Tabular Display of Data

Gary W. Oehlert

School of Statistics University of Minnesota

January 29, 2007

Gary W. Oehlert

Tabular Display of Data

Introduction Rules and Advice Summary

Outline

- Introduction
- 2 Rules and Advice
- 3 Summary

- Introduction
- 2 Rules and Advice
- Summary

Gary W. Oehlert

Tabular Display of Data

Introduction Rules and Advice Summary

What we're up against

Getting information from a table is like extracting sunlight from a cucumber.

Farquhar and Farquhar, 1891, p55

Perhaps not that bad, but a challenge.

Our discussion follows Ehrenberg (1977, **JRSSA**) and Wainer (1997, **JEBS**).

Types of tables

Wainer lists four types of tables:

- Exploration
- Communication
- Storage
- Decoration

Gary W. Oehlert

Tabular Display of Data

Introduction Rules and Advice Summary

Communication

From Ehrenberg (1977 The Statistician)

Consumers' (C) and Retailers' (R) ratings of the nutritional and economic values of different foods

omic
R
11
95
51
52
18
10
25
21

^{*}In decreasing order of Consumers' Nutritional Ratings.

Exploration

Table of residuals from additive fit in unreplicated two-way design, rows and columns sorted by increasing marginal effects:

966	878	482	-74	-2251	
871	790	320	-112	-1868	
449	405	793	803	-2449	
-2286	-2072	-1594	-617	6569	
	871 449	871 790 449 405	871 790 320 449 405 793	871 790 320 -112 449 405 793 803	871 790 320 -112 -1868 449 405 793 803 -2449

What does it tell us about data?

Gary W. Oehlert

Tabular Display of Data

Introduction Rules and Advice Summary

Archiving

From Wainer (1997 **JEBS**). Illegibility is practically the point.

		Grade 8 - 1992										
PUBLIC SCHOOLS	Graduate	Graduated College Some Education						Did Not Finish High School		I Don't Know		
	Percentage of Students	Average Proficiency	Percentage of Students	Average Proficiency	Percentage of Students	Average Proficiency	Percentage of Students	Average Proficiency	Percentage of Students	Average Proficienc		
NATION	40 (1.4)	279 (1.4)	18 (0.6)	270 (1.2)	25 (0.8)	256 (1.4)	8 (0.6)	248 (1.8)	9 (0.5)	251 (1.7)		
Northeast	38 (3.1)	282 (4.2)	18 (1.1)	267 (3.0)	26 (2.2)	259 (4.2)	8 (0.9)	246 (4.2)	10 (1.2)	250 (3.3)		
Southeast	35 (1.9)	270 (1.9)	17 (0,8)	263 (2.0)	28 (1.4)	249 (1.9)	12 (1.6)	246 (4.2)	8 (1.0)	248 (4.3)		
Central	42 (2.7)	283 (2.9)	20 (1.4)	273 (1.6)	26 (1.7)	264 (2.3)	4 (0.7)	(***)	7 (0.8)	258 (3.8)		
West	43 (2.9)	279 (2.6)	18 (1,2)	274 (2.6)	19 (1,5)	252 (2.9)	9 (1.1)	248 (2.4)	11 (0.9)	248 (2.9)		
STATES							- 1,	(,	(0.0)	2-0 (2.0)		
Alabama	33 (1.6)	261 (2.5)	18 (0.7)	258 (2.0)	29 (1.1)	244 (1.8)	13 (0,9)	239 (2.0)	7 (0.6)	237 (2.9)		
Arizona	36 (1.5)	277 (1.5)	22 (1.0)	270 (1.5)	21 (0.9)	256 (1.6)	10 (0.7)	245 (2.5)	12 (0.8)	248 (2.7)		
Arkansas	30 (1.1)	264 (1.9)	20 (0.8)	264 (1.7)	31 (1,1)	248 (1.6)	11 (0.7)	246 (2.4)	8 (0.6)	245 (2.7)		
California	39 (1.8)	275 (2.0)	18 (1.0)	266 (2.1)	17 (0.9)	251 (2.1)	10 (0.9)	241 (2.2)	16 (1.1)	240 (2.9)		
Colorado	46 (1.2)	282 (1,3)	19 (0.9)	276 (1.6)	21 (0.9)	260 (1,5)>		250 (2.4)	7 (0.5)	252 (2.6)		
Connecticut	47 (1.3)	288 (1.0) >		272 (1.8)	22 (0.9)	260 (1.8)	6 (0.6)	245 (3.3)	9 (0.6)	251 (2.4)		
Delaware	39 (1.2)	274 (1,3)	18 (1.0)	268 (2.3)	30 (1.0)	251 (1.7)	6 (0.5)	248 (4.0)	8 (0.9)	248 (3.4)		
Dist. Columbia	32 (1.0)	244 (1.7)	17 (0.8)	240 (1.9)	29 (0.8)	224 (1.6)	9 (0.7)	225 (3.2)	12 (0.6)	229 (2.2)		
Florida	39 (1.5)	268 (1.9)	19 (0,7)	266 (1.9)	24 (1.1)	251 (1.8)	8 (0.7)	244 (2.7)	10 (0.7)	244 (3.2)		
Georgia	35 (1.7)	271 (2.1)	18 (0.7)	264 (1.7)	30 (1.2)	250 (1.3)	11 (0.8)	244 (2.2)	6 (0.6)	245 (2.6)		
Hawaii	38 (1.1)	267 (1.5)	15 (0.9) <	266 (1.9)	25 (1.0)	246 (1.8)	6 (0.5)	242 (3.5)	16 (0.8)	246 (2.1)		
Idaho	48 (1.2)	281 (0.9)	20 (0.8)	278 (1.3)	19 (0.9)	268 (1.4) >		254 (2.3)	6 (0.5)	254 (2.8)		
Indiana	33 (1.5)	283 (1.5)	21 (0.9)	275 (1.9)	32 (1.1)	260 (1.6)	8 (0.6)	250 (2.6)	6 (0.5)	249 (3.3)		
lowa	44 (1.4)	291 (1.2)>		285 (1.5)	25 (1.1)	273 (1.3)	4 (0,4)	262 (2.4)	5 (0.4)	266 (2.8)		
Kentucky	28 (1.4)	278 (1.6) 2		267 (1.6)	32 (0.9)	254 (1.6)	15 (0.9)	246 (1.7)	6 (0.4)	242 (2.8)		
Louisiana	32 (1.4)	256 (2.5)	20 (0.9)	259 (1.8)	30 (1,3)	242 (1.6)	10 (0.7)	237 (2.4)	7 (0.6)	236 (3.7)		
Maine	40 (1.5)	288 (1.4)	22 (1.0)	281 (1.5)	26 (1.1)	267 (1.1)	6 (0.5)	259 (2.7)	5 (0.5)	266 (2.6)		
Maryland	44 (1.7)	278 (1.8)	18 (0.9)	266 (1.9)	25 (1.2)	250 (1.8)	6 (0.8)	240 (3.7)	7 (0.5)	245 (3.8)		
Massachusetts	48 (1.5)	284 (1.3)	17 (0.8)	272 (1.8)	21 (1.0)	261 (1.4)	7 (0.6)	248 (3.2)	7 (0.6)	248 (2.6)		
Michigan	38 (1.6)	277 (2.2)	23 (0.9)	271 (2.0)	26 (0.9)	257 (1.7)	6 (0.5)	249 (2.0)	7 (0.6)	248 (3.0)		
Minnesota	48 (1.3)>			284 (1.8)	22 (0.9) ≪	270 (1.8) >	3 (0.4)	256 (4.2)	7 (0.6)	268 (3.0)		
Mississippi	36 (1.7)	254 (1.6)	16 (0.7)	256 (2.0)	29 (1.4)	239 (1.6)	13 (0.8)	234 (1.8)	7 (0.6)	231 (2.8)		
Missouri	36 (1.3)	280 (1.7)	22 (0.9)	275 (1.5)	29 (1.0)	264 (1.6)	8 (0.7)	254 (2.4)	6 (0.5)	252 (2.9)		
Nebraska	46 (1.5)	287 (1.2)	20 (1.0)	280 (1.6)	24 (1.2)	267 (1.7)	4 (0.5)	247 (3.3)	6 (0.6)	256 (3.8)		

Introduction Rules and Advice Summary

Or computer files

- # Number of hawks responding to the "alarm" call
- # Variables are year (1999 or 2000), season (courtship,
- # nestling, fledgling), distance in meters between the
- # alarm call and the nest, number of hawks responding,
- # and number of.

year	season	distance	respond	trials
1	1	100	1	4
1	1	150	2	4
1	1	225	1	4
1	1	325	2	2
2	1	100	6	8

. .

Should be labeled and annotated.

Gary W. Oehlert

Tabular Display of Data

Introduction Rules and Advice Summary

Eye on the ball

Most displays only do one thing well.

To build any effective display we must have a firm notion of purpose. We cannot know what the best answers are unless we know what the questions are. Thus we must first understand what questions will be asked of data. Any discussion of data display in the abstract is pointless.

Wainer (1997 JEBS)

We will concentrate on communication.

Introduction Rules and Advice Summary

Decoration

May have data, but also draws eye

PARTY PATROL CITATIONS

In fall 2006, the party patrol issued a total of 984 citations. The following is a breakdown of those citations:

CRIME	NUMBER OF CITATIONS
Consumption of alcohol by minor	404
Noisy assembly	377
Noisy assembly and minor consumption	181
Loitering with open bottle	9
Disorderly conduct	4
Obstruction of law enforcement	2
Moving violation	2
Public urination	1
Flee from officer, obstruction and disconduct	1
Public urination and minor consumption	1
Littering and minor consumption	1
Loitering with open bottle and minor consumpt	ion 1
TOTAL CITATIONS	984

SOURCE: UMPD

(MN Daily, Jan 19, 2007)

Gary W. Oehlert

Tabular Display of Data

Introduction Rules and Advice Summary

Back to communication

A display for communication should

- Target an audience
- Have a goal (tell a story)
- Make the story obvious
- Be uncluttered
- Cause no pain

It's a lot like oral communication!

Gary W. Oehlert Tabular Display of Data

Gary W. Oehlert

Outline

- Introduction
- 2 Rules and Advice
- Summary

Gary W. Oehlert

Tabular Display of Data

Introduction Rules and Advice Summary

Ehrenberg's Criteria

Strong Criterion for Good Table

The patterns and exceptions in a table should be obvious at a glance.

Weak Criterion for Good Table

The patterns and exceptions in a table should be obvious at a glance once one has been told what they are.

Always meet the weak criterion.

Rules for Communication

Ehrenberg, Wainer, and many others give rules/advice.

We illustrate with examples from their papers.

Remember, we want to communicate, to show a story, which could be

- Big picture
- Trends
- Comparisons
- Typical values
- Atypical values

Gary W. Oehlert

Tabular Display of Data

Introduction Rules and Advice Summary

Before (Ehrenberg)

UK Merchant Vessels in Service							
	1962	1967	1973				
Number							
All vessels	2,689	2,181	1,776				
Passenger	242	173	122				
Dry cargo	1,847	1,527	1,165				
Tankers	600	481	489				
Thousand deadweight tons							
All vessels	26,577	27,488	46,763				
Passenger	1,467	919	349				
Dry cargo	13,990	14,362	20,115				
Tankers	11,120	12,167	26,299				

Gary W. Oehlert Tabular Display of Data

Introduction Rules and Advice Summary

Before (Ehrenberg)

After (Ehrenberg)

UK Merchant Vessels in Service

	<u> </u>		1070
Vessels over 500 tons	1962	1967	1973
Number			
Passenger	240	170	120
Tankers	600	480	490
Dry cargo	1,800	1,500	1,200
All vessels	2,700	2,200	1,800
Deadweight tons (thousands)			
Passenger	1,500	920	350
Tankers	11,000	12,000	26,000
Dry cargo	14,000	14,000	20,000
All vessels	26,000	27,000	47,000

Gary W. Oehlert

Tabular Display of Data

Introduction Rules and Advice Summary

After (Ehrenberg)

Correlation among TV audiences

Programmes		WoS	MoD	GrS	PrB	RgS	24H	Pan	ThW	Tod	LnU
World of Sport	ITV		.6	.6	.5	.3	.1	.2	.1	.1	.1
Match of the Day	BBC	.6		.6	.5	.3	.1	.1	.1	.0	.0
Grandstand	BBC	.6	.6		.5	.3	.1	.2	.1	.1	.1
Prof. Boxing	ITV	.5	.5	.5		.3	.1	.2	.1	.1	.1
Rugby Special	BBC	.3	.3	.3	.3		.1	.1	.1	.1	.1
24 Hours	BBC	.1	.1	.1	.1	.1		.5	.4	.2	.2
Panorama	BBC	.2	.1	.2	.2	.1	.5		.4	.2	.2
This Week	ITV	.1	.1	.1	.1	.1	.4	.4		.3	.2
Today	ITV	.1	.0	.1	.1	.1	.2	.2	.3		.2
Line Up	BBC	.1	.0	.1	.1	.1	.2	.2	.2	.2	

Correlation among TV audiences

		PrB	ThW	Tod	WoS	GrS	LnU	MoD	Pan	RgS	24H
ITV	PrB	1.000	0.106	0.065	0.505	0.474	0.092	0.473	0.168	0.309	0.124
,,	ThW	0.106	1.000	0.270	0.142	0.132	0.189	0.082	0.352	0.064	0.395
,,	Tod	0.065	0.270	1.000	0.093	0.070	0.155	0.038	0.200	0.051	0.244
"	WoS	0.505	0.147	0.093	1.000	0.622	0.079	0.581	0.187	0.297	0.140
BBC	GrS	0.474	0.132	0.070	0.622	1.000	0.085	0.593	0.181	0.341	0.142
"	LnU	0.092	0.189	0.155	0.079	0.085	1.000	0.049	0.197	0.097	0.266
"	MoD	0.473	0.082	0.039	0.581	0.593	0.049	1.000	0.131	0.327	0.122
"	Pan	0.168	0.352	0.200	0.187	0.181	0.197	0.131	1.000	0.147	0.524
"	RgS	0.309	0.064	0.051	0.296	0.341	0.097	0.326	0.147	1.000	0.121
"	24H	0.124	0.395	0.244	0.140	0.142	0.266	0.122	0.524	0.121	1.000

Gary W. Oehlert

Tabular Display of Data

Introduction Rules and Advice Summary

Round Drastically

Use two significant figures

- Don't usually understand more than two digits Budget is \$27,329,681 versus budget is 27 million dollars.
- Rarely justify more than two digits statistically God gave us $1/\sqrt{n}$, but how big must n be for that third digit?
- We rarely care Life expectance 67.14 years; .01 year is about 4 days.

Gary W. Oehlert Tabular Display of Data Gary W. Oehlert Tabular Display of Data

Ehrenberg original

Ehrenberg after rounding

Unemployment in Great Britain (thousands)

	1966	1968	1970	1973
Total unemployed	330.9	549.4	582.2	597.9
Males	259.6	460.7	495.3	499.4
Females	71.3	88.8	86.9	98.5

Unemployment in Great Britain (thousands)

1 7	•	,		
	1966	1968	1970	1973
Total unemployed	330	550	580	600
Males	260	460	500	500
Females	71	89	87	98

Gary W. Oehlert

Tabular Display of Data

Gary W. Oehlert

Tabular Display of Data

Introduction Rules and Advice Summary

Order Rows/Columns Sensibly

Introduction Rules and Advice Summary

Wainer before ordering

Helps organize and facilitate comparison

- Alphabetical (Alabama first!) almost never correct
- Could be by size
- Could be a natural order, such as time
- By interest (rows or columns to compare should be adjacent)

Battery Life in Hours

Battery Ene in Floars									
Battery	Cassette			Portable					
Brand	Player	Radio	Flashlight	Computer					
Constant Charge	5	19	10	3					
Electro-Blaster	10	26	15	4					
Never Die	8	28	16	6					
PowerBat	7	24	13	5					
Servo-Cell	4	21	12	2					

Gary W. Oehlert Tabular Display of Data

Gary W. Oehlert

Wainer after ordering

Battery Life in Hours

	Datter	y Life ili i lot	113	
Battery			Cassette	Portable
Brand	Radio	Flashlight	Player	Computer
Never Die	28	16	8	6
Electro-Blaster	26	15	10	4
PowerBat	24	13	7	5
Servo-Cell	21	12	4	2
Constant Charge	19	10	5	3

Gary W. Oehlert

Tabular Display of Data

Introduction Rules and Advice Summary

Ehrenberg with Summaries

Unemployment in Great Britain (thousands)

(, , , , , , , , , , , , , , , , , , ,					
	1966	1968	1970	1973	Ave.
Total unemployed	330	550	580	600	520
Males	260	460	500	500	430
Females	71	89	87	98	86

Row/Column Summaries

Give a standard for comparison

- Could be mean/median/total/etc
- Give a visual focus
- Provide a standard of "usual"
- An overall summary can also help
- Can highlight for emphasis

Gary W. Oehlert

Tabular Display of Data

Introduction Rules and Advice Summary

Wainer with Summaries

Battery Life in Hours

Dattery Life in Flours						
Battery			Cass.	Port.	Brand	
Brand	Radio	Flash.	Player	Comp.	Averages	
Never Die	28	16	8	6	15	
Electro-Blaster	26	15	10	4	14	
PowerBat	24	13	7	5	12	
Servo-Cell	21	12	4	2	10	
Constant Charge	19	10	5	3	9	
Usage averages	24	13	7	4	12	

Down Columns

Ehrenberg after Transposition

It's easier to compare numbers down columns.

- Numbers are closer
- Digits line up

Gary W. Oehlert

Tabular Display of Data

Introduction Rules and Advice Summary

Layout/Spacing

- Remove excess lines/boxing
- Use space to emphasize groups/gaps
- Excess space breaks adjacency

What is a stem and leaf plot, but a severely rounded table with meaningful spacing?

Unemployment in Great Britain (thousands)

			`	
Year	Male	Female	Total	
1966	260	71	330	
1968	460	89	550	
1970	500	87	580	
1973	500	99	600	
Average	430	86	520	

Gary W. Oehlert

Tabular Display of Data

Introduction Rules and Advice Summary

Wainer with Summaries

Battery Life in Hours

	Datte	JI y LIIC I	ii i louis		
Battery			Cass.	Port.	Brand
Brand	Radio	Flash.	Player	Comp.	Averages
Never Die	28	16	8	6	15
Electro-Blaster	26	15	10	4	14
PowerBat	24	13	7	5	12
Servo-Cell	21	12	4	2	10
Constant Charge	19	10	5	3	9
Usage averages	24	13	7	4	12

Introduction Rules and Advice Summary

40 (1.4) 38 (3.1) 35 (1.9) 42 (2.7) 43 (2.9)

261 (2.5) 277 (1.5) 264 (1.9) 275 (2.0) 282 (1.3) 288 (1.0) > 274 (1.3) 244 (1.7) 268 (1.9) 271 (2.1) 267 (1.5) 281 (0.9)

281 (0.9) 283 (1.5) 291 (1.2) > 278 (1.6) > 256 (2.5) 288 (1.4) 277 (2.2) 290 (1.0) > 254 (1.6) 280 (1.7)

18 (0.7) 22 (1.0) 20 (0.8) 18 (1.0) 19 (0.9) 16 (0.8) 18 (1.0) 17 (0.8) 19 (0.7) 15 (0.9) 21 (0.8) 20 (0.8) 20 (0.8) 22 (1.0) 18 (0.9) 22 (1.0) 18 (0.9) 22 (1.0) 18 (0.9) 17 (0.9) 22 (1.0) 18 (0.9) 17 (0.9) 22 (0.9) 22 (0.9) 22 (0.9) 22 (0.9) 22 (0.9) 22 (0.9) 259 (2.0) 270 (1.5) 264 (1.7) 266 (2.1) 276 (1.6) 272 (1.8) 266 (1.9) 266 (1.9) 266 (1.9) 269 (1.9) 269 (1.9) 278 (1.3) 275 (1.9) 285 (1.5) 267 (1.6) 269 (1.9) 271 (2.0) 284 (1.8) 271 (2.0) 284 (1.8) 272 (2.0) 275 (1.6) 286 (2.0) 275 (1.6)

Gary W. Oehlert

Wainer's Grades

Student	Score
Α	88
В	65
С	91
D	36
Е	72
F	57
G	50
Н	85
I	62
J	48

Student	Score
С	91
Α	88
Н	85
Ε	72
В	65
	62
F	57
G	50
J	48
D	36

NOT!

Add

Simplicity

Gary W. Oehlert

Tabular Display of Data

Tabular Display of Data

244 (1.8) 256 (1.8) 258 (1.6) 251 (2.1) 260 (1.5) 260 (1.5) 260 (1.8) 251 (1.8) 251 (1.8) 252 (1.3) 254 (1.6) 273 (1.3) 254 (1.6) 273 (1.3) 254 (1.6) 273 (1.3) 274 (1.6) 275 (1.7) 270 (1.8) 275 (1.7) 270 (1.8) 275 (1.6) 276 (1.6) 277 (1.7)

19 (0.9)
32 (1.1)
25 (1.1)
32 (0.9)
30 (1.3)
26 (1.1)
25 (1.2)
21 (1.0)
26 (0.9)
22 (0.9)
22 (0.9)
29 (1.4)
29 (1.0)
24 (1.2)

6 (0.5) 9 (0.7) 8 (0.7) 11 (0.8) 6 (0.5) 7 (0.5) 8 (0.6) 15 (0.9) 10 (0.7) 6 (0.5) 6 (0.8) 7 (0.6) 8 (0.6) 8 (0.6) 13 (0.8) 8 (0.7) 4 (0.4)

Introduction Rules and Advice Summary

Avoid

- Multidimensional tables
- Multivariate tables
- Too many rows or columns

Introduction Rules and Advice Summary

NATION
Northeast
Southeast
Contral
Southeast
Contral
STATES
Alabama
Anzona
Anzona
Arkansz
California
Colorado
Connecticu
Delaware
Dist. Colum
Florida
Georgia
Hawaii
Idaho
Indiana
Iowa
Kentucky
Louisiana
Maine
Maryland
Massachuse
Michigan
Minnesota
Mississippi
Missouri

Labels

• Good titles and explanatory text

The table with its labels, title, and accompanying text should stand alone and be comprehensible.

Gary W. Oehlert

Tabular Display of Data

Gary W. Oehlert

Point out unusual values

	UBLIC CHOOLS	Graduated College	Some Education After High School	Graduated High School	Did Not Finish High School	i Don't Know	Mean
	Nation	279	270	256	248	251	267
	States						
1	lowa	291	285	273	262	266	283
2	North Dakota	289	283	271	259	272	283
3	Minnesota	290	284	270	256	268	282
4	Maine	288	281	267	259	266	278
5	Wisconsin	287	282	270	254	255	278
6	New Hampshire	287	280	267	259	262	278
7	Nebraska	287	280	267	247 -	256	277
8	Idaho	281	278	268	254	254	274
9	Wyoming	281	278	266	258	260	274
10	Utah	280	278	258	254	258	274
11	Connecticut	288	272	260	245 -	251	273

Gary W. Oehlert Tabular Display of Data

Introduction Rules and Advice Summary

Summary

- Design for purpose and audience
- Round!
- Organize
- Simplify
- Add summaries
- Good title/labels
- Clean layout/proper spacing

Introduction

2 Rules and Advice

Summary

Gary W. Oehlert