

Intro to Data Visualization



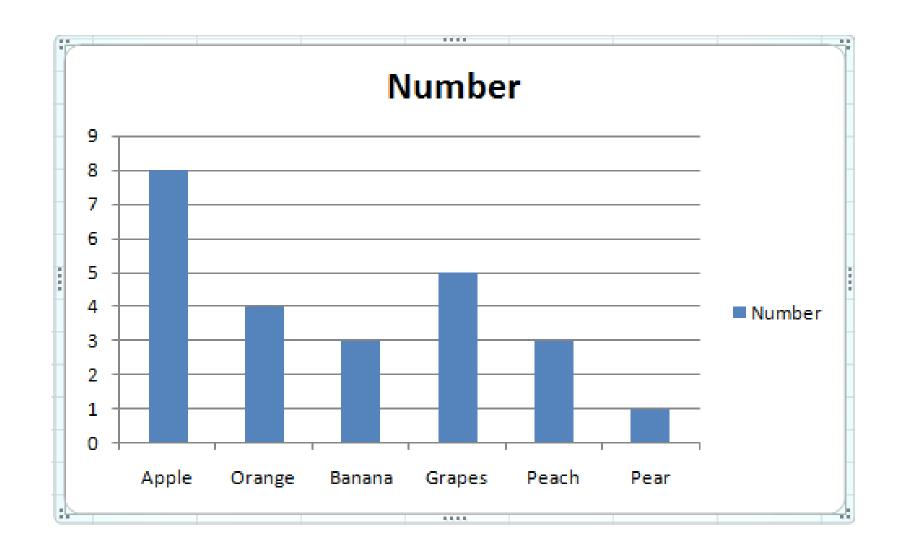


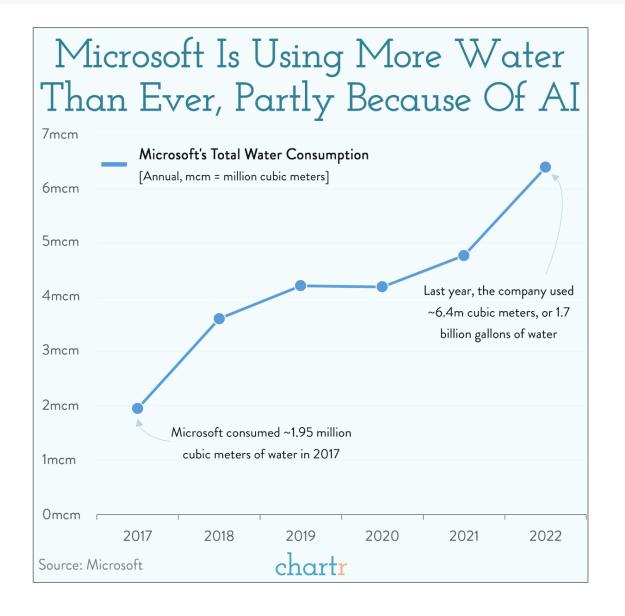
What is data visualizaton?

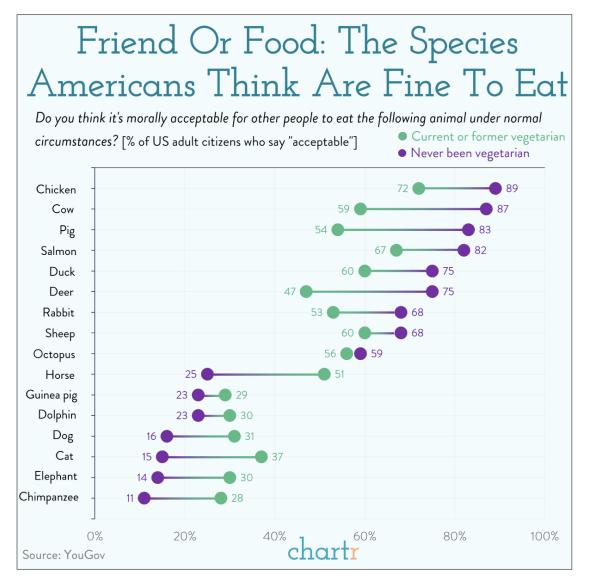


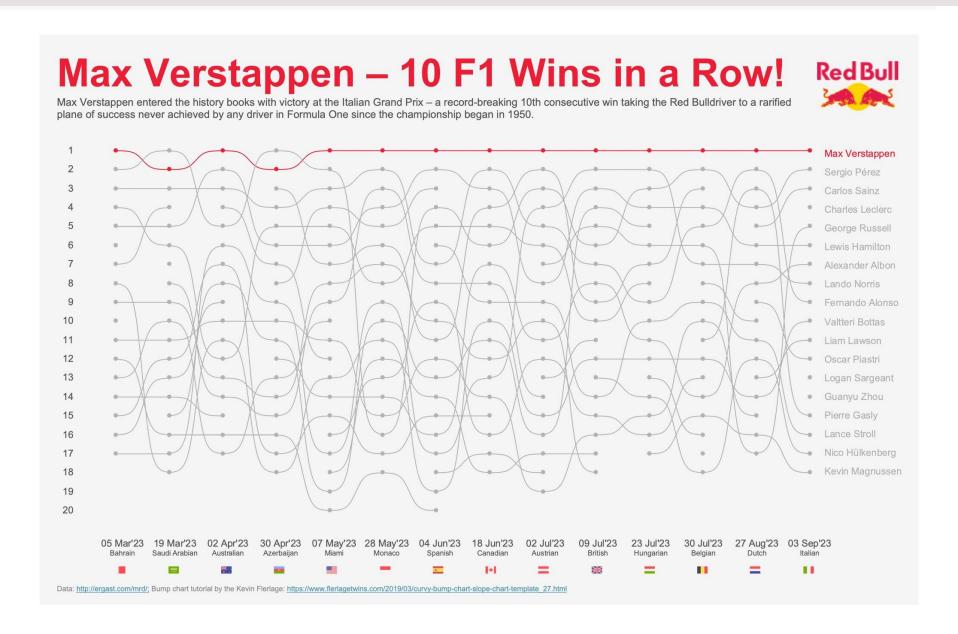


Branch	Regi	on Brand	Showroon	n Ab.	In	Out	Cou	nt Av	. ring		Ring A	v. dur.	Dur	Av. est.	Est.
4elbourne	East	Gadgets	Sales	923	8173	7236	154	09	00:14	1.08	:17:36	02:41	28.20:24:0	\$1.45	\$10,479.12
Perth	West	Gizmos	Showroom	659	5302	4613	99	15	00:14	21:	:17:34	02:41	18.11:56:4	5 \$1.39	\$6,425.26
Total				1582	13475	11849	2532	4 (00:14	2.05:3	35:10	02:41	47.08:20:49	\$1.43	\$16,904.38
Departme	ent	Manager	Ab.	In	Out	Count	Av.	ring		Ring	Av. du	ır.	Dur.	Av. est.	Est.
Administrat	tion	Peter	492	4115	3625	7740		00:14	16	5:24:22	02:	40	14.08:38:21	\$1.47	\$5,331.78
Developme	ent	Ralph	594	5202	4581	9783		00:14	20	:23:37	02:	42	18.08:24:11	\$1.40	\$6,407.26
Sales		Cyril	496	4158	3643	7801		00:15	16	5:47:11	02:	42	14.15:18:17	\$1.42	\$5,165.34
Total			1582	13475	11849	25324	0	0:14	2.05	35:10	02:4	41 4	7.08:20:49	\$1.43	\$16,904.38
Extension	n		Color	Last call	Ab.	In	Out	Count	Av. r	ing	Ring	Av. du	ır. Du	ır. Av. est.	Est.
PABX-MEL	- 100 -	Reception	Blue	05-Oct-12	91	721	666	1387	00	1:14	02:51:19	02:	40 2.13:41:	00 \$1.53	\$1,016.52
PABX-MEL	- 101 -	Ralph Hosking	g 2 Red	05-Oct-12	80	704	629	1333	00	:15	02:50:22	02:	38 2.10:51:	41 \$1.48	\$934.04
PABX-MEL	- 102 -	John Duncan		05-Oct-12	83	761	664	1425	00	1:14	02:59:14	02:	44 2.17:14:	50 \$1.52	\$1,008.67
PABX-MEL	- 103 -	Dougo Demo		05-Oct-12	89	771	705	1476	00	1:14	03:05:35	02:	42 2.18:39:	43 \$1.39	\$976.59
PABX-MEL	- 104 -	Ralph Hosking	g*	05-Oct-12	90	780	660	1440	00	1:14	02:59:50	02:	40 2.16:20:	43 \$1.37	\$902.76
PABX-MEL	- 105 -	Trevor Samm	is	05-Oct-12	81	744	649	1393	00	1:14	02:50:18	02:	40 2.14:09:	24 \$1.45	\$941.22
PABX-MEL	- 106 -	Susan Andrev	WS	05-Oct-12	82	747	643	1390	00):15	03:01:22	02:	42 2.14:37:	27 \$1.45	\$929.58
PABX-MEL	- 107 -	Sam Salesma	in	05-Oct-12	94	720	652	1372	2 00	:15	02:56:57	02:	37 2.11:58:	27 \$1.42	\$927.82
PABX-MEL	- 108 -	David Hosking	g	05-Oct-12	63	713	641	1354	00	14	02:43:27	02:	44 2.14:00:	08 \$1.46	\$935.44
PABX-MEL	- 109 -	Gillian Trento	n	05-Oct-12	83	711	646	1357	00	1:14	02:45:03	02:	42 2.13:22:	27 \$1.37	\$886.22
PABX-MEL	- 110 -	Geoff Oden		05-Oct-12	87	801	681	1482	2 00	15:15	03:14:09	02:	43 2.19:28:	14 \$1.50	\$1,020.28
PABX-PER	- 100 -	Charlie Demo	Blue	26-Sep-12	63	497	405	902	2 00	15:15	02:00:40	02:	36 1.15:19:	40 \$1.47	\$593.88
PABX-PER	- 101 -	John Smith	Blue	26-Sep-12	52	450	433	883	00	14	01:46:34	02:	38 1.14:50:	51 \$1.43	\$619.97
PABX-PER	- 102 -	Donald Wilsor	n Red	26-Sep-12	64	486	401	887	00	14	01:56:38	02:	40 1.15:31:	49 \$1.50	
PABX-PER	- 103 -	Geoff Oden	Red	26-Sep-12	59	496	427	923	00	14	01:59:35	02:	40 1.17:08:	30 \$1.30	\$555.48
		Gillian Trento	n	26-Sep-12	67	484	430	914		:15	01:59:42				
		Live Demo		26-Sep-12	57	517	411	928		15:15	02:06:20				
		Sam Salesma	an	26-Sep-12	57	461	398	859		15:15	01:54:01				
PABX-PER	- 107 -	John Duncan		26-Sep-12	51	464	425	889	00	14	01:50:59	02:	48 1.17:35	42 \$1.29	\$548.00

















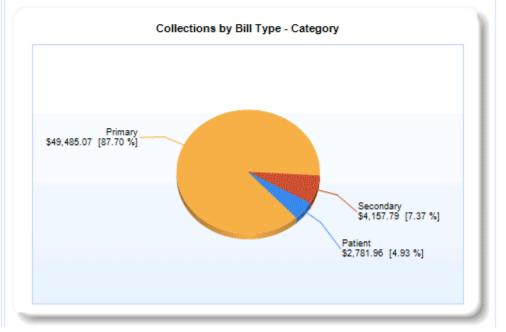
Total Cash Collections Projection

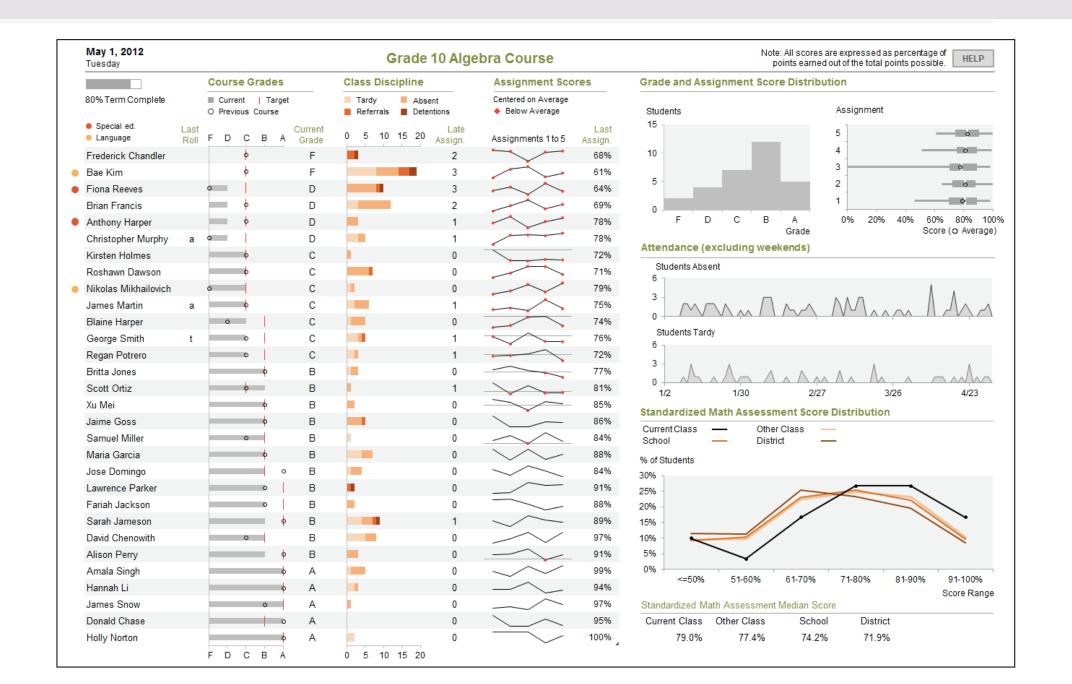


Value as of Post Date:
Target at Period End:
Amount needed to reach target:
Number of Days Remaining in Period:
Amount Needed per Day to Reach Goal:
Average Daily Amount:
Projected Value at Period End:
Projected Shortfall or Surplus:

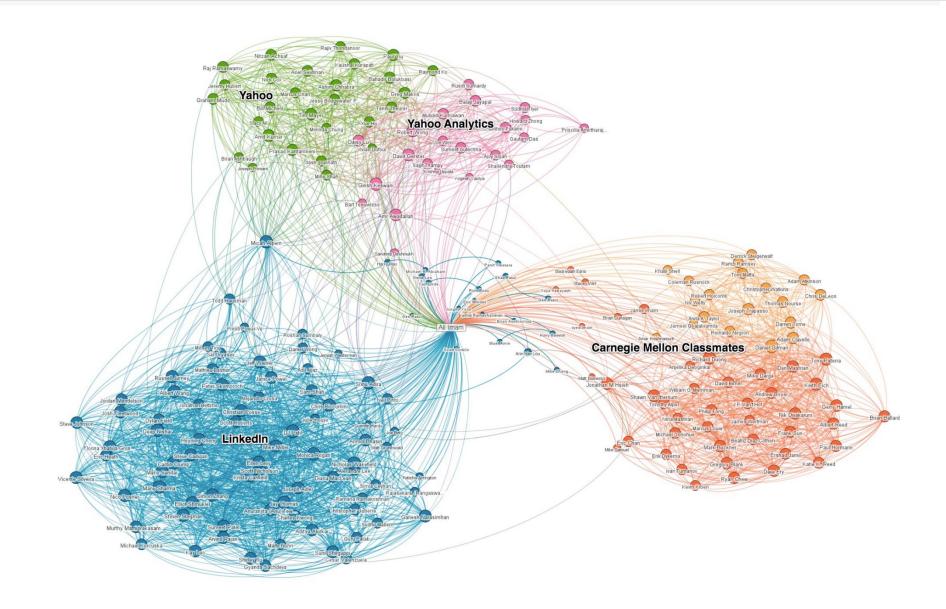
\$56,424.82 \$170,000.00 \$113,575.18 8 \$14,196.90 \$3,761.65 \$86,518.06 (\$83,481.94)

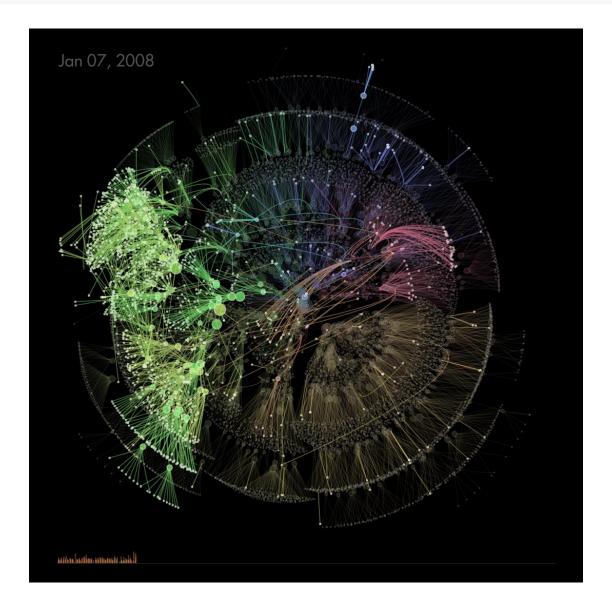
Collections By Bill Type







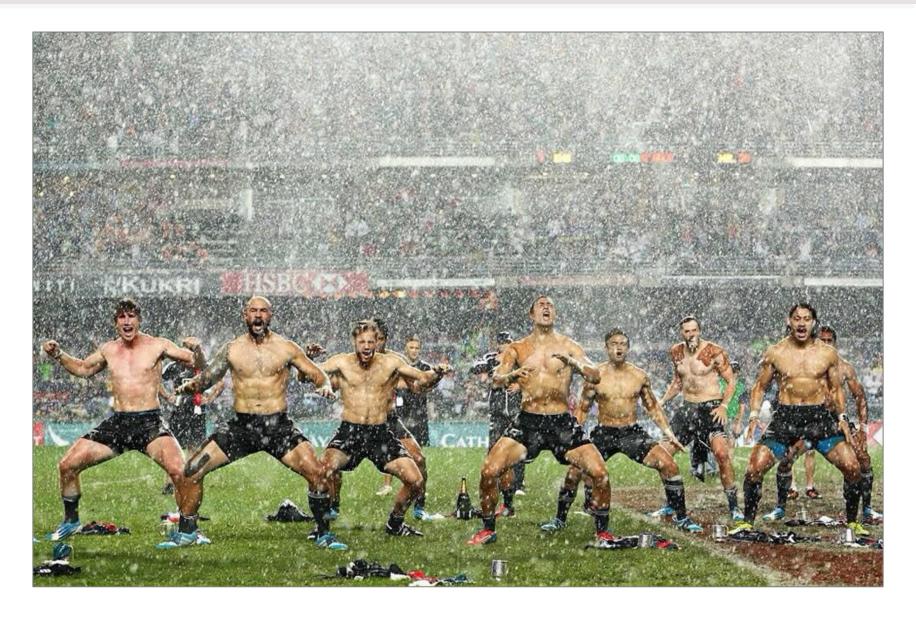




www.youtube.com/watch?v=mkJ-Uy5dt5g

Why visualize data?

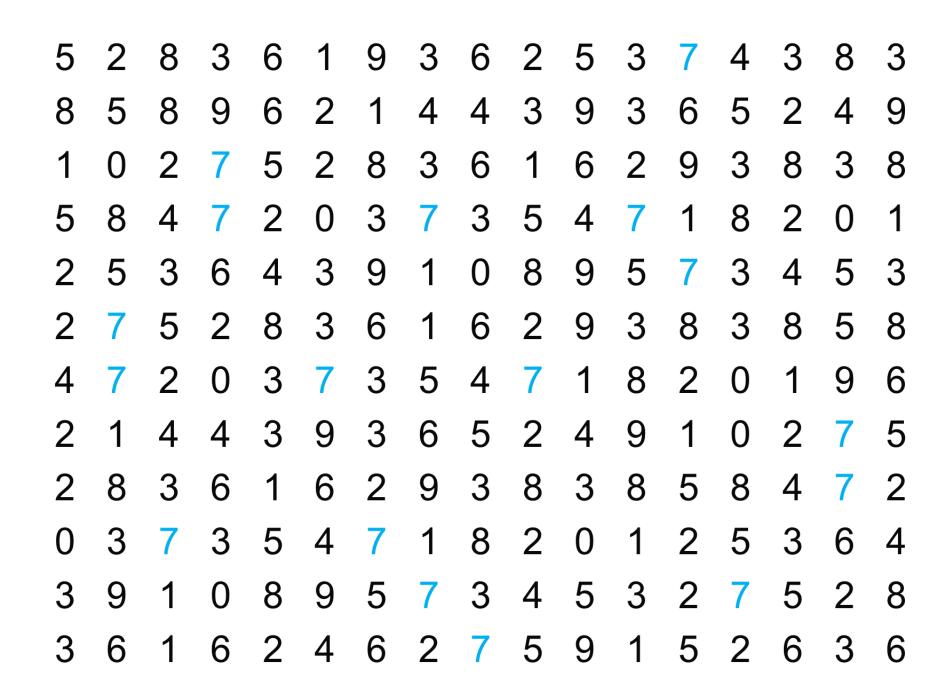
Why visualize data?



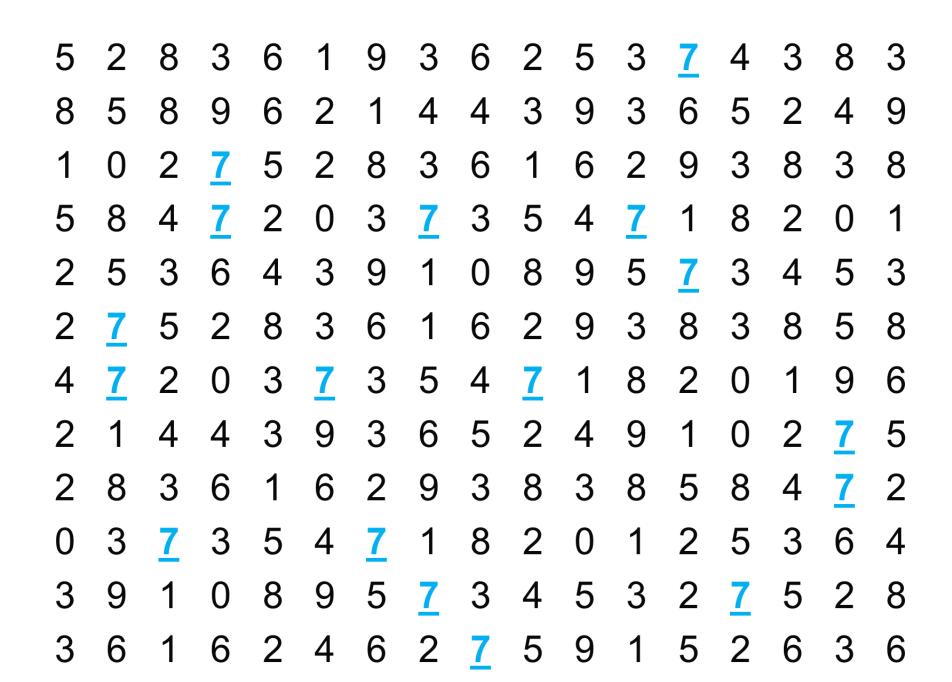
How many times does the digit 7 appear?

How many times does the digit 7 appear?

5	2	8	3	6	1	9	3	6	2	5	3	7	4	3	8	3
8	5	8	9	6	2	1	4	4	3	9	3	6	5	2	4	9
1	0	2	7	5	2	8	3	6	1	6	2	9	3	8	3	8
5	8	4	7	2	0	3	7	3	5	4	7	1	8	2	0	1
2	5	3	6	4	3	9	1	0	8	9	5	7	3	4	5	3
2	7	5	2	8	3	6	1	6	2	9	3	8	3	8	5	8
4	7	2	0	3	7	3	5	4	7	1	8	2	0	1	9	6
2	1	4	4	3	9	3	6	5	2	4	9	1	0	2	7	5
2	8	3	6	1	6	2	9	3	8	3	8	5	8	4	7	2
0	3	7	3	5	4	7	1	8	2	0	1	2	5	3	6	4
3	9	1	0	8	9	5	7	3	4	5	3	2	7	5	2	8
3	6	1	6	2	4	6	2	7	5	9	1	5	2	6	3	6







56789

56789 color

56789 size

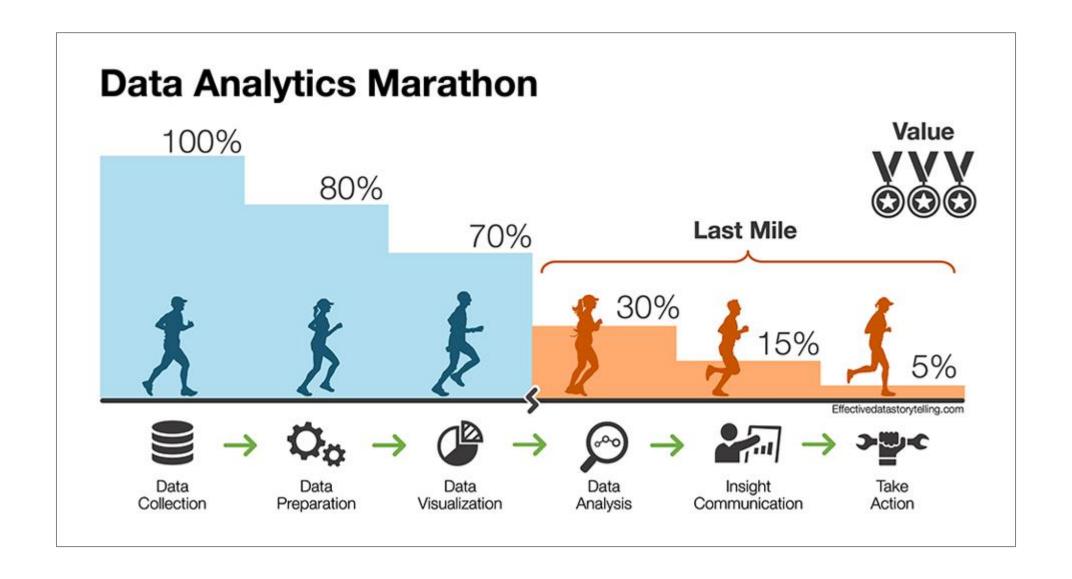
56789 orientation

56 789 texture

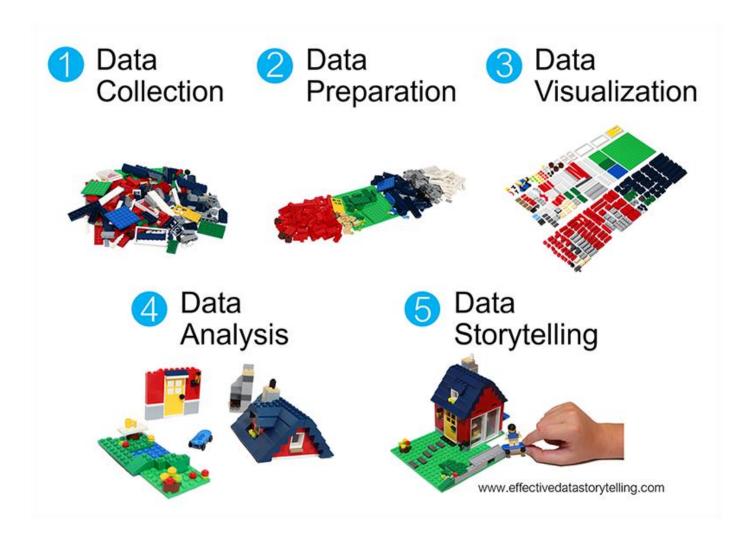
5 2 8 3 6 1 9 3 6 2 5 3 7 4 3 8 3 8 5 8 9 6 2 1 4 4 3 9 3 6 5 2 4 9 0 2 7 5 2 8 3 6 1 6 2 9 3 8 3 5 8 4 7 2 0 3 7 3 5 4 7 1 8 2 0 2 5 3 6 4 3 9 1 0 8 9 5 7 3 4 5 3 2 7 5 2 8 3 6 1 6 2 9 3 8 3 8 5 8 4 7 2 0 3 7 3 5 4 7 1 8 2 0 1 9 4 4 3 9 3 6 5 2 4 9 1 0 2 7 2 8 3 6 1 6 2 9 3 8 3 8 5 8 4 7 2 3 7 3 5 4 7 1 8 2 0 1 2 5 3 6 3 9 1 0 8 9 5 7 3 4 5 3 2 7 5 2 8 1 6 2 4 6 2 7 5 9 1 5 2 6 3 6

Analytics lifecycle

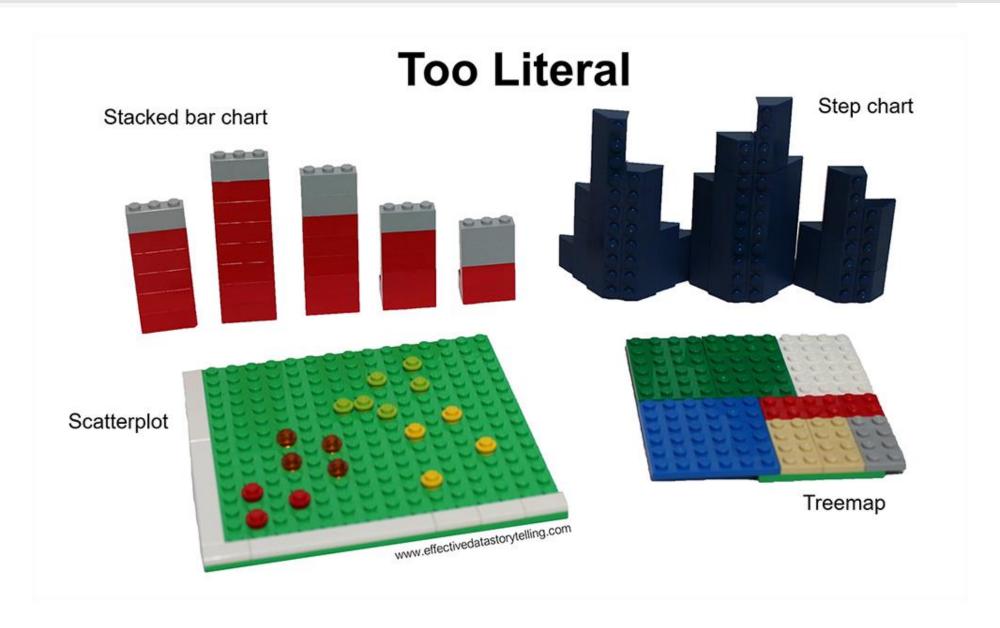
Lifecycle



Lifecycle



Lifecycle



Fundamentals

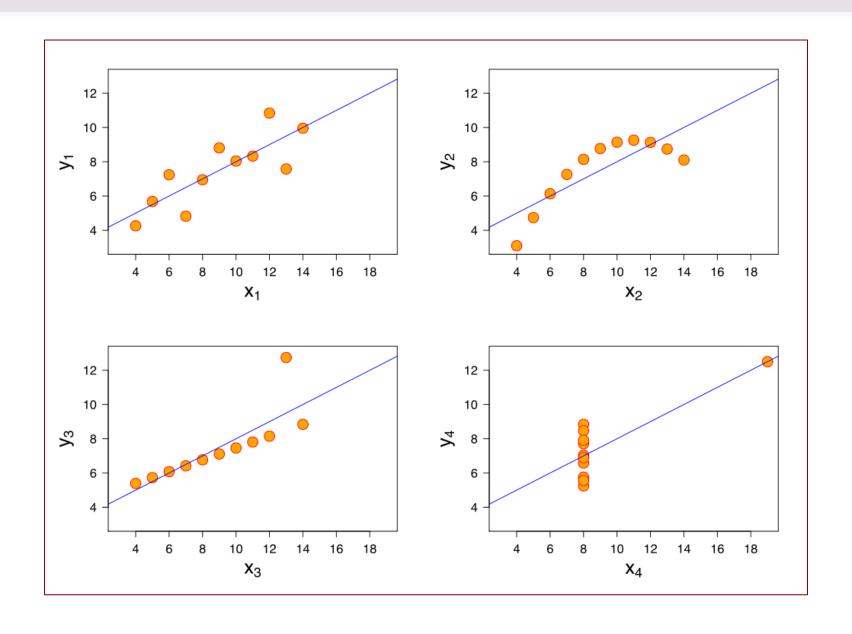
Why?

Anscombe's quartet										
	I	I	I	I	Ш	IV				
x	У	X	у	X	У	X	У			
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58			
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76			
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71			
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84			
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47			
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04			
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25			
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50			
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56			
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91			
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89			

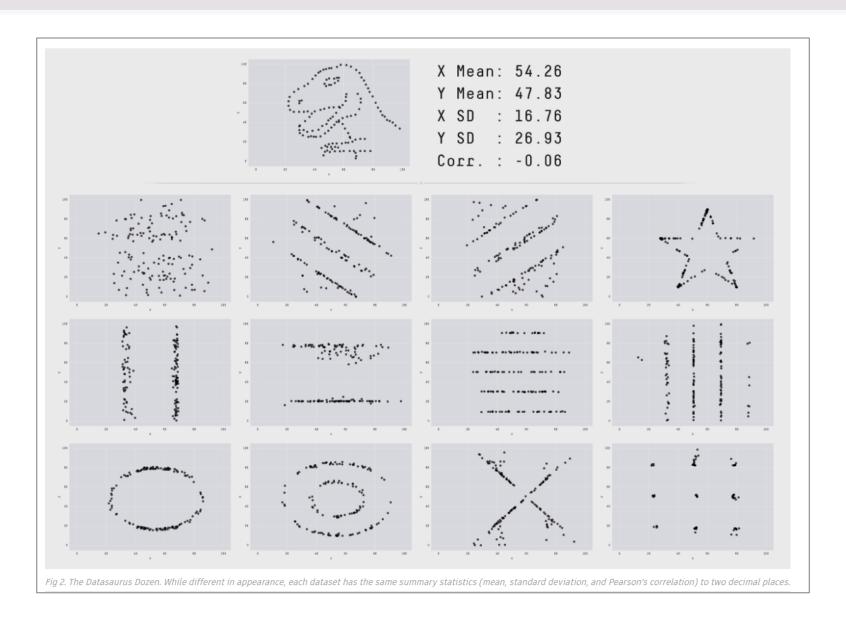
Why?

Property	Value
Mean of x in each case	9 (exact)
ample variance of x in each case	11 (exact)
ean of <i>y</i> in each case	7.50 (to 2 decimal places)
ample variance of <i>y</i> in each case	4.122 or 4.127 (to 3 decimal places)
orrelation between <i>x</i> and <i>y</i> in each	0.816 (to 3 decimal places)
near regression line in each case	y = 3.00 + 0.500x (to 2 and 3 decimal places, respectively)

Why?



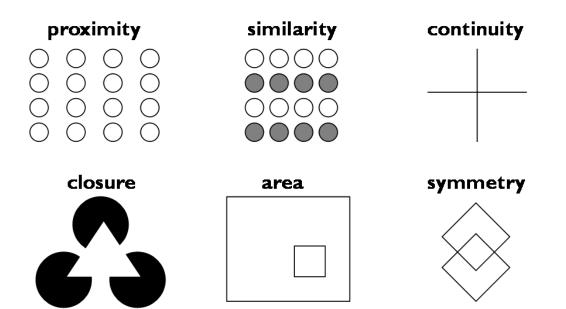
The Datasaurus Dozen



Gestalt

The Gestalt Principles of Grouping

 Gestalt principles explain how eye creates a whole (gestalt) from parts



Spring 2008

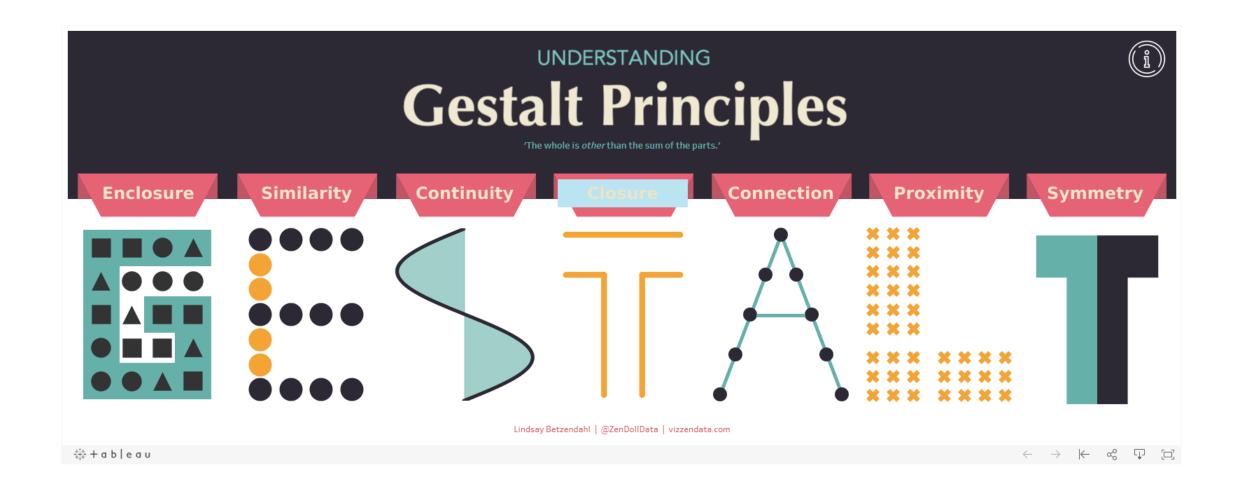
6.831 User Interface Design and Implementation

24

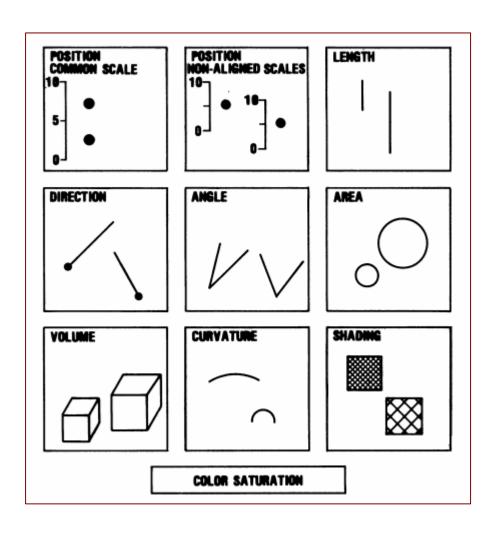
Proximity	Objects that are close together are perceived as a group.		000
Similarity	Objects that share similar attributes (e.g., color or shape) are perceived as a group.		
Enclosure	Objects that appear to have a boundary around them (e.g., formed by a line or area of common color) are perceived as a group.		0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

Closure	Open structures are perceived as closed, complete, and regular whenever there is a way that they can be reasonably interpreted as such.	
Continuity	Objects that are aligned together or appear to be a continuation of one another are perceived as a group.	Curves are perceivedas this,not this.
Connection	Objects that are connected (e.g., by a line) are perceived as a group.	••
		• •

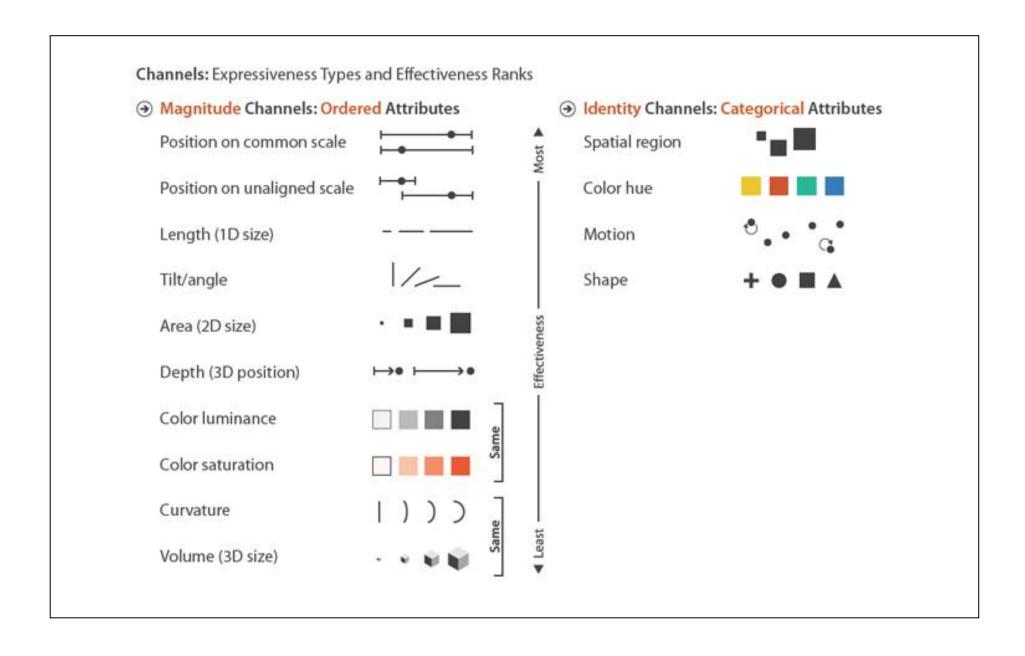
Gestalt



Visual encoding of data



Visual encoding of data



Visual encoding of data

Order of effectiveness

- 1. Position along a common scale (bar chart)
- 2. Positions along nonaligned, identical scales (*Small multiples*)
- 3. Length, direction, angle (pie chart)
- 4. Area (treemap)
- 5. Volume, curvature (3-D bar charts, area charts)
- 6. Shading, color saturation (heat maps, choropleth maps)

Dataviz best practices

Best practices

 There are many different best practices in the field of data visualization based on decades of theory and practice

Some examples

- Chart selection which chart to use for a specific goal
- Effective layouts position the charts and elements properly
- Provide context help your audience to understand your chart
- Color usage Selecting your colors wisely (and consider color blindness)
- Data density use your space smartly, dont waste your ink

Best practices

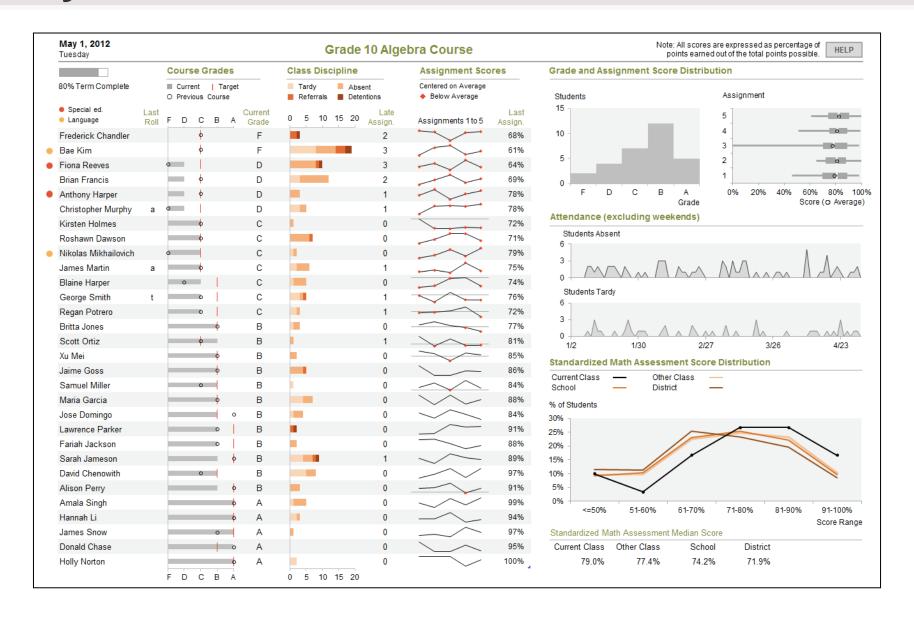
A few key concepts

- Data density use the space effectively to show meaningful data
- Data-ink ratio use the "ink" ONLY the show meaningful data
- No chart junk remove the excess decorations which contains no useful information

Data density



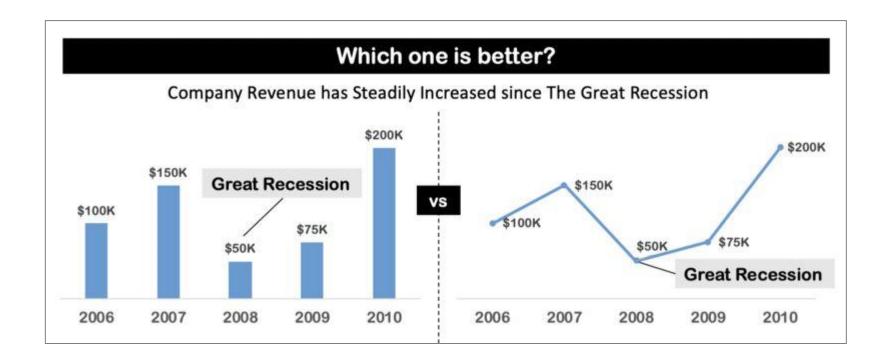
Data density



High data density, much more data visible, effective space usage

Data-Ink Ratio

- Edward Tufte popularized the concept of the "Data-Ink" ratio
 - Maximize the amount of ink used to represent data.
 - Minimize the amount of ink used to draw any unnecessary elements.



Data-ink ratio

Role	Name	Year of the	Debut	Number of Fans	Takedown Rate
Face (The Hero)	The Ultimate Warrior	Tiger	May-2011	97320.00	86.2
Face (The Hero)	Hulk Hogan	Oxen	Jan-2008	988551.00	61.978
Face (The Hero)	Macho Man Randy Savage	Monkey	Feb-2008	157618.00	59.29
Face (The Hero)	Hacksaw Jim Duggan	Pig	Mar-2008	30300.00	53.4332
Face (The Hero)	Superfly Jimmy Snuka	Dragon	Mar-2008	12341.00	52.7
Heel (The Bad Guy)	Rowdy Roddy Piper	Rooster	Jun-1968	71645.00	45.4
Heel (The Bad Guy)	The Million Dollar Man Ted DiBiase	Rat	Apr-1975	449342.00	43.7689
Heel (The Bad Guy)	Mr. Perfect Curt Henning	Rat	May-1980	13773.00	38
Heel (The Bad Guy)	Jake the Snake Roberts	Snake	Jul-1975	5609.00	37.99
Jobber (The Unknown)	Brad Smith	Sheep	Aug-2008	1103.00	36.316
Jobber (The Unknown)	Ted Duncan	Sheep	Aug-2008	200.00	33.61
Jobber (The Unknown)	Joey the Uber Nerd Cherdarchuk	Snake	Aug-2008	5.00	21.0196

Remove to improve the data tables edition

Before

Role	Name	Year of the	Debut	Number of Fans	Takedown Rate
Face (The Hero)	The Ultimate Warrior	Tiger	May-2011	97320.00	86.2
Face (The Hero)	Hulk Hogan	Oxen	Jan-2008	988551.00	61.978
Face (The Hero)	Macho Man Randy Savage	Monkey	Feb-2008	157618.00	59.29
Face (The Hero)	Hacksaw Jim Duggan	Pig	Mar-2008	30300.00	53.4332
Face (The Hero)	Face (The Hero) Superfly Jimmy Snuka		Dragon Mar-2008		52.7
Heel (The Bad Guy) Rowdy Roddy Piper		Rooster	Jun-1968	71645.00	45.4
Heel (The Bad Guy)	The Million Dollar Man Ted DiBiase	Rat Apr-1975		449342.00	43.7689
Heel (The Bad Guy)	Mr. Perfect Curt Henning	Rat	May-1980	13773.00	38
Heel (The Bad Guy)	Jake the Snake Roberts	Snake	Jul-1975	5609.00	37.99
Jobber (The Unknown)	Brad Smith	Sheep	Sheep Aug-2008		36.316
Jobber (The Unknown)	Ted Duncan	Sheep	Aug-2008	200.00	33.61
Jobber (The Unknown)	Joey the Uber Nerd Cherdarchuk	Snake	Aug-2008	5.00	21.0196

Created by **Darkhorse Analytics**

www.darkhorseanalytics.com

After

Role	Name	Year of the	Debut	Thousands of Fans	Takedown Rate
Face (The Hero)	The Ultimate Warrior	Tiger	May-2011	97.3	86.2
	Hulk Hogan	Oxen	Jan-2008	988.6	62.0
	Macho Man Randy Savage	Monkey	Feb-2008	157.6	59.3
	Hacksaw Jim Duggan	Pig	Mar-2008	30.3	53.4
	Superfly Jimmy Snuka	Dragon	Mar-2008	12.3	52.7
Heel (The Bad Guy)	Rowdy Roddy Piper	Rooster	Jun-1968	71.6	45.4
	The Million Dollar Man Ted DiBiase	Rat	Apr-1975	449.3	43.8
	Mr. Perfect Curt Henning	Rat	May-1980	13.8	38.0
	Jake the Snake Roberts	Snake	Jul-1975	5.6	38.0
Jobber (The Unknown)	Brad Smith	Sheep	Aug-2008	1.1	36.3
	Ted Duncan	Sheep	Aug-2008	0.2	33.6
	Joey the Uber Nerd Cherdarchuk	Snake	Aug-2008	0.0	21.0

Created by Darkhorse Analytics

www.darkhorseanalytics.com

Thank You

