Be sure to document all cells with formulas that you type using the =VFORMULA() function.

1. Use the data in 2014HWK1.XLSX, worksheet Data 1, and answer the steps below.

    a. Calculate the summary statistics
    b. Estimate trend regressions
    c. Forecast variables for 5 years using trend regression
    d. Calculate the correlation matrix for the variables
    e. Calculate the covariance matrix for the variables
    f. Calculate the square root of the correlation matrix; call it R
    g. Transpose the R matrix
    h. Calculate the product of RR', check if it equals the original correlation matrix
    i. Sort the original data on column 2 from low to high, carrying along all the other columns
    j. Convert the matrix of data to a column vector
    k. Next reverse the steps in j and create a matrix from the column vector
    l. Create a diagonal matrix (10 x 10) of 1's

2. Use the data in 2014HWK1.XLSX, workbook Data 3.
    Use Theta for the univariate parameter estimator to calculate the parameters for the Y variable. Add the EMP distribution and be sure to add in the common random number in the EMP function. Simulate the distributions for 500 iterations and then use CDFDEV to determine the best distribution to simulate the random variable Y. Report the CDFDEV stats and the UPES parameters in a table on one page.

3. Use the data in 2014HWK1.XLSX, workbook Data 2 and run the regression below.

    a. Estimate the multiple regression model using the Simetar function
    b. Forecast the dependent variable for three years given these X's and report the Y values:

	Year 1	Year 2	Year 3
X_1	4	3	3.5
X_2	74	75	76
X_3	40	41	42
X_4	9.7	9.7	9.7
X_5	3.9	2.9	2.7

    c. Simulate the dependent variables for 3 years using the Y-hats in step b using the std. dev. of the residuals, assuming the residuals are normally distributed. Program your own formulas using Y-hat + Std Dev * SND, i.e., do not use the random value cells (NORM) in the regression, I want you to do it on your own.

d. Repeat step c using an approximation of the standard error of prediction, calculated as:
   $\text{Y-tilda}_t = \text{Y-hat}_t + \text{std. dev.} * (\text{Y-hat}_t / \text{mean}) * \text{SND}$.
e. Report the summary statistics for the 6 simulated variables in steps c and d in a table that is easy to read.
f. Report the PDF and CDF of the third year's values in steps c and d.

4. Develop a simple simulation model to estimate the PDF and CDF of total net returns for a business. This is a fun example of what can be done easily with simulation. Assume it is a farm with three crops. Since mean and std dev are given for the random variables use TNORM when a minimum and/or maximum are provided.

	X_1	X_2	X_3
Acres Planted to each Crop			
	25	50	40
Production (bushels/acre)			
Mean	60	20	40
Std. dev.	30	5	12
Minimum	0	0	0
Price			
Mean	7.00	6.00	4.50
Std. dev.	1.50	0.5	0.55
Minimum	0	0	4.00
Maximum			6.00
Variable Cost/Bushel			
	3.00	6.00	4.00
Fixed Cost ($/acre)	100.	90.	30.

KOVs:
Calculate total revenue
Calculate total costs
Calculate total net returns

Simulate the model and report summary statistics for all stochastic variables and the KOVs. Present the CDF and PDF for the net returns distribution.

5. I want you to test the power of Latin hypercube simulation for reducing the number of iterations necessary to reduce the number of iterations required to get an accurate sample. Do the following 5 simulations using UNIFORM(0,1) and RAND():
   - 100 iterations
   - 200 iterations
   - 500 iterations
   - 1000 iterations
   - 5000 iterations

For each sample print the summary statistics and a CDF of the results one the same sheet of paper. Actually if you can put two or more samples on one sheet I would prefer it.