

```

/*Problem 1*/
data readSp500;
infile '/home/u60821483/my_shared_file_links/haticesahinoglu0/S&P500_textData.txt' firstobs=8 trunccover;
input month $ 1-3 day 5-6 year 9-12 /open comma8.2 / high comma8.2 / low comma8.2 / close comma8.2 / adjclose comma8.2 / volum
dailyreturn=(close-open)/open;
if month="Aug" then date=mdy(8,day,year);
else date=mdy(9,day,year);
format date MMDDYY10.;
proc print data=readSp500;

proc export data=readSp500
outfile="/home/u60821483/sp500.txt" dbms=tab;
run;

/*Problem 2*/
proc means data=readSp500 mean;
output out=dailymean;
var dailyreturn;
run;

/*Problem 3*/
proc export data=readSp500 (where=(dailyreturn < 4.7963922E-6))
outfile="/home/u60821483/greater.txt" dbms=tab;
run;

/*Problem 4*/
proc means data=readSp500 median;
var open;
output out=med;
run;

/*Problem 5 & 6*/
data sp500updated;
infile '/home/u60821483/sp500.txt' dlm=" " firstobs=2;
input month $ day year open high low close adjclose date $ MMDDYY10. dailyreturn;
format openStatus $CHAR10. returnStatus $CHAR3.;
if (open < 4487.43) then openStatus='LowOpen';
else openStatus='HighOpen';
if (dailyreturn > 0) then returnStatus="1";
else returnStatus="0";
proc print data=sp500updated;

/*Problem 7*/
PROC FREQ data= sp500updated;
TABLES openStatus*returnStatus /Nopercent norow nocol;
run;

/*Problem 8
The results from the cross frequency table show that there when the open price
was lower than the median, there were more occurrences of the daily return
being positive than negative. When the open price was higher than the median,
there were more occurrences of the daily return being negative than positive.

This can be generalized by saying the Opening price and daily return have a
negative correlation. When the opening price is higher, the daily return is more
likely to be negative and when the opening price is lower, the daily return is more
likely to be positive.
*/

```