

UCLA Statistical Consulting

Phil Ender

UCLA Statistical Consulting Group
IDRE Technology Services
www.ats.ucla.edu/stat/seminars/special/special_stat.pdf

June 11, 2014

Who do we help?

- Members of the UCLA research and teaching community...
- Faculty
- Staff Researchers
- Graduate Students (including Medical Students)
- Physicians, Residents, Interns, etc.

What do we charge?

- We don't charge for our services,
- It's free...
- Zero...
- Zip...
- Zilch...
- Bupkis...
- Naught...
- Nil...
- Nada.

Who are we?

- 4 Full-time Consultants
- 2 Half-time Consultants

What can we help with?

- Stat package use
- Applied statistics
- Data analysis consulting
- Research design
- Research planning
- Data management
- Graphical presentation of data and results

What are our services?

- Webpages
- Walk-in consulting
- Periodic presentations

Webpages

- www.ats.ucla.edu/stat/
- Online classes and seminars (including movies)
- Frequently Asked Questions
- Data Analysis Examples
- Annotated Output
- News, What's New on Our Web, blog
- Over 45M hits in 10 years
- Approx 1M hits per month

Walk-in consulting

- Math Science 4919
- Monday 1:00-4:00
- Tuesday 1:00-4:00
- Wednesday 1:00-4:00
- Thursday 1:00-4:00
- Friday no walk-in consulting
- Clients are restricted to two visits per week

Supported stat packages

- The big four: SAS, SPSS, Stata, R
- Specialized software: Mplus, HLM, SUDAAN, others
- StatTransfer
- We do not support Excel other than for data input

Things we won't do

- We do not do any type of homework or classwork
- We cannot work on any file containing patient identifying information;
- i.e., no patient names or patient id numbers
- We do not assist non-UCLA people
- We do not do data analysis for clients

Our statistical consulting model

- Help the client to do their own data analyses
- Get them started
- Show examples
- Work through one analysis with them
- Have them do one analysis on their own
- Give them homework

Please Note:

We cannot help you in walk-in consulting if your data are not cleaned, organized and ready to move into a statistics package.

Also, consulting progresses much more smoothly if you know and understand your data.

1st Commandment of Data Analysis

Know thy data.

Biggest consulting problems

- The question the client asks is not the "real" question
- Figuring out what the "real" question is
- Clients not knowing their own data or their own research
- Getting clients to switch to a better package for their specific analysis needs
- Clients with tight deadlines ... "I need this analysis today!"

Contact UCLA Statistical Consulting

- Visit: Math Sciences 4919
- Web: www.ats.ucla.edu/stat

Concerning Excel

- Data are best analyzed using a statistics package
- Excel is not a statistics package
- In general, we don't know how to do data analysis in Excel
- Excel can be good for data entry
- Excel can also be evil at times

Good Excel Practices

- Columns are variables; rows are observations
- Observations can be subjects or patients
- Put variable names in first row for each column
- Variable names need to be unique
- Use one word variable names (use underscore if needed)
- Try to use only numeric data, no text or string data
- do not mix numeric and text data in the same column
- Do not include any kind of patient identifier
- including patient id, name, ssn, etc
- Create dummy or fake ids

More Good Excel Practices

- Don't use colors to indicate groups
- Have all data for all groups on one sheet
- Don't include graphs
- Don't include averages, totals or subtotals
- Code variables with 2 levels as 0 or 1
- Name binary variables by the 1 category; example female
- Leave cells with missing data blank
- webpage with Excel tips:
- `www.ats.ucla.edu/stat/mult_pkg/faq/general/Excel_file_set_up.htm`

Demonstration Time

- Spreadsheet Examples

Suboptimal Spreadsheet

id#	sex	gender	heart rate	anx score	Q#1	Question 2	1st test
Control Group							
104	2	female	73	12.1	1	10	78
27	1	Male	80	9.8	3	8	82
Smith, Bob	2	female	69	11	5	7	80
1102	1	male	84	9.2	4	7	91
87	2	woman	64	6.05	2	6	77
276	2	Female	70	15	2	9	57
Average			73.33	10.53			
Treatment Group							
Mary Brown	2	_female_	82	13.1	-9	5	92
892	2	female	85	8.7	3	6	95
415	1	male	112	18.4	2	99	89
127	2	fem	71	12.6	1	4	84
B329	1	Male	90	final exam	3	6	81
67	unk	N/A	77	17.8	3	7	79
104	2	Female	71	8.1	4	8	90
78	1	man	90	7.7	1	9	95
Average			84.75	12.34			

Approved Spreadsheet

id	female	gender	heart_rate	anx_score	Q1	Question_2	test1	treatment
104	0	female	73	12.1	1	10	78	0
27	1	male	80	9.8	3	8	82	0
612	0	female	69	11	5	7	80	0
1102	1	male	84	9.2	4	7	91	0
87	0	female	64	6.05	2	6	77	0
276	0	female	70	15	2	9	57	0
613	1	female	82	13.1		5	92	1
892	1	female	85	8.7	3	6	95	1
415	0	male	112	18.4	2		89	1
127	1	female	71	12.6	1	4	84	1
B329	0	male	90		3	6	81	1
67			77	17.8	3	7	79	1
104	1	female	71	8.1	4	8	90	1
78	0	male	90	7.7	1	9	95	1