

# Week 2 - Lab 2: Spreadsheet Design Rules

# Introduction

- Basic spreadsheet design rules
- Components of a well-designed spreadsheet

# Basic spreadsheet design rules

- Design the spreadsheet on paper first
- Test and edit your calculations
- Keep the components of a calculation visible
- Be aware of the "space" or "geography" of the spreadsheet

# Well-designed Spreadsheet

- Components of a Well-designed Spreadsheet
  - Introduction
  - Model and Assumptions
  - Data Dictionary
  - Raw Data
  - Calculated Data
  - Presentation/Dashboard

# Spreadsheet Component: Introduction

- What is this spreadsheet or workbook about?
  - Title
  - Purpose
  - Author
  - Creation
  - Revision dates, etc.

# Example: Introduction

The screenshot shows a Microsoft Excel spreadsheet with the following content:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	<b>Introduction: Political Control of the Economy</b>																
2																	
3	This spreadsheet is based on a Data Set by Edward R Tufte gathered from U.S. Presidential elections between 198 and 1976																
4																	
5	Using Tufte's data we look at Incumbent Advantage in terms of Change in Real Disposable Income																
6																	
7	<b>Reference:</b>	"Political Control of the Economy"															
8		By Edward, R. Tufte															
9		Princeton University Press, 1978															
10		pp. 121 to 123															
11																	
12	<b>Data Source:</b>	Table 5.5 from above reference, pp. 121															
13																	
14																	
15																	
16																	
17																	
18																	
19																	
20																	
21																	
22																	
23																	
24																	
25																	
26																	
27																	
28																	
29																	
30																	

The spreadsheet is displayed in the 'Introduction' worksheet, which is highlighted with a red box in the bottom-left corner of the Excel window. The status bar at the bottom indicates the current zoom level is 120%.

# Model & Assumptions

- Justify the calculations you are using
  - any models
  - summary statistics
  - or calculated variables, etc.

# Example : Model & Assumptions

The screenshot shows a Microsoft Excel window with the following content:

Worksheet: **Model & Assumptions**

1 **Model & Assumptions**

2

3 We developed a linear Model with "IncumbentAdvantage" as a Function of "ChangeDisposableIncome"

4

5 **Model Formula is:**  $\text{IncumbentAdvantage} = 0.1456\text{ChangeDisposableIncome} - 0.081$

6 **Model R-Squared is** 0.1435

7

8 Assumptions

9

10 We Assume a linear Model is Appropriate

11 Given the dispersion of the data, and low R-Squared -- this assumption may not be valid

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

The interface includes the ribbon (Home, Insert, Page Layout, Formulas, Data, Review, View), the formula bar (E28), and the status bar (Ready, 100%).



# Data Dictionary

- For every variable 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

# Example: Raw Data

AnalysisExamples\_PresidentsDataAndAnalysis\_4\_20080306 [Compatibility Mode] - Microsoft Excel

Home Insert Page Layout Formulas Data Review View

Normal Page Layout Page Break Preview Custom Views Full Screen

Workbook Views Show/Hide

Zoom 100% Zoom to Selection

New Window Arrange All Freeze Panes Unhide

View Side by Side Synchronous Scrolling Reset Window Position Window

Save Workspace Switch Windows

Macros

R13 =AVERAGE(R2:R9)

Year	Incumbent	ChangeDisposibleIncome	VoteForIncumbent	IncumbentAdvantage	Model1PredictionsforIncumbentAdvantage	Model2PredictionsforVoteForIncumbent	Model1Errors	Model2Errors	Model1Errors Squa	
1948	Truman	3.4	52.3	0.09	0.41404	54.82322	-0.32404	-2.52322	0.105	
1952	Stevensor	1.1	44.6	-0.41	0.07916	49.23813	-0.48916	-4.63813	0.23928	
1956	Eisenhow	2.6	57.8	1.15	0.29756	52.88058	0.85244	4.91942	0.72665	
1960	Nixon	0	49.9	0.37	-0.081	46.567	0.451	3.333	0.2034	
1964	Johnson	5.6	61.3	1.04	0.73436	60.16548	0.30564	1.13452	0.09342	
1968	Humphrey	2.8	49.6	-0.35	0.32668	53.36624	-0.67668	-3.76624	0.4579	
1972	Nixon	3.3	61.8	0.9	0.39948	54.58039	0.50052	7.21961	0.25052	
1976	Ford	3.3	48.9	-0.221	0.39948	54.58039	-0.62048	-5.68039	0.385	
							SumOfErrors	-0.00076	-0.00143	
								MeanSquaredError	0.30765	

Ready

# Calculated Data

- Summary Statistics
  - Usually Summary statistics result from calculations across rows for a single column.
- Derived variables are often based on calculations across columns for a row.

# Presentation

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

# Example: Presentation

