Data Visualisation

(CMP020L013S)

Week 3: Design & Principals

Mohammad Javaheri

(Dr Mohammad Ali Javaheri Javid)

Agenda

- Review
- Visualising Data: What, Why, and How to Start
  - What makes a data visualisation?
  - Why they can be effective?
  - How to plan?
- Identifying the Story in Data

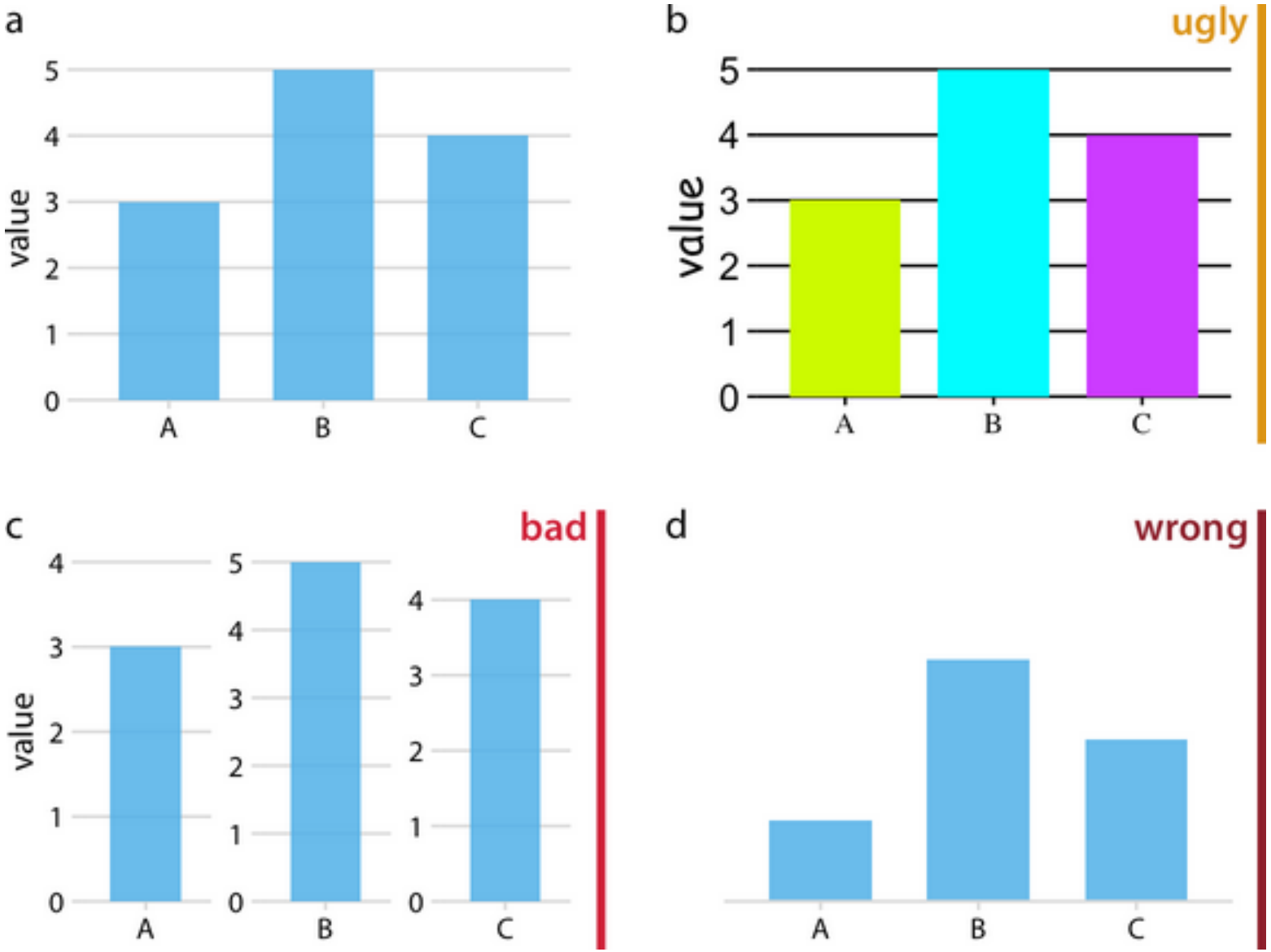
Review: Use Cases of Data Visualisation

- Exploratory
  - helps discover trends and patterns interactive
  - quick and dirty
  - early stages of research
- Explanatory
  - conveys a clear message
  - makes a point or answers a question
  - polished, publication/presentation ready
  - final stages of research

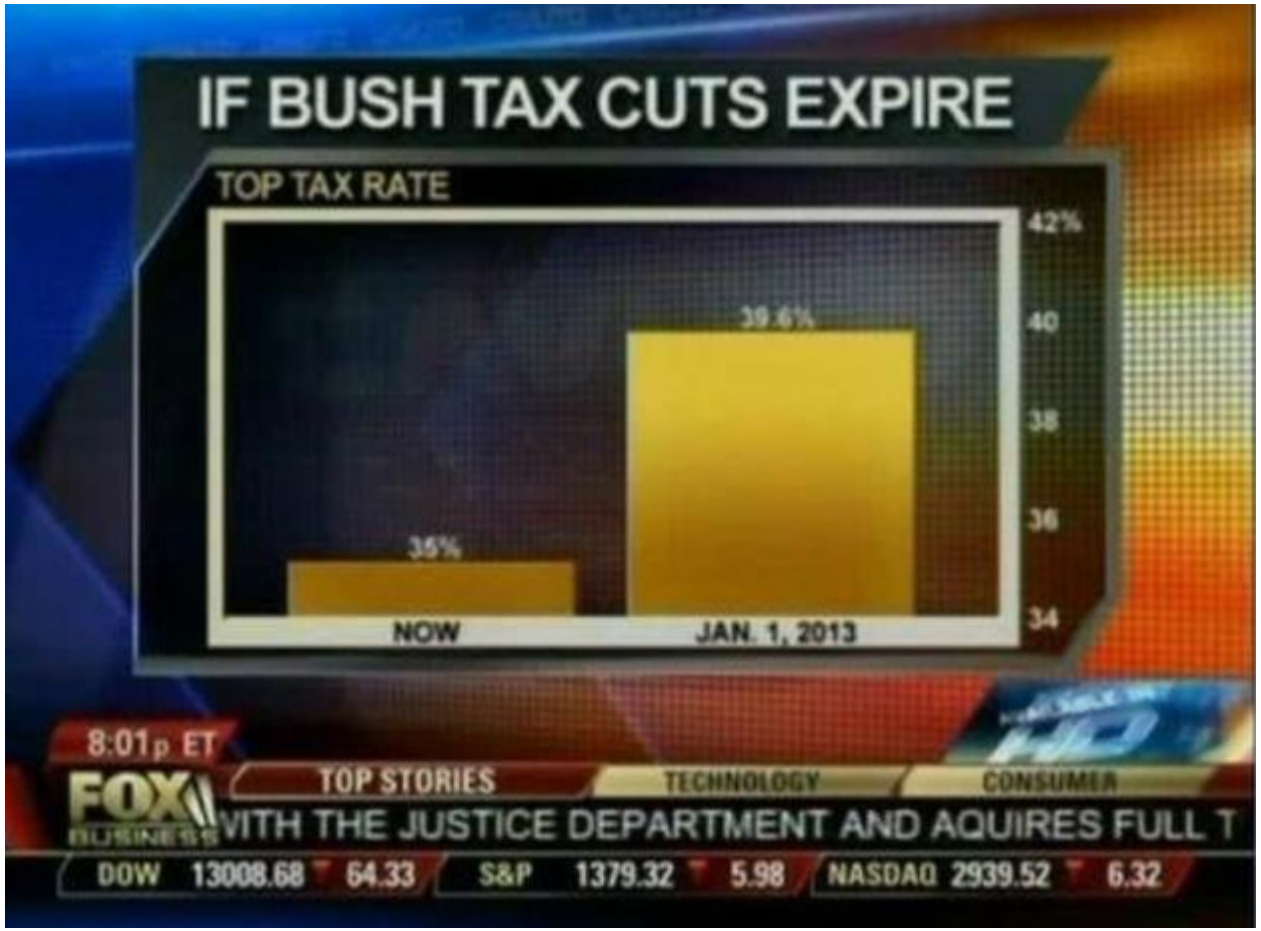
Review: Good Data Visualization

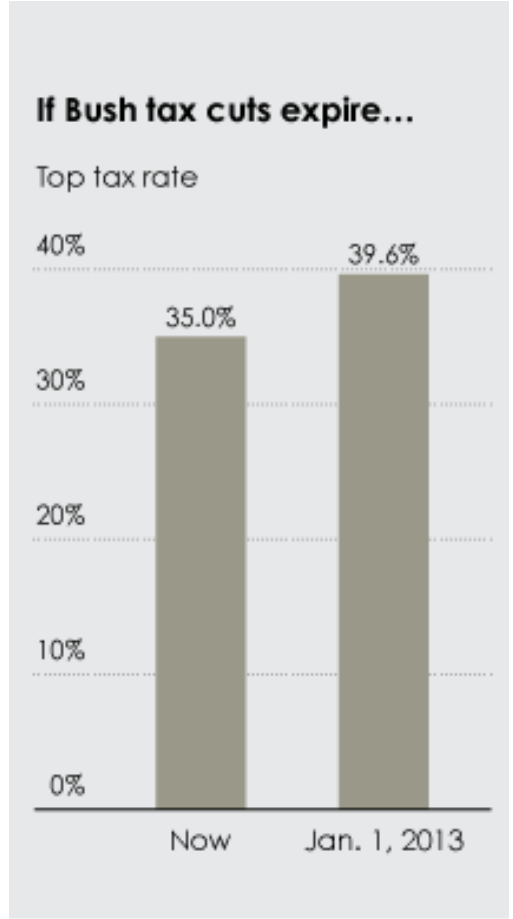
- Makes data accessible
- Combines strengths of humans and computers
- Enables insight
- Communicates effectively
- A visualisation is more effective than another visualisation if the information conveyed by one visualisation is more readily perceived than the information in the other visualisation.

Review: A Case for Ugly Visualisations



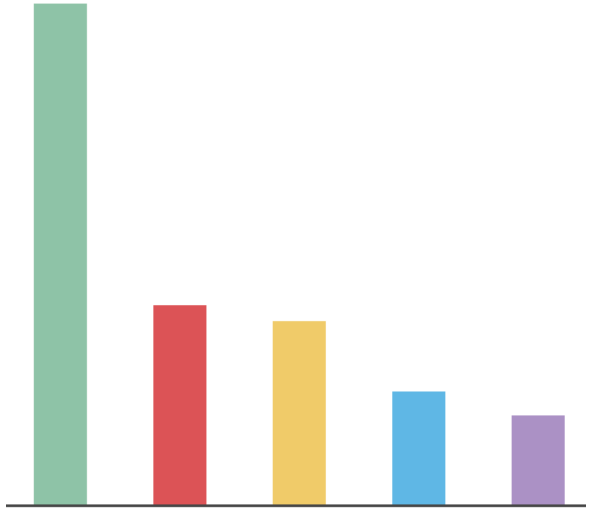
Example: Misleading Bar Chart

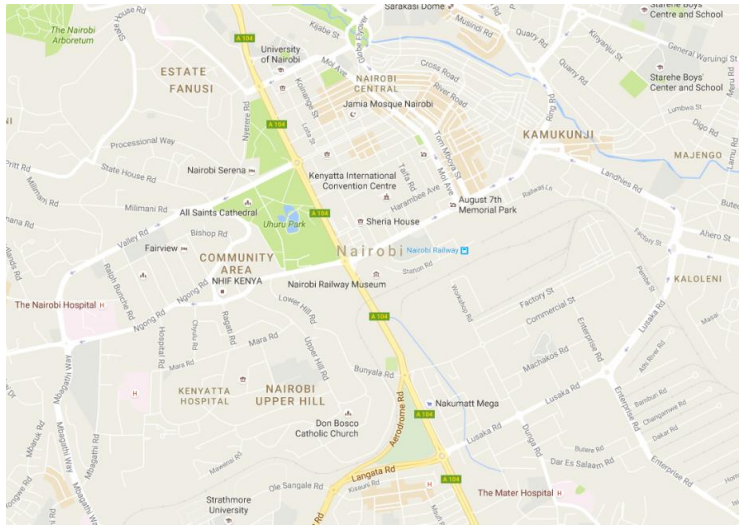




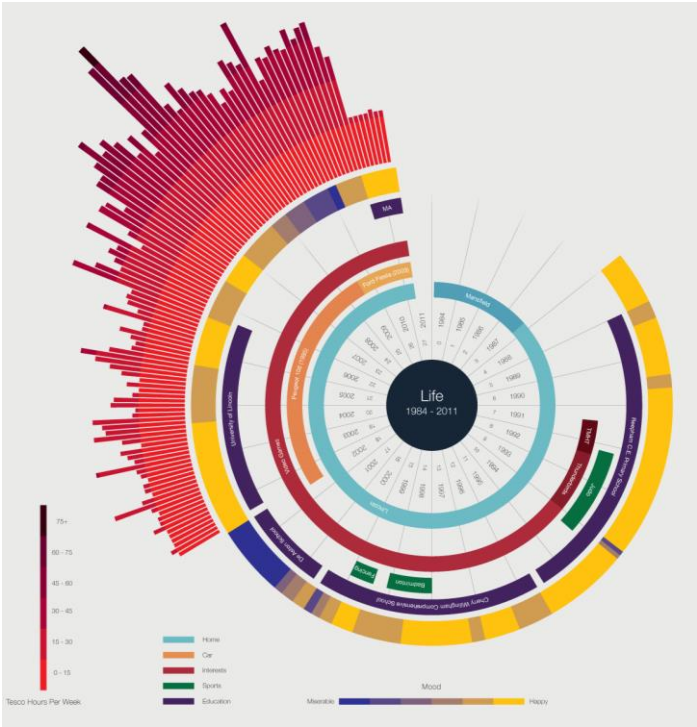
SOURCE: http://flowingdata.com/category/statistics/mistaken-data/

Which one is a Data Visualisation?





58%



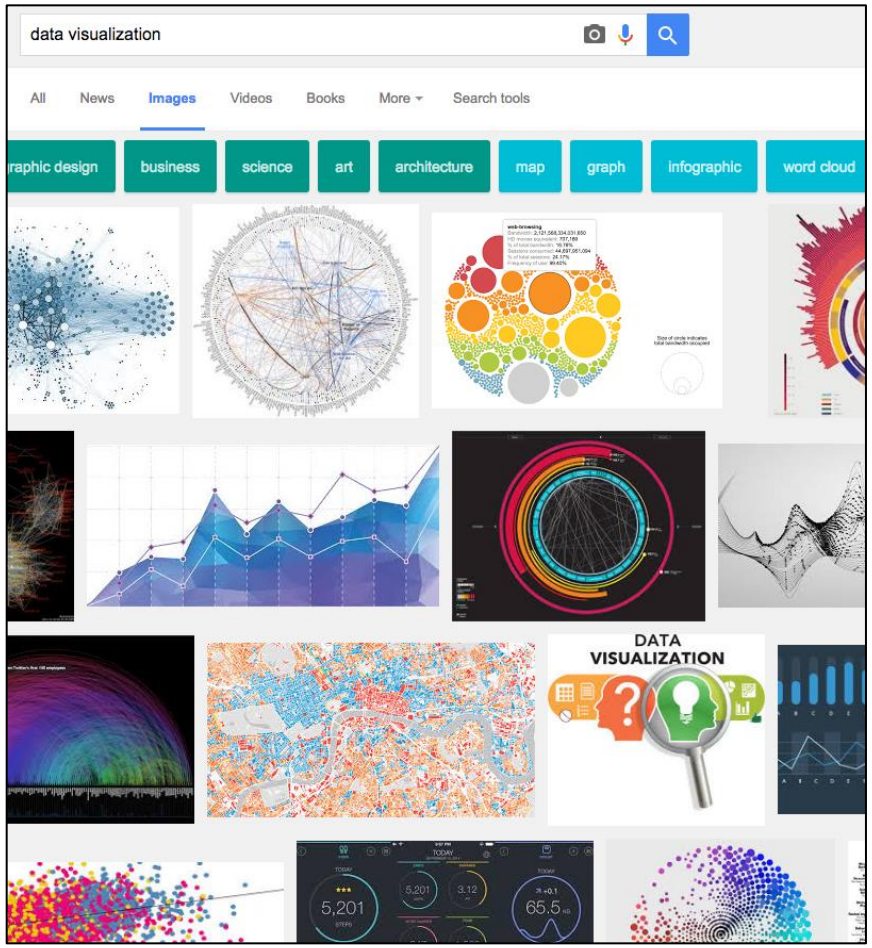
- "Nicely designed posters with a few numbers on them aren't really data viz."
  - Amanda Cox, who does a lot of data vis work at *The New York Times*

- This quote nails one thing: you must have numbers but as numbers are not enough you need to enhance user cognition using graphs and effectively communicate a message.

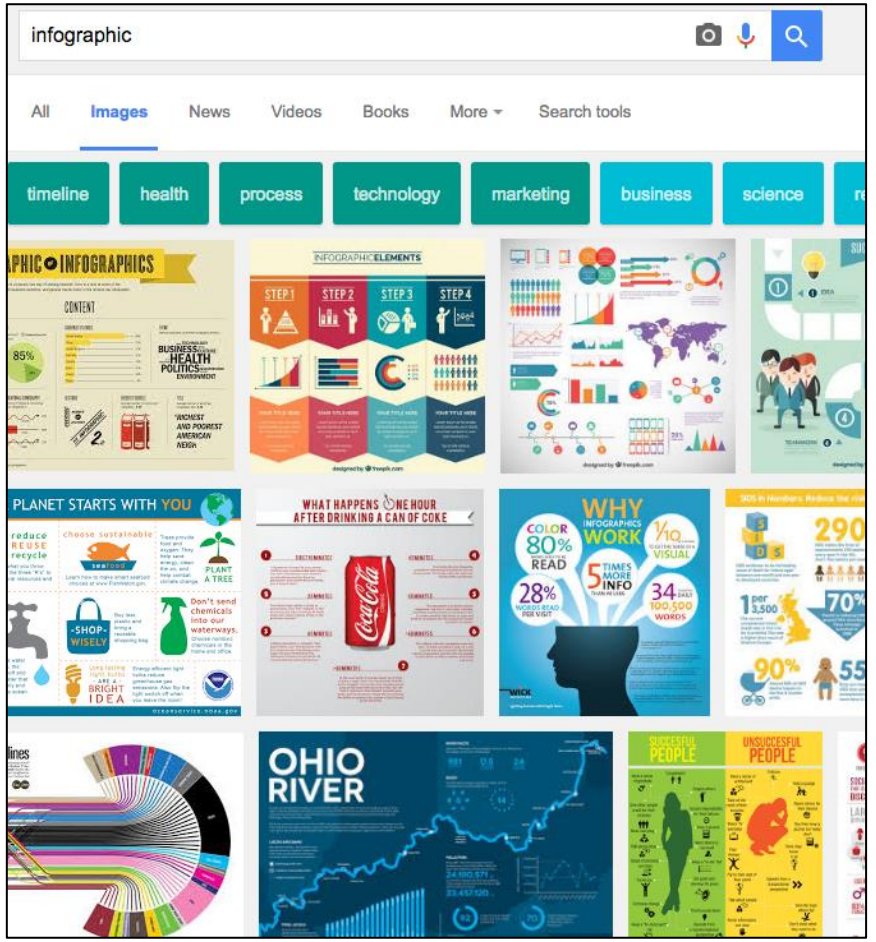
Data Visualisation vs. Infographic

- "Infographic" is a term often used interchangeably with "data viz."
- You're going to find a lot of opinions out there on what constitutes a data viz versus an infographic.
- Just remember that you need data plus a visual representation of the data to make a data vis!
- To make the data easier to understand, we do not just arrange the data nicely on a screen.
- Many infographics do not, in fact, include data.

What Google Thinks

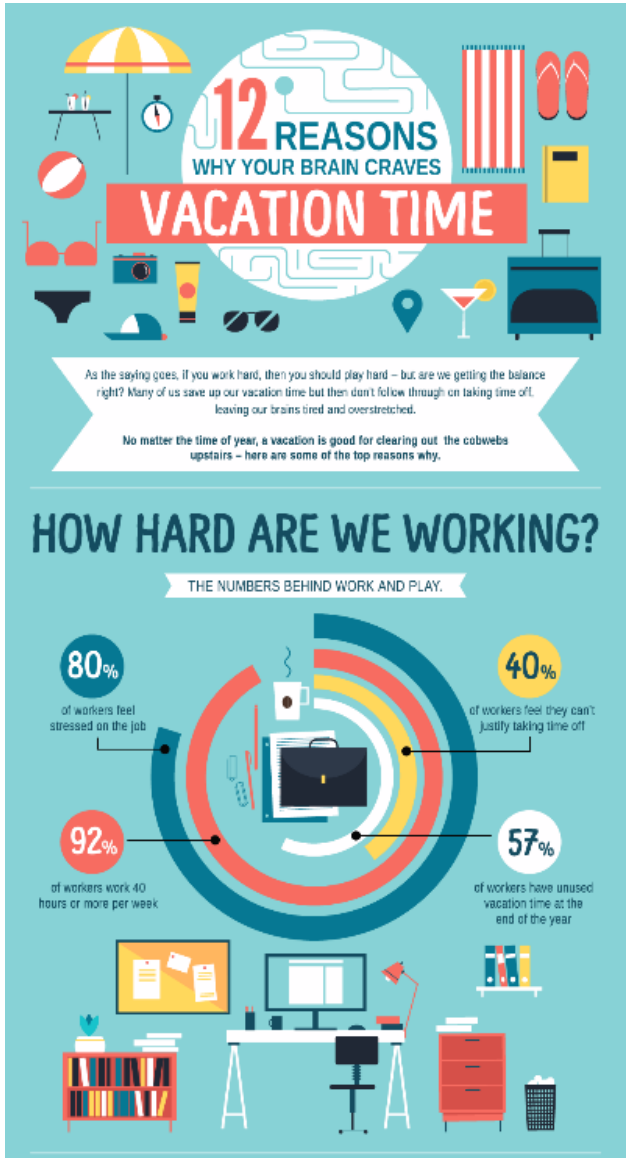


Graphs, geometric shapes, color coded, number driven



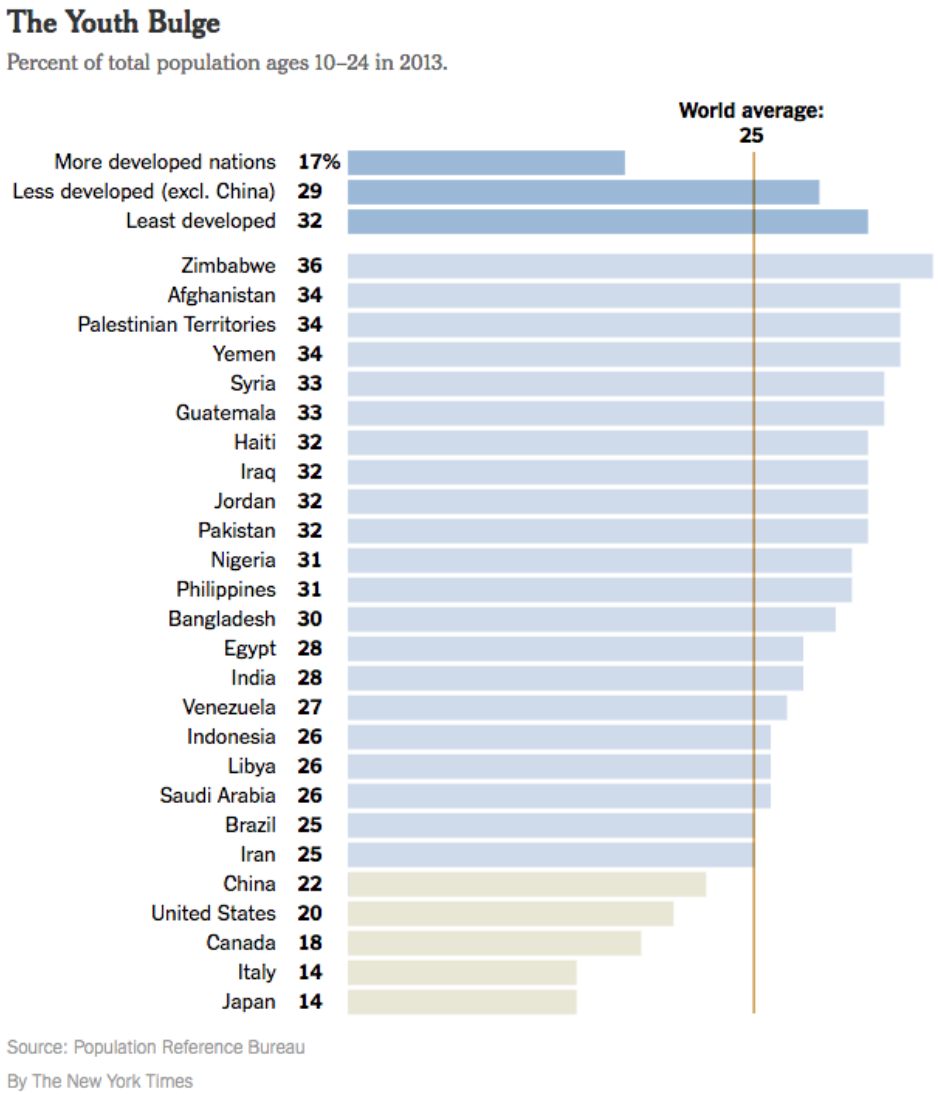
Collection of facts, icons, pictorials, text driven

Classical Example of Infographic



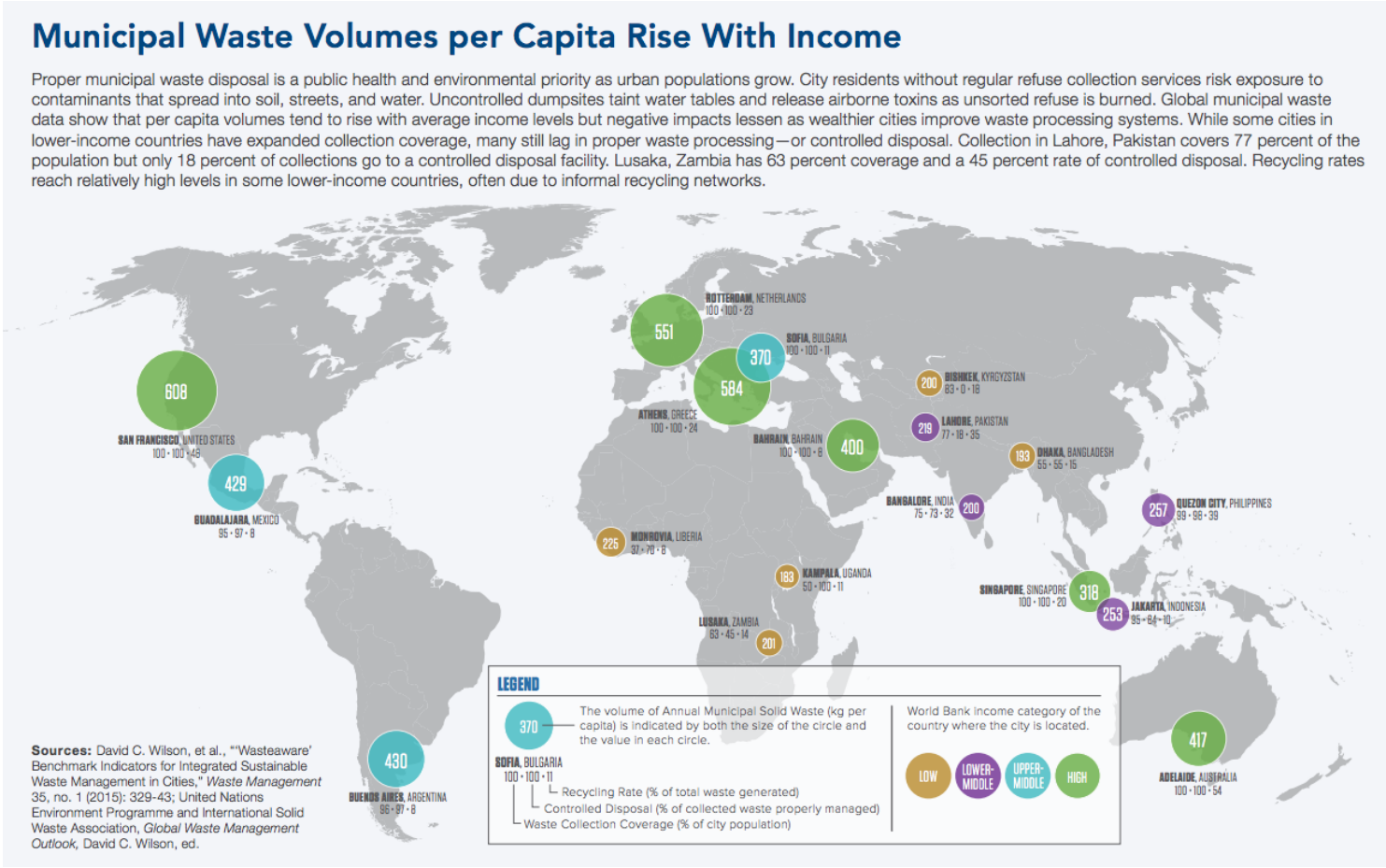
Data Visualization Example with a Message

- Its main purpose is to show how these numbers relate to each other in size.
- There is much less text here.
- Your eye is simply drawn to the lengths of the bars.
- In this example, we can see the percentage of a country's population between the ages of 10 and 24.
- The vertical line shows the world average is 25%, so you can tell which countries are above or below the average.
- The variation in the colour of the bars helps you quickly see that the top section is regions and the bottom section is countries below the world average.



Data Visualisation Example with a Message 2

- This example shows how municipal waste increases when cities have higher incomes.
- The bubbles are sized proportionally according to the volume of waste and colour-coded to show relative income level.
- we do have more text, so we're not excluding text as a possible component of DV.
- But the main focus here is the map, which makes the data visual through the location of the numbers on the map, the size of the shape they are on, and the colour.



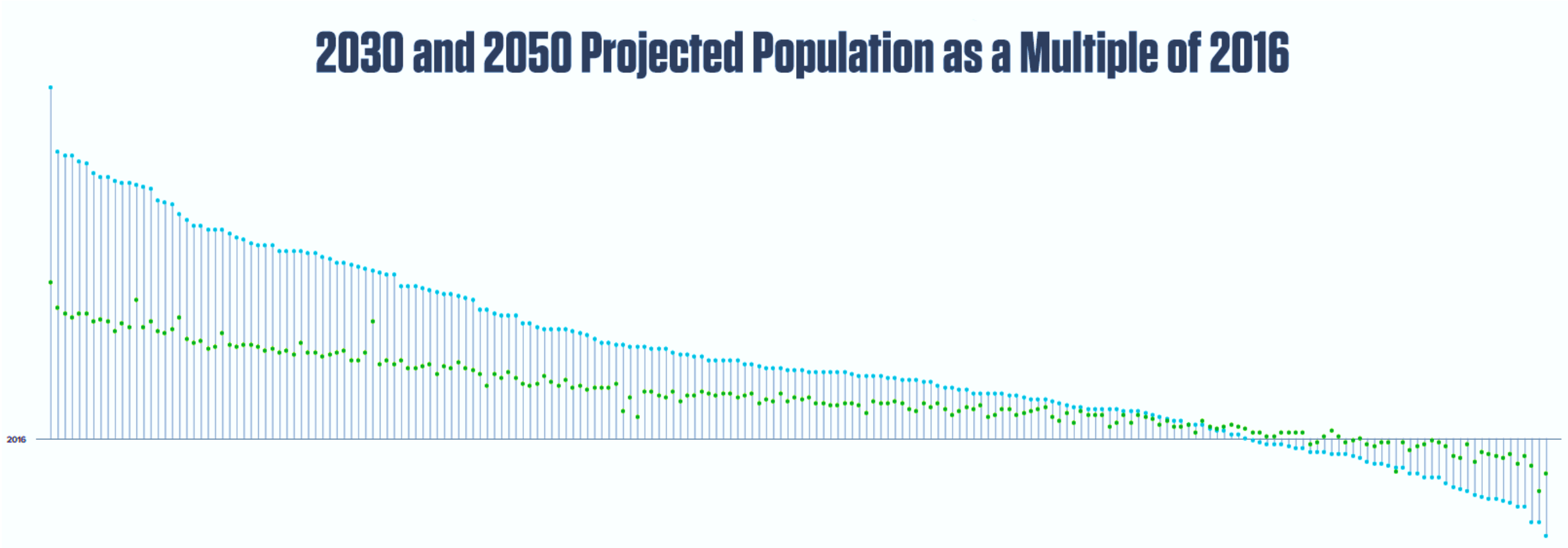
Added Value of Visualisations

- "visualizations that do some work; they persuade, they explain, they explore. Beauty is desirable, but I'm not excited by pretty charts that are also pretty useless."
  - Graham Wills, Data Science and Visualization Expert at IBM
- The idea is that data vis does more for us than looks pretty.

World Population Data Before Visualization

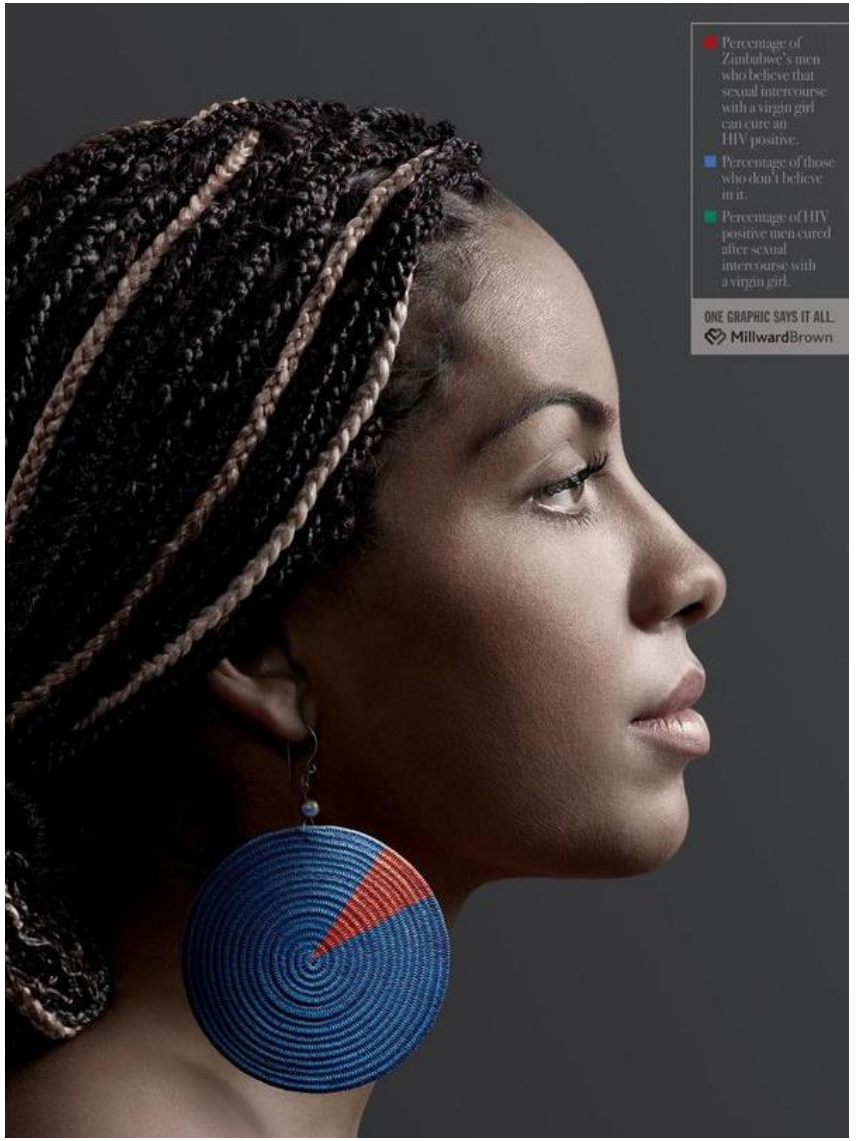
	POPULATION, HEALTH, AND ENVIRONMENT DATA AND ESTIMATES FOR THE COUNTRIES AND REGIONS OF THE WORLD													
	Population mid-2016 (millions)	Births per 1,000 Population	Deaths per 1,000 Population	Net Migration Rate per 1,000	Population mid-2030 (millions)	Population mid-2050 (millions)	Infant Mortality Rate[a]	Total Fertility Rate[b]	Percent of Population Ages <15	Percent of Population Ages 65+	GNI per Capita ($US) 2015	Percent Urban	Percent of Married Women 15-49 Using Contraception[c] All Methods	Percent of Married Women 15-49 Using Contraception[c] Modern Methods
WORLD	**7,418**	**20**	**8**	**—**	**8,539**	**9,869**	**36**	**2.5**	**26**	**8**	**15,415**	**54**	**62**	**56**
MORE DEVELOPED	**1,254**	**11**	**10**	**3**	**1,298**	**1,322**	**5**	**1.7**	**16**	**18**	**39,963**	**78**	**70**	**62**
LESS DEVELOPED	**6,164**	**22**	**7**	**-0**	**7,241**	**8,548**	**39**	**2.6**	**28**	**7**	**10,214**	**49**	**61**	**55**
LESS DEVELOPED (Excl. China)	**4,778**	**24**	**7**	**-0**	**5,821**	**7,195**	**43**	**2.9**	**31**	**5**	**8,936**	**47**	**54**	**46**
LEAST DEVELOPED	**962**	**33**	**9**	**-1**	**1,318**	**1,923**	**59**	**4.3**	**41**	**4**	**2,424**	**32**	**36**	**32**
AFRICA	**1,203**	**36**	**10**	**-1**	**1,681**	**2,527**	**57**	**4.7**	**41**	**4**	**4,802**	**41**	**35**	**30**
SUB-SAHARAN AFRICA	**974**	**37**	**11**	**-0**	**1,388**	**2,128**	**62**	**5.0**	**43**	**3**	**3,606**	**39**	**31**	**26**
NORTHERN AFRICA	**229**	**29**	**6**	**-1**	**293**	**400**	**29**	**3.4**	**32**	**5**	**9,798**	**51**	**52**	**46**
Algeria	40.8	26	5	0	50.6	63.2	21	3.1	29	6	14,280	71	57	48
Egypt	93.5	31	6	-1	121.6	168.8	22	3.5	31	4	10,690	43	58	57
Libya	6.3	20	5	-10	7.4	8.4	23	2.4	30	5	15,140	79	42	20
Morocco	34.7	20	6	-2	38.1	38.8	24	2.4	25	6	7,680	60	67	57
Sudan	42.1	37	8	-3	61.7	105.0	51	5.2	43	3	4,080	34	12	12
Tunisia	11.3	20	7	-1	13.1	14.9	17	2.4	24	8	11,060	68	62	50
Western Sahara[e]	0.6	18	6	8	0.7	0.9	35	2.1	26	3	—	81	—	—
WESTERN AFRICA	**359**	**39**	**11**	**-1**	**515**	**800**	**64**	**5.4**	**44**	**3**	**4,135**	**45**	**18**	**13**
Benin	10.8	36	9	-0	16.6	24.0	66	4.7	45	3	2,100	44	18	12
Burkina Faso	19.0	41	10	-1	28.2	47.0	65	5.7	49	3	1,640	30	21	20
Cape Verde	0.5	21	5	-4	0.6	0.7	19	2.3	28	6	6,390	66	61	57
Côte d'Ivoire	23.9	37	13	0	33.0	50.1	69	4.9	42	3	3,240	54	18	12
Gambia	2.1	41	9	-1	3.2	5.1	45	5.6	46	2	1,580	60	9	8
Ghana	28.2	33	8	-1	37.1	50.4	41	4.2	39	5	4,070	54	35	29
Guinea	11.2	37	10	-0	18.3	27.5	67	5.1	43	3	1,120	37	6	5
Guinea-Bissau	1.9	37	12	-1	2.5	3.6	88	4.9	43	3	1,450	49	16	14
Liberia	4.6	35	9	-1	6.4	9.4	54	4.7	42	3	720	50	20	19
Mali	17.3	44	13	-3	26.1	43.6	56	6.0	47	3	2,360	40	16	15
Mauritania	4.2	31	8	-1	5.7	8.0	72	4.2	40	3	3,710	60	11	10
Niger	19.7	49	9	-0	34.3	68.9	56	7.6	50	3	950	22	14	12
Nigeria	186.5	39	13	-0	261.9	397.5	69	5.5	43	3	5,800	48	15	10
Senegal	14.8	38	6	-1	21.6	34.4	39	5.0	44	4	2,390	45	23	21
Sierra Leone	6.6	37	14	-1	8.9	12.6	89	4.9	42	3	1,560	40	17	16
Togo	7.5	36	9	-0	10.9	17.4	47	4.7	42	3	1,320	38	20	17
EASTERN AFRICA	**394**	**36**	**9**	**-1**	**561**	**844**	**51**	**4.8**	**43**	**3**	**2,019**	**26**	**40**	**37**
Burundi	11.1	42	10	0	17.2	30.4	63	6.1	46	2	730	12	32	30
Comoros	0.8	33	7	-3	1.0	1.4	36	4.3	40	3	1,430	28	19	14
Djibouti	0.9	25	9	-3	1.1	1.2	53	3.2	33	4	—	77	19	18
Eritrea	5.4	34	7	-5	7.3	10.4	43	4.2	43	3	—	23	8	7
Ethiopia	101.7	30	7	-0	132.9	168.6	47	4.2	41	3	1,620	20	37	36

World Population Data After Visualisation

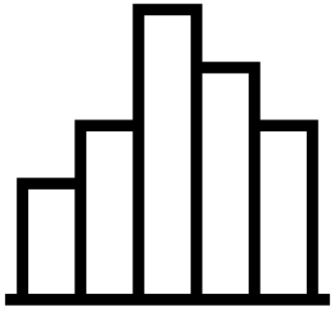


Appeal

- We could list percentages for stats.
- In this unusual example, we have a photograph playing a central role.
- The woman's earring is a pie chart with the red segment showing the percent of Zimbabwe's men who believe having sex with a virgin girl can cure HIV.
- The blue portion represents those who don't believe this.
- The legend has also added a third category not shown in the pie chart—men who have actually been cured of HIV through this method—to make the point that it's not a real solution.
- They've taken what could simply be a summary of percentages or a basic pie chart and turned it into something memorable.
- This example shows how data vis can draw attention in ways that text alone cannot.

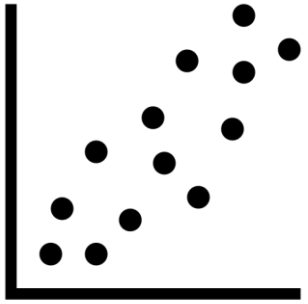


Recall on DV Tools

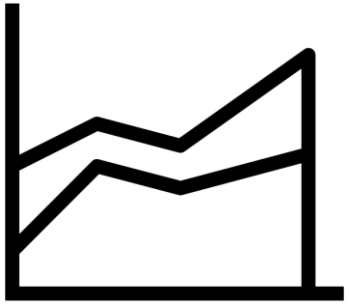


Bar chart

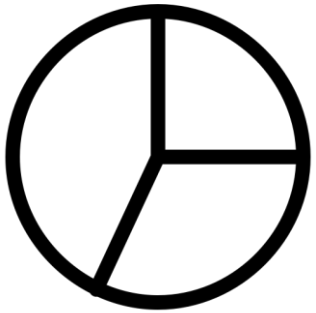
Line chart



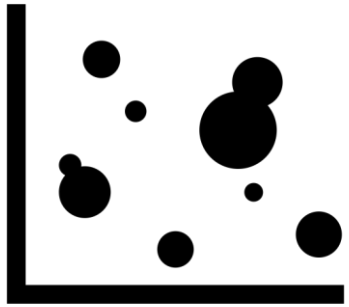
Scatter plot



Area chart



Pie chart



Bubble chart

Heat map

Planning for DV

- A Good Data Visualisation:

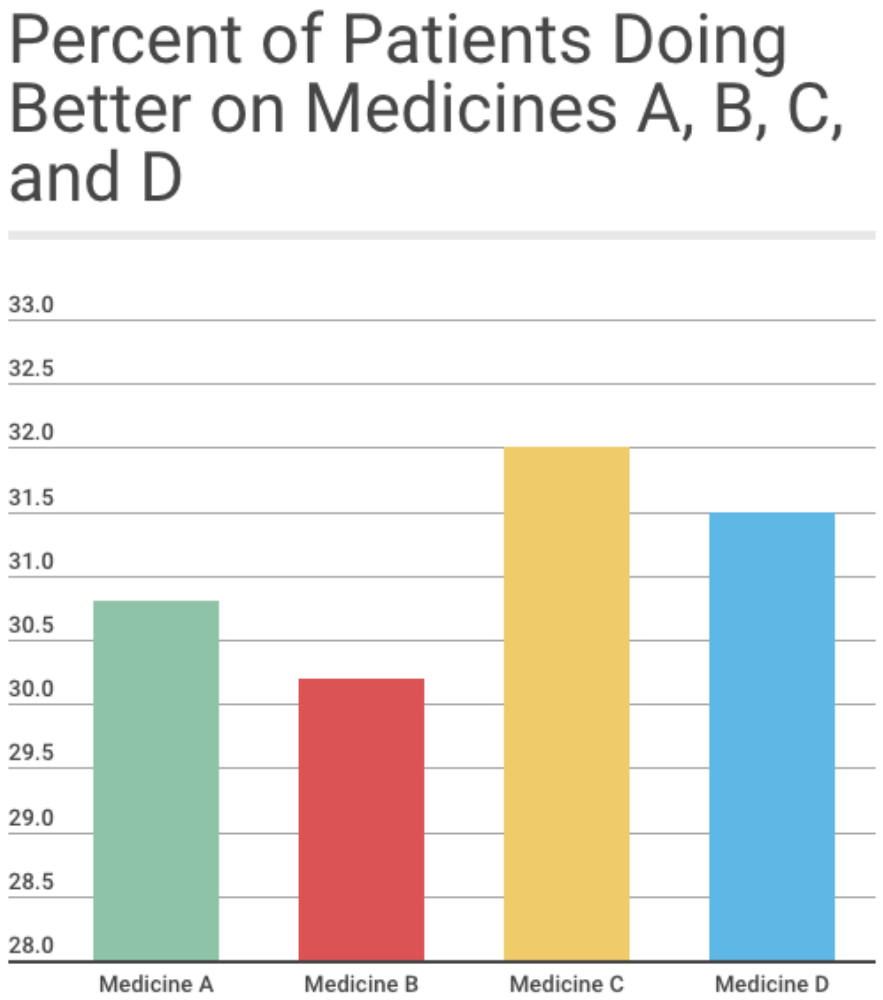
1. Must be accurate and not misrepresent data
   - Accuracy includes not misrepresenting data.
   - People should have an understanding of the truth, so make sure your visualisation is being truthful.
2. Must be easy to understand
   - If it requires several minutes to comprehend/if it's confusing, then it's not doing its job.
3. Relates to your audience
   - Consider your audience's point of view. Whenever you create a visualisation, it should answer the question, "So what?" so that your audience cares about the information being presented (processed).
4. Only shows what's necessary
   - You want everything streamlined and to the point. If you need to add annotations (text etc) or sth visual to make it clearer, then do it.
   - You should achieve a balance between minimising visual clutter and giving enough information.

Common Mistakes With Data Accuracy

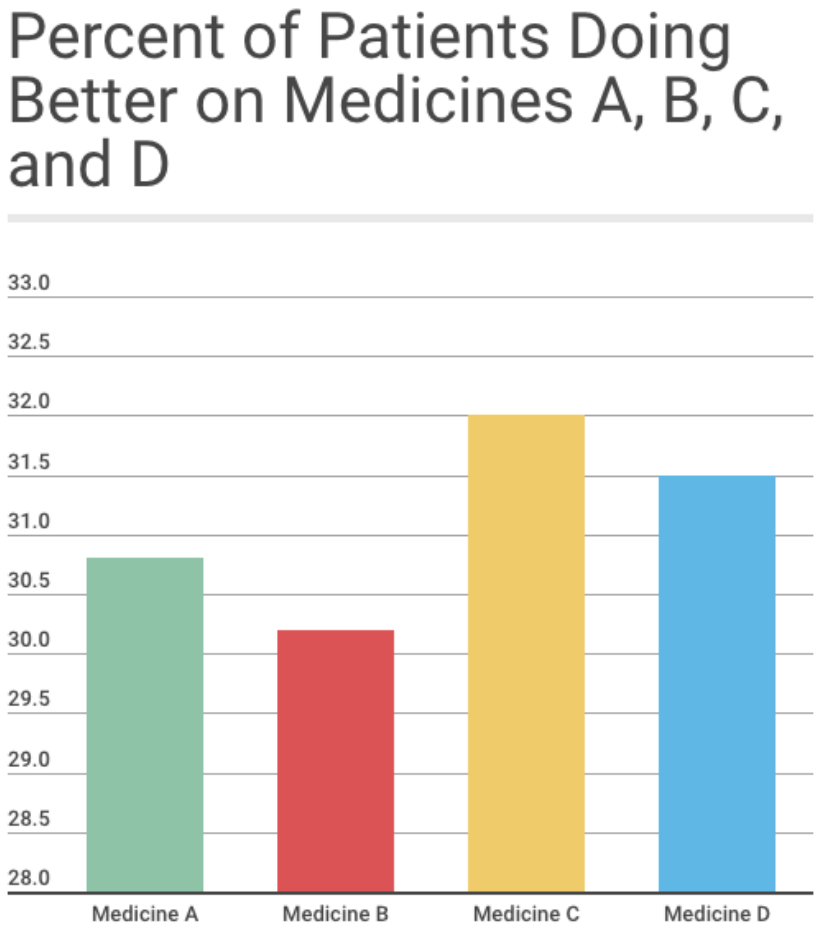
- Mixing up percent change and percentage point change
- Pie chart does not add up to 100 percent
- Proportionally sized bubbles using diameter instead of area for sizing
- Oversimplifying data so you don’t get enough context
- Truncating y-axis

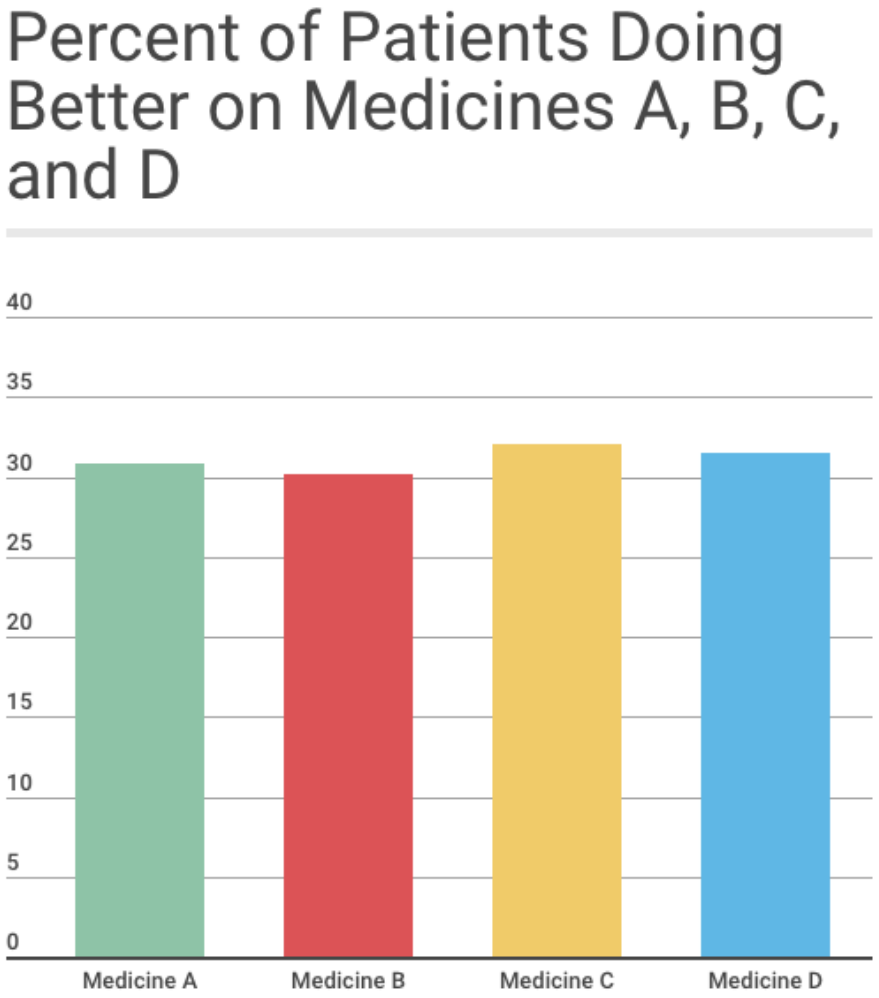
Truncated Axis (before)

- This graph shows the percentage of patients doing better on Medicines A through D.
- The y-axis is the percentage of people doing better on that particular medicine.
- It looks like the share of patients doing better on Medicine C is much greater than the others.
- But look closely at the numbers. We have a truncated axis here.
- Instead of starting at zero, it starts at 28% and goes up to 33%, which makes the difference between the four medicines look bigger than it actually is.

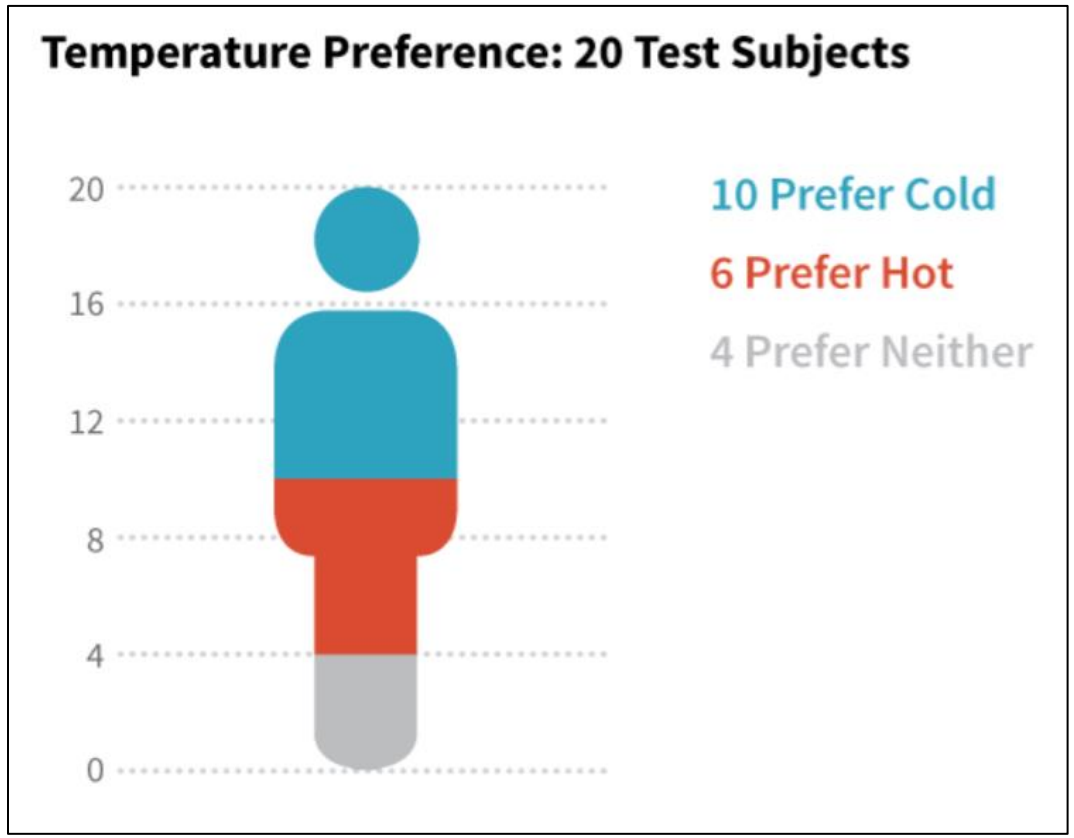


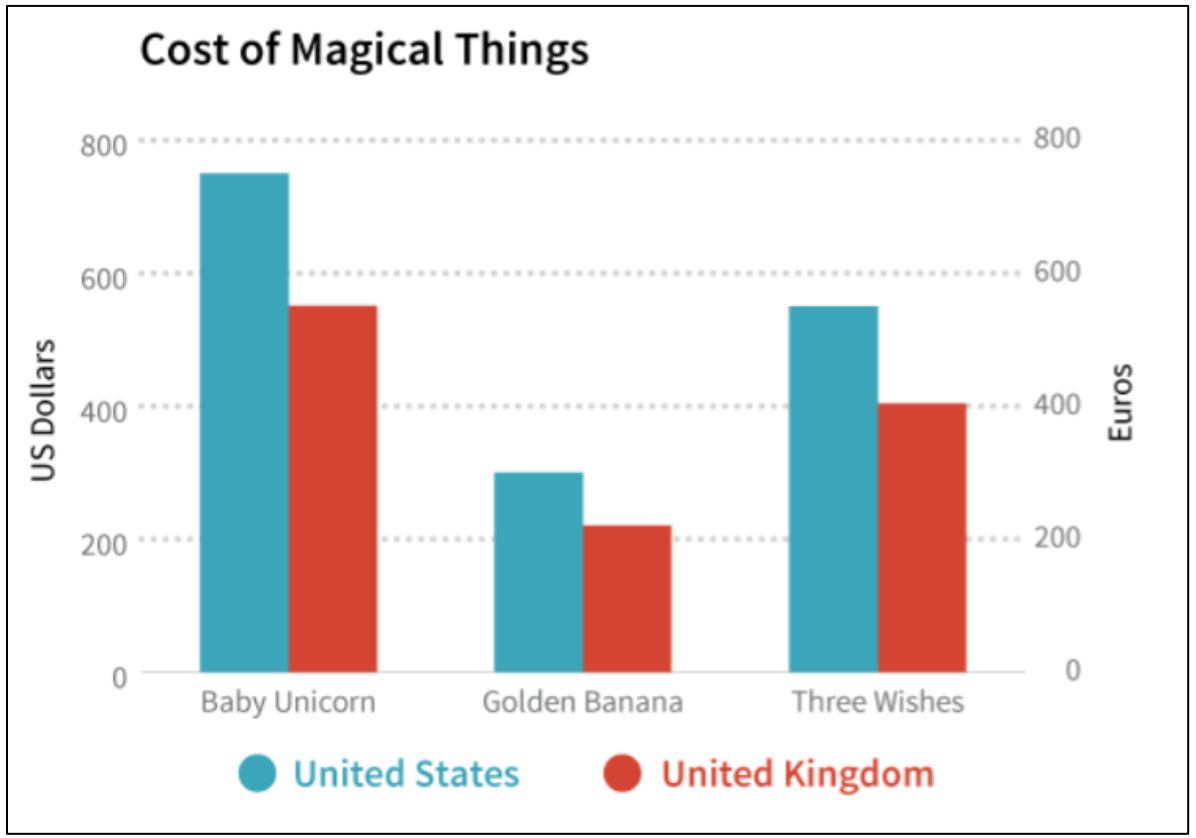
Truncated Axis (after)





Activity: What's Wrong Here?





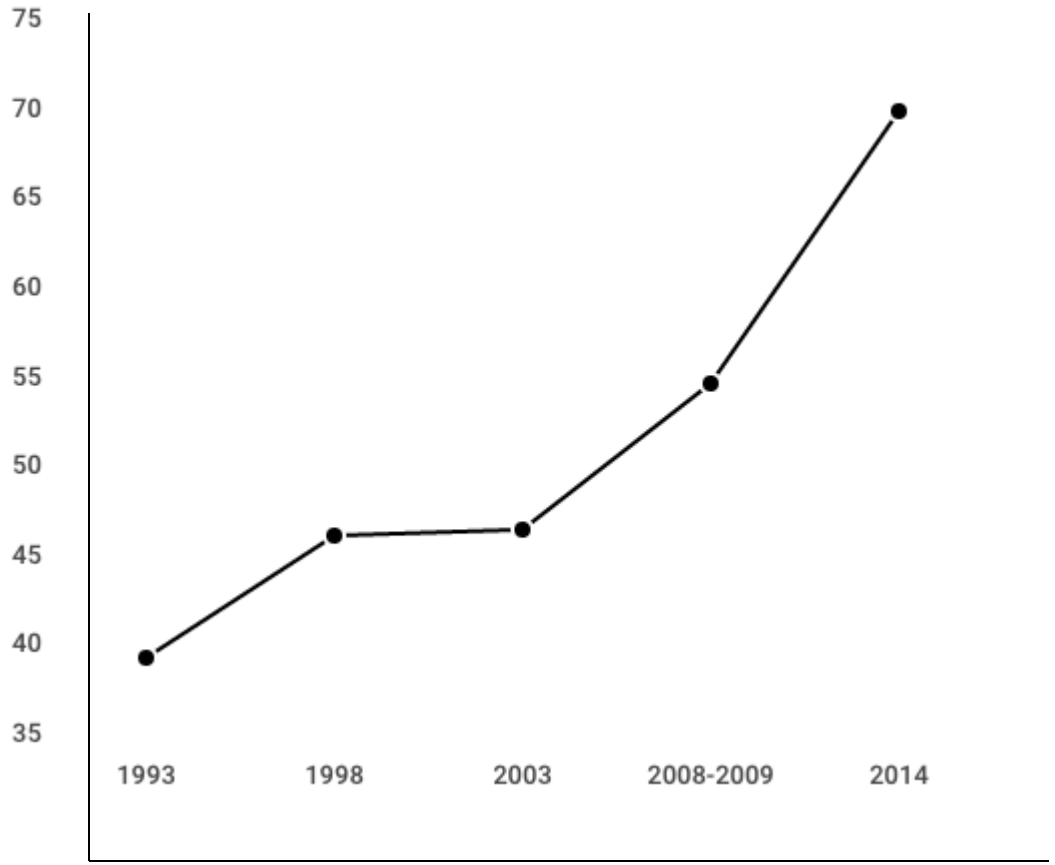
Identifying the Story in Data

Kenya

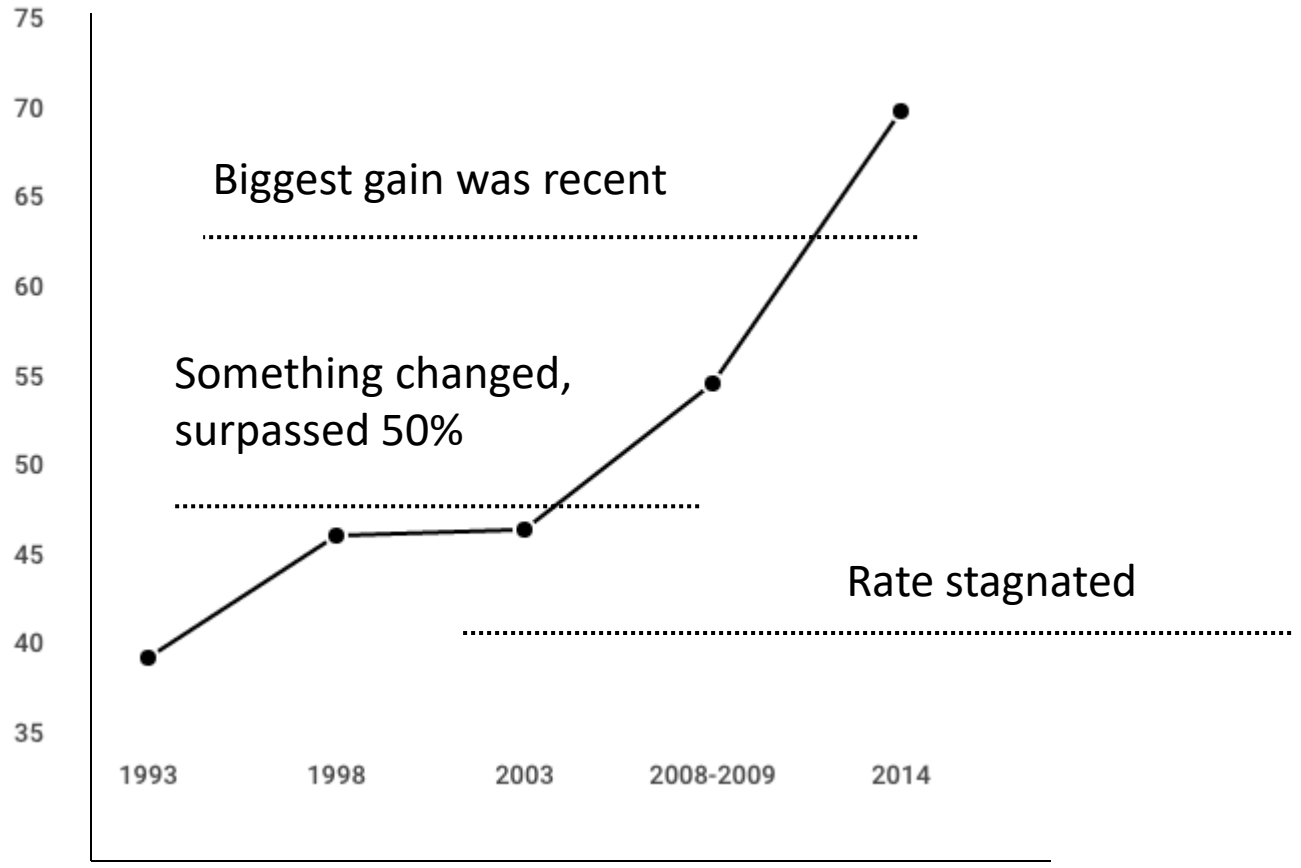
DHS Survey Year	% of Demand for Family Planning Satisfied with Modern Methods
2014	70.7
2008-2009	55.5
2003	47.3
1998	47.0
1993	40.1

Source: Demographic and Health Surveys, Various Years

What's the Story?



What's the Story?

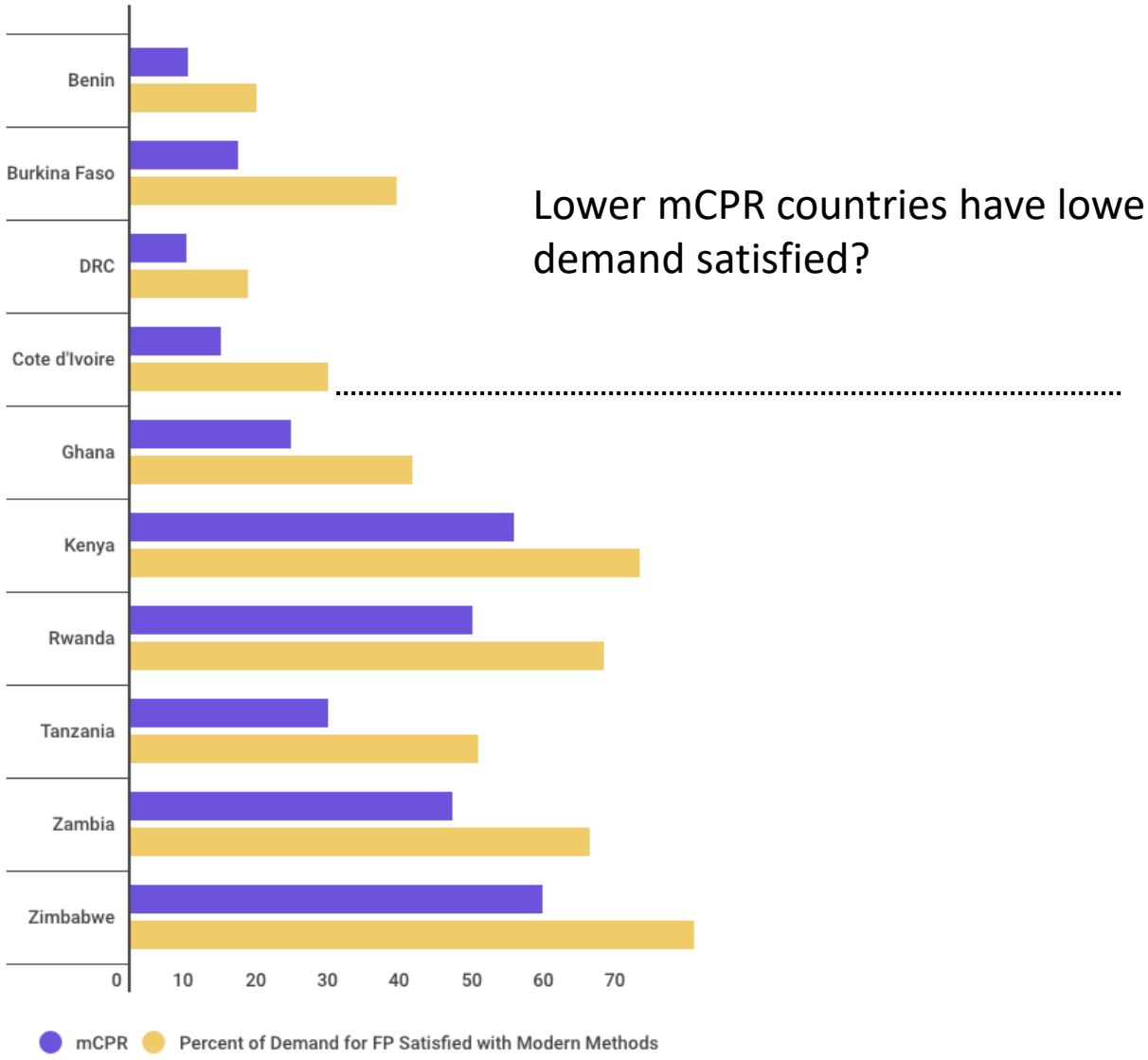


Identifying the Story in Data

Country	Survey Year	mCPR	% of Demand for Family Planning (FP) Satisfied with Modern Methods
Benin	2011-2012	7.9	17.4
Burkina Faso	2010	15.0	36.9
DRC	2013-2014	7.8	16.3
Cote d'Ivoire	2011-2012	12.5	27.5
Ghana	2014	22.2	39.2
Kenya	2014	53.2	70.7
Rwanda	2014-2015	47.5	65.8
Tanzania	2010	27.4	48.3
Zambia	2013-2014	44.8	63.8
Zimbabwe	2010-2011	57.3	78.3

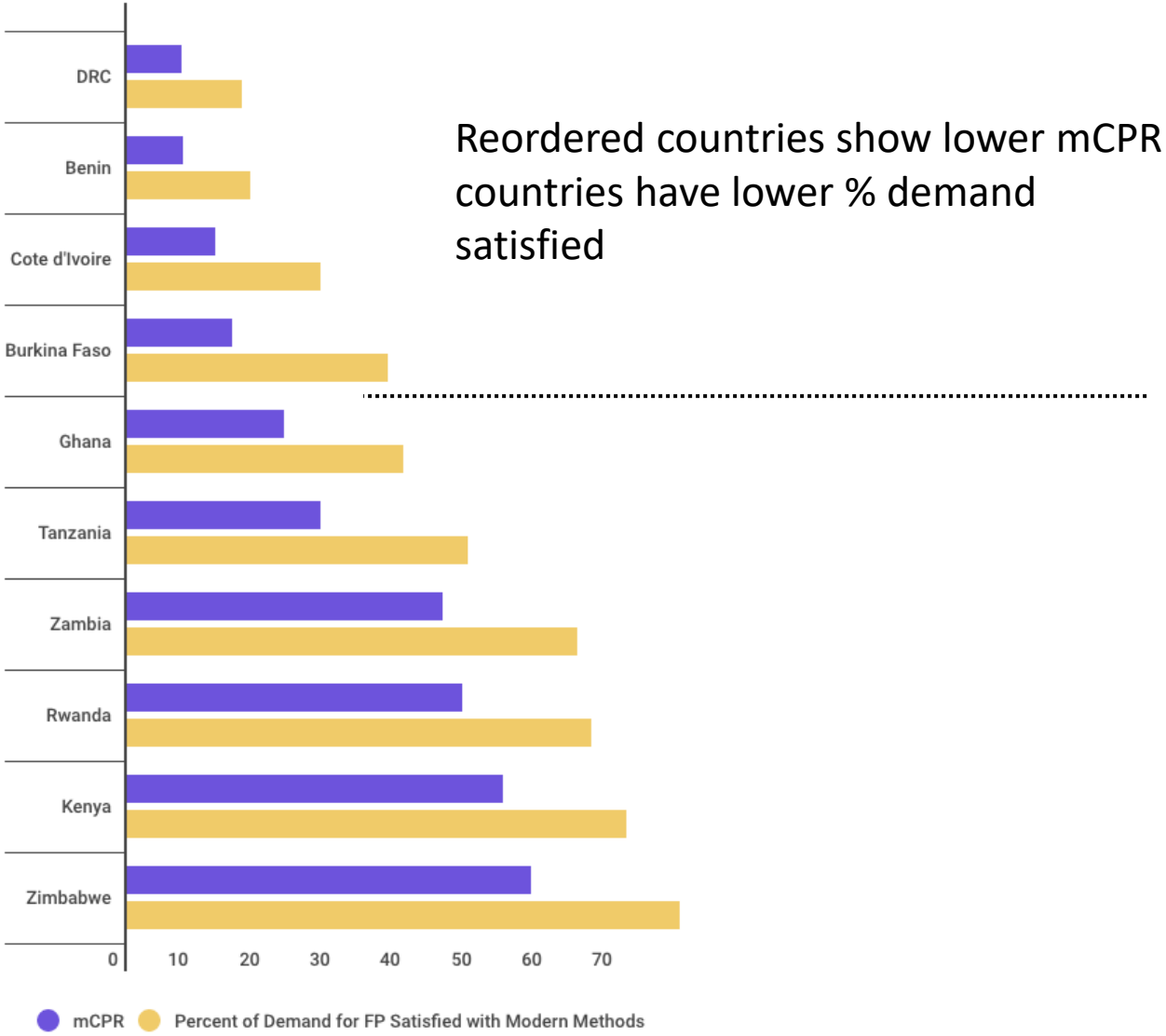
Source: Demographic and Health Surveys, Various Years

What's the Story?



Lower mCPR countries have lower % demand satisfied?

What's the Story?



Reordered countries show lower mCPR countries have lower % demand satisfied