## Getting means (& 95% CI) plots from SPSS via Excel

prepared grudgingly by Robert Grant (robertgrantstats.co.uk) who dislikes both packages

When you have a continuous outcome variable and are comparing it across various groups, it's often informative to plot the mean and its confidence interval in each group. However, this is not directly possible in SPSS. If you go to ANOVA:







because it doesn't show confidence intervals: how uncertain are those means? Especially when some groups are smaller than others



## So, I would click on Data – Split file first:

ta Split File	×
Current Status: Organize ou	<ul> <li>Analyze all cases, do not create groups</li> <li><u>Compare groups</u></li> <li><u>Organize output by groups</u></li> <li><u>Groups Based on:</u></li> <li><u>Groups Based on:</u></li> <li><u>Sort the file by grouping variables</u></li> <li><u>File is already sorted</u></li> <li>the the file by grouping the the the file by group</li> </ul>

then whatever you do, it will repeat for each of the categories in the predictor variable. Don't forget to turn off split file when you're finished!

Frequencies		×
group	Variable(s):	Statistics Charts Eormat Style Bootstrap
Display frequency tables	3	
ОК	easte Reset Cancel He	elp

Then I would go to Analyze – Descriptive – Frequencies:

## and click on statistics

Frequencies: Statistics	×								
Percentile Values Quartiles Cyt points for: 10 equal groups Percentile(s): Add Change Remove	Central Tendency Mean Megian Mode Sum								
	Values are group midpoints								
Dispersion	Distribution								
Std. deviation 🕅 Minimum	🔲 Ske <u>w</u> ness								
🔲 <u>V</u> ariance 📄 Ma <u>x</u> imum	🔲 <u>K</u> urtosis								
🔲 Ra <u>n</u> ge 🛛 S. <u>E</u> . mean									
Cancel Help									

then 'continue', and then click on 'Bootstrap' (do you have a bootstrap button here? If not, either hand over the \$\$\$ to get it from IBM, or just ignore the bootstrap stuff that follows (or get better stats software!))

The Bootstrap										
Perform bootstrapping										
Number of samples: 1000										
Set seed for Mersenne Twister										
See <u>d:</u> 2000000										
Confidence Intervals										
Level(%): 95										
Percentile										
◎ <u>B</u> ias corrected accelerated (BCa)										
Sampling										
◎ Si <u>m</u> ple										
◎ Stratified										
Variables: Strata Variables:										
scale										
►										
Continue Cancel Help										

Tick 'Perform bootstrapping' and then Continue. Don't forget to untick 'Display frequency tables'! Then when you click OK, you will get the mean, its standard error (from which you can calculate a classical approximate 95% CI) and the bootstrap CI (which helps when you have non-normal data with a ceiling effect).

Now copy those numbers into Excel, and add a formula to get the classical CI like this:

	🕱 🕞 🕫 🖓 🕫 🖓 🐨 🖓 🐨 🖓 🐨 🖓 🐨 🖾 🔤 🔤 🔂											• X			
F	le H	ome Inse	rt Page	Layout F	ormulas	Data F	Review	View Fox	t Reader PD	DF				۲ م	- 6 X
	<b>X</b>	Calibri	- 11	• A A	= =	= %		General	Ŧ	Condit	ional Formatti	ing 🔹 🖁	a•⊐ Insert →	Σ·Α	A
Par	- 🗈 🗝	D 7 11		A. A.				- N -	≪-0 .00	Format	as Table 👻	3	👫 Delete 👻	Sort 8	Eind &
1	° 🛷 –	na <u>n</u>	* <u>-</u>	ON TA T	= =		- <u></u>	· %	.00 ⇒.0	🖳 Cell Sty	les *	l	Format 👻	∠ Filter	Select ▼
Clip	board 🕞		Font	5	i A	lignment	Es.	Number	Fa		Styles		Cells	Editi	ng
	VLOOKUP → ★ ★ ★ =D3-(1.96*E3)														
	А	В	С	D	E	F	G	Н	1	J	К	L	М	N	0
1			bootstrap			classical									
2		upper Cl	lower CI	mean	SE	lower CI	upper Cl								
3	A	3.47	2.33	2.93	0.284	.96*E3)	3.4866	4							
4	В	3.4	2.4	2.9	0.277	2.35708	3.4429	2							
5	С	3.06	2.12	2.59	0.243	2.11372	3.0662	8							
6															
-															
/															

the lower CI is mean – (1.96\*SE) and the upper CI is mean + (1.96\*SE). Note the odd way of writing upper CI, then lower CI, then mean. That's how Excel will want it for the chart. If the classical and bootstrap options are notably different, you may find that the bootstrap avoids the upper or lower limit of the outcome (floor/celiling effect), which is good. But if you have very few data or some suspicious outlier, it might be best to ignore the bootstrap. Otherwise, the choice doesn't matter much. If you're worried about these CIs, talk to a statistician (and I mean statistician here, not any old computer science graduate friend with a hip job title like Data Samurai).

Now, highlight the numbers in upper, lower and mean, go to Insert and choose other chart types, where you should see the "high-low-close" option



## This will read the data in columns but we need to switch it to rows:



Right click on the chart, go to Select data, and then click 'Switch Row/Column'. Now it should look pretty good and just needs a bit of formatting:



Mission accomplished; you now have a plot of the mean and its 95% CI in each group.