STAT 35A HW1 Solutions

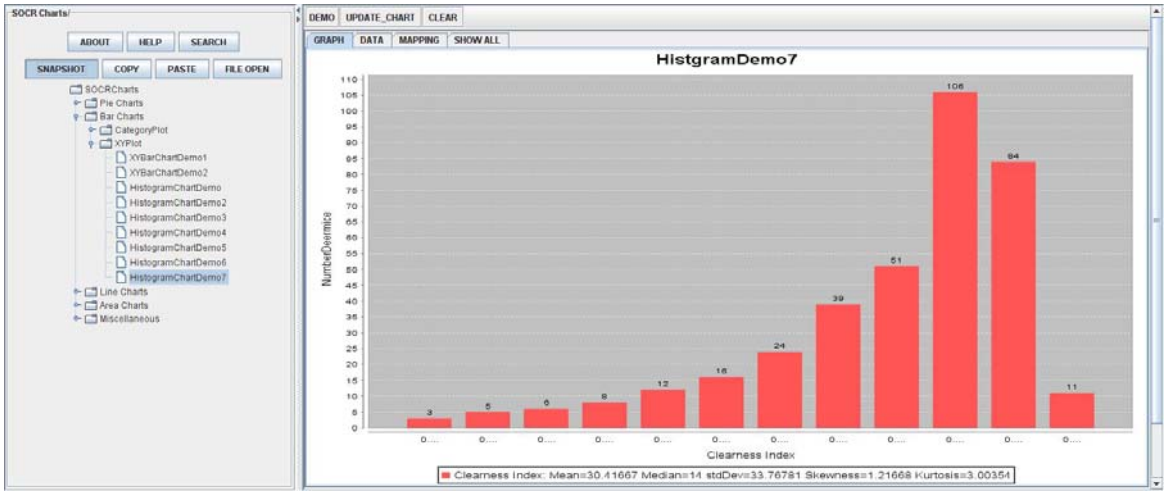
http://www.stat.ucla.edu/~dinov/courses_students.dir/08/Spring/STAT35.dir

Problem 1

- Determine the Relative Frequency and the Cumulative Relative Frequency

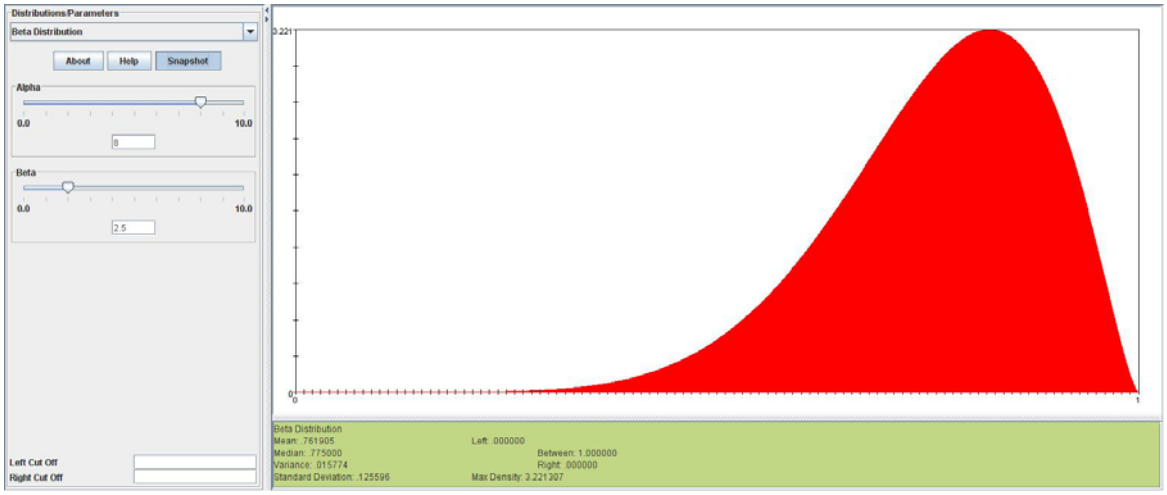
Class interval for the clearness index	Number of days	Relative Frequency	Cumulative Relative Frequency
0.16-0.20	3	0.0082	0.0082
0.21-0.25	5	0.0137	0.0219
0.26-0.30	6	0.0164	0.0384
0.31-0.35	8	0.0219	0.0603
0.36-0.40	12	0.0329	0.0932
0.41-0.45	16	0.0438	0.1370
0.46-0.50	24	0.0658	0.2027
0.51-0.55	39	0.1068	0.3096
0.56-0.60	51	0.1397	0.4493
0.61-0.65	106	0.2904	0.7397
0.66-0.70	84	0.2301	0.9699
0.71-0.75	11	0.0301	1

- Sketch the Relative Frequency histogram



- Visually choose a model distribution for these data using SOCR Distributions, compute and enter the corresponding model probabilities for each range in the last column

  There is no wrong answer to this question, one possible solution is the beta distributionas shown below setting α=8 and β=2.5



- Cloudy days are those with the clearness index < 0.35. What proportion of the days were cloudy? How different are the data and model probabilities?
  6.03% of days were cloudy. Answer varies depending on the model used.

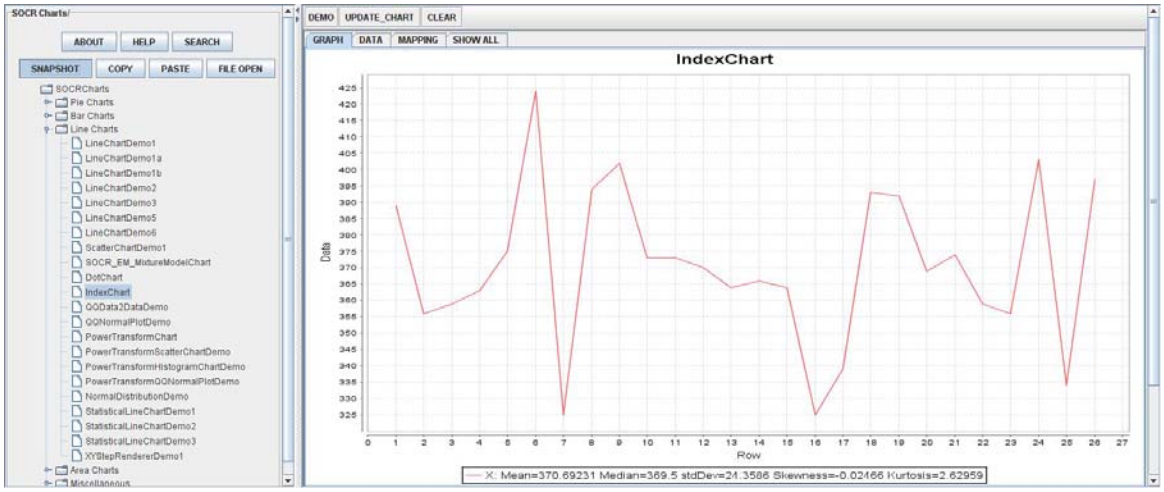
- Clear days are those for which the clearness index is at least 0.66. What proportion of the days were clear?
  26.03% of days were clear.

Problem 2

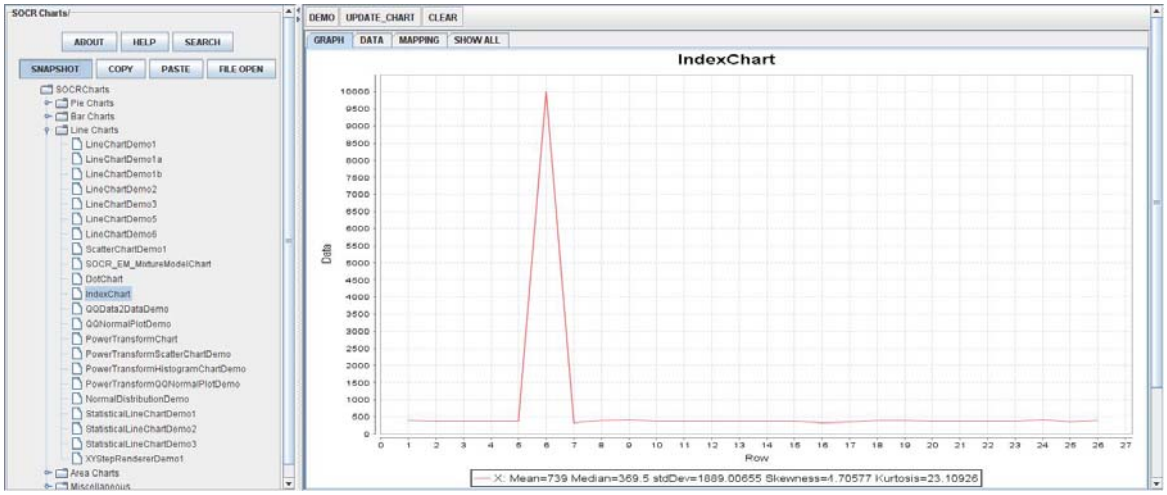
- How different are the sample mean and median?
  Mean = 370.69, Median = 369.5. The mean and median are almost the same.

- By how much should the largest time be increased so that the sample median is half the sample mean?
  The largest time needs to be increased by (739 - 370.69) × 26 = 9576.

- Use SOCR Charts to interactively compute/validate your answers



After changing the largest time (sixth observation) from 424 to 10000.

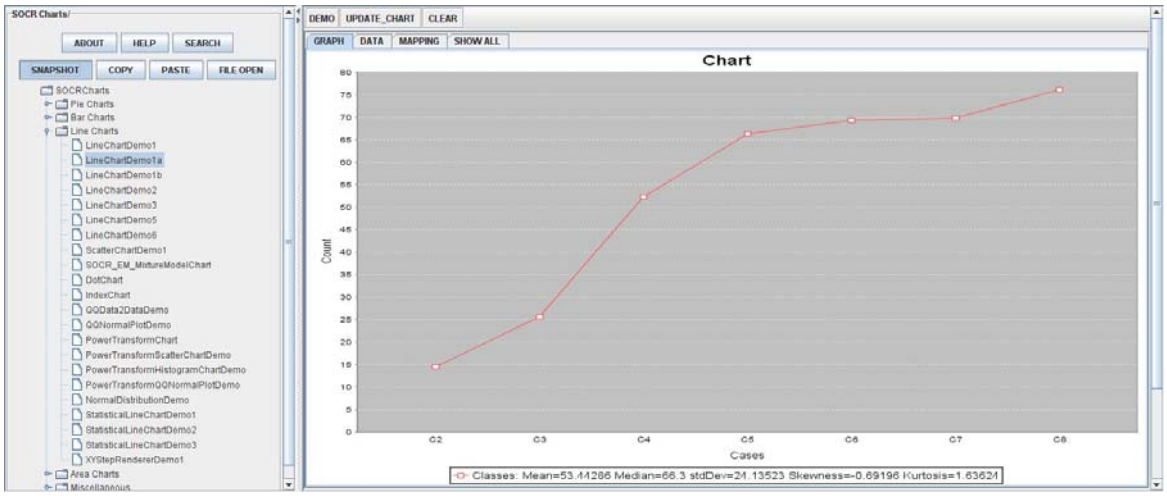


Problem 3

- What is the five-number summary for this data?
  (14.5, 25.6, 66.3, 69.8, 76.2).

- Calculate the following sample measures of spread: variance, standard deviation and the mean-absolute-deviation.
  Variance=582.51, Standard Deviation=24.14, Mean Absolute Deviation=19.38.

- Validate your results using SOCR Charts.



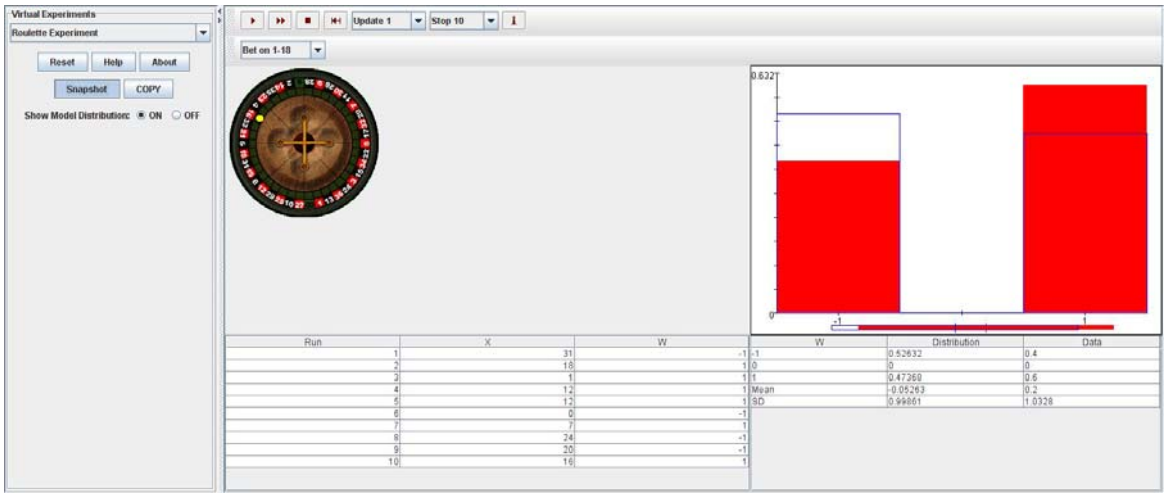
Problem 4

- Should all 38 possible outcomes occur the same number of times? Why?
  It is not possible for all 38 number to appear in just 10 trials, so cannot be the same.

- Does it appear as if some outcomes are just too frequent and some are too rare?
  It should not.

- How large is the difference between the even and odd outcomes in your 10 experiments? Is this expected to vary for different students?
  Answer varies, should be in the range of 0-6.

- Would the answers to the above questions change if we did 1,000 experiments, instead of just 10?
  If we run 1,000 trails, the 38 possible outcomes should appear roughly equal number of times, there shouldn't be any number that occur significantly more (or less) number of times than the rest. The difference between even and odd occurrence should approximately be in the range of 0-40.



Using the Spinner Experiment, you can propose a loaded-Roulette experiment. The image below shows a loading of the probabilities of the small numbers (1, 2 and 3) and the result of 10 trials: