

```

1  *// Open 2012 MAP-Works data
2  use "/Users/Brad/Documents/SAU Materials/MATH 300/M300 2013/Student
   Activities/Site/17 data/mapworks.dta"
3  describe
4
5  *// histogram and kernel density plot of HSGPA
6  histogram s1_p47hsgpa
7  dotplot s1_p47hsgpa
8  kdensity s1_p47hsgpa
9  graph box s1_p47hsgpa
10
11 *// Get summary statistics
12 summarize s1_p47hsgpa, detail
13 robmean s1_p47hsgpa
14
15 *// Proportion of students with HSGPA > 3.5
16 tabulate s1_p47hsgpa if s1_p47hsgpa>3.5
17
18 *// What's the probability we have a student with HSGPA > 3.5?
19 *// Best estimate = 221/580 = 38.1%
20 *// Normal approximation = .33.9%
21 mata
22   1-normal((3.5-3.270719)/.5527017)
23 end
24
25
26   *// Let's take 5,000 samples of size n=5 and calculate the mean of
   each sample.
27 bootstrap (location: mean=r(mean) median=r(p50)) (scale: sd=r(sd)
   range=(r(max)-r(min))), reps(5000) nodots size(5) saving(/Users/Brad
   /Documents/SAU Materials/MATH 300/M300 2013/Student Activities/Site/
   17 data/size5.dta, replace) bca nodrop nowarn : summarize
   s1_p47hsgpa, detail
28
29   *// Open all those sample means and graph
30 use "/Users/Brad/Documents/SAU Materials/MATH 300/M300 2013/Student
   Activities/Site/17 data/size5.dta", clear
31 kdensity location_mean
32
33 *// Is this sampling distribution approximately normal?
34 histogram location_mean, normal
35 qnorm location_mean
36
37 *// Does the CLT hold in terms of mean and standard error?
38 *// They should be: 3.270719 and 0.5527017/SQRT(5) = 0.247176
39 summarize location_mean, detail
40
41 *// What's the probability we have an average HSGPA > 3.5 for 5

```

```

students?
42 *// Best estimate = whatever we get below / 580
43 tabulate location_mean if location_mean>3.5
44 *// Normal approximation = 17.7%
45 mata
46   1-normal((3.5-3.270719)/.247176)
47 end
48
49
50 *// What do the other sampling distributions look like?
51   kdensity location_median
52   kdensity scale_sd
53   kdensity scale_range
54
55
56 *// Let's take 5,000 samples of size n=144 and calculate the mean
of each sample.
57 use "/Users/Brad/Documents/SAU Materials/MATH 300/M300 2013/Student
Activities/Site/17 data/mapworks.dta", clear
58 bootstrap (location: mean=r(mean)), reps(5000) nodots size(144)
saving(/Users/Brad/Documents/SAU Materials/MATH 300/M300 2013/
Student Activities/Site/17 data/size25.dta, replace) bca nodrop
nowarn : summarize s1_p47hsgpa
59
60 *// Open all those sample means and graph
61 use "/Users/Brad/Documents/SAU Materials/MATH 300/M300 2013/Student
Activities/Site/17 data/size25.dta", clear
62 kdensity location_mean
63
64 *// Is this sampling distribution approximately normal?
65 histogram location_mean, normal
66 qnorm location_mean
67
68 *// Does the CLT hold in terms of mean and standard error?
69 *// They should be: 3.270719 and  $0.5527017/12 = 0.046$ 
70 summarize location_mean, detail
71
72 *// What's the probability we have an average HSGPA > 3.5 for 5
students?
73 *// Best estimate = 0
74 tabulate location_mean if location_mean>3.5
75 *// Normal approximation = .0000033%
76 mata
77   1-normal((3.5-3.270719)/.046)
78 end
79
80

```