Bessler's Recommendations for Using Color in Web Pages, Graphs, Tables, Maps, Text, and Print

LeRoy Bessler PhD

Bessler Consulting and Research

Visual Data Insights™

Strong Smart Systems™

BesslerLeRoy2024@gmail.com

Visual Data Insights and Strong Smart Systems trademarks of LeRoy Bessler PhD

About the Speaker

LeRoy Bessler PhD is a data artist, the world's longest serving advisor to SAS users on best practices for graphic design and use of color, and author of Visual Data Insights Using SAS® ODS Graphics: A Guide to Communication-Effective Data Visualization. His principles explained and demonstrated in the book are useful for any graphics software.

For More Information

See Chapter 2:
Principles of Communication-Effective Use of Color in
"Visual Data Insights Using SAS® ODS Graphics: A Guide to Communication-Effective Data Visualization"
This book contains hundreds of examples, most of them using color. See how it's done. For questions about what you read in the book, email me.

Use Color to Communicate, Not to Decorate

This is the cover of a flyer that I picked up in the lobby of St James's Church, Piccadilly when in London for the 2001 or 2003 VIEWS Conference for SAS users.

The Healing Power of

COLOUR

Awakening, strengthening and inspiring the self

Workshops, talks, courses and healing

including

Aura-Soma Colour Therapy

Colour Therapy Readings and Certificated Courses

ANN LLOYD

8 Rosslyn Hill   Hampstead
LONDON NW3 1PH

Tel/Fax: (020) 7794 7064

This is NOT intended as a VERY belated (or tardy) referral to Ms. Lloyd.

But I do hope that my advice may help you enhance your graphs' aura and soma!

Let <u>me</u> be your
Color Therapist

I am not a dietary advisor, but I have a weakness for sweet things (esp. Carol Bessler ♥).
So I need to share some dietary information—i.e., statistics on doughnuts (aka donuts).
OK, I don't NEED to,
but I WANT to.

Awkward Design for Doughnut Stats found at https://aytm.com/blog/doughnuts-survey/ split into three pieces and pasted in here

Photo Credit: Doughnuts by the dozen from Flickr

What do you want to know? If you need some consumer insights on a part comments below and we'll consider it for an upcoming survey post.

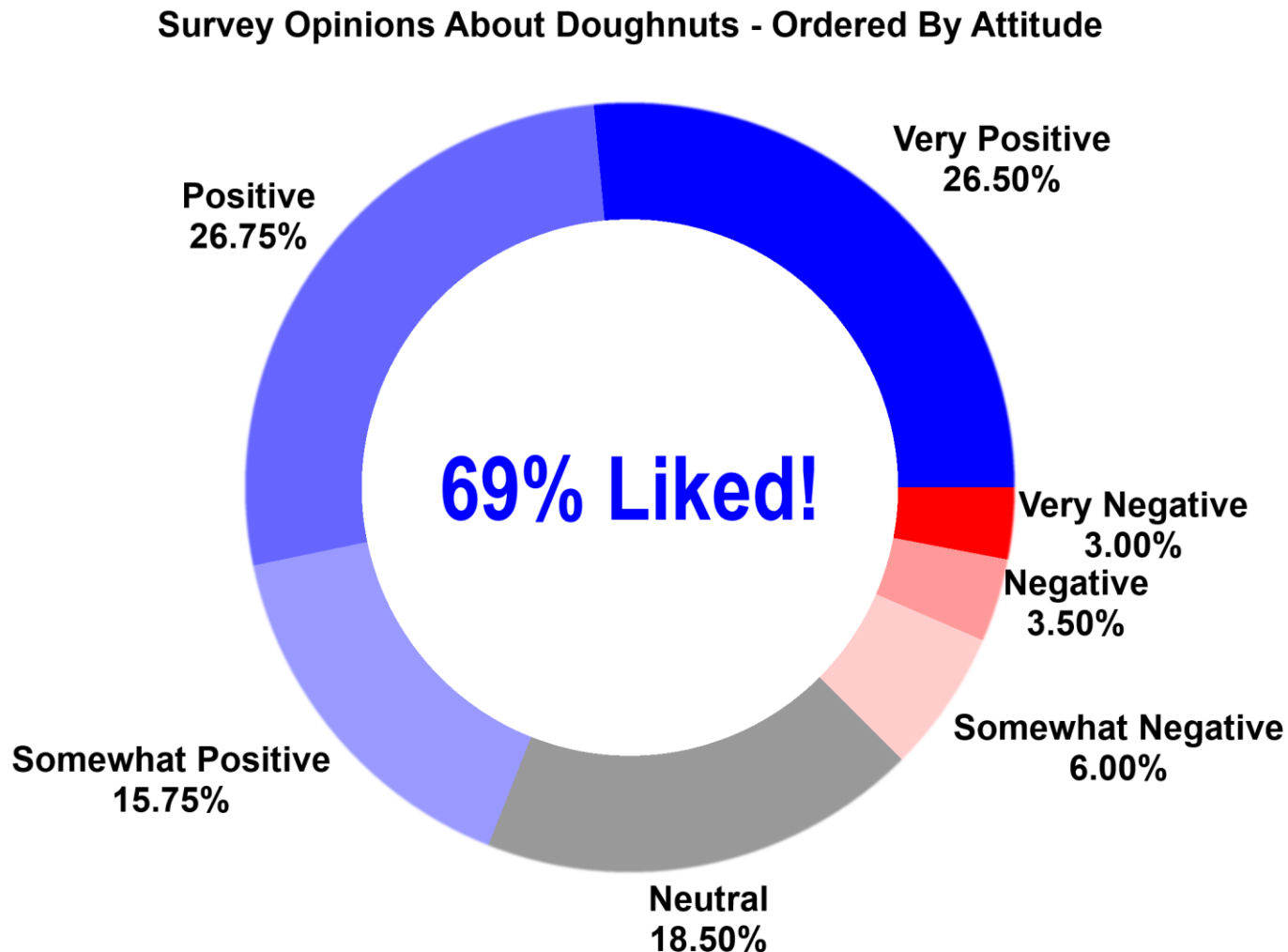**Q1** How would you describe your opinion of doughnuts?
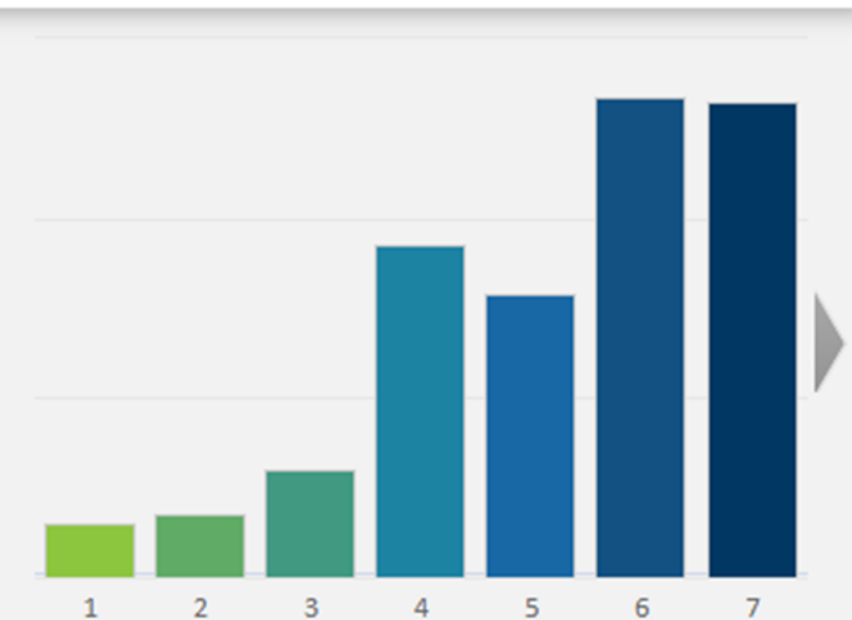
OPEN FULL STATS REPORT	
A1: Very negative	3%
A2: Negative	3.5%
A3: Somewhat negative	6%
A4: Neutral	18.5%
A5: Somewhat positive	15.75%
A6: Positive	26.75%
A7: Very positive	26.5%

Ever Heard of the Famous Four Color Map Problem?
The Donut Chart Seven Color Problem has been solved!
Gray with shades of Blue and Red (neutral, good, bad)

Survey Opinions About Doughnuts - Ordered By Attitude

Very Positive
26.50%

Positive
26.75%

69% Liked!

Somewhat Positive
15.75%

Neutral
18.50%

Somewhat Negative
6.00%

Negative
3.50%

Very Negative
3.00%

Data Source: https://aytm.com/blog/doughnuts-survey/

My Seven Color Palette
Is a Successful Substitute for
of Traffic Lighting Color Coding
(TLCC danger is discussed later)

When you evaluate the color distinguishability on these slides—

In slides, color distinguishability is better than when slides are printed.

- Projection magnifies the color areas.

- Printed color does not shine like your laptop monitor color.

About the readability of **these** slides

It is delivered with the Rockwell font.

I prefer Rockwell font because I find its readability to be greater than that of **<u>ANY</u>** typical font.

Assure colors are distinguishable

Difficult/Impossible to distinguish:

➢ thin or tiny text

➢ for graphs:
  ➢ thin lines
  ➢ too small plot markers
  ➢ too small legend color swatches

Typical Uses of Color in Graphs

➢ Color coding

   (legends, traffic lighting*)

➢ Emphasis

➢ Decoration

*Traffic Lighting should not be used. It is a communication problem for some viewers—details later.

Common Bad Uses of Color

➢ Textured Backgrounds

➢ Color Gradient Backgrounds

➢ Image Backgrounds

ALL impair readability of foreground text

ALL add ZERO communication value

Common Bad Use of Color in graphs

➢ **graph or map legend with a color gradient rather than discrete colors**

➢ The human eye can reliably distinguish only a limited number of shades of one hue (e.g., shades of green), plus black and white.

➢ Being able to <u>SEE</u> MILLIONS of colors does not mean that it's easy to distinguish Color N from Color N + K

for small values of K.

➢ Color Gradient Legends discussed later

Emphasis Options for Colorless Text

➢ **Bold**

➢ ***Bold Italics***

➢ *Italics*

➢ <u>Underline</u> (if not a web context)
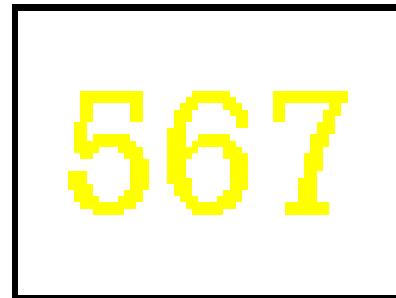
➢ (used sparingly) ALL CAPS

A Popular
Color Communication Failure

It once was a fad.

It might have mercifully fallen out of favor.

Traffic Lighting

333	333
567	567
999	999

Commonest Color Blindness:
Red = Green

Some Recommended Effective Alternatives

333	333	-222	-222
567	567	-111	-111
999	999	000	000
		+111	+111
		+222	+222

Alternative to "Traffic Lighting"

Color-Coding the Sex of Seven Students in SASHELP.CLASS

Name	Sex	Age	Height	Weight
Alfred	M	14	69.0	112.5
Alice	F	13	56.5	84.0
Barbara	F	13	65.3	98.0
Carol	F	14	62.8	102.5
Henry	M	14	63.5	102.5
James	M	12	57.3	83.0
Jane	F	12	59.8	84.5

NOTE: Use light shades of red and blue with black text. Full strength is OK with white text.

Pop. Density (per sq. km.) Change in Counties of Republic of Ireland, 1901-1996 Highlighted Spatial Population Table of County Name, Density Change, & 1996 Density
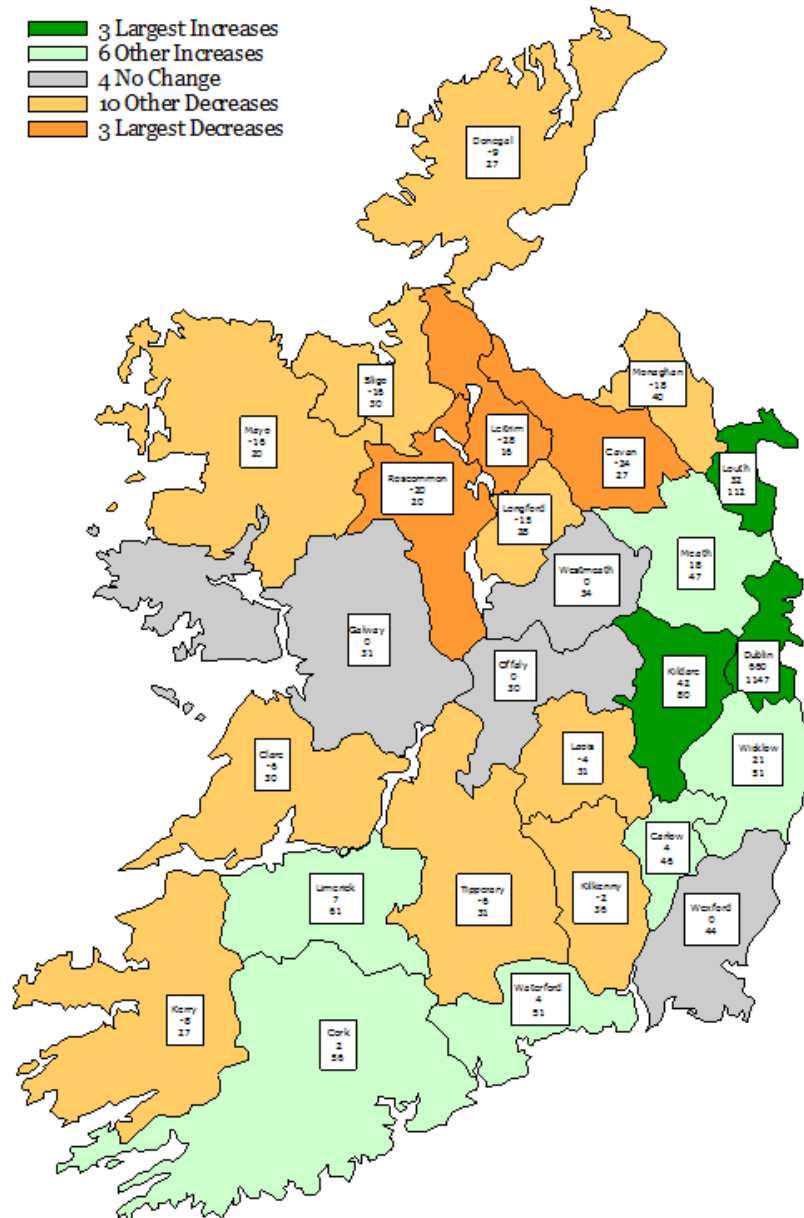
An Alternative to Traffic Lighting: augmented with light shades of the signal colors.
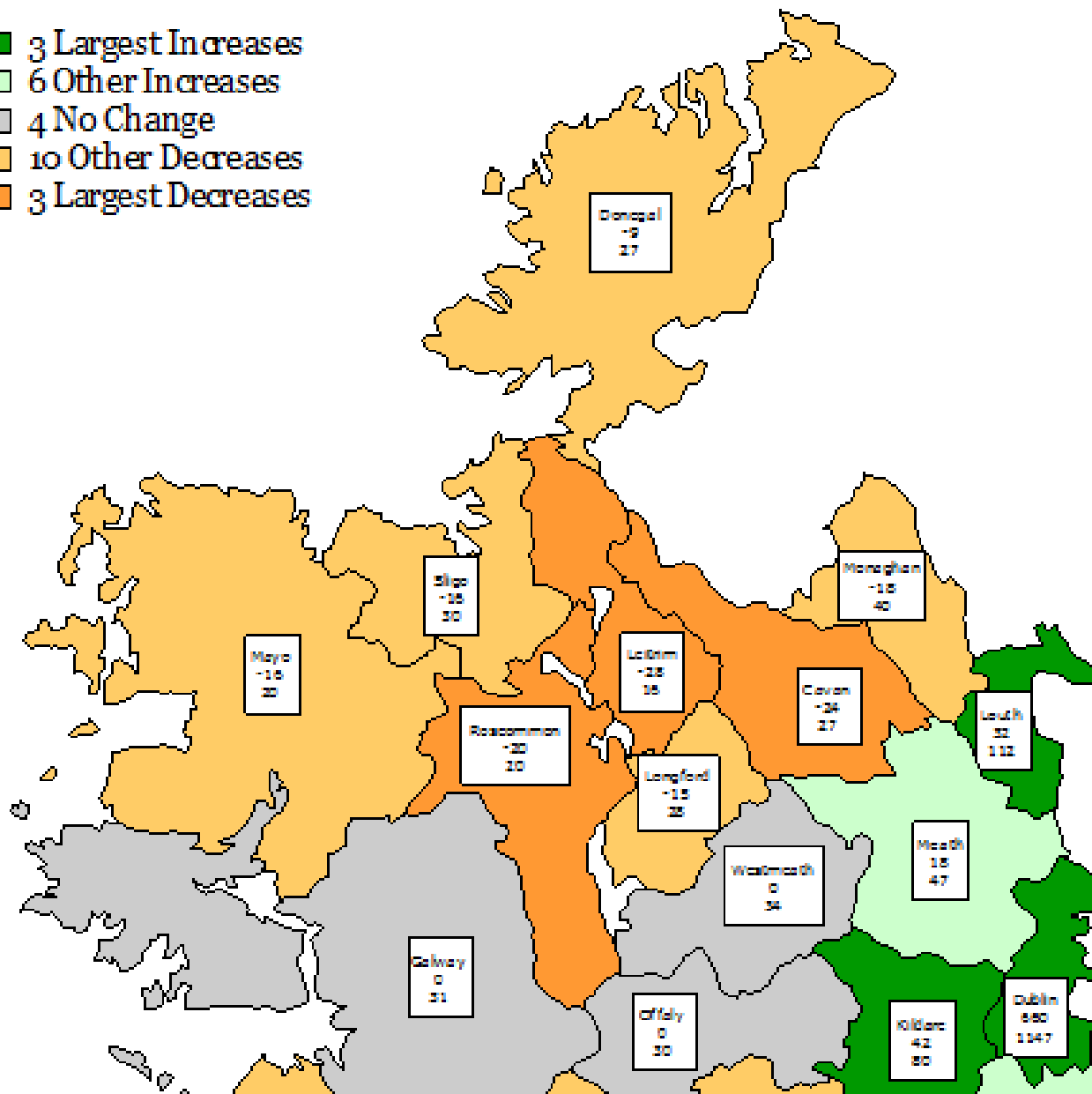
Pop. Density (per sq. km.) Change in Counties of Republic of Ireland, 1901-1996
Highlighted Spatial Population Table of County Name, Density Change, & 1996 Density

Legend:
- 3 Largest Increases
- 6 Other Increases
- 4 No Change
- 10 Other Decreases
- 3 Largest Decreases

Donegal
-9
27

Monaghan
-15
40

Sligo
-18
30

Mayo
-15
20

Leitrim
-25
16

Cavan
-24
27

Louth
52
112

Roscommon
-20
20

Longford
-15
22

Meath
15
47

Westmeath
0
34

Galway
0
31

Offaly
0
30

Kildare
42
80

Dublin
680
1347

Color Marker Distinguishability
Color Text Readability

Dots were ORIGINALLY at 28 pt, 28 pt bold, 56 pt bold. The images are magnified 117.6% of the original size.
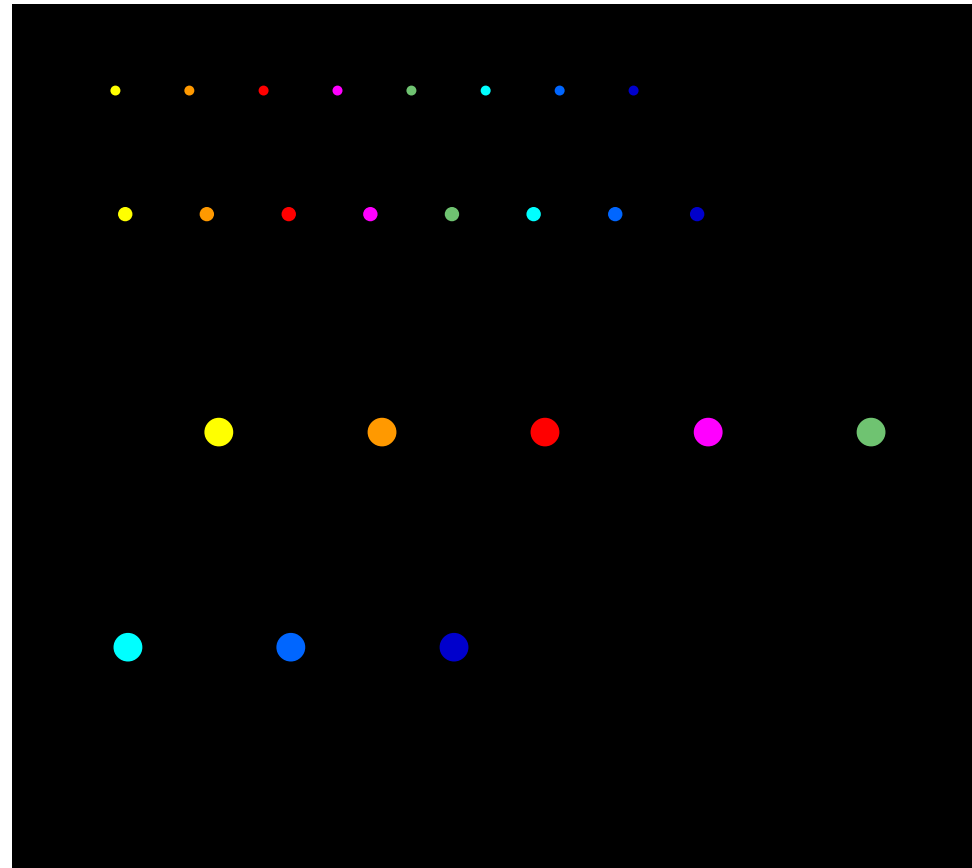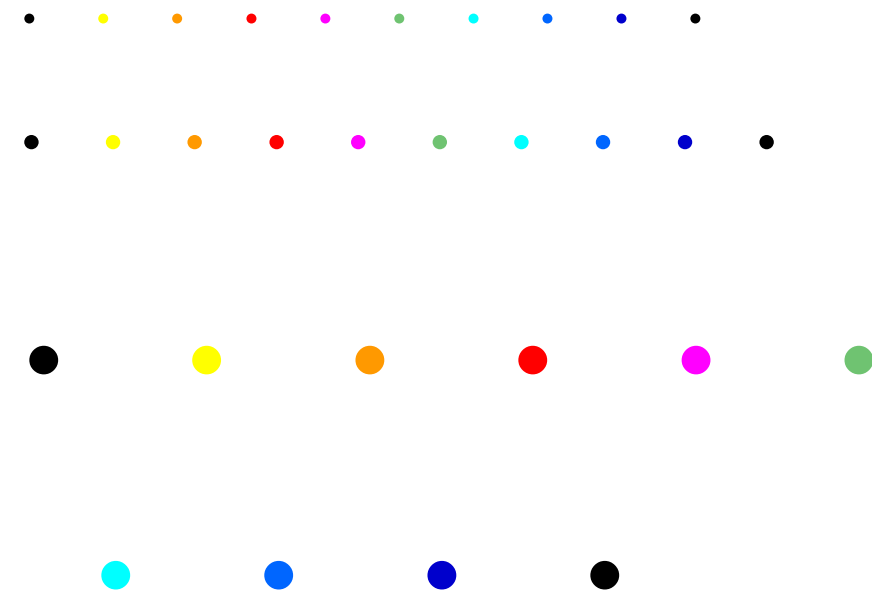
Make Colored Text Thicker

➢ this is black Arial text

➢ this is red Arial text

➢ this is bold red Arial text

➢ **this is Arial Black text**

➢ **this is bold Arial Black text***

*Arial Black is the boldest that any Arial can be

Compare Readability

white   grey

yellow   orange   red   magenta

green

cyan   blue   medium blue   dark blue

Compare Readability

grey black

yellow orange red magenta

green

cyan blue medium blue dark blue

It is generally believed that the human eye and the brain cannot reliably distinguish more than fives shades of one hue.

Though regrettably popular, a continuous color gradient is definitely useless for distinguishing all of its hues.

If five ranges of response are suitable to characterize your data, then a five color legend is necessary and sufficient.

Data has five clusters. Use five colors.

Five-Cluster Distribution of 532 VIEWS Members by County in 1999

Legend:
- 1-10
- 14-24
- 29-30
- 43-49
- 61
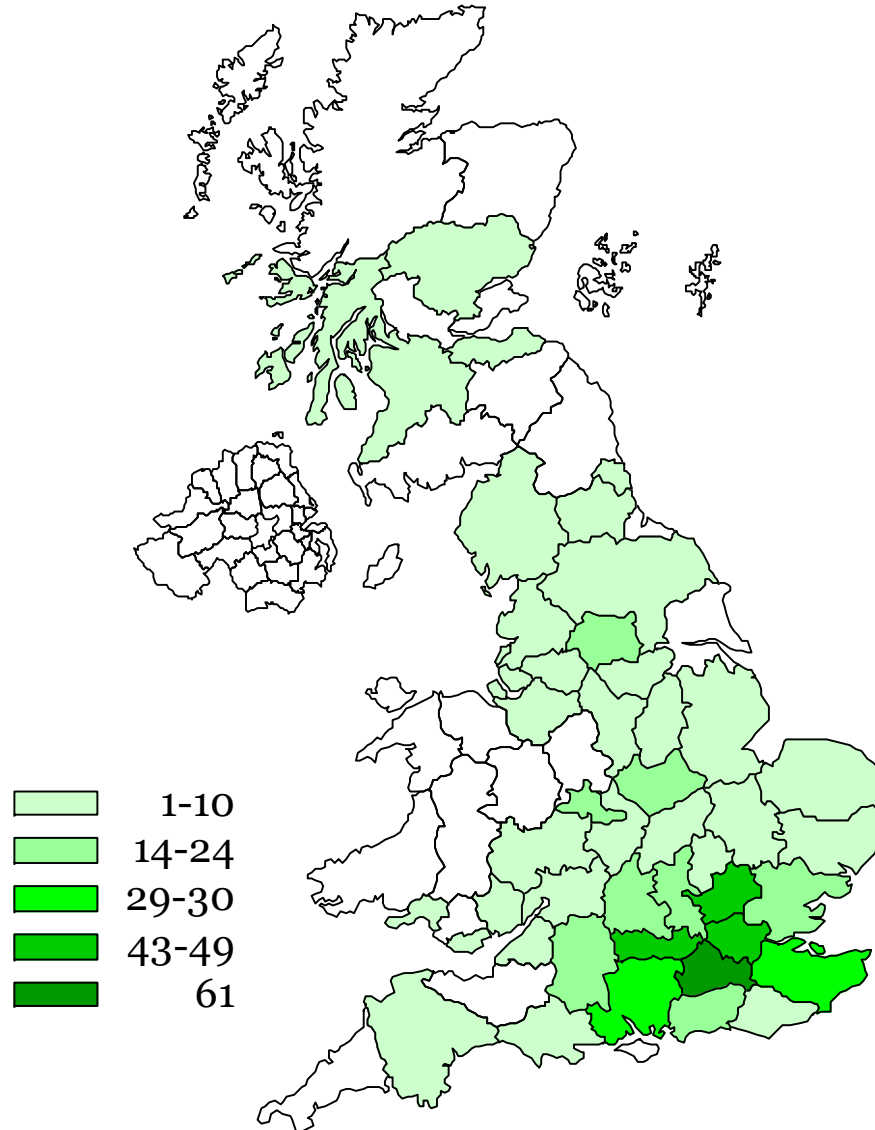
Use Color to Communicate, Not to Decorate

► If your slide does not <u>require</u> color
(for a graph legend to distinguish
different plot lines, pie slices, bars,
area color fills, etc.),
the best color combination
is ALWAYS black & white

► If not using white, try yellow

► If you don't like yellow, try cream

Color to Communicate, Not to Decorate

It's not an accident

that books & newspapers

are printed black on white

Benefit: Always readable

Side-Benefit: No confusion as to purpose

or significance of color palette

Color Use

necessary

harmless

confusing
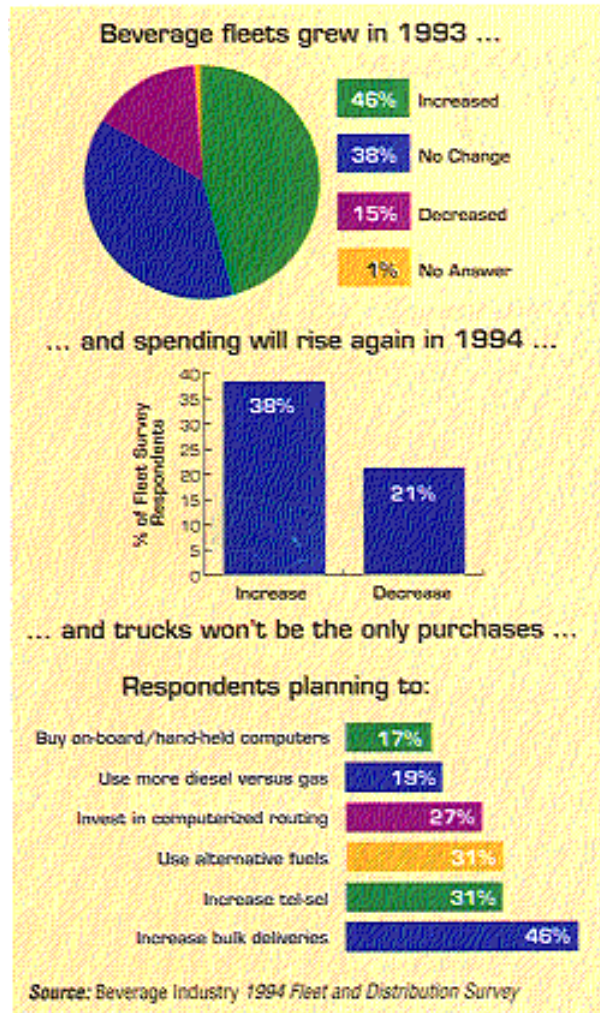
What do the same color bars have in common?

Respondents planning to:

Buy on-board/hand-held computers	17%
Use more diesel versus gas	19%
Invest in computerized routing	27%
Use alternative fuels	31%
Increase cabsel	31%
Increase bulk deliveries	46%

Source: Beverage Industry 1994 Fleet and Distribution Survey

Absolutely nuthin'
(Same answer as to "War, what is it good for?")

What do the same color bars have in common?

Respondents planning to:

Buy on-board/hand-held computers	17%
Use more diesel versus gas	19%
Invest in computerized routing	27%
Use alternative fuels	31%
Increase tel-sel	31%
Increase bulk deliveries	46%

Source: Beverage Industry 1994 Fleet and Distribution Survey
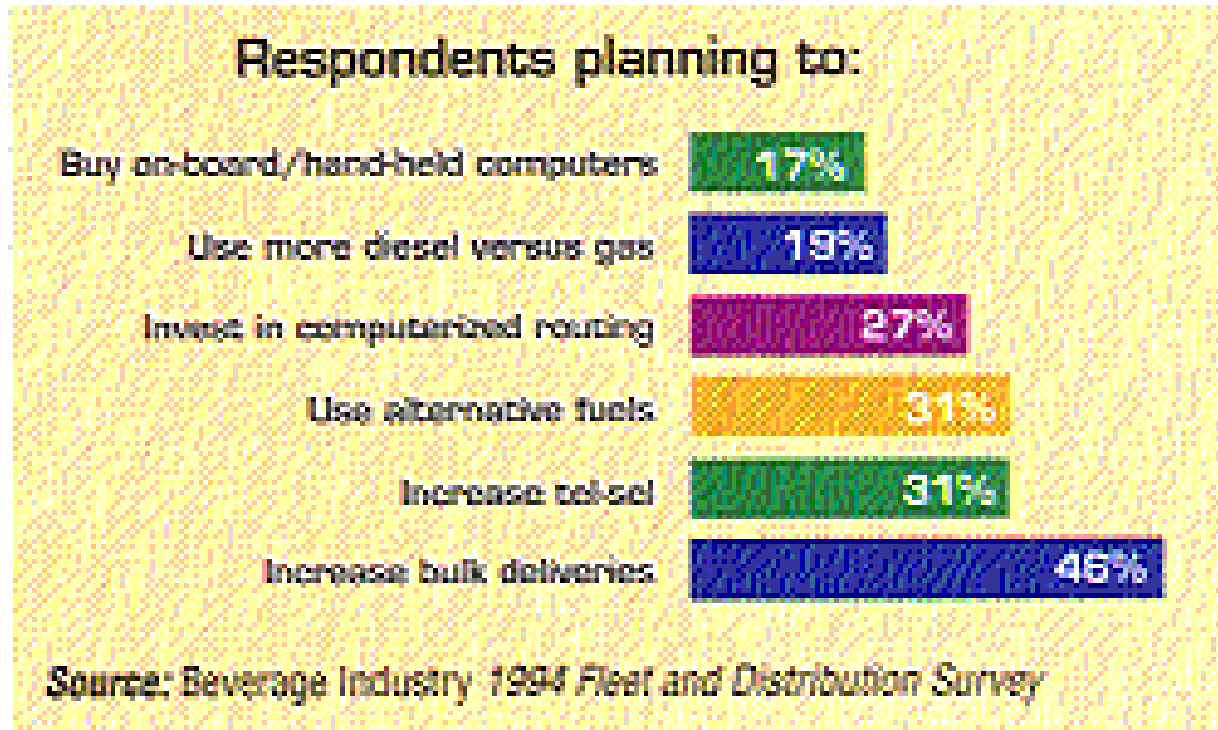
After further examination of the article that contained this puzzling graph, I noticed that every graph in the article was limited to this same four-color palette. This was a Color Palette Decision, not a communication design.

What do YOU see?

Carol sees brown. LeRoy sees green.

- **Differences in human perception:**
  She sees brown. He sees green. Notice how hard it is to assess the color of that thin text?

- Here is a solid block of it:

Color Variance Due To Technology

➢ Monitor uses RGB colors, Printer uses CMYK

➢ "transmitted" color vs. reflected color

➢ Monitor vs. Projector

➢ LED vs. CRT Projector

➢ Projectors require color tuning

➢ PC color vs. Mac color

If possible, preview your presentation on the intended projector

➢ In the preview of a slide about unreliability of color names, I found that my demo of a light shade versus the true shade (which actually was rendered lighter than light) was going to present shades of gray

➢ At a different conference, everybody's red and blue were rendered as orange and purple

Unexpected, until one day I noticed it!

➢ **If you tilt the screen on your laptop computer, the colors will change**

➢ Can you mandate that viewers of your work will use the same equipment and set it to the same angle? Of course, No!

➢ OK, What to do? Not much I suppose.

➢ Perhaps this suggests that fussing with exotic colors isn't worth the effort—their exoticism, if "migrated" by a viewer's different screen angle, might go unappreciated

When to Use Color

➢ No data levels or categories: black & white

➢ Few levels or categories: gray shades maybe However, picking out a palette with fewer colors will be easy.

➢ Many levels or categories: color is necessary

Benefits of "Boring" Black & White

➤ Faster, cheaper, more reliable

➤ Easier to use

   ➤ simpler equipment

   ➤ no agonizing over color strategy

➤ More copyable

(more, cheaper, faster copiers):

Useful Documents Get Copied

How or whether to use color:

What purpose
does this color serve?

How to use color

Unless for a legend* or color coding,
the best color combination
is ALWAYS black & white.

*When doing a multi-line plot without an X axis
table (shown by me later),
it can be adequately rendered in black & white
with differently shaped plot symbols, but color
is "nicer". Prettier?

All About Non-Gradient Legends

Legend Color Swatch Size Defects

When Color Swatches Are Too Small—It's Difficult or Impossible to Distinguish the Colors

Anti-Communicative Bubble Plot
(Default Color Palette) Default Legend has color swatches too small!

Avg Price and Avg Miles Per Gallon (City) by Car Origin and Type
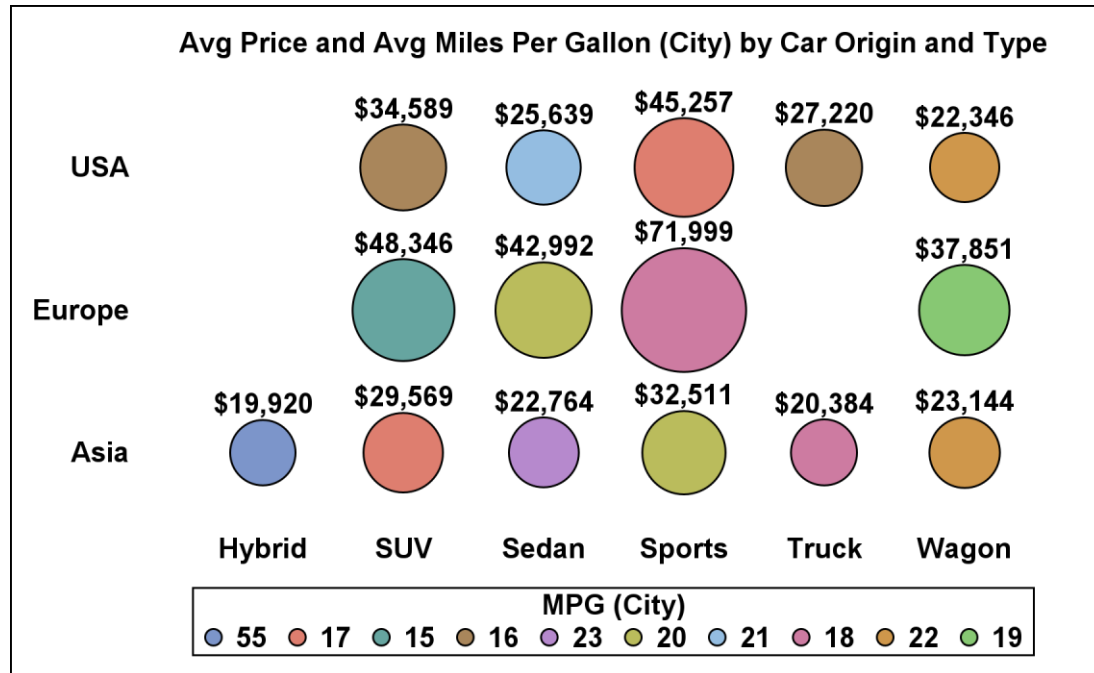
Communicative Bubble Plot
Same Color Palette, But Swatch Size Maximized To Fit In Space Available

Avg Price and Avg Miles Per Gallon (City) by Car Origin and Type

Default Legend

Avg Price and Avg Miles Per Gallon (City) by Car Origin and Type

	Hybrid	SUV	Sedan	Sports	Truck	Wagon
USA		$34,589	$25,639	$45,257	$27,220	$22,346
Europe		$48,346	$42,992	$71,999		$37,851
Asia	$19,920	$29,569	$22,764	$32,511	$20,384	$23,144

MPG (City)
● 55 ● 17 ● 15 ● 16 ● 23 ● 20 ● 21 ● 18 ● 22 ● 19

Custom Legend

Avg Price and Avg Miles Per Gallon (City) by Car Origin and Type

	Hybrid	SUV	Sedan	Sports	Truck	Wagon
USA		$34,589	$25,639	$45,257	$27,220	$22,346
Europe		$48,346	$42,992	$71,999		$37,851
Asia	$19,920	$29,569	$22,764	$32,511	$20,384	$23,144

MPG
● 55
● 23
● 22
● 21
● 20
● 19
● 18
● 17
● 16
● 15

Anti-Communicative Excel Pie Chart

In a Word Document or Printed, unusable legend: BrandB or BrandE?

Figure 7: Excel Pie Chart With Legend.
Legend Color Samples Are Too Small.
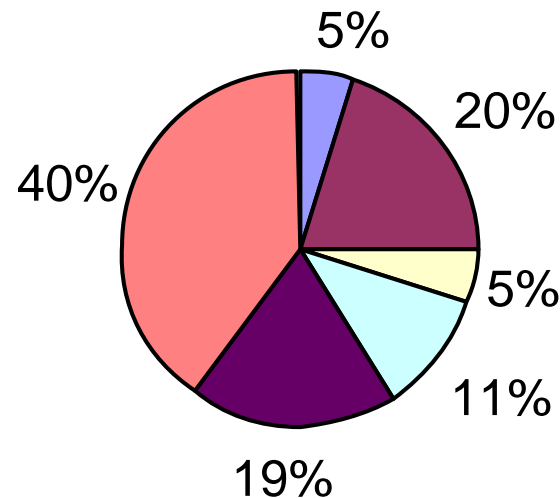
Anti-Communicative Excel Pie Chart

Excel does not (or back then, did not) support legend color size adjustment

Figure 7: Excel Pie Chart With Legend.
Legend Color Samples Are Too Small.
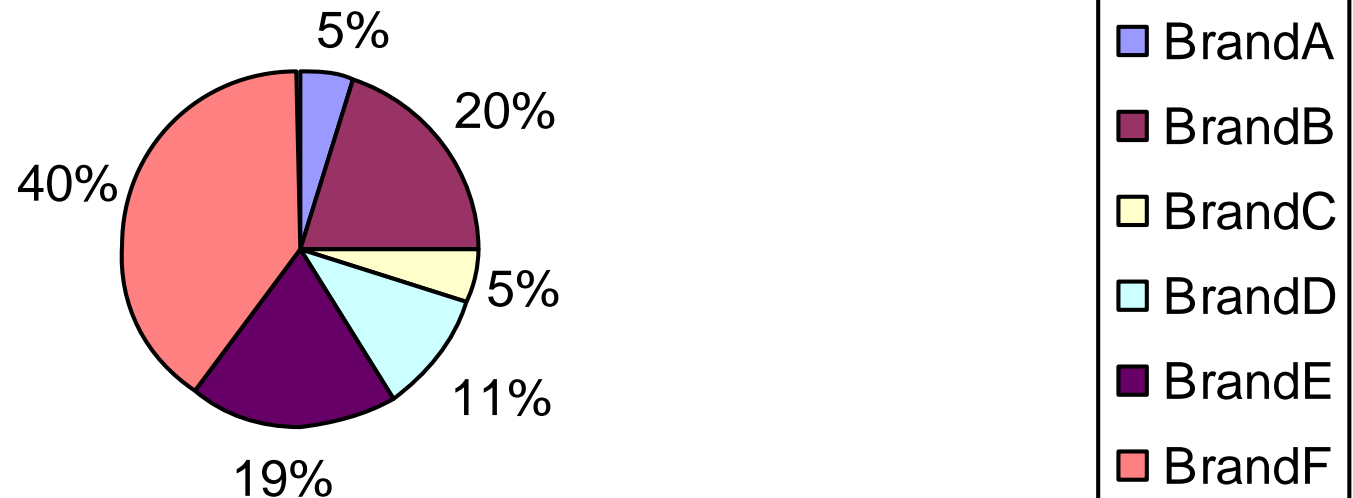
# Best Pie Chart Legend
# Using SAS ODS Graphics
FILLASPECT=GOLDEN yields color
swatch 1.618 times area of a square
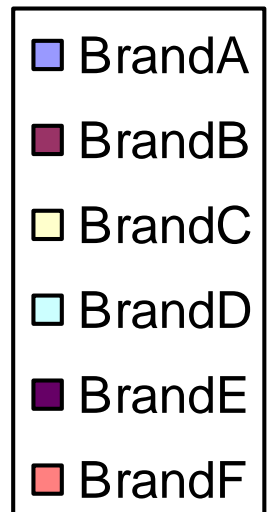
Shoe Sales and Percent Share By Product - Total = $33,851,566

- ■ 1 Men's Casual $7,933,707 23.4%
- ■ 2 Women's Dress $6,226,475 18.4%
- ■ 3 Slipper $6,175,834 18.2%
- ■ 4 Men's Dress $5,507,243 16.3%
- ■ 5 Women's Casual $4,137,861 12.2%
- ■ 6 Boot $2,350,543 6.9%
- ■ 7 Sandal $868,436 2.6%
- ■ 8 Sport Shoe $651,467 1.9%

Similar Legend: FILLASPECT=GOLDEN yields color swatch 1.618 times area of a square

Monthly Open for Three Stocks in 2000

Month	IBM	Intel	Microsoft
Jan	112	83	117
Feb	112	98	99
Mar	102	113	90
Apr	120	131	94
May	113	128	73
Jun	108	126	64
Jul	109	134	80
Aug	112	67	70
Sep	133	76	70
Oct	113	43	61
Nov	98	44	69
Dec	94	39	58

IBM   Intel   Microsoft
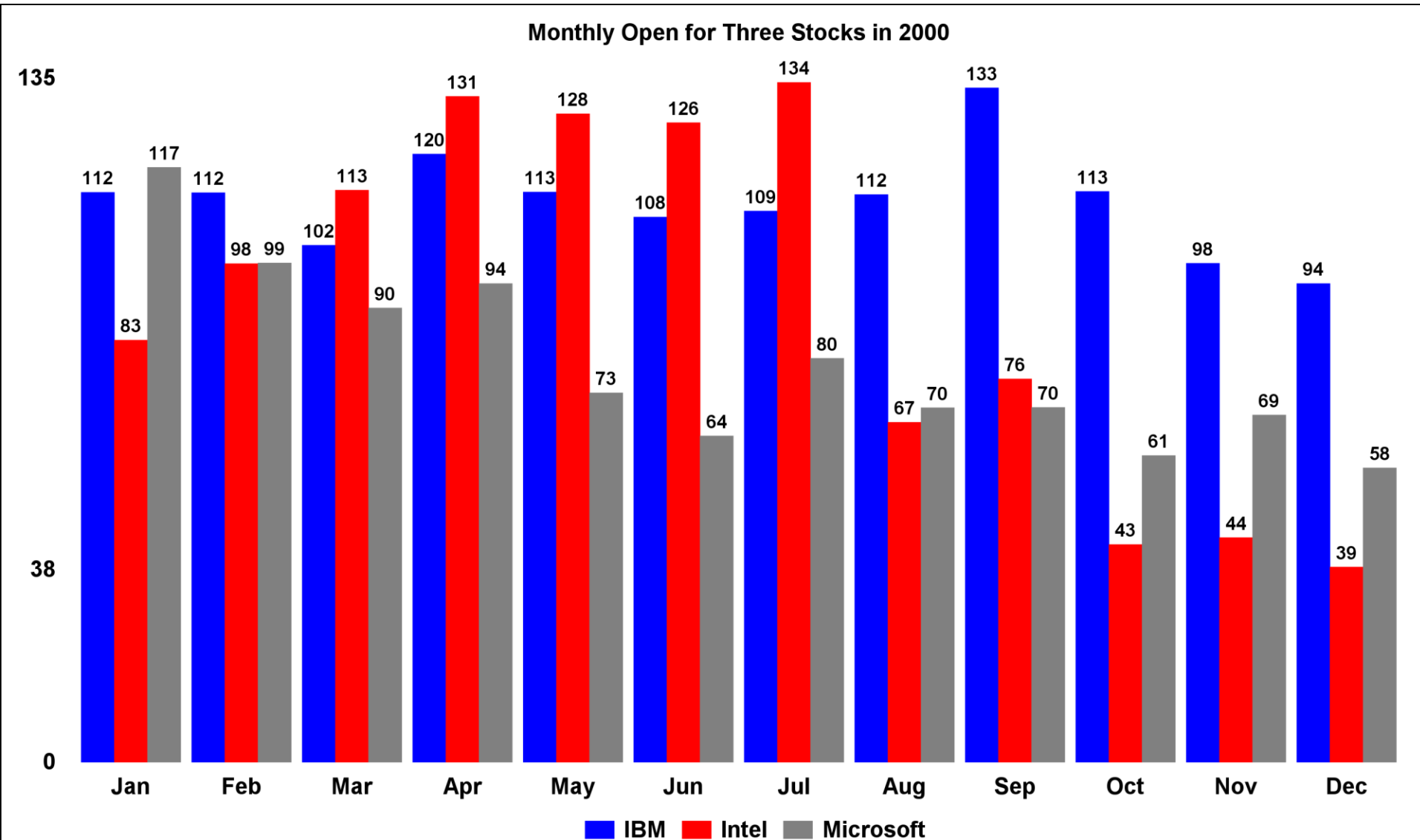
My Favorite Multi-line Plot Solution Using an X Axis Table

► Image for quick, easy inference

► Precise Numbers for reliable inference

► No Need for a Legend

► Color-coded, color-coordinated

   plot with on-chart table = Legend

A Magnificent Solution

No Legend Needed
Y Values Conveniently Close to Dates

Monthly Open Three Tech Stocks in 2000

135

38

0

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
IBM	112	112	102	120	113	108	109	112	133	113	98	94
Intel	83	98	113	131	128	126	134	67	76	43	44	39
Microsoft	117	99	90	94	73	64	80	70	70	61	69	58
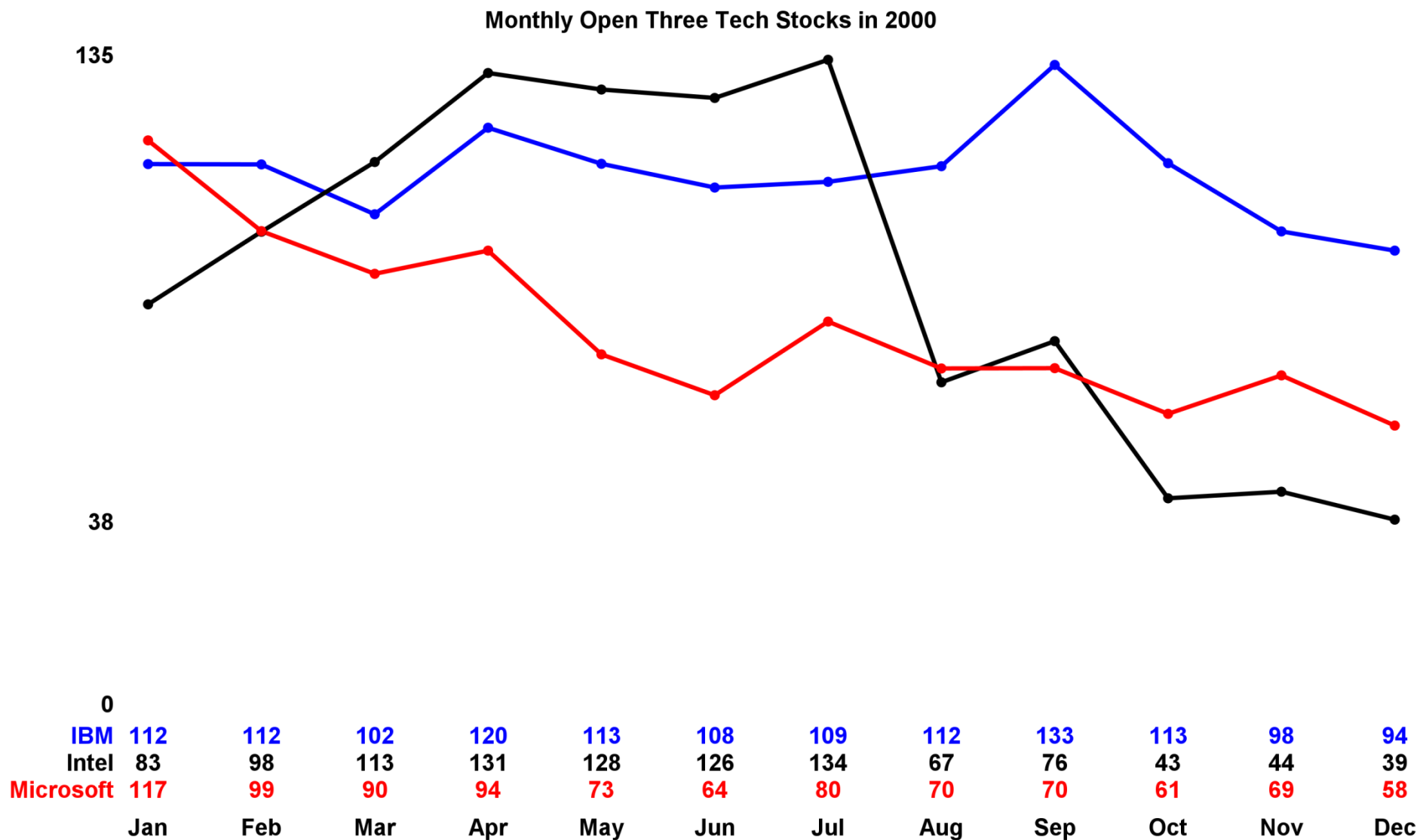
Beware of Color Names

Verify that the result is what you expect.

Misleading color names are common in a variety of software tools.

Color codes (when you know what the color LOOKS like) are more reliable.

For Color Names & Color Codes

In the SAS documentation, see:

Predefined Colors at

https://documentation.sas.com/doc/en/pgmsas
cdc/9.4_3.2/grstatug/n0w92nymec38can1c86rg
36cotkc.htm

Color-Naming Schemes at

https://documentation.sas.com/doc/en/pgmsas
cdc/9.4_3.5/grstatproc/p0edl20cvxxmm9n1i9ht
3n21eict.htm

Color Gradient Legends
A Problem with ANY Graphics Software

Next slides are SAS-ODS-Graphics-Specific, but are representative of the inherent problems of Color Gradient Legends

Suited for Rough Comparison Only

► The MOST that gradient color coding can do is tell the viewer that a large response is bigger than a small response.

► It cannot tell how MUCH bigger.

► For responses with smaller differences, it may not even make it clear that they ARE different.

This is the SAS ODS Graphics ThreeColorModel

Shoe Sales By Region and Product

What is the value for Africa Slipper? For ANY cell?

Shoe Sales By Region and Product
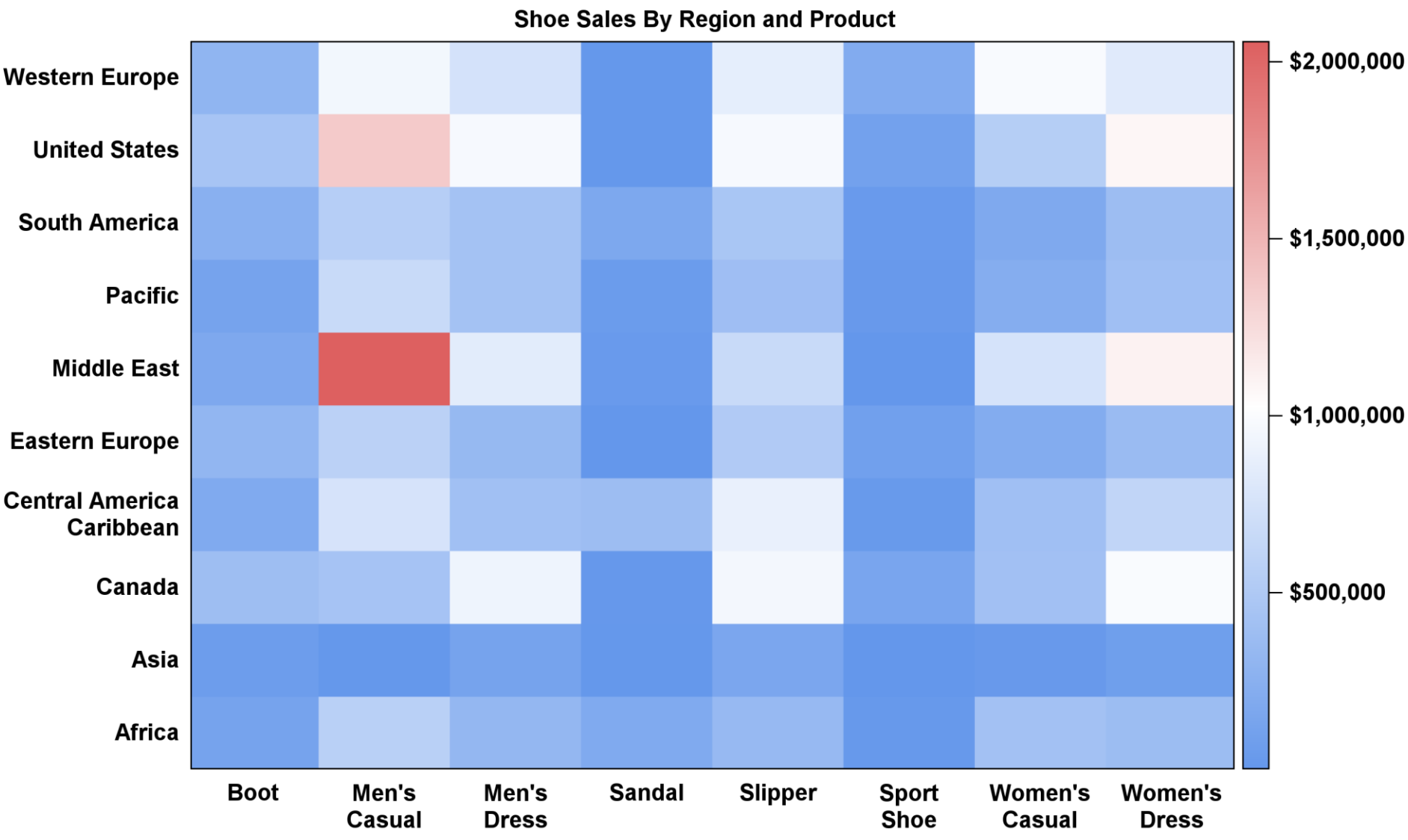
► **You cannot match any cell color to an exact point on the legend (or with certainty to another cell's color).**
► **If you really think that you found it at the legend, the legend can't tell you the precise value.**

Shoe Sales By Region and Product
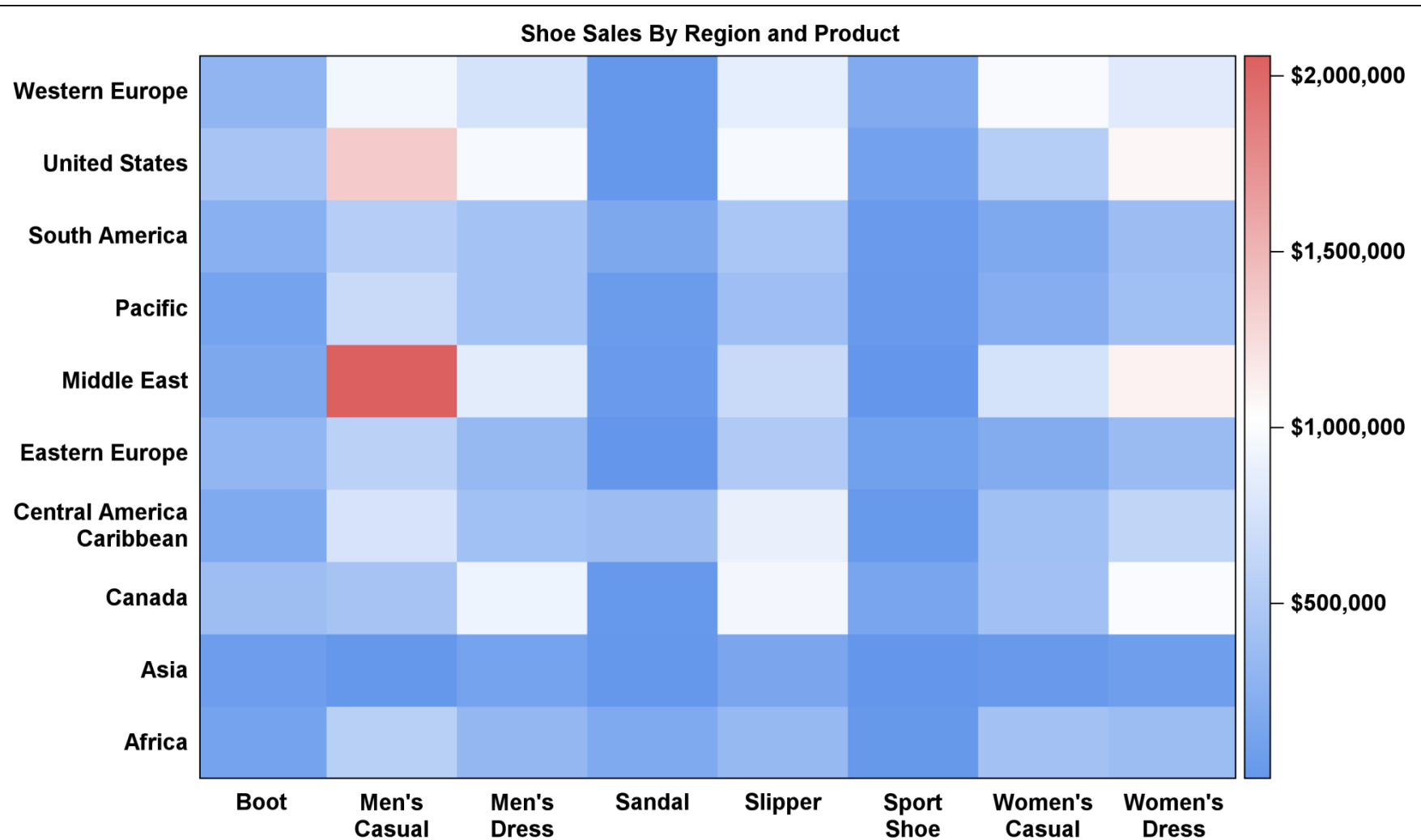
TwoColorModel Gradient Legend
Usable Only Because I Annotated It

Counts of Car Origin and Type

	Hybrid	SUV	Sedan	Sports	Truck	Wagon
USA		25	90	9	16	7
Europe		10	78	23		12
Asia	3	25	94	17	8	11
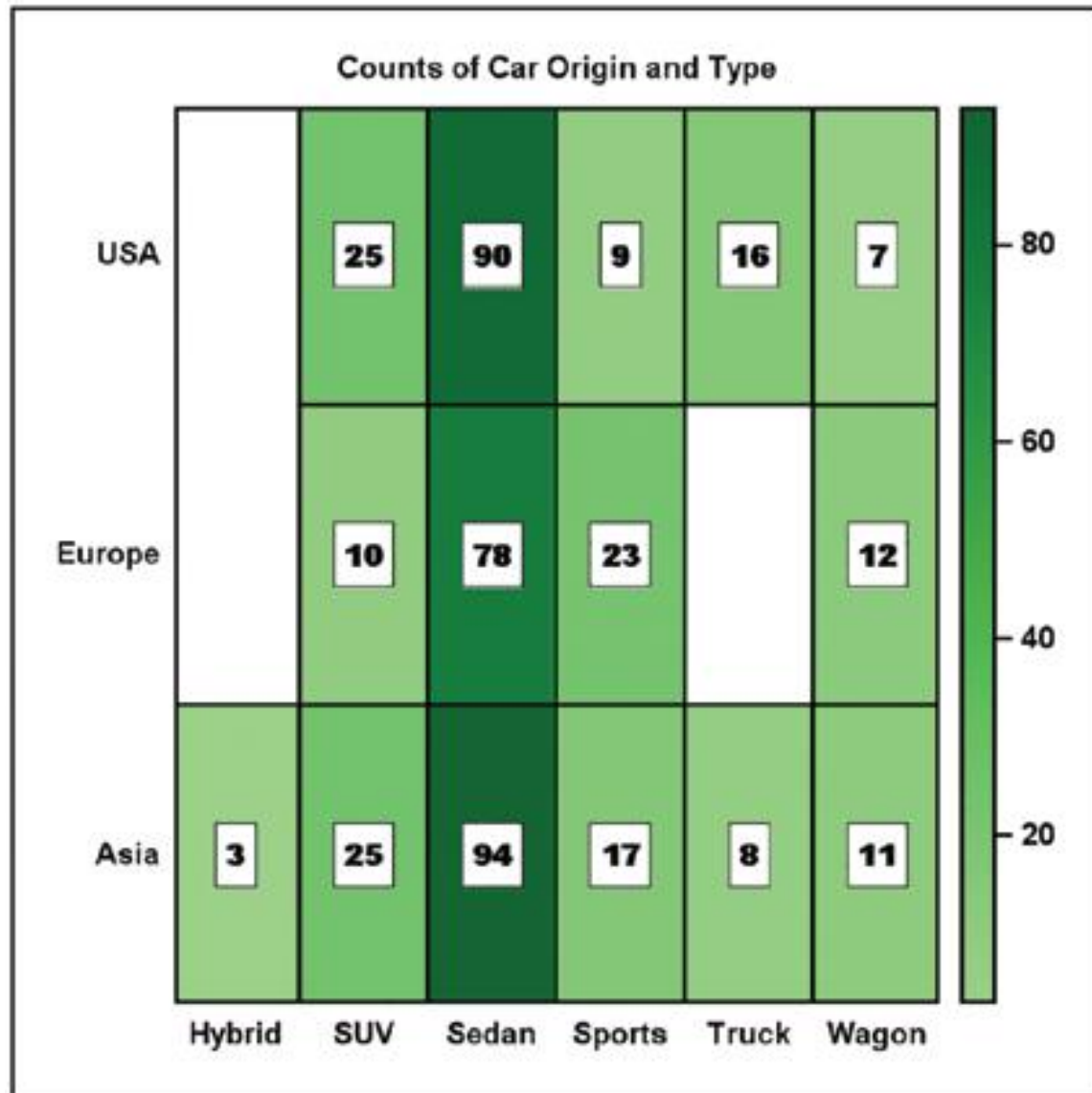
Five-Color Discrete Range Legend
Annotated for Precise Number (In Range)

Counts of Car Origin and Type

	Hybrid	SUV	Sedan	Sports	Truck	Wagon
USA		25	90	9	16	7
Europe		10	78	23		12
Asia	3	25	94	17	8	11
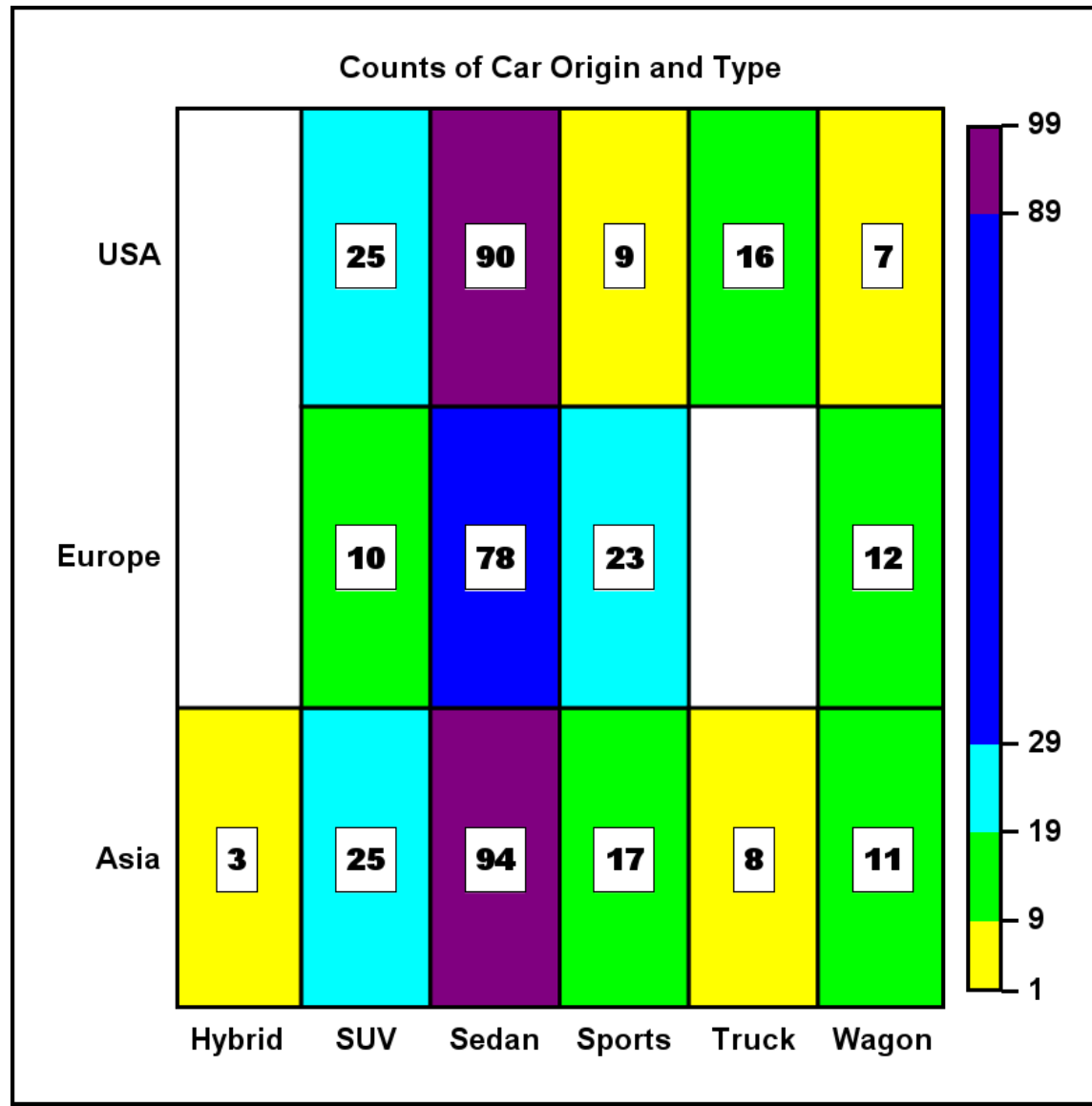
Legend:
- 99
- 89
- 29
- 19
- 9
- 1

Red-Blue Two-Gradient Legend
Really Usable Only Due to Annotation
Color Gradient for Comparison Only
Esp. Useful If Only a Few Shades Appear

**Average Blood Pressure Status (Scale 1 to 3) and Count
by Cholesterol Status and Weight Status
BP Status: 1=Optimal, 2=Normal, 3=High**

Cholesterol Status

	Underweight	Normal	Overweight
High	31 at 2.00	403 at 2.17	1354 at 2.51
Borderline	61 at 1.87	518 at 2.03	1281 at 2.40
Desirable	84 at 1.86	509 at 1.89	810 at 2.26

Average BP Status (scale: 3, 2, 1)
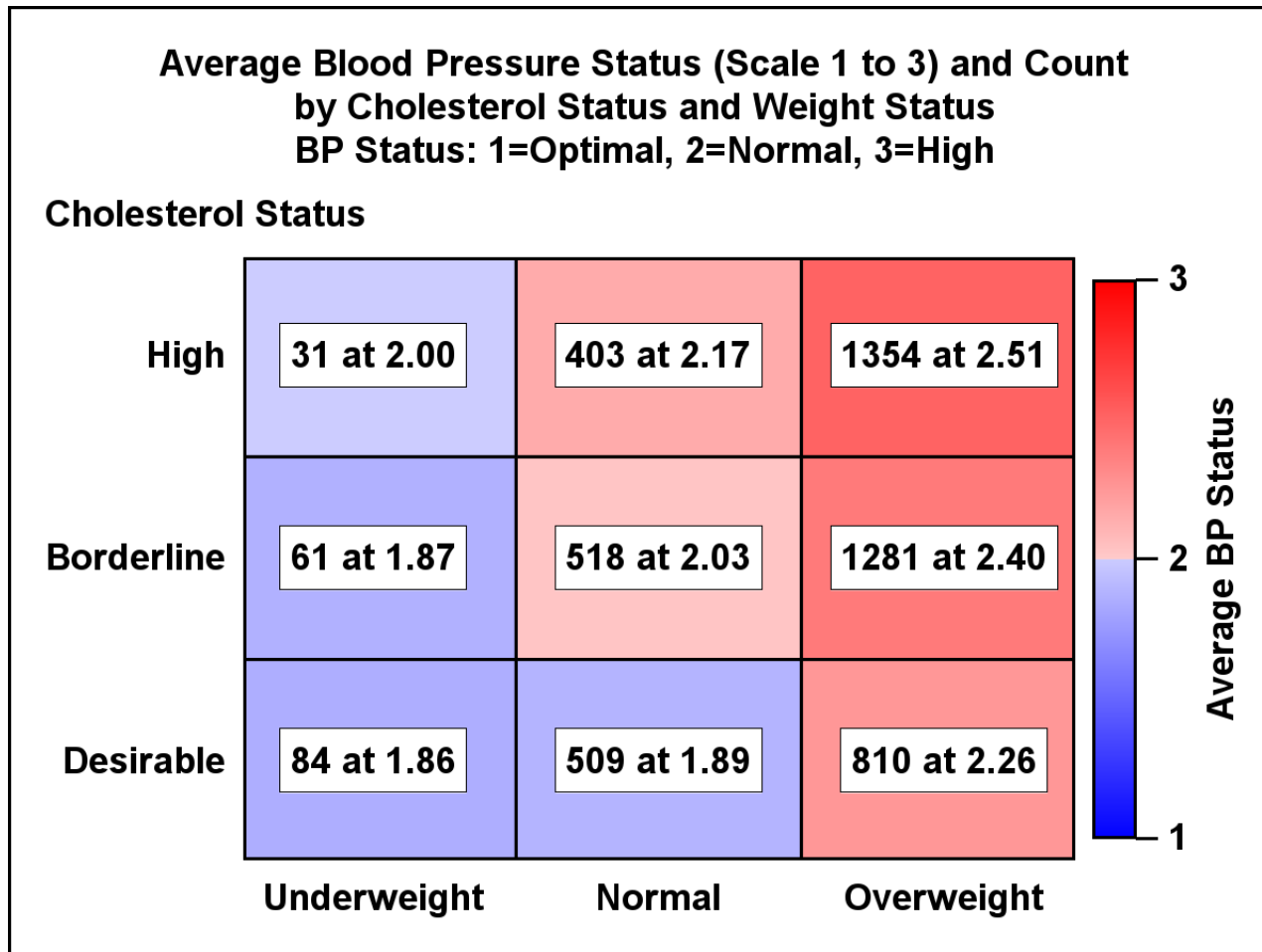
Three One-Color Gradients
For Ranges with Significance

Average Diastolic Blood Pressure
By Average Weight in pounds & Average Height in inches
For 5199 Observations in 4 weight Bins, 13 height Bins, & 42 Cells
Range of Avg Diastolic is 70-113 with Mean 88.8 & Standard Deviation 9.6
Frequency Counts in Weight-Height Cells and Color Gradient for BP
All bins are equal width, but axis values are averages, not bin midpoints.
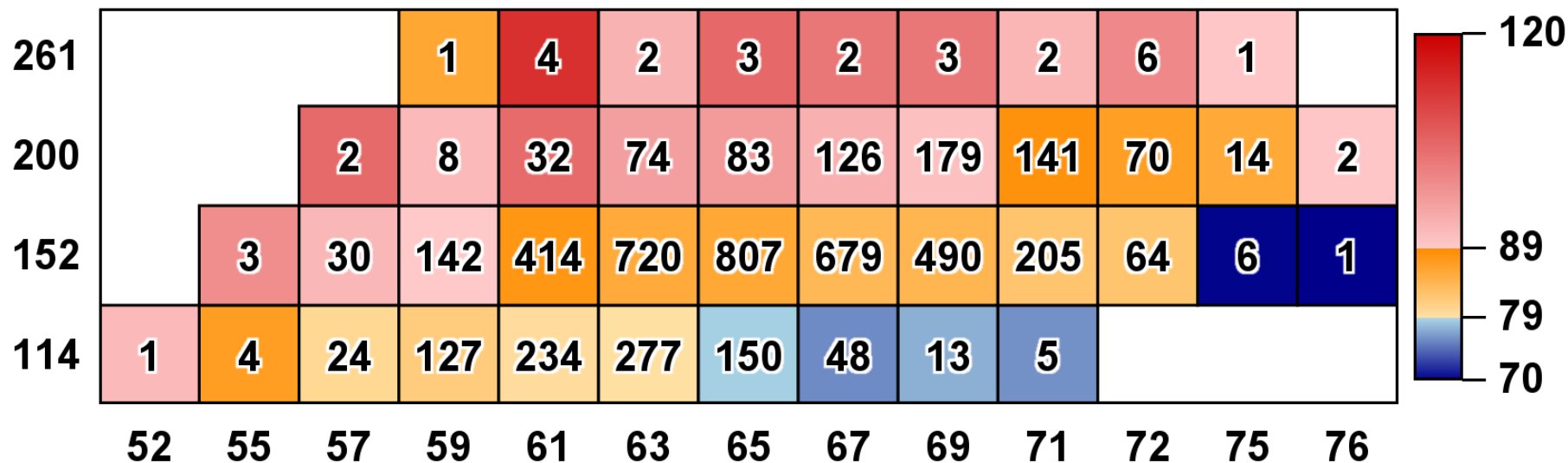So the increment between axis values along an axis can vary.

Weight	52	55	57	59	61	63	65	67	69	71	72	75	76
261				1	4	2	3	2	3	2	6	1	
200			2	8	32	74	83	126	179	141	70	14	2
152		3	30	142	414	720	807	679	490	205	64	6	1
114	1	4	24	127	234	277	150	48	13	5			

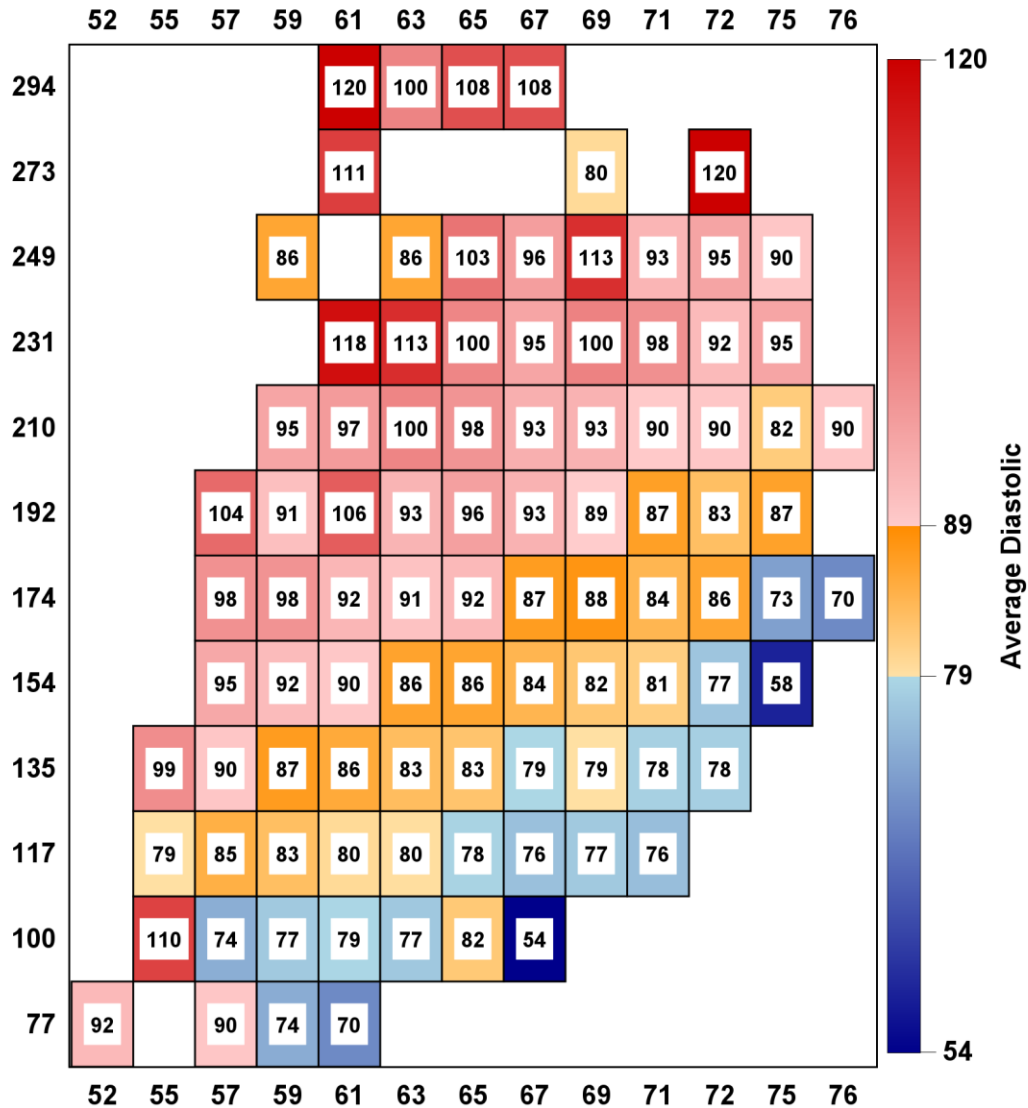Color scale: 120 — 89 — 79 — 70

Blue [to 79] is Normal Diastolic BP & Orange [to 89] is Hypertension Stage 1
Red [to 120] is Hypertension Stage 2 & Purple [> 120] is Hypertensive Crisis
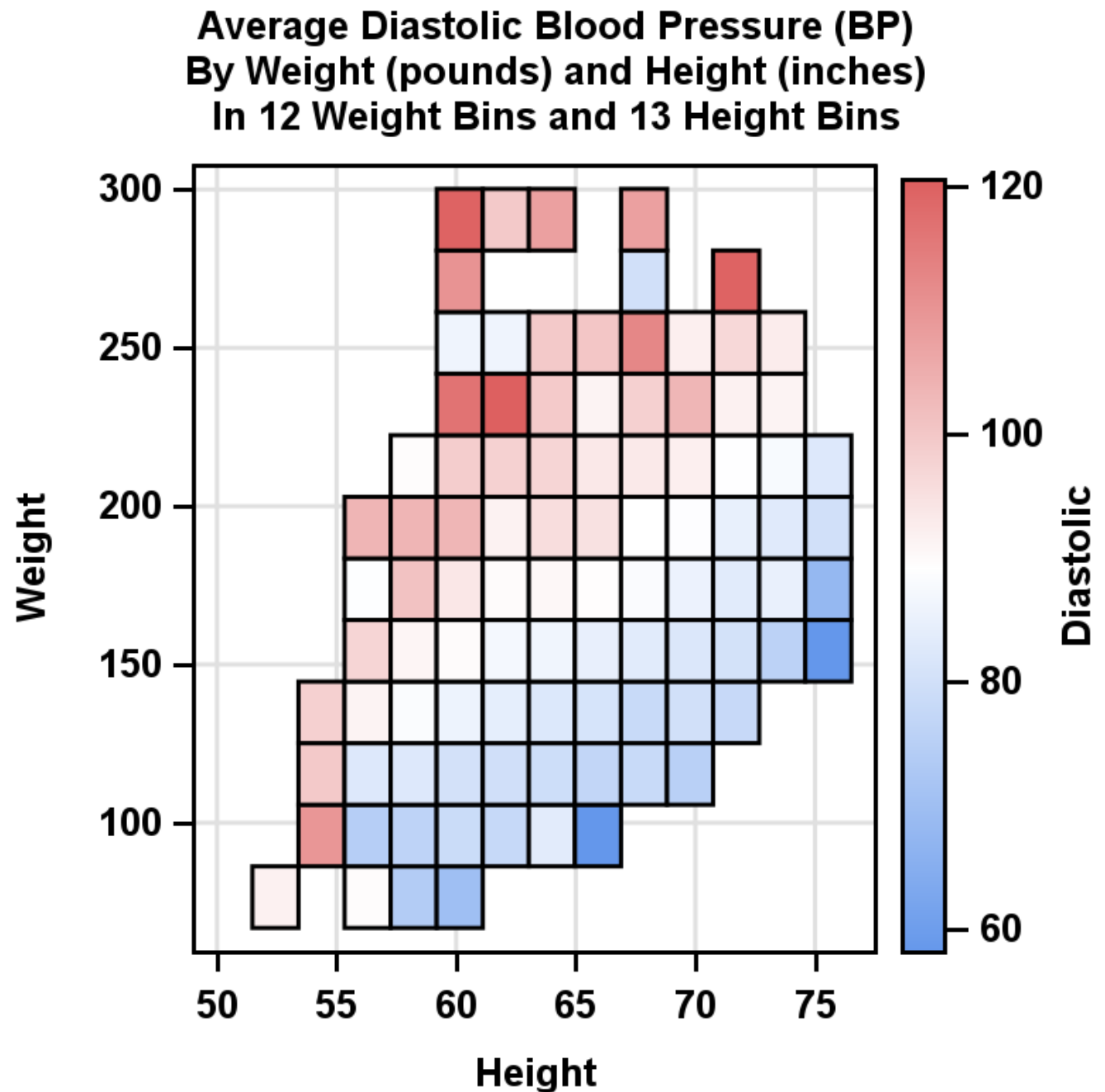
The Unsimplified Version of the Preceding Heat Map

Average Diastolic Blood Pressure
By Average Weight in pounds & Average Height in inches
For 5199 Observations in 12 weight Bins, 13 height Bins, & 94 Cells
Range of Avg Diastolic is 54-120 with Mean 89.4 & Standard Deviation 12.1

All bins are equal width, but axis values are averages, not bin midpoints.
So the increment between axis values along an axis can vary.

	52	55	57	59	61	63	65	67	69	71	72	75	76
294					120	100	108	108					
273					111				80		120		
249			86		86	103	96	113	93		95	90	
231					118	113	100	95	100	98	92	95	
210				95	97	100	98	93	93	90	90	82	90
192			104	91	106	93	96	93	89	87	83	87	
174			98	98	92	91	92	87	88	84	86	73	70
154			95	92	90	86	86	84	82	81	77	58	
135		99	90	87	86	83	83	79	79	78	78		
117		79	85	83	80	80	78	76	77	76			
100		110	74	77	79	77	82	54					
77	92		90	74	70								

Average Diastolic — scale from 120 (red) through 89, 79, to 54 (purple/blue).

Blue [to 79] is Normal Diastolic BP & Orange [to 89] is Hypertension Stage 1
Red [to 120] is Hypertension Stage 2 & Purple [> 120] is Hypertensive Crisis

The
Default
Version
of That
Heat
Map

Average Diastolic Blood Pressure (BP)
By Weight (pounds) and Height (inches)
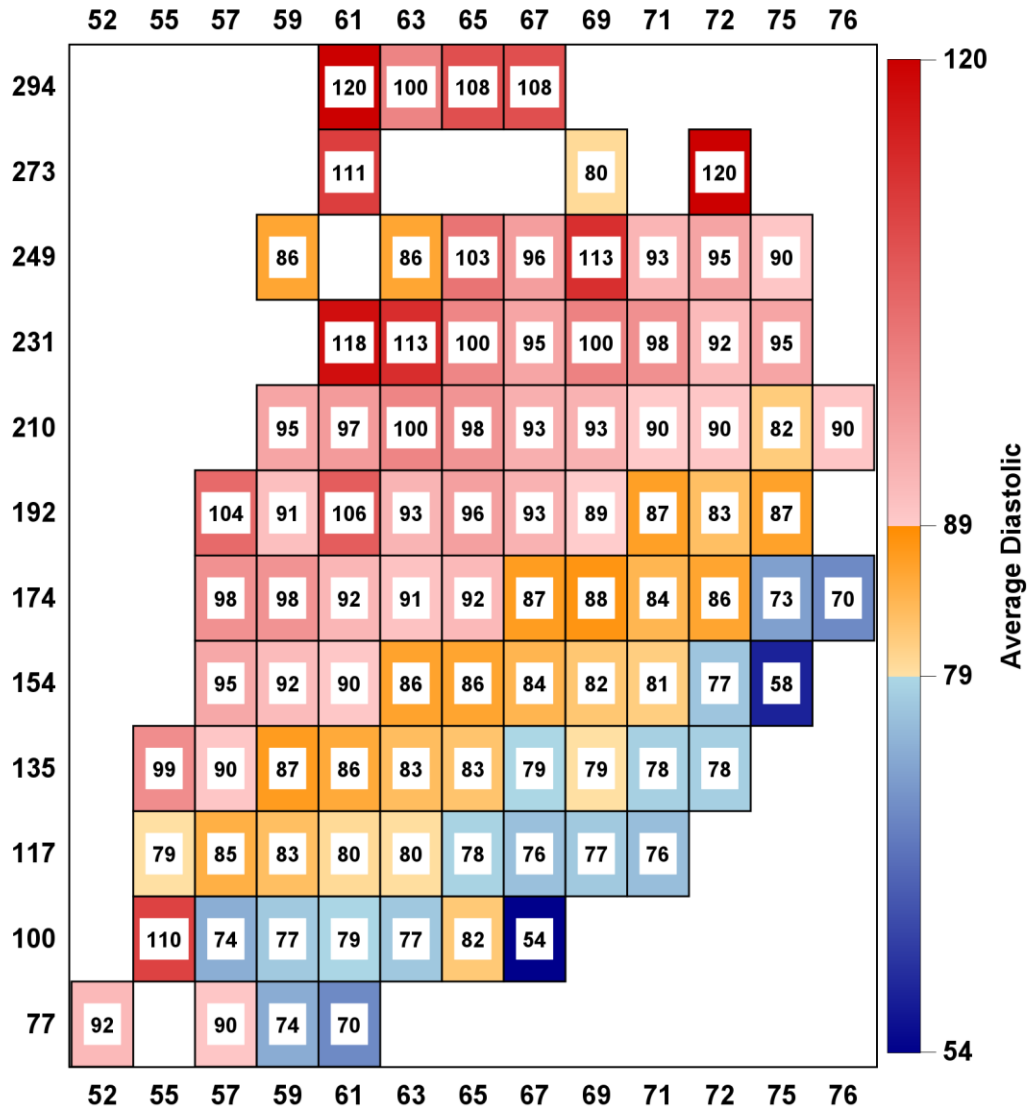**In 12 Weight Bins and 13 Height Bins**

The Custom Heat Map Again

Meaningful Gradient Legend

Cell Values are Frequency Counts

Average Diastolic Blood Pressure
By Average Weight in pounds & Average Height in inches
For 5199 Observations in 12 weight Bins, 13 height Bins, & 94 Cells
Range of Avg Diastolic is 54-120 with Mean 89.4 & Standard Deviation 12.1

All bins are equal width, but axis values are averages, not bin midpoints.
So the increment between axis values along an axis can vary.

	52	55	57	59	61	63	65	67	69	71	72	75	76
294					120	100	108	108					
273					111				80		120		
249			86		86	103	96	113	93		95	90	
231					118	113	100	95	100	98	92	95	
210				95	97	100	98	93	93	90	90	82	90
192			104	91	106	93	96	93	89	87	83	87	
174			98	98	92	91	92	87	88	84	86	73	70
154			95	92	90	86	86	84	82	81	77	58	
135		99	90	87	86	83	83	79	79	78	78		
117		79	85	83	80	80	78	76	77	76			
100		110	74	77	79	77	82	54					
77	92		90	74	70								

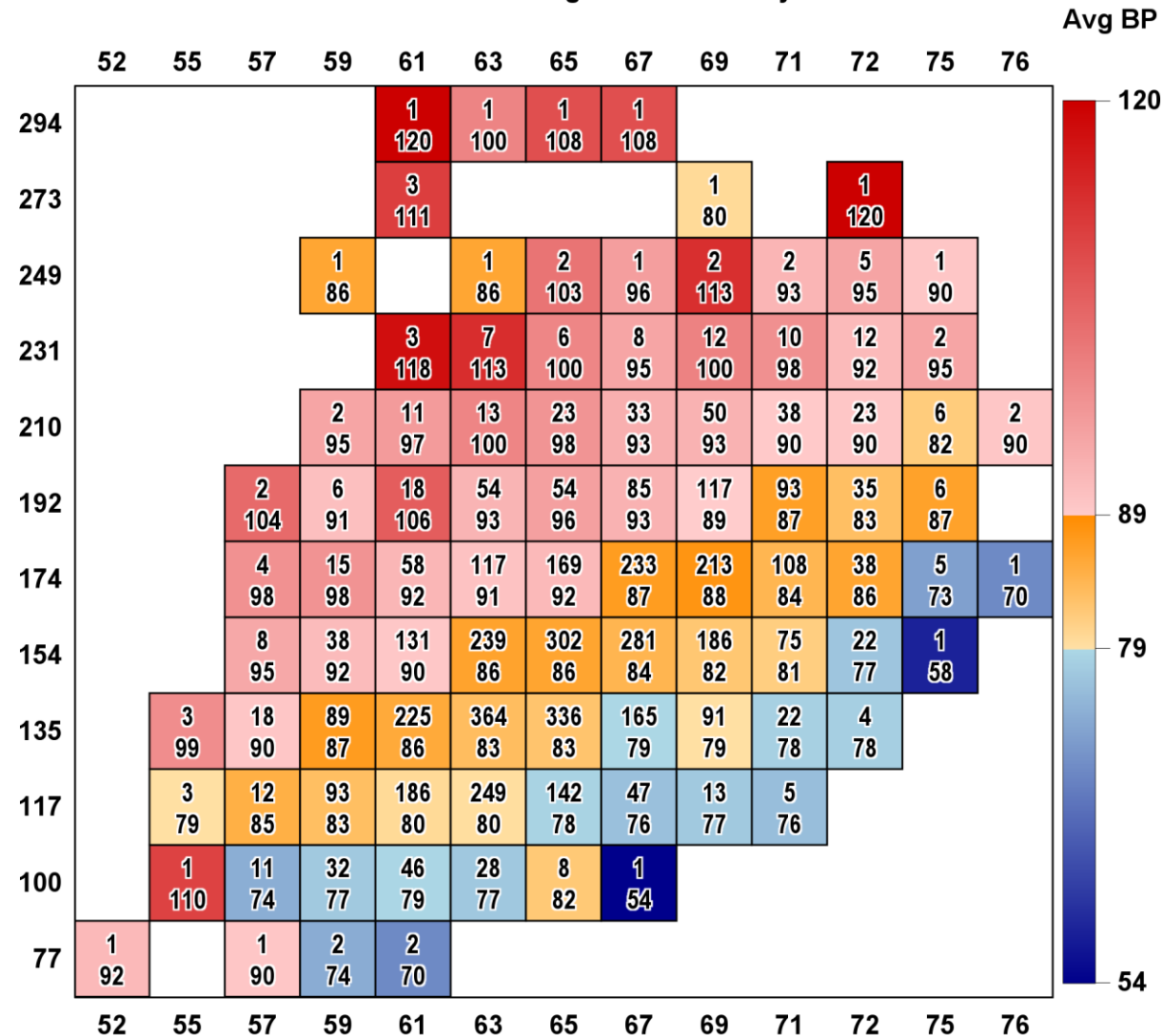| | 52 | 55 | 57 | 59 | 61 | 63 | 65 | 67 | 69 | 71 | 72 | 75 | 76 |

Average Diastolic

120

89

79

54

Blue [to 79] is Normal Diastolic BP & Orange [to 89] is Hypertension Stage 1
Red [to 120] is Hypertension Stage 2 & Purple [> 120] is Hypertensive Crisis

More Informative Custom Heat Map

- Annotate with Precise Numbers for Avg BP & Freq
- Subranges Gradient Legend for Quick Easy Category ID

Average Diastolic Blood Pressure
By Average Weight in pounds & Average Height in inches
For 5199 Observations in 12 weight Bins, 13 height Bins, & 94 Cells
Range of Avg Diastolic is 54-120 with Mean 89.4 & Standard Deviation 12.1
Labels Frequency above BP in Weight-Height Cells, Color Gradient for BP
All bins are equal width, but axis values are averages, not bin midpoints.
So the increment between axis values along an axis can vary.

Avg BP

Blue [to 79] is Normal Diastolic BP & Orange [to 89] is Hypertension Stage 1
Red [to 120] is Hypertension Stage 2 & Purple [> 120] is Hypertensive Crisis

With More Cells, Maximum Frequency Is Lower—
The Discrete Legend Can Be Used for Cell Values
Cell Values are Average Blood Pressure

Average Diastolic Blood Pressure
By Weight in pounds & Height in inches
For 318 Observations in 12 weight Bins, 18 height Bins, & 89 Cells
Range of Avg Diastolic is 60-130 with Mean 87.5 & Standard Deviation 13.1
BP in Weight-Height Cells and Color Gradient for Frequency Counts
Each bin is for only one rounded value of height or weight
Each cell is only one height-weight, but may be for multiple data points.
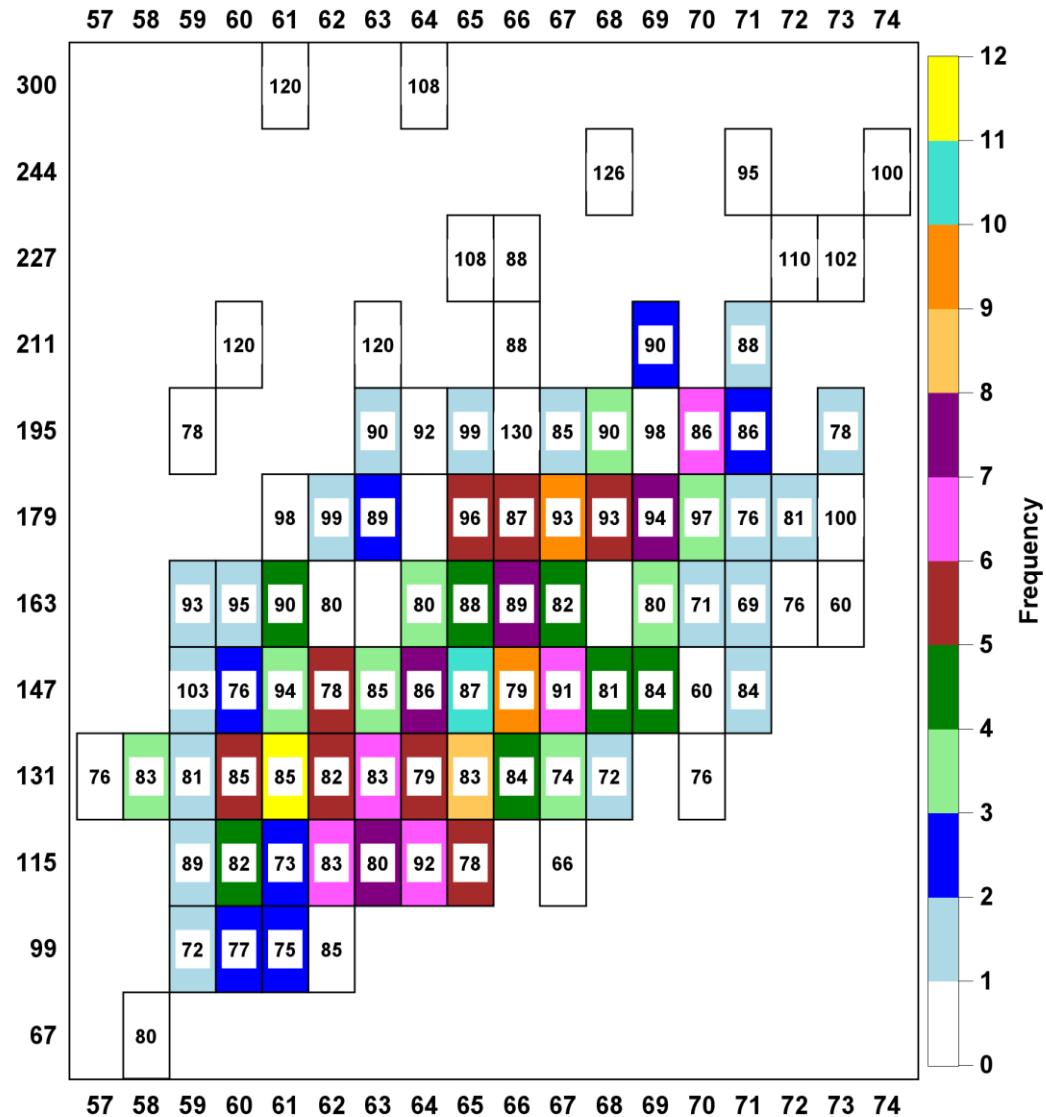
Under 80 is Normal BP and 80-89 is Hypertension Stage 1
90-120 is Hypertension Stage 2 and Over 120 is Hypertensive Crisis

For More Information

See Chapter 2:
Principles of Communication-Effective
Use of Color in
"Visual Data Insights Using SAS ODS
Graphics: A Guide to Communication-
Effective Data Visualization"
This book contains hundreds of
examples, most of them using color.
See how it's done. For questions about
what you read in the book, email me.

Your questions, comments, and ideas about communicating with color, or graphic design, or using ODS Graphics are always welcome.

LeRoy Bessler PhD
Bessler Consulting and Research
Visual Data Insights™
Strong Smart Systems™
Valparaiso, Indiana, USA
BesslerLeRoy2024@gmail.com