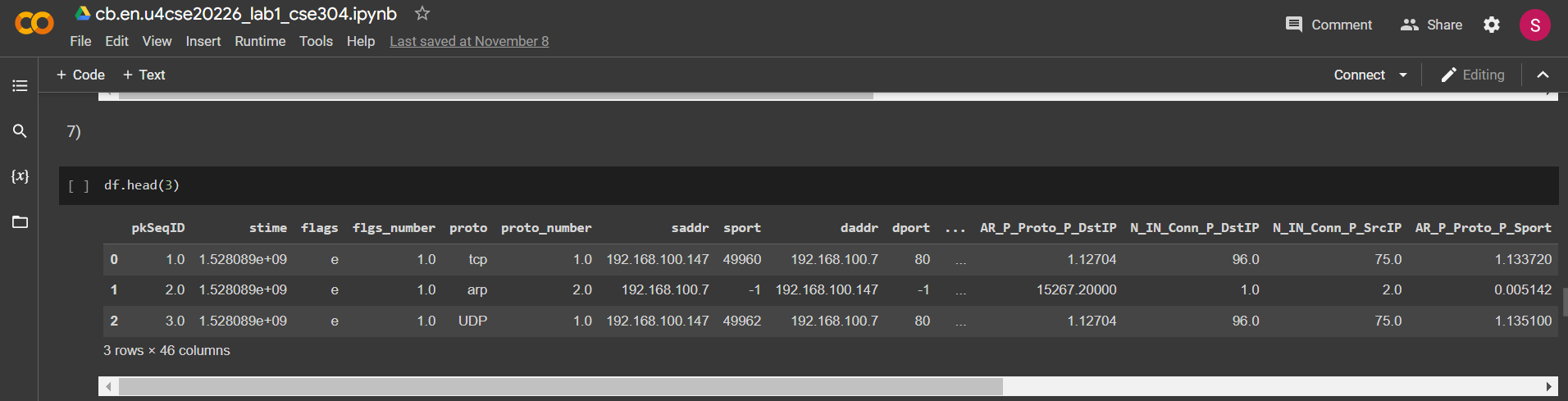
5) Do a correlation analysis of any two variables. (2)



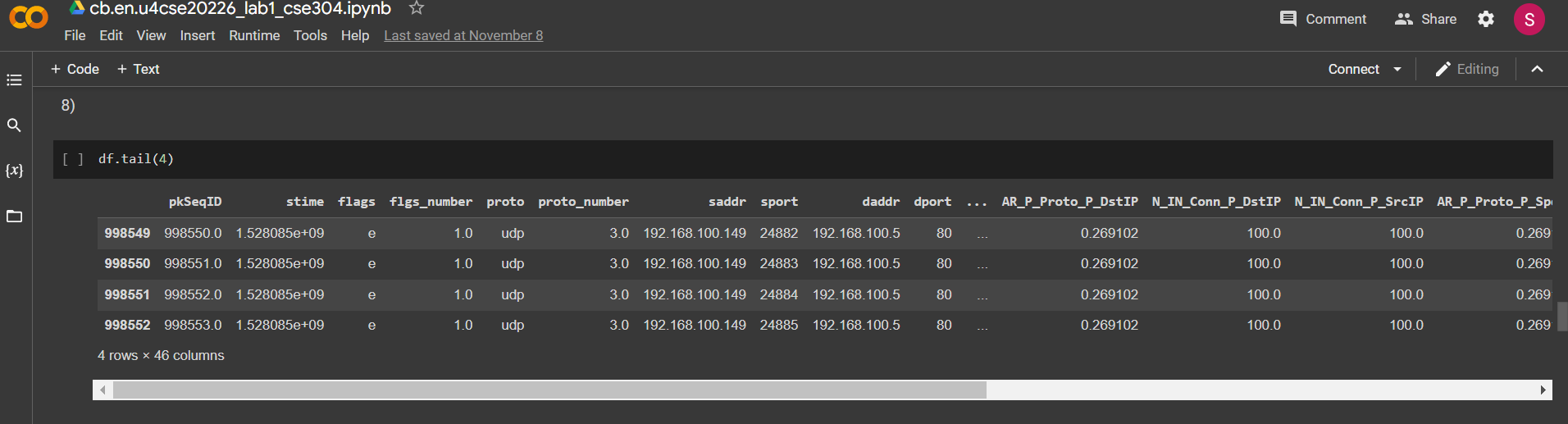
6) Describe the data (1).



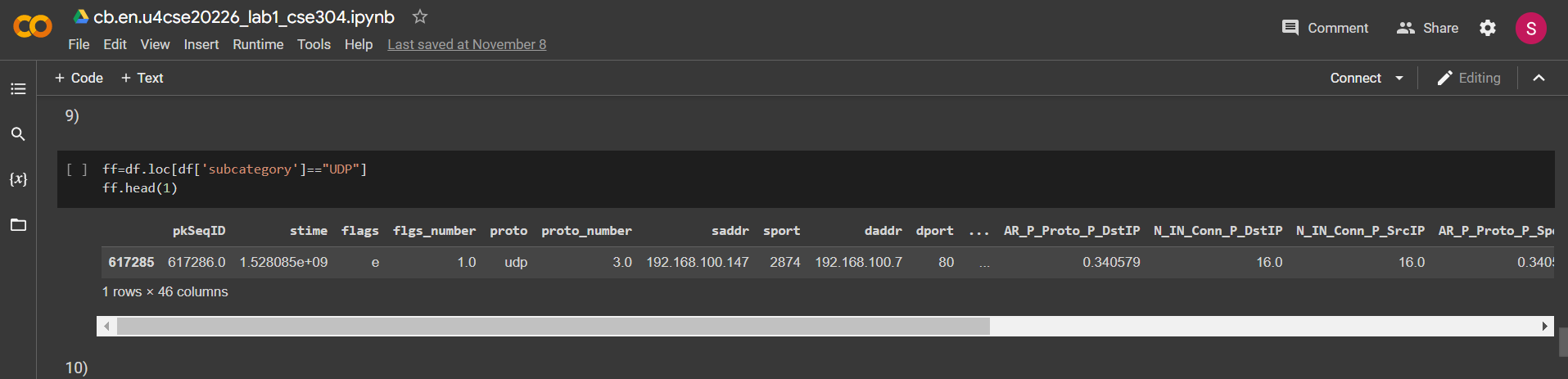
7) Display the first 3 rows (1).



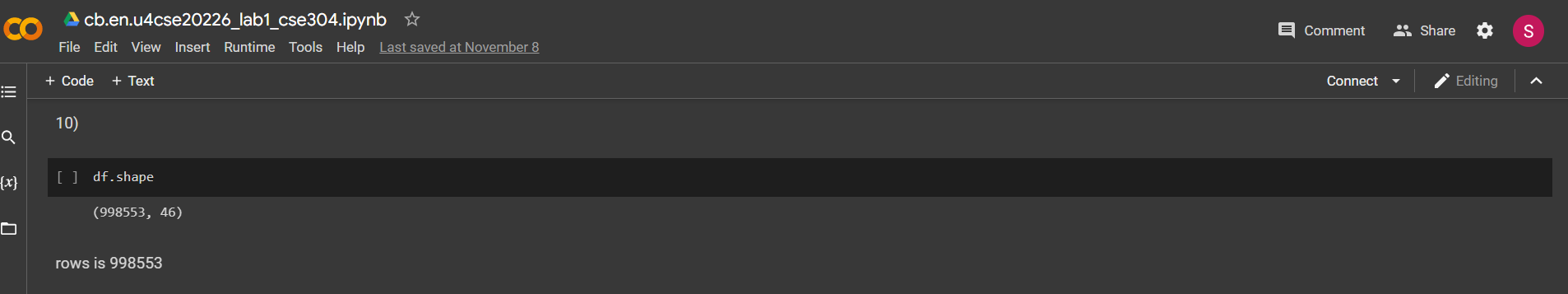
8) Display the last 4 rows(1)



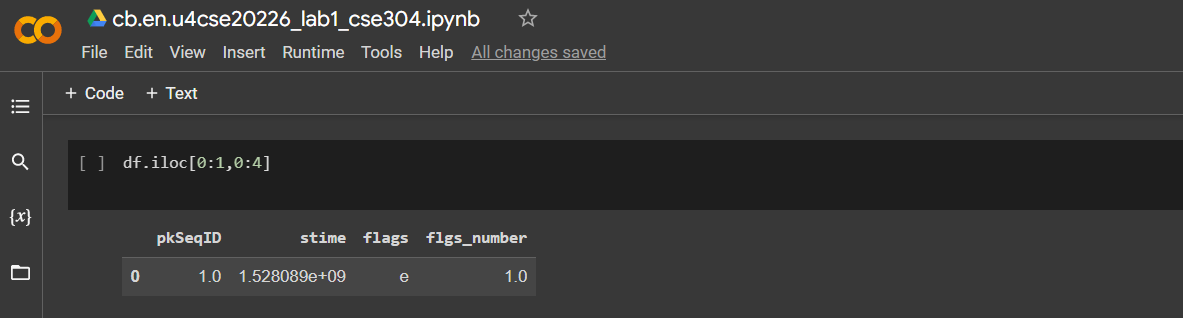
9) Locate the first row corresponding to the proto as “UDP” (2)



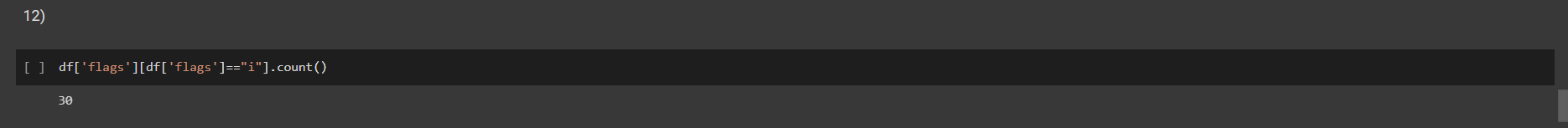
10) How many rows of data do you have? (1)



11) Display the first 5 columns of the first row (2)



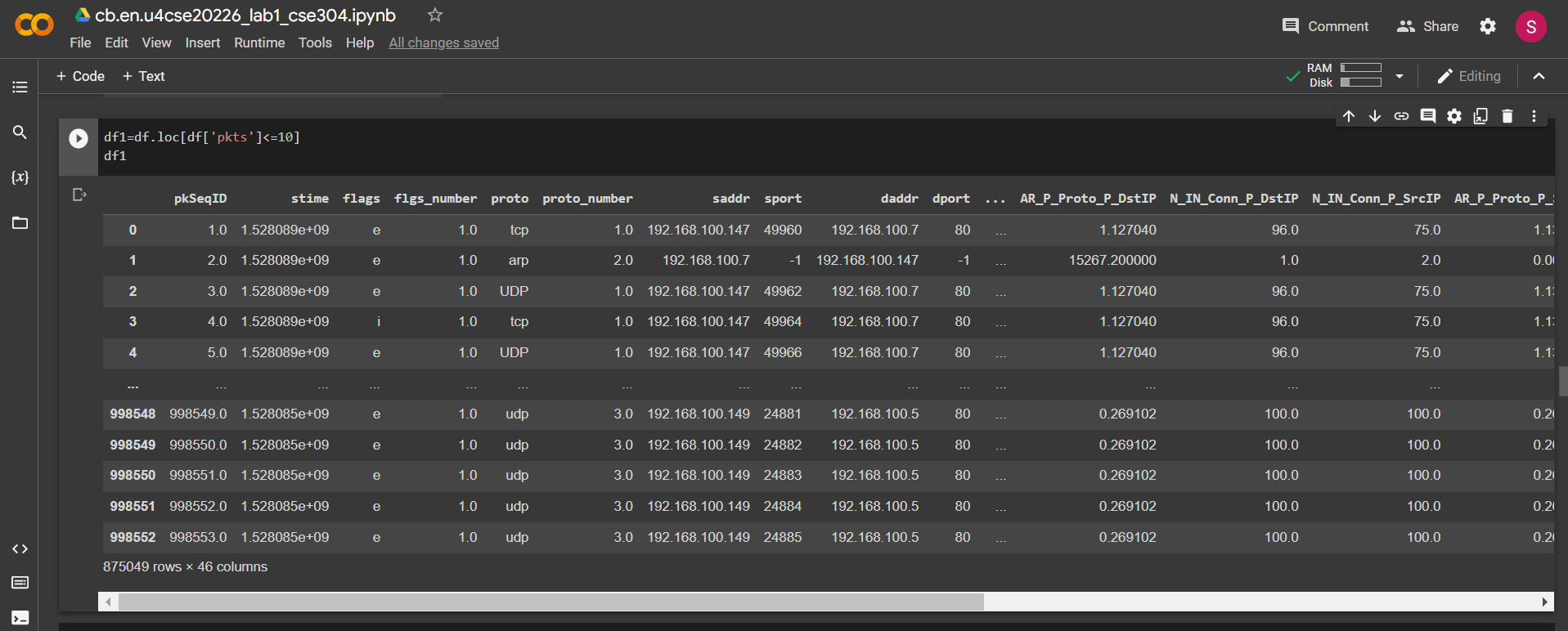
12) Count the number of packets affected based on flags as “i”(2)

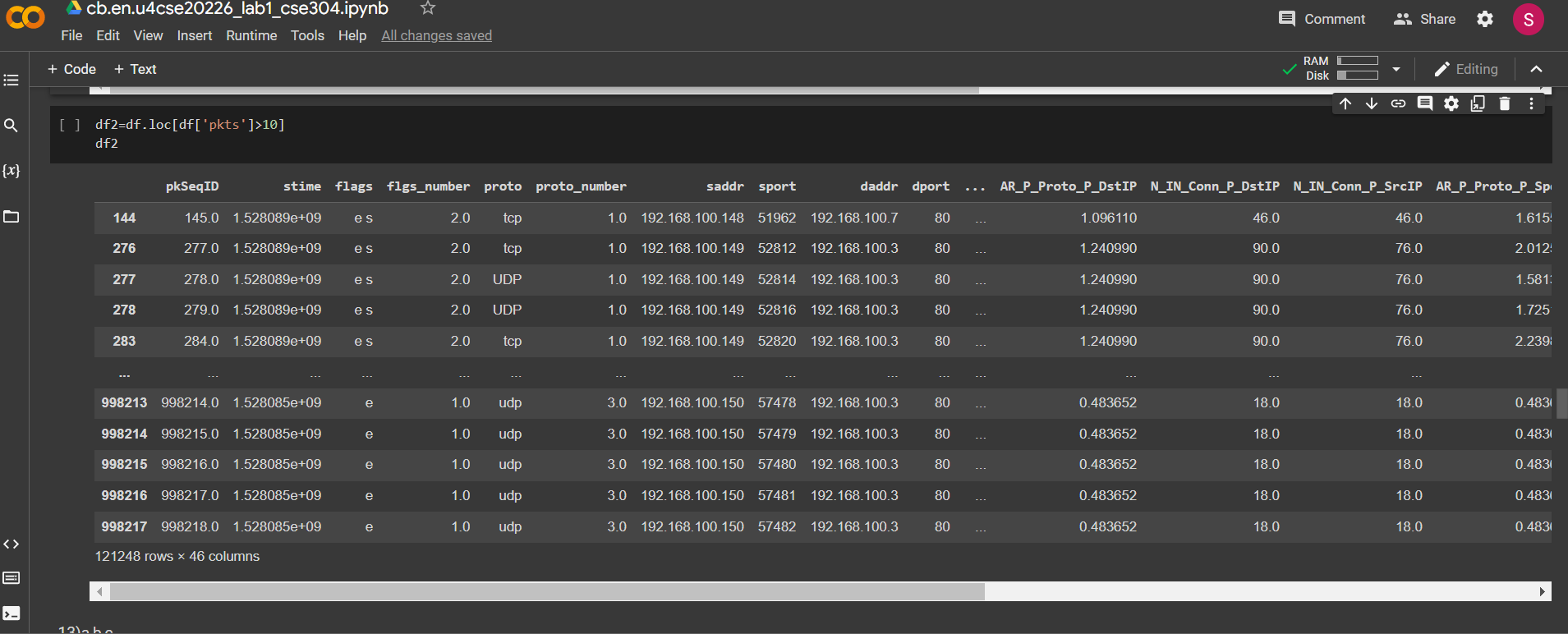


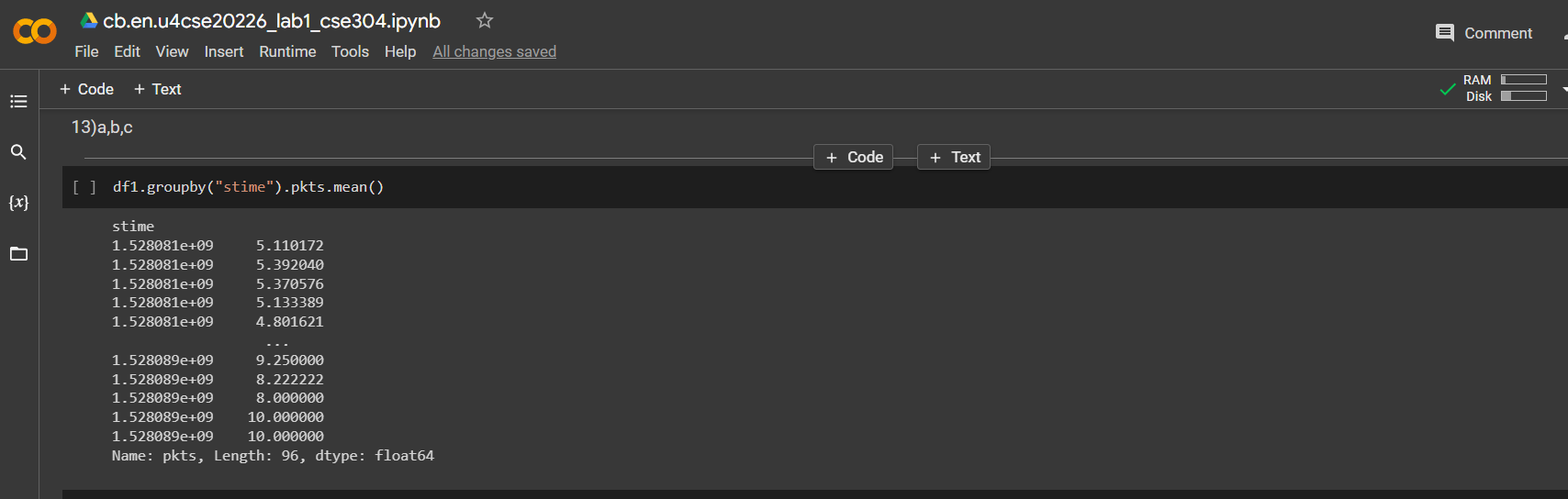
13.Split the traffic based on traffic into two groups Group 1: pkts<=10 and Group 2:pkts>10 (2)

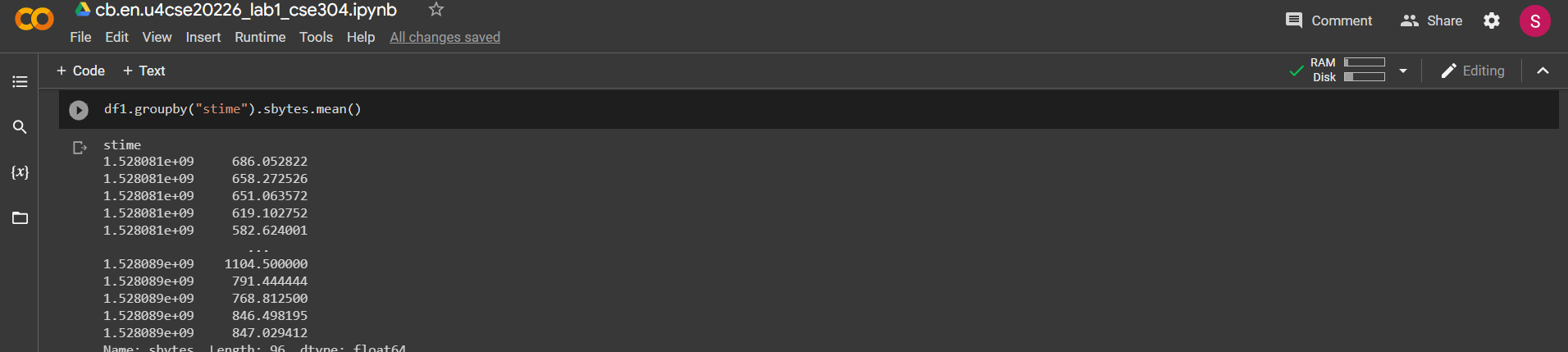
13Evaluate the characteristics of the packets distribution with mean and, deviation

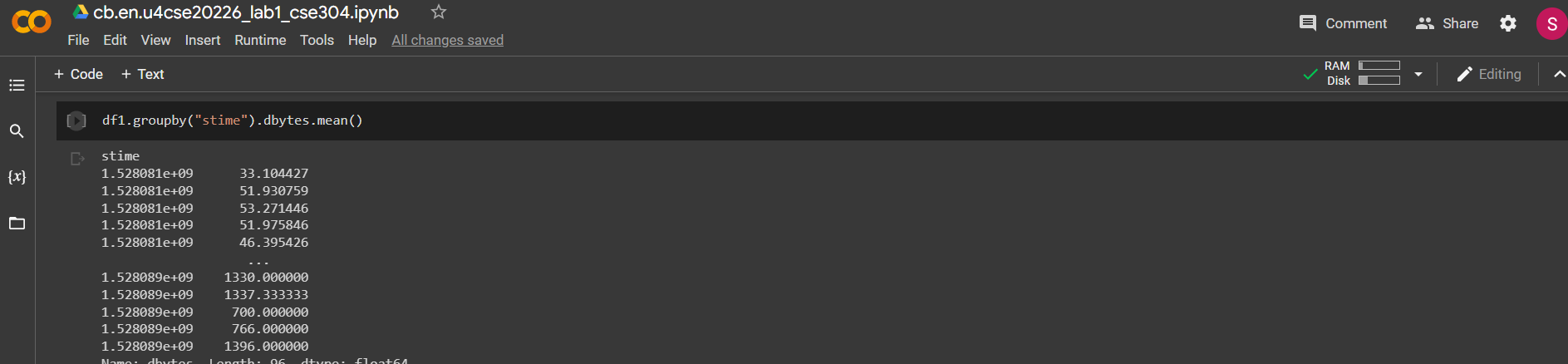
1. Average packets for a given stime (b) Average sbytes for a given stime (c) Average dbytes for a given stime (2)



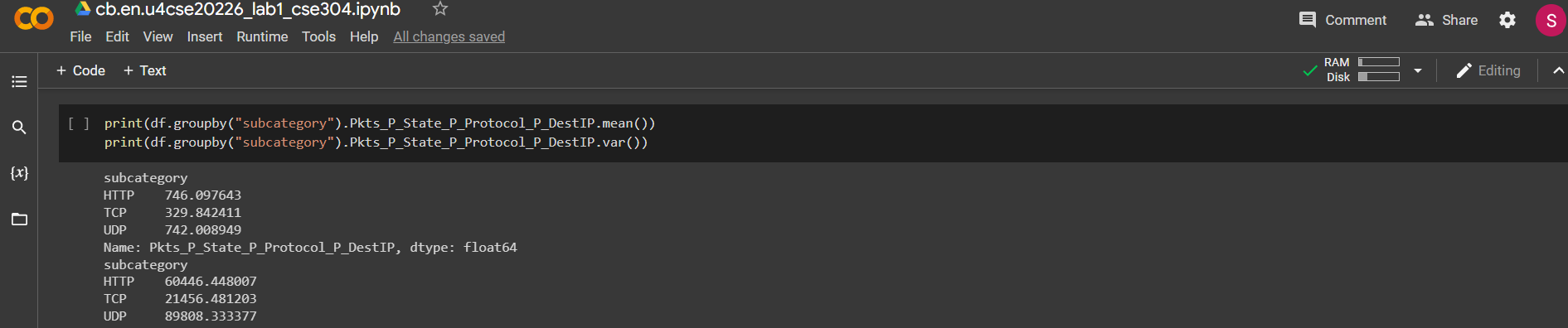








14) Compute the mean and the variance of " Pkts\_PState\_P\_Protocol\_P\_DestIP” for subcategory as HTTP and TCP (1).



15) Draw a histogram for packets with stime as 1528088521, 1528088522., 1528088523. Histogram should be step-filled with bin of size 20. Find the skew of this distribution, and comment if it is positive or negative (5).

