Quiz Day 25, MAT 128, Spring 2019 Name:

You are allowed one 8 1/2" x 11" sheet of paper with hand-written notes on both sides.

1) The CSV file `cunyHistEnrollment.csv` contains the historical enrollment for several CUNY colleges.

`cunyHistEnrollment.csv`:

```
The data in this file comes from the CUNY data book:,,,,
http://www.cuny.edu/irdatabook/rpts2_AY_current/ENRL_0023_UGGR_GEN_HIST.rpt.pdf,,,,
Year,Brooklyn,City,Hunter,Lehman
1990,12000,11206,14866,8620
1995,11332,10686,13980,8022
2000,10094,8155,15421,6922
2005,11364,9418,15631,8442
2010,12804,12263,15684,9841
2015,14207,13201,16550,10800
2017,14689,13210,16844,11978
```

(a) Write a piece of code to read in this file and compute the correlation matrix of all columns.

(b) The correlation matrix for the columns `Brooklyn`, `City`, `Hunter`, and `Lehman` is shown below.

	Brooklyn	City	Hunter	Lehman
Brooklyn	1.000	0.954	0.728	0.989
City	0.954	1.000	0.522	0.921
Hunter	0.728	0.522	1.000	0.766
Lehman	0.989	0.921	0.766	1.000

(i) Which two columns are the most correlated?

(ii) Which two columns are the least correlated?

2) Suppose the mean weight of a certain breed of dog is 25.4 lbs with a standard deviation of 2.7. A researcher randomly selects a sample of 40 dogs of that breed, weighs them, and computes their mean weight. What distribution does this sample mean come from and why? Be as precise as possible.