Week 2 - Lab 2: Spreadsheet Design Rules

Prepared by: Ealaf Selim

CPSC 203 - T16

Winter 2009



Overview

- This Tutorial includes the following topics:
 - Some basic rules about design
 - Parts of a well designed spreadsheet
 - More Examples

Basic Design Rules

1. Design first on paper:

Graph paper often works well.

2. Test and edit your calculations:

- use intermediate calculations
- check-sums to ensure calculations are correct

3. Keep the components of a calculation visible:

- No "magic numbers".
- Place fixed numbers used in a calculation in their own cell with a descriptive title.

4. Be aware of the geography of the spreadsheet:

 Arrange your information so that it is well spaced and easy to take in at a glance.

Well-Designed Spreadsheet

Components of a well-designed spreadsheet:

– Introduction:

- What is this spreadsheet or workbook about?
- Note the title, purpose, author, creation and revision dates etc.

– Model and Assumptions:

 Justify any models, summary statistics, or calculated variables you are using.

Well-Designed Spreadsheet

Components of a well-designed spreadsheet:

– Data Dictionary:

- For every <u>variable</u> in the spreadsheet note: its
 - Location (cell range),
 - Name,
 - The **Data Class** it is (Raw Data, Statistical Summary, Calculated Variable, Score etc.),
 - Data Type (e.g. Integer, Text, Currency, Date, etc.) and
 - Description (a description of the data or what it's 'purpose' is).

– Raw Data:

 Present your raw data in tabular form -- with columns representing variables and rows representing cases.

Well-Designed Spreadsheet

Components of a well-designed spreadsheet:

– Calculated Data:

- <u>Summary Statistics</u>:
 - Usually Summary statistics result from calculations across rows for a single column.
- <u>Derived variables</u>:
 - Often based on calculations across columns for a row.

– Presentation:

- Emphasize the final information you wish to show without excessive background details.
- Use <u>charts</u> wherever appropriate to summarize large volumes of data.

Example

<u>Media:AnalysisExamples</u>
<u>PresidentsDataAndAnalysis 4 20080306.xls</u>



Example – Introduction

	А	В	С	D	Е	F	G	Н	I	J	K
1	Introduction: Po	olitical C	ontrol of	the Eco	nomy						
2											
3	This spreadsheet is I	based on a	Data Set b	y Edward F	R Tufte gath	ered from U	J.S. Preside	ential electio	ons between	n 198 and 1	1976
4											
5	Using Tufte's data we	e look at Inc	umbent Ad	vantage in t	terms of Ch	ange in Rea	al Disposal	ole Income			
6											
7	Reference:	"Political C	Control of th	e Economy	,''						
8		By Edward	d, R. Tufte								
9		Princeton	University F	ress, 1978							
10		pp. 121 to	123								
11											
12	Data Source:	Table 5.5 f	from above	reference,	pp. 121						
13											
14											
15											
16											

Example – Model and Assumptions

			_		_	_	_				1.4				
	А	В	C	D	E	F	G	Н		J	K	L	M	N	
1	Model 8	& Assum _i	ptions												
2															
3	We develo	ped 2 linea	r Modelsto	examine th	e explanato	ory power o	f Change in	Disposable	e Income (v	hich was u	sed as the	independer	nt variable i	n both mode	els).
4															
5	Model1 F	ormula is:	Incumbent	Advantage	= 0.1456 (0	ChangeDisp	osablelnco	me) - 0.081							
6	Model 1	R-Squared	0.1435	_											
7															
8	Model2 F	ormula is:	VoteForInd	cumbent = 2	2.4283(Cha	ngeDisposa	ableIncome) + 46.567							
9	Model 2	R-Squared	0.4131												
10															
11	Assumpti	ons													
12															
		me a linear l													
14	That assu	mption seer	ms valid for	model 2, le	ess valid for	model 1									
15															
16	Sums of E	rrors and M	/leanSquare	dErrors we	re calculate	ed to further	examine e	ach model							
17		<u>]</u>													
18															
19															

Example – Data Dictionary

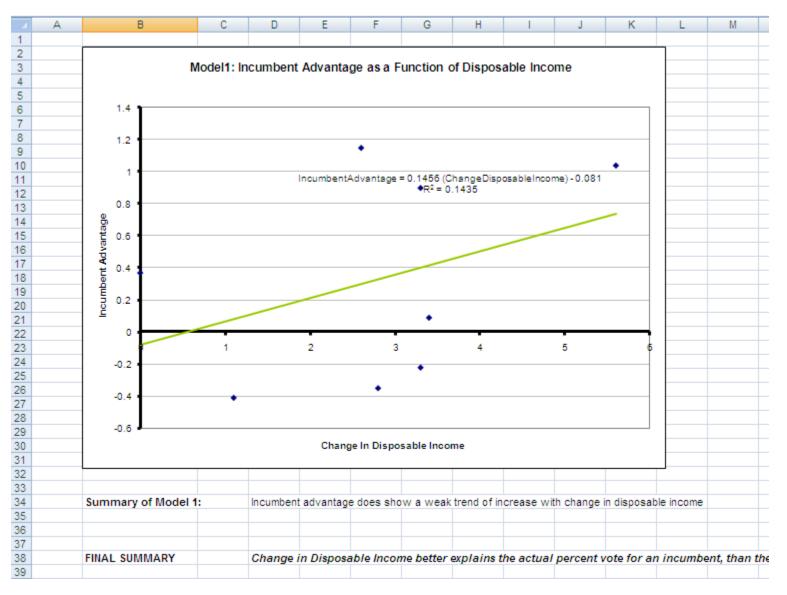
4	А	В	С	D	Е	F	G	Н		J	
1	Location	Name	DataClass	Data Type	Descript	tion					
2	Data!A1:A9	Year	Raw	Date	Year of Ele	ection					
3	Data!B1:B9	Incumbent	Raw	Text	Name of si						
4	Data!C1:C9	VoteForIncumbent	Raw	Percentage	Percentage	e of vote for	incumbent				
5	Data!D1:D9	ChangeDisposableIncome	Raw	Percentage	Yearly cha	nge in real	disposable	income pe	r capita		
6		IncumbentAdvantage	Raw	Score	Net preside	ential candi	date advan	tage. Positi	ve values ir	ndicate the	non-i
7	Data!G1:H9	Model1PredictionsforIncumbentAdvantage	Calculated	Numerical	Incumbent	Advantage	as predicte	ed by Mode	l1		
8	Data!H1:H9	Model2Predictionsfor VoteForIncumbent	Calculated	Numerical	Vote For In	cumbent a	s predicted	by Model2			
9	Data!L1:L9	Model1Errors	Calculated	Numerical	Incumbent	Advantage -	- Model1Pr	edictionsfor	Incumbent/	Advantage	
10	Data!M1:M9	Model2Errors	Calculated	Numerical	VoteForInc	umbent - N	/lodel2Pred	ictionsforVo	teforIncum	bent	
11	Data!:P1:P9	Model1ErrorsSquared	Calculated	Numerical	Square of I	Model1Erro	rs				
12	Data!R1:R9	Model2ErrorsSquared	Calculated	Numerical	Square of I	Model2Erro	rs				
13	Data!!11:M11	SumOfErrors	Statistical Summary	Numerical	Sum of Err	ors for Mod	lel1Errrors	and Model2	Errors		
14	Data!M13:R13	MeanSquaredError	Statistical Summary	Numerical	Mean of the	e Squared	Error for Mo	odel1Errors	Squared an	d Model2E	rrors!
15											

Example – Data

	Α	В	С	D	Е	F	G	
1	Year	Incumben	ChangeDisposableIncome	VoteForIncumben	IncumbentAdvantage		Model1PredictionsforIncumbentAdvantage	Mode
2	1948	Truman	3.4	52.3	0.09		0.41404	
3	1952	Stevensor	1.1	44.6	-0.41		0.07916	
4	1956	Eisenhow	2.6	57.8	1.15		0.29756	
5	1960	Nixon	0	49.9	0.37		-0.081	
6	1964	Johnson	5.6	61.3	1.04		0.73436	
7	1968	Humphrey	2.8	49.6	-0.35		0.32668	
8	1972	Nixon	3.3	61.8	0.9		0.39948	
9	1976	Ford	3.3	48.9	-0.221		0.39948	
10								
4.4								

Н	- 1	J	K	L	M	N	0	Р	Q
2Predictionsf	or VoteFor	Incumben	t	Model1Errors	Model2Err	ors		Model1Err	ors Squa
54.82322				-0.32404	-2.52322			0.105	
49.23813				-0.48916	-4.63813			0.23928	
52.88058				0.85244	4.91942			0.72665	
46.567				0.451	3.333			0.2034	
60.16548				0.30564	1.13452			0.09342	
53.36624				-0.67668	-3.76624			0.4579	
54.58039				0.50052	7.21961			0.25052	
54.58039				-0.62048	-5.68039			0.385	
	SumOfErr	ors		-0.00076	-0.00143				
					MeanSqua	aredErro	r	0.30765	

Example – Presentation



Example – Presentation

