HOW TO MAKE DECISIONS IN DATA VISUALIZATION

Alberto Cairo

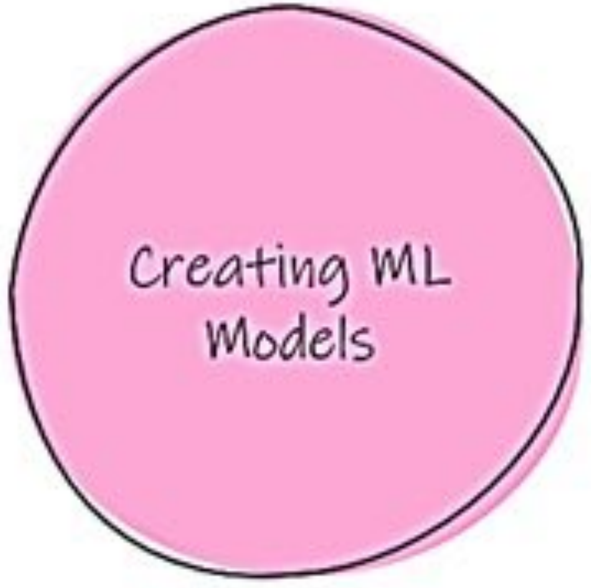
Banco de Portugal, 2020

Time spent in a life of a Data Scientist

@datavizzdom

Gulrez

Perception

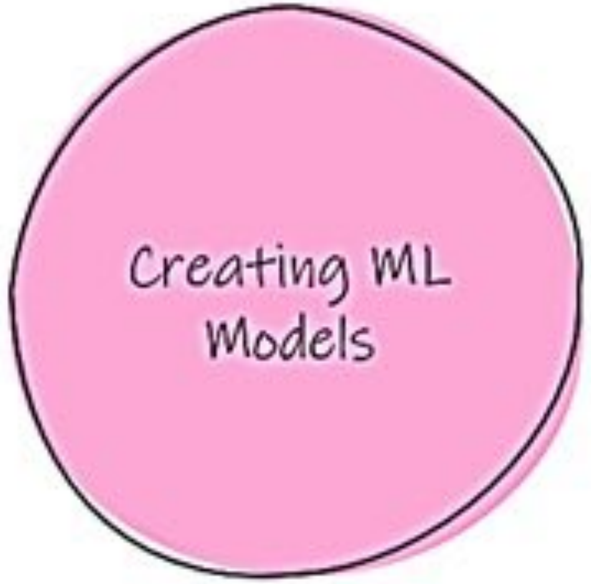


Time spent in a life of a Data Scientist

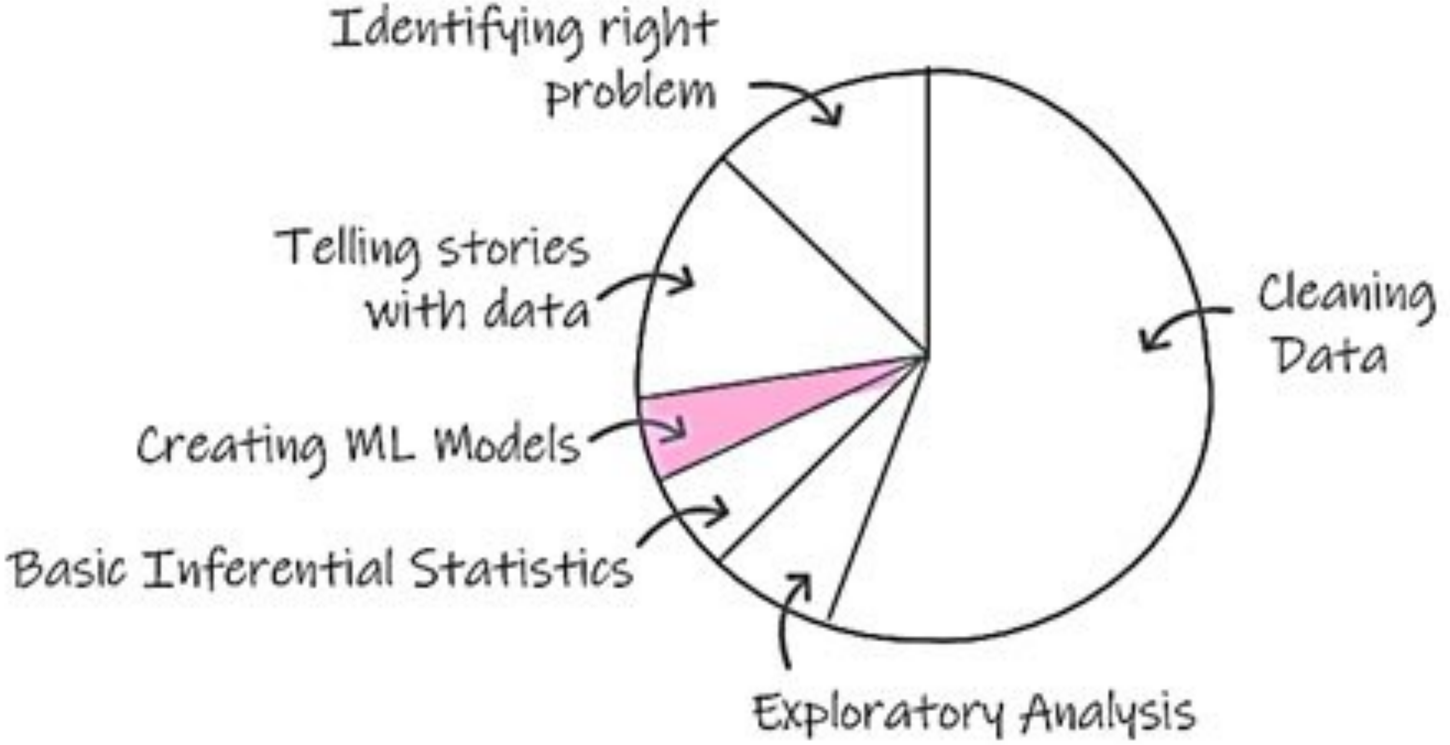
@datavizzdom

Gulrez

Perception



Reality

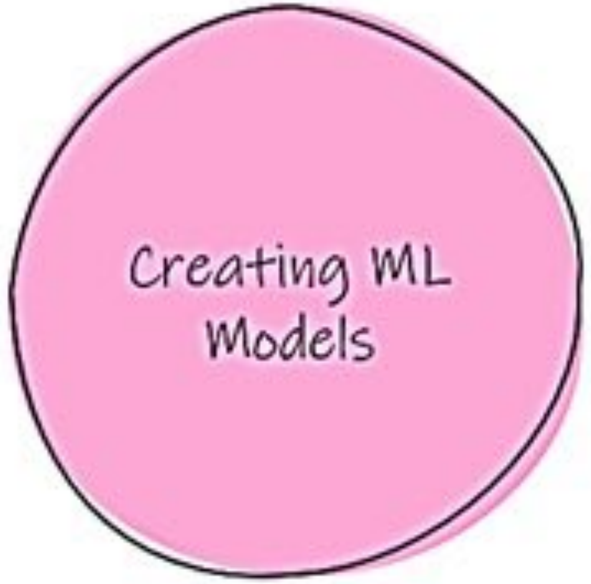


Time spent in a life of a Data Scientist

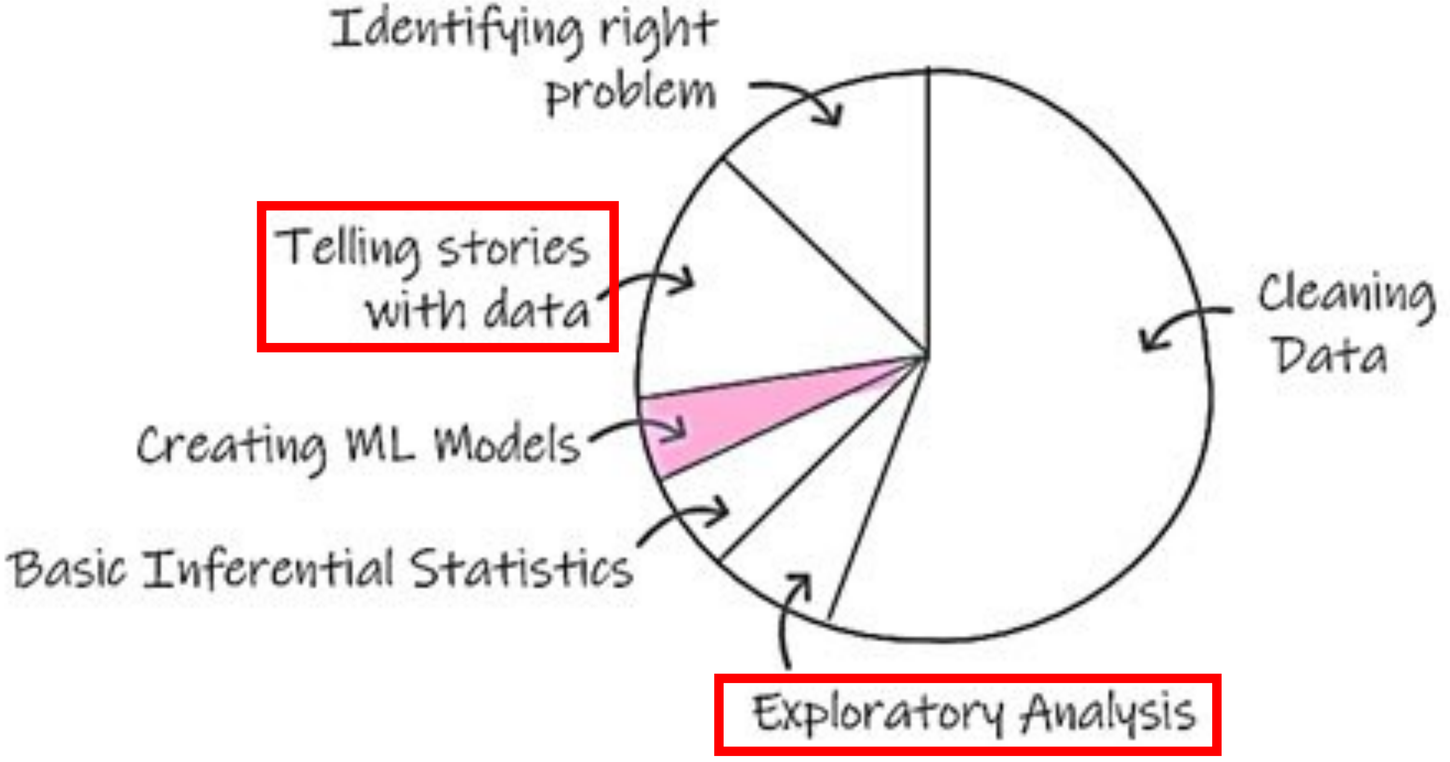
@datavizzdom

Gulrez

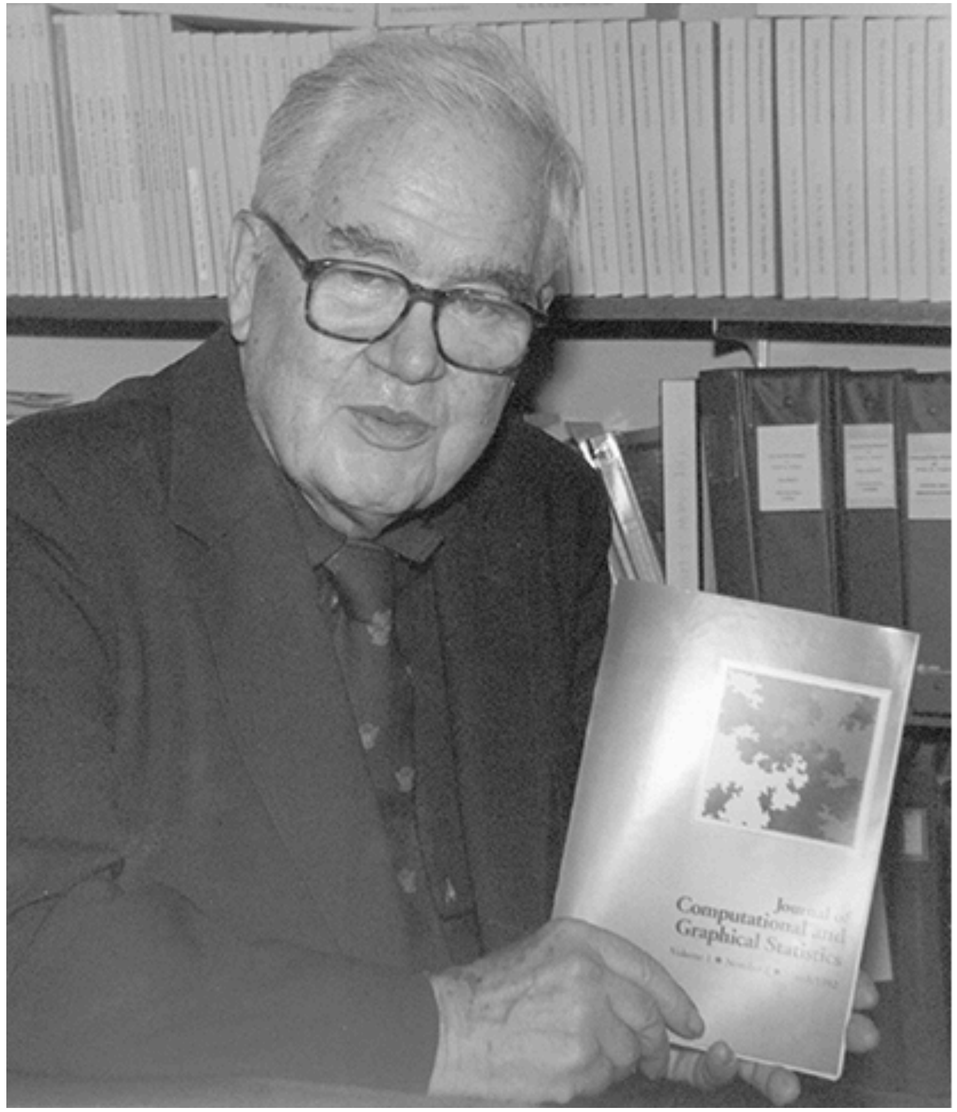
Perception

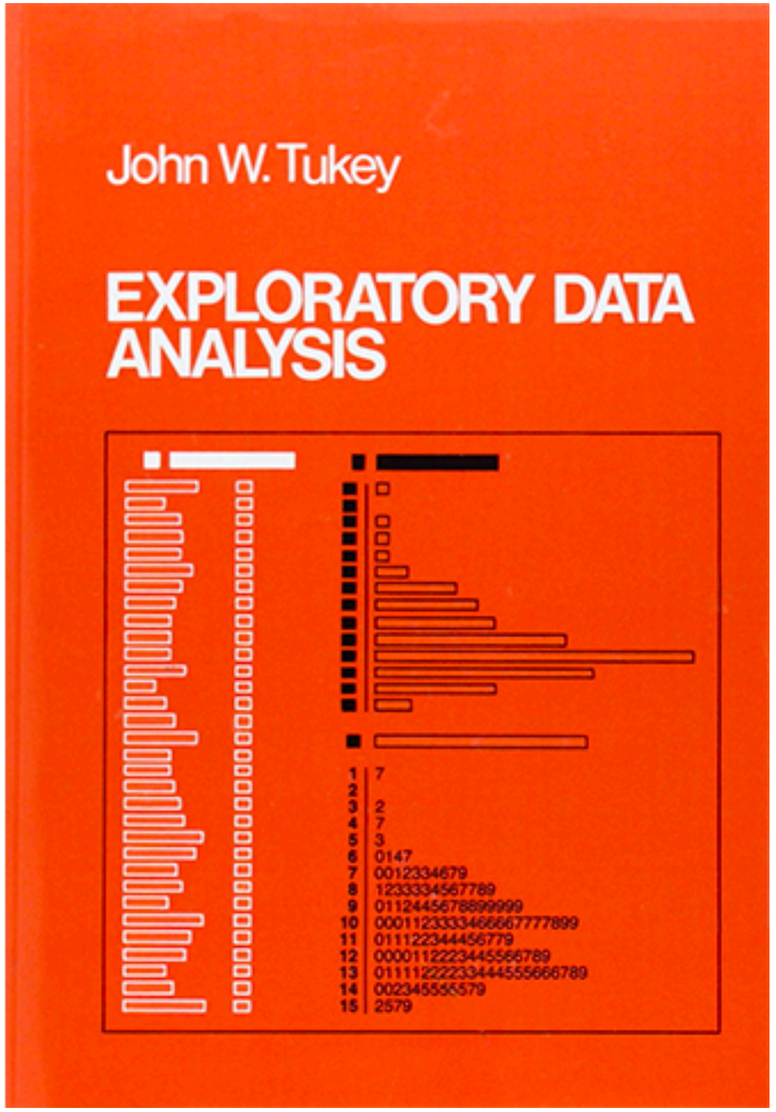


Reality



We are living in a Golden Age of visualization





"The greatest value of a picture is when it forces us to notice what we never expected to see."

John W. Tukey

We are living in a Golden Age of visualization

We are living in a Golden Age of visualization

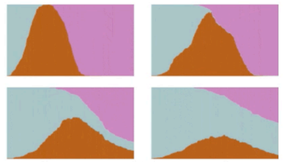
Ed Hawkins's 'Warming stripes' (read more: https://chezvoila.com/blog/warmingstripes/)

We are living in a Golden Age of visualization

Sections

The Washington Post
Democracy Dies in Darkness

Alberto Cairo To...



Health

Why outbreaks like coronavirus spread exponentially, and how to "flatten the curve"

By Harry Stevens March 14, 2020

PLEASE NOTE

The Washington Post is providing this story for free so that all readers have access to this important information about the coronavirus. For more free stories, sign up for our daily Coronavirus Updates newsletter.

2.6k

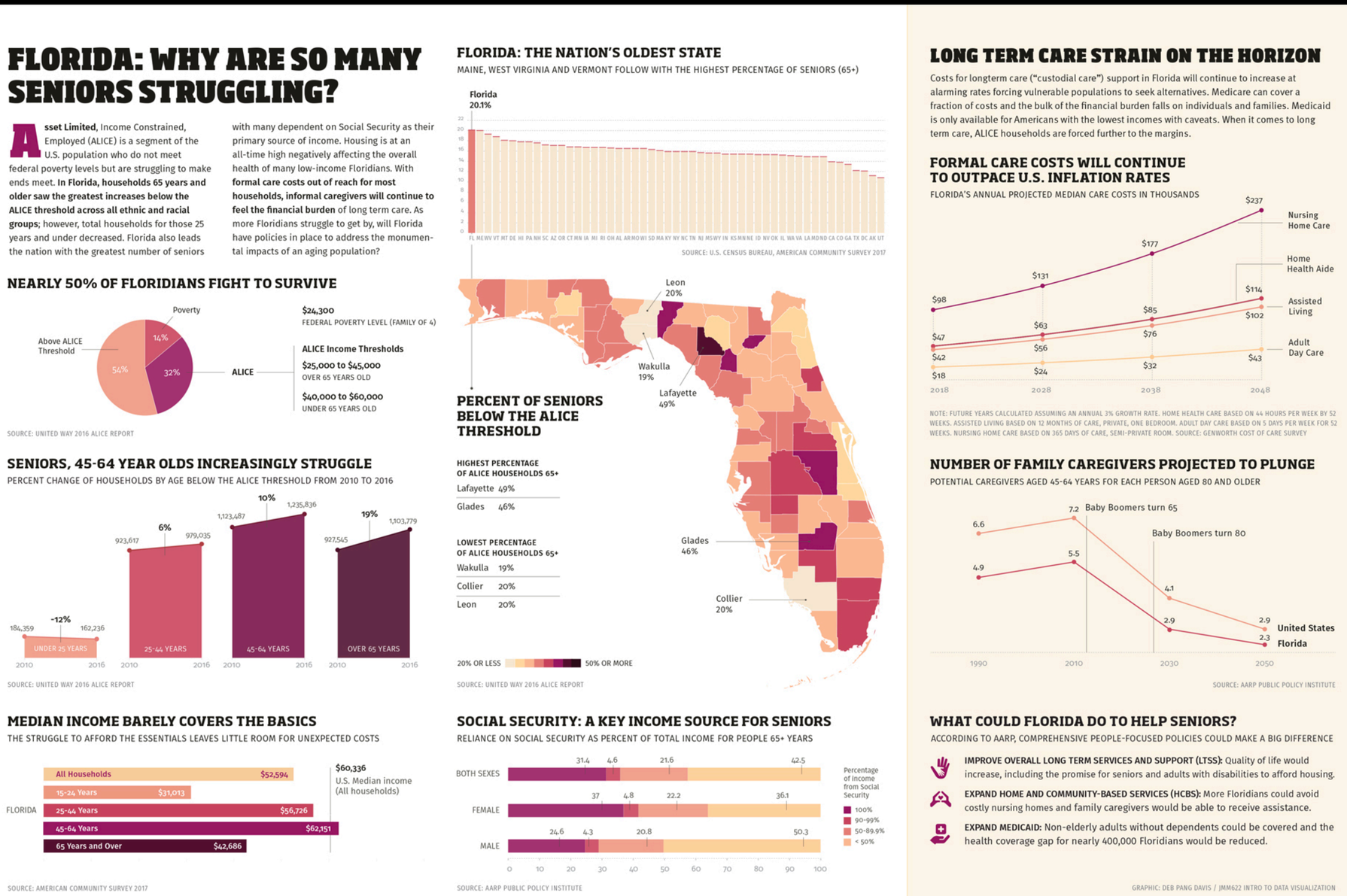
https://www.washingtonpost.com/graphics/2020/world/corona-simulator/

We are living in a Golden Age of visualization

The Post's visual journalism, which involves staff throughout the newsroom, has attracted large audiences and contributed to record subscriber growth. Six of the seven most visited stories in The Washington Post's history have been graphics, including the coronavirus simulator that became the most visited article in The Post's history, with more than three times as many visits as the second. It also includes this year's Democratic candidate quiz, which set the record for converting readers to subscribers.

https://www.washingtonpost.com/pr/2020/06/26/washington-post-expand-graphics-design-teams-with-14-new-positions/

We are living in a Golden Age of visualization



ANYONE can learn to make data visualizations.

Example of my students' work:
https://www.deb.is/

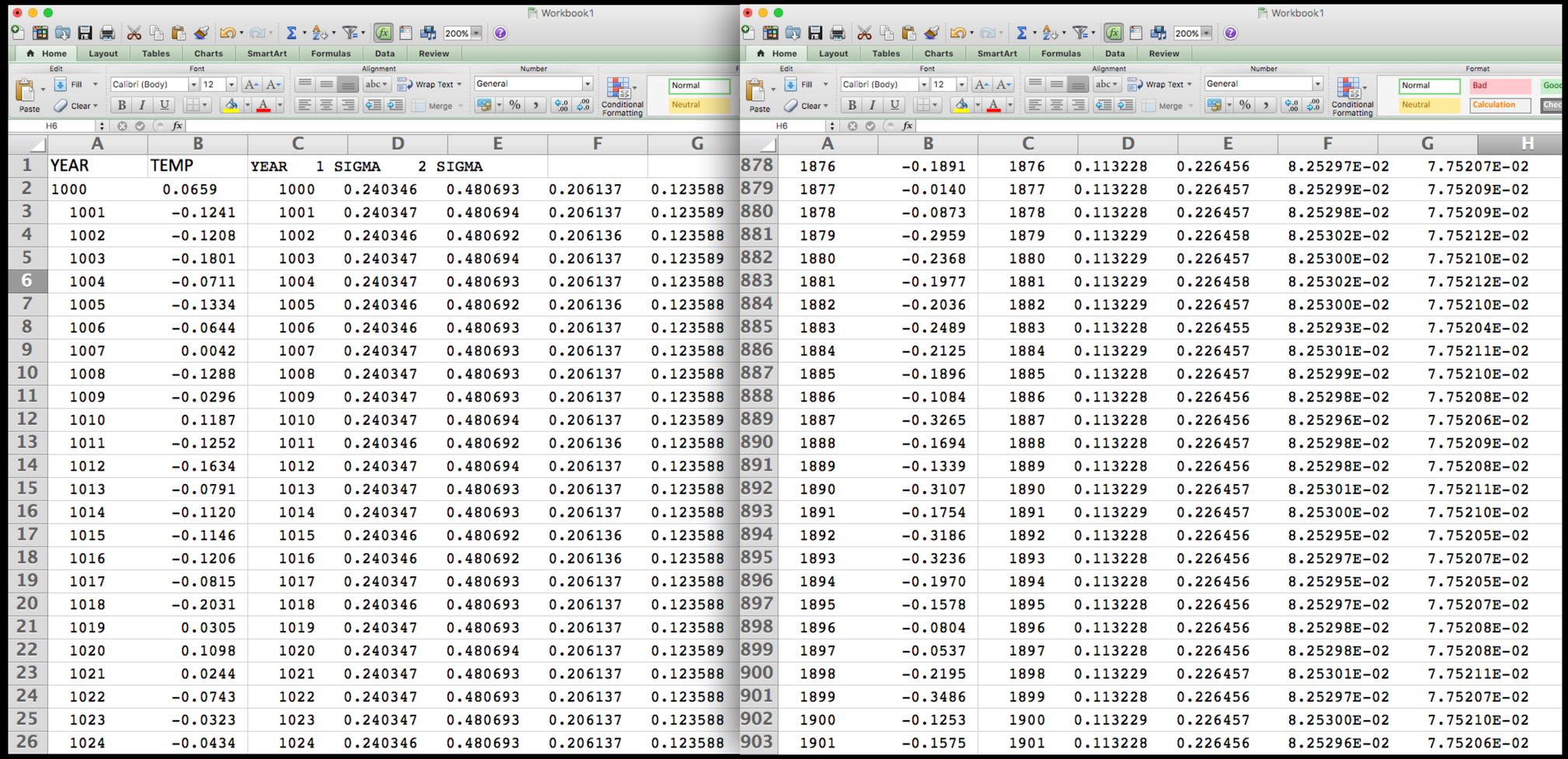
How to teach or learn data visualization?

Visualization is a bit like writing: beyond some conventions and constraints regarding symbols, visual grammar, perception, and cognition, visualization **can't be based on *"rules"* that are set in stone**.

Instead, when designing visualizations, we need to be guided by **reasoned, justifiable choices**.

1. Why should my visualization exist?

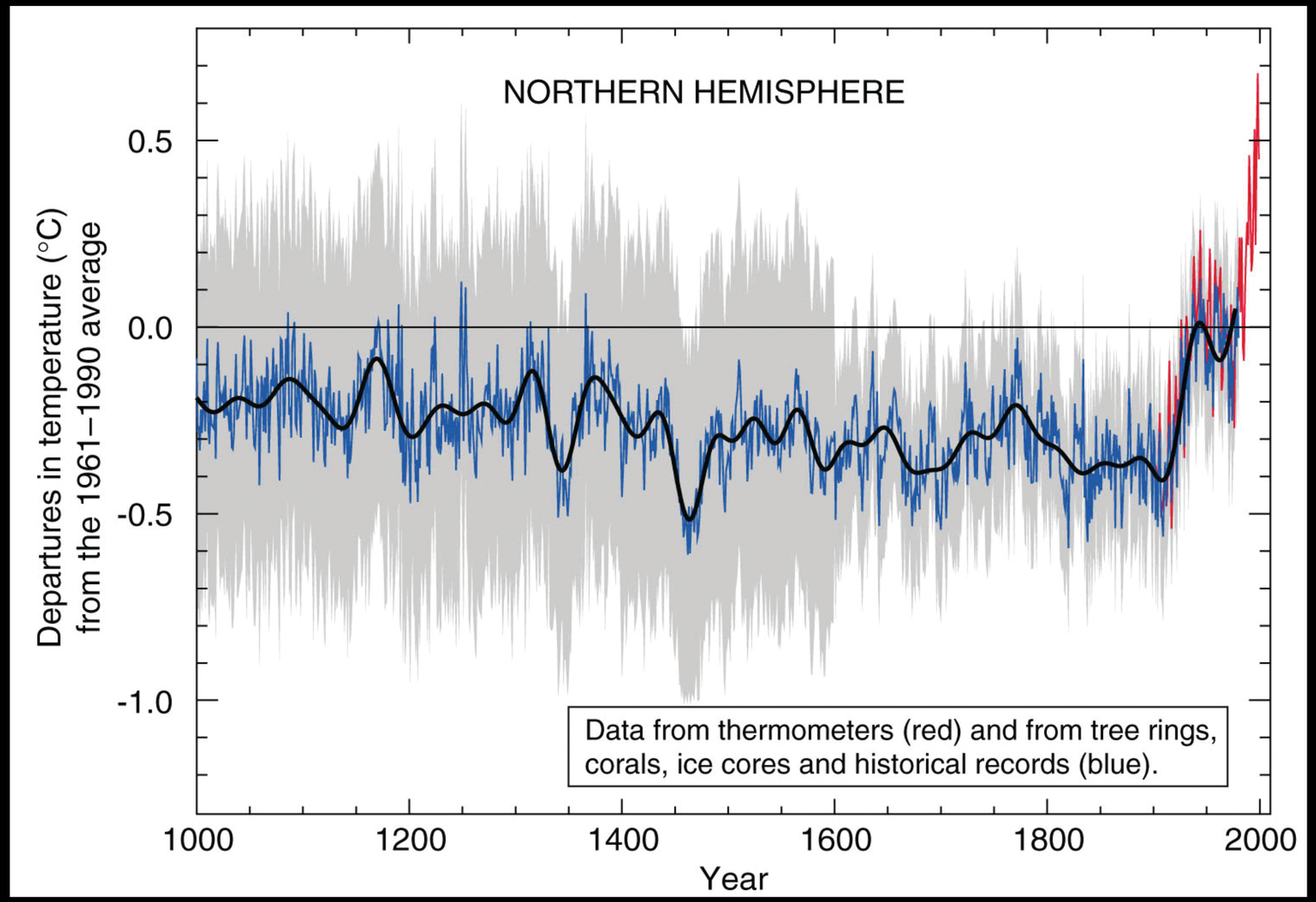
What are tables useful for?



	A	B	C	D	E	F	G
1	YEAR	TEMP	YEAR	1 SIGMA	2 SIGMA		
2	1000	0.0659	1000	0.240346	0.480693	0.206137	0.123588
3	1001	-0.1241	1001	0.240347	0.480694	0.206137	0.123589
4	1002	-0.1208	1002	0.240346	0.480692	0.206136	0.123588
5	1003	-0.1801	1003	0.240347	0.480694	0.206137	0.123589
6	1004	-0.0711	1004	0.240347	0.480693	0.206137	0.123588
7	1005	-0.1334	1005	0.240346	0.480692	0.206136	0.123588
8	1006	-0.0644	1006	0.240346	0.480693	0.206137	0.123588
9	1007	0.0042	1007	0.240347	0.480693	0.206137	0.123588
10	1008	-0.1288	1008	0.240347	0.480693	0.206137	0.123588
11	1009	-0.0296	1009	0.240347	0.480693	0.206137	0.123588
12	1010	0.1187	1010	0.240347	0.480694	0.206137	0.123589
13	1011	-0.1252	1011	0.240346	0.480692	0.206136	0.123588
14	1012	-0.1634	1012	0.240347	0.480694	0.206137	0.123588
15	1013	-0.0791	1013	0.240347	0.480693	0.206137	0.123588
16	1014	-0.1120	1014	0.240347	0.480693	0.206137	0.123588
17	1015	-0.1146	1015	0.240346	0.480692	0.206136	0.123588
18	1016	-0.1206	1016	0.240346	0.480692	0.206136	0.123588
19	1017	-0.0815	1017	0.240347	0.480693	0.206137	0.123588
20	1018	-0.2031	1018	0.240346	0.480693	0.206137	0.123588
21	1019	0.0305	1019	0.240347	0.480693	0.206137	0.123588
22	1020	0.1098	1020	0.240347	0.480694	0.206137	0.123589
23	1021	0.0244	1021	0.240347	0.480693	0.206137	0.123588
24	1022	-0.0743	1022	0.240347	0.480693	0.206137	0.123588
25	1023	-0.0323	1023	0.240347	0.480693	0.206137	0.123588
26	1024	-0.0434	1024	0.240346	0.480693	0.206137	0.123588

	A	B	C	D	E	F	G	H
878	1876	-0.1891	1876	0.113228	0.226456	8.25297E-02	7.75207E-02	
879	1877	-0.0140	1877	0.113228	0.226457	8.25299E-02	7.75209E-02	
880	1878	-0.0873	1878	0.113228	0.226457	8.25298E-02	7.75209E-02	
881	1879	-0.2959	1879	0.113229	0.226458	8.25302E-02	7.75212E-02	
882	1880	-0.2368	1880	0.113229	0.226457	8.25300E-02	7.75210E-02	
883	1881	-0.1977	1881	0.113229	0.226458	8.25302E-02	7.75212E-02	
884	1882	-0.2036	1882	0.113229	0.226457	8.25300E-02	7.75210E-02	
885	1883	-0.2489	1883	0.113228	0.226455	8.25293E-02	7.75204E-02	
886	1884	-0.2125	1884	0.113229	0.226457	8.25301E-02	7.75211E-02	
887	1885	-0.1896	1885	0.113228	0.226457	8.25299E-02	7.75210E-02	
888	1886	-0.1084	1886	0.113228	0.226456	8.25298E-02	7.75208E-02	
889	1887	-0.3265	1887	0.113228	0.226456	8.25296E-02	7.75206E-02	
890	1888	-0.1694	1888	0.113228	0.226457	8.25298E-02	7.75209E-02	
891	1889	-0.1339	1889	0.113228	0.226456	8.25298E-02	7.75208E-02	
892	1890	-0.3107	1890	0.113229	0.226457	8.25301E-02	7.75211E-02	
893	1891	-0.1754	1891	0.113229	0.226457	8.25300E-02	7.75210E-02	
894	1892	-0.3186	1892	0.113228	0.226456	8.25295E-02	7.75205E-02	
895	1893	-0.3236	1893	0.113228	0.226456	8.25297E-02	7.75207E-02	
896	1894	-0.1970	1894	0.113228	0.226456	8.25295E-02	7.75205E-02	
897	1895	-0.1578	1895	0.113228	0.226456	8.25297E-02	7.75207E-02	
898	1896	-0.0804	1896	0.113228	0.226456	8.25298E-02	7.75208E-02	
899	1897	-0.0537	1897	0.113228	0.226456	8.25298E-02	7.75208E-02	
900	1898	-0.2195	1898	0.113229	0.226457	8.25301E-02	7.75211E-02	
901	1899	-0.3486	1899	0.113228	0.226456	8.25297E-02	7.75207E-02	
902	1900	-0.1253	1900	0.113229	0.226457	8.25300E-02	7.75210E-02	
903	1901	-0.1575	1901	0.113228	0.226456	8.25296E-02	7.75206E-02	

Visualization is about patterns, trends, the big picture



Michael E. Mann, Raymond S. Bradley, and Malcolm K. Hughes

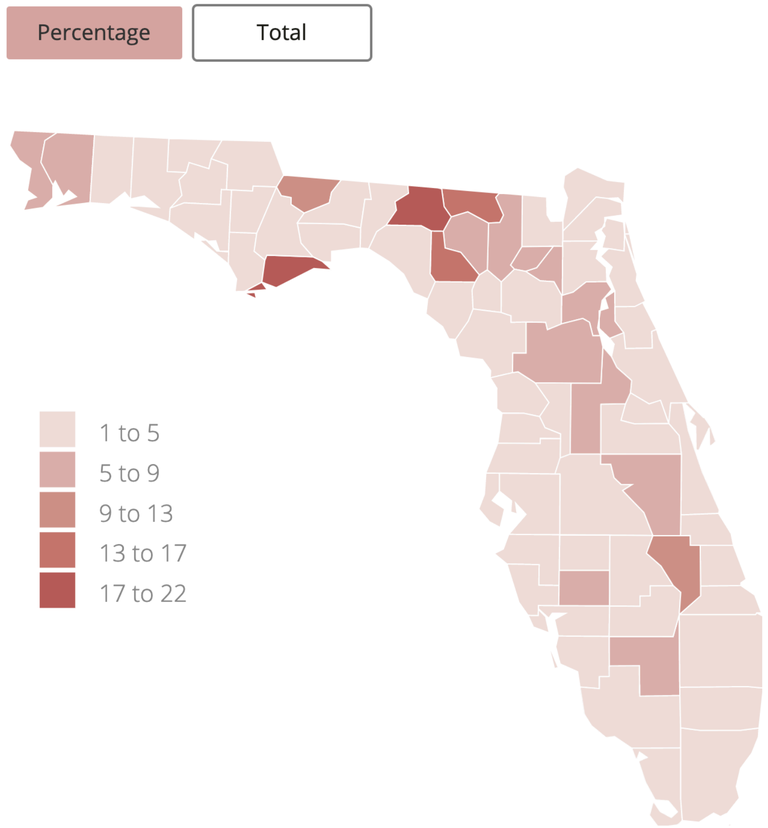
Intergovernmental Panel on Climate Change (IPCC), Third Report, 2001

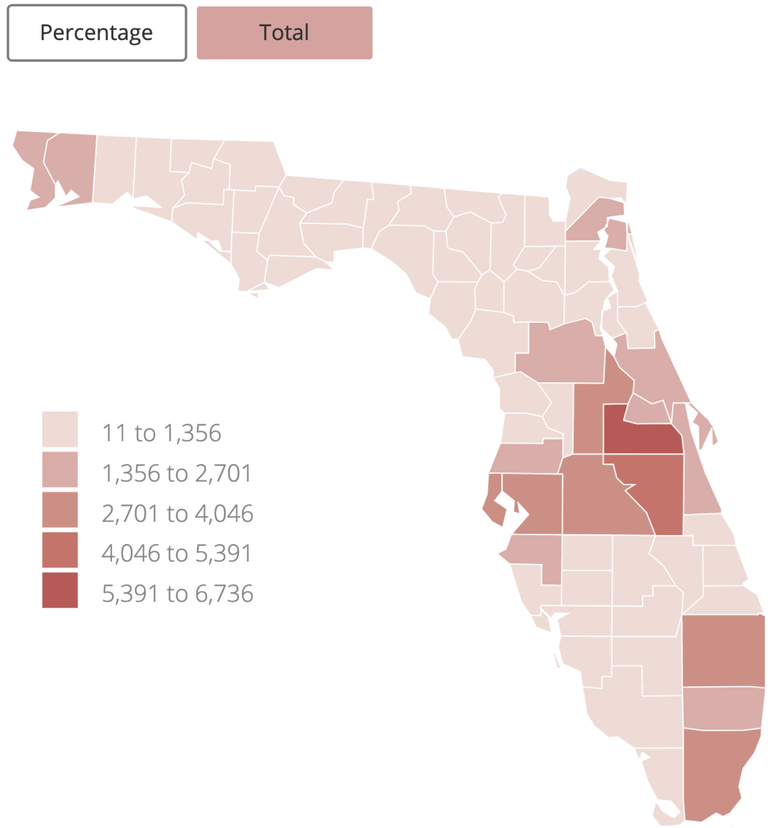
2. What to visualize?

AT SCHOOL WITHOUT A ROOF

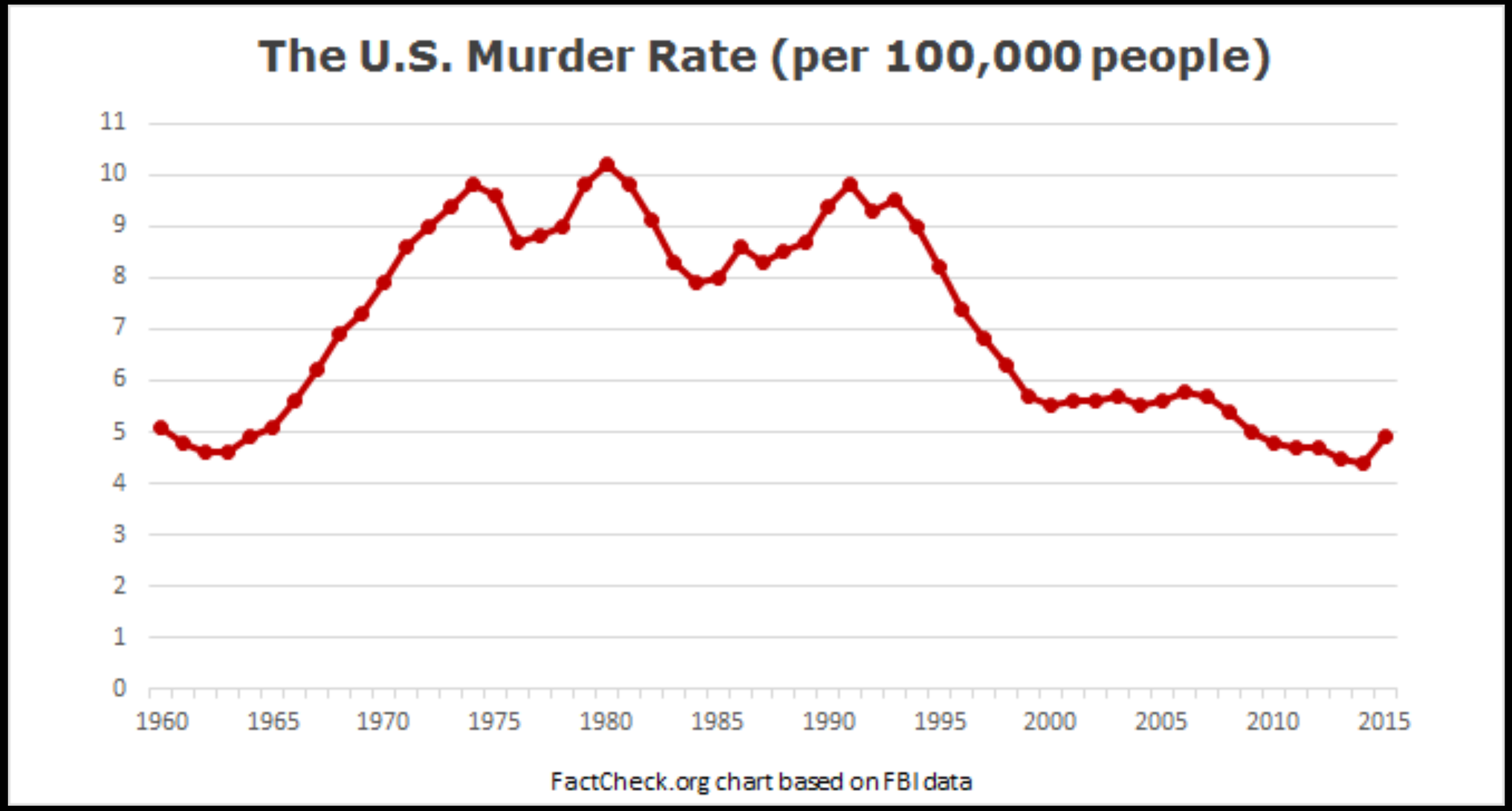
http://www.lmelgar.me/without-a-roof/

In Florida more than 71,000 students are homeless. During the last decade, this population rocketed as a result of the recession and how hard it has become for the poorest families to find affordable housing.

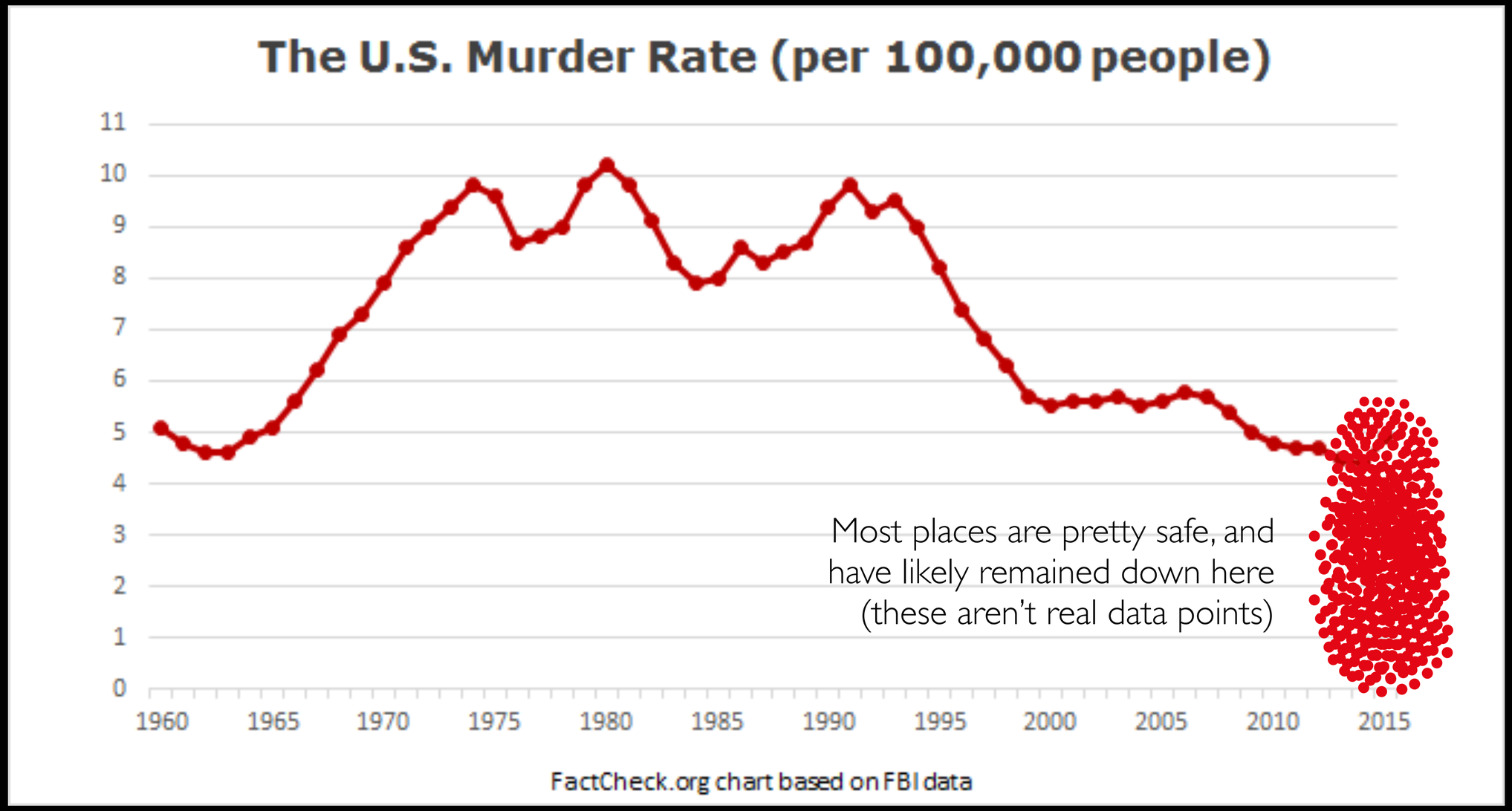




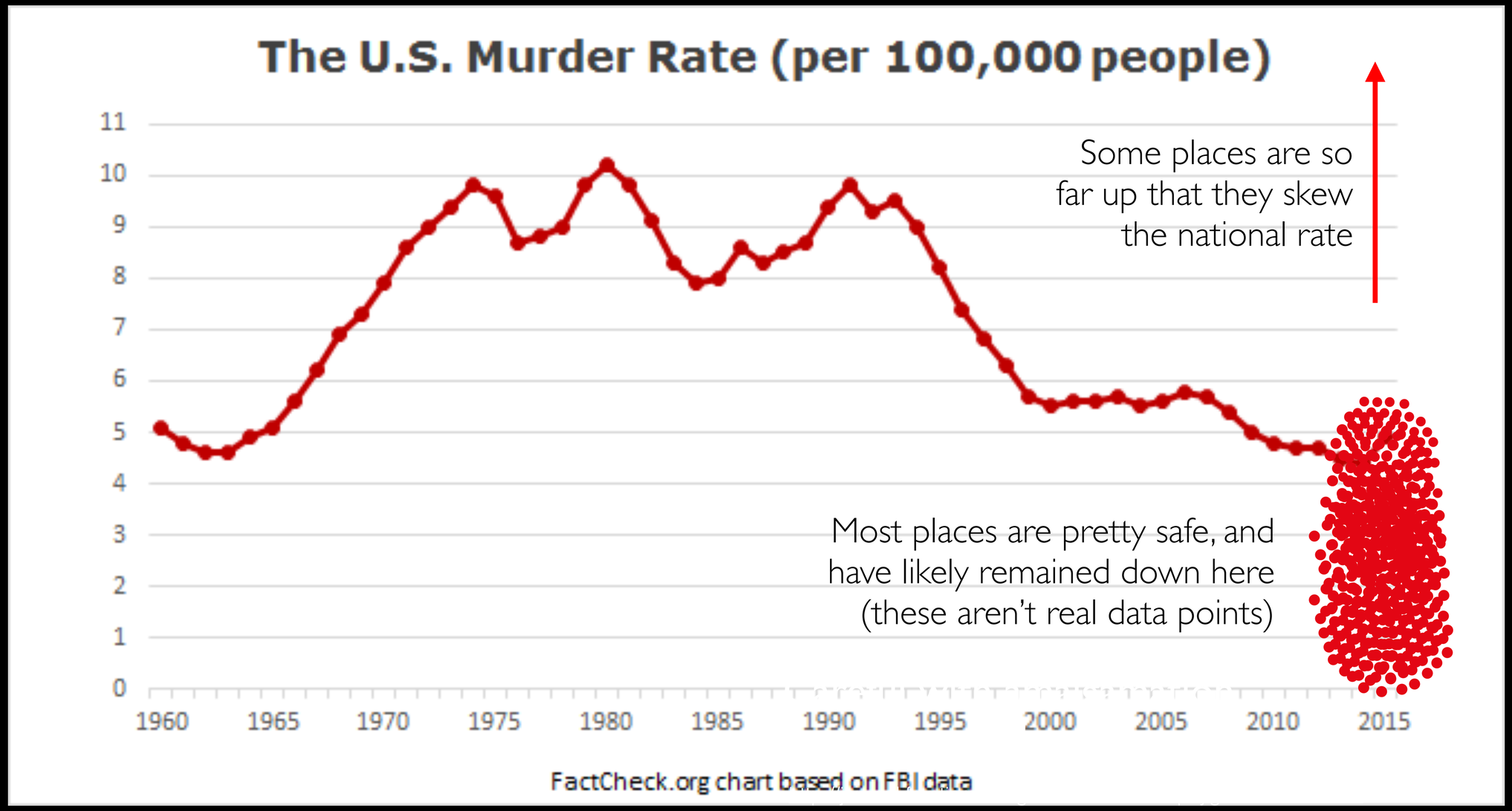
3. How much to visualize?


The U.S. Murder Rate (per 100,000 people)
11
10
9
8
7
6
5
4
3
2
1
0
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
2015
FactCheck.org chart based on FBI data

The danger of aggregating data too much, and presenting just averages and other statistical summaries



The danger of aggregating data too much, and presenting just averages and other statistical summaries



4. How to visualize it?

Figure 2 - Main nationalities of arriving migrants – 2016

Greece

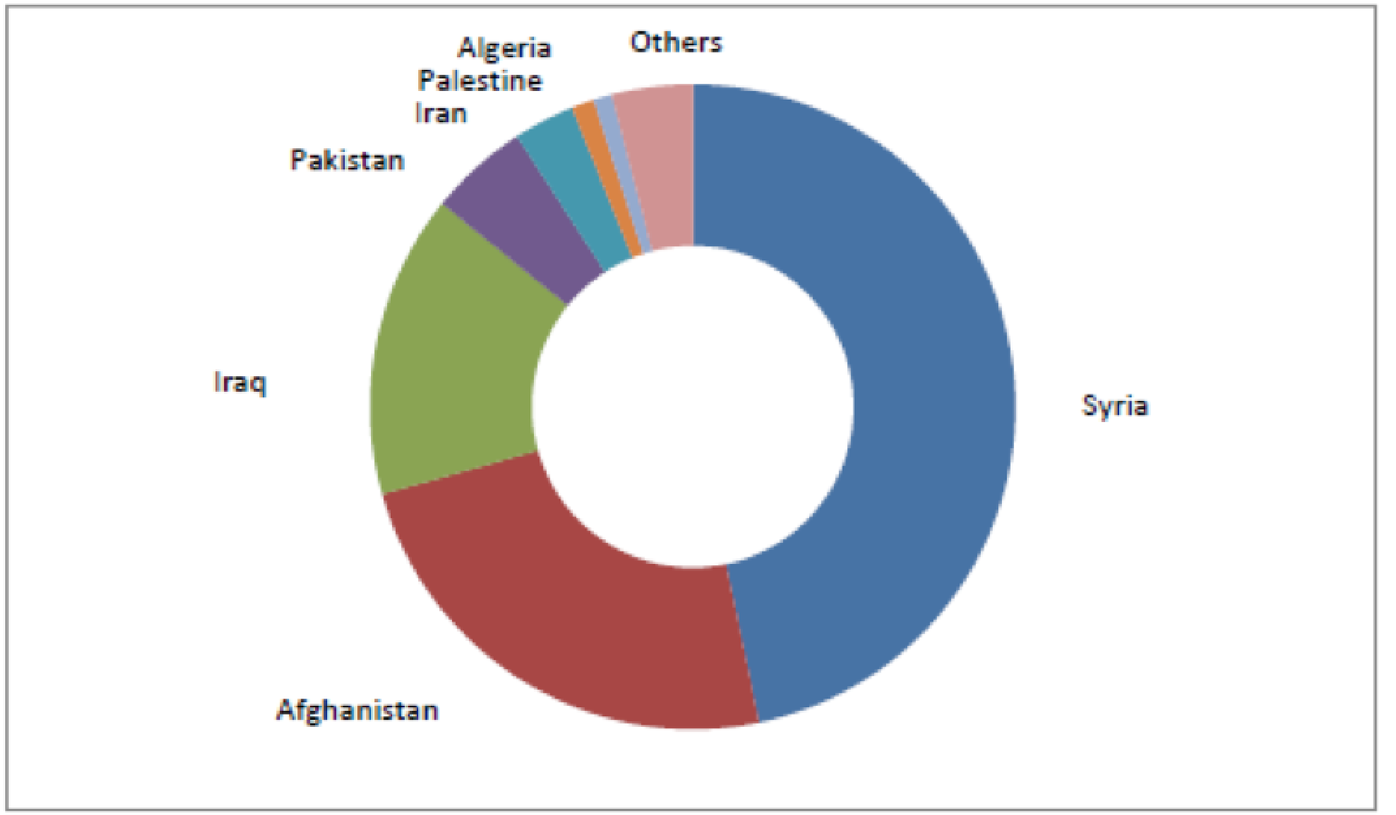
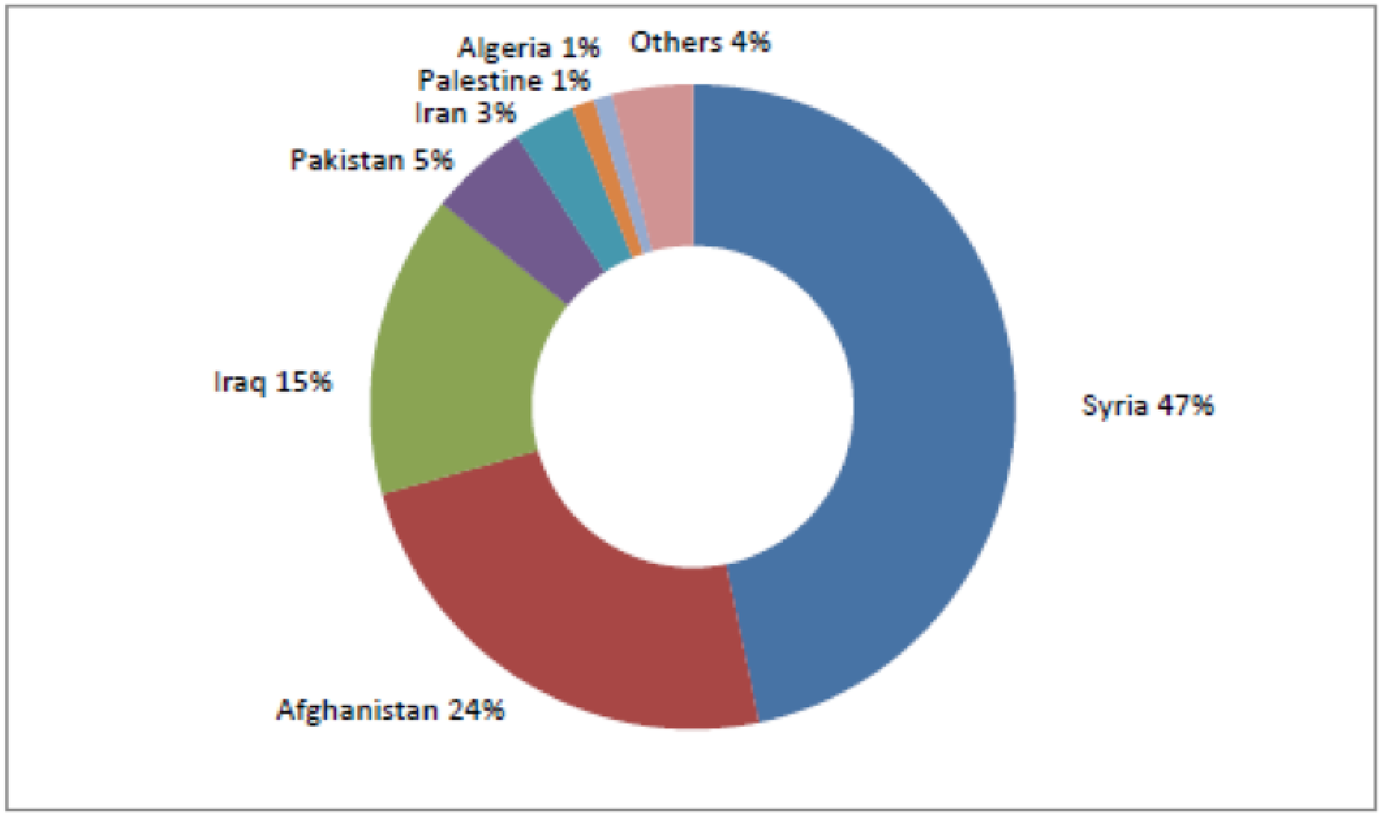
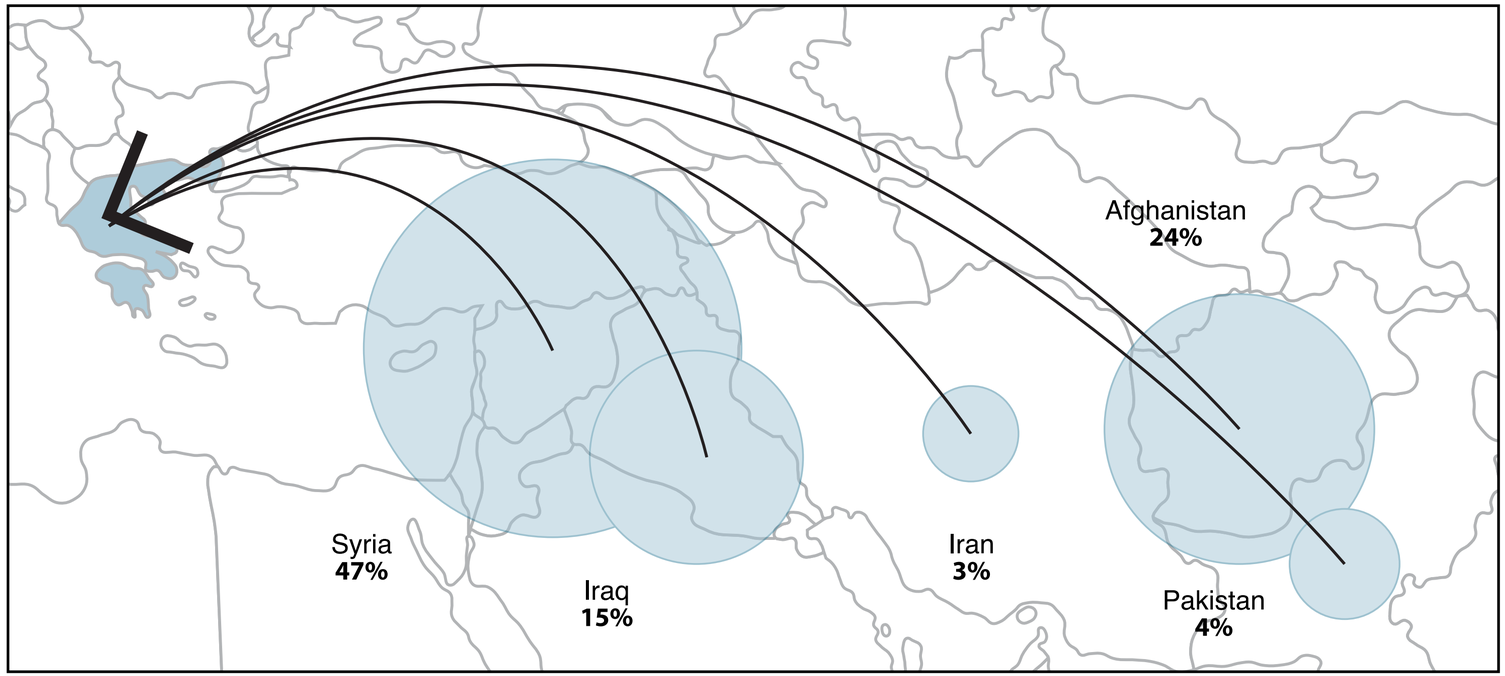
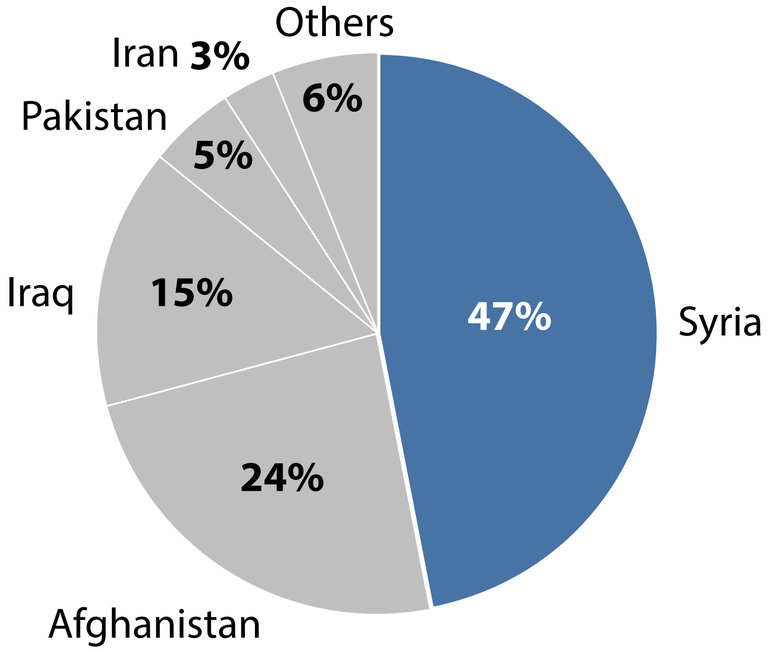


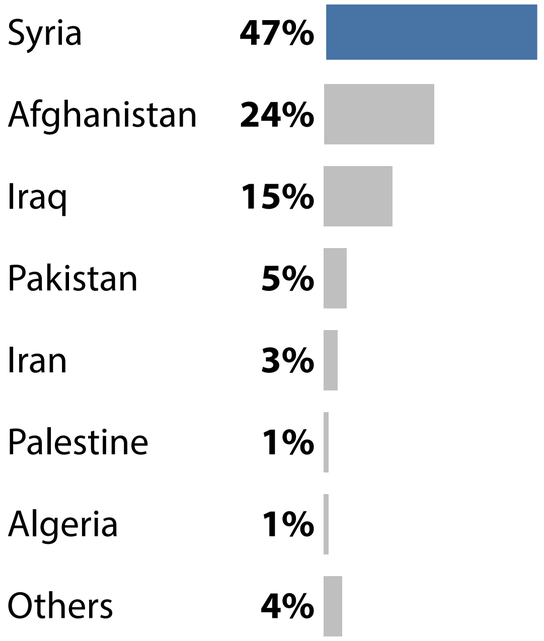
Figure 2 - Main nationalities of arriving migrants – 2016

Greece

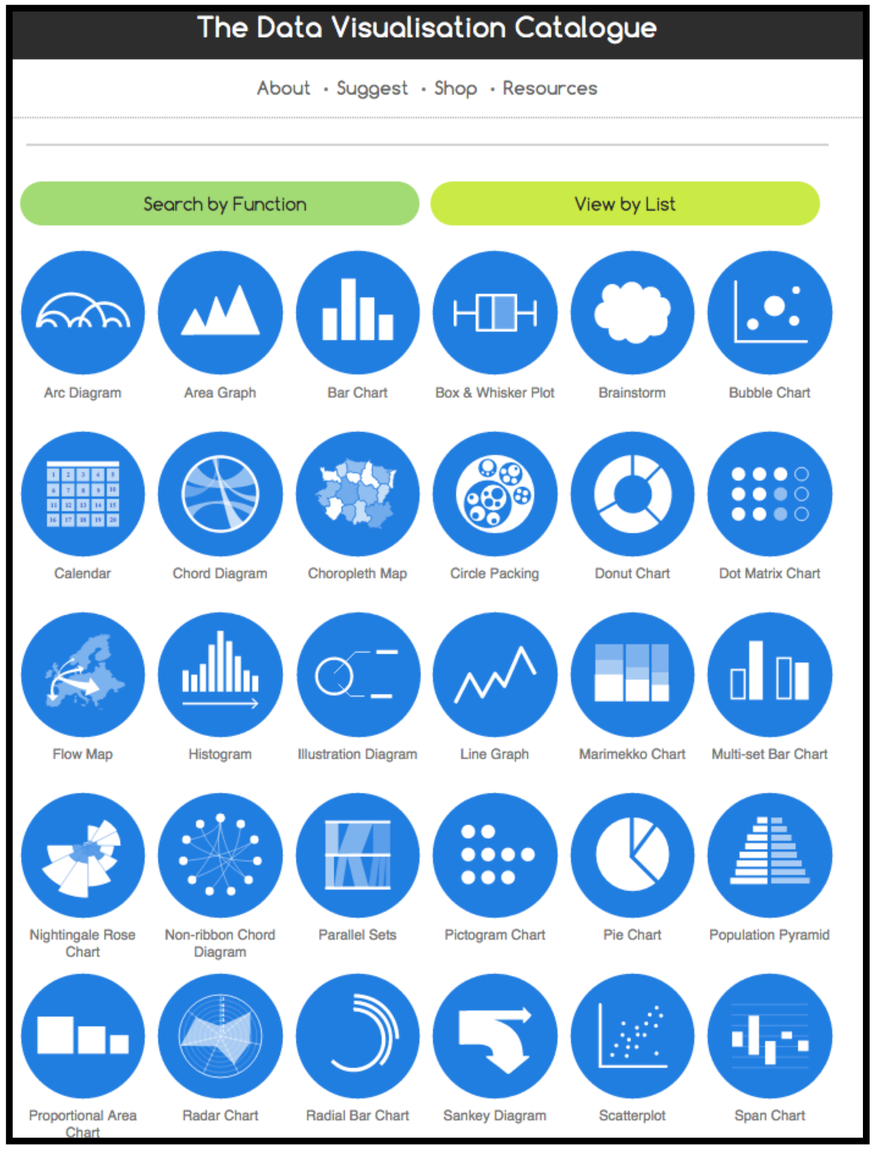


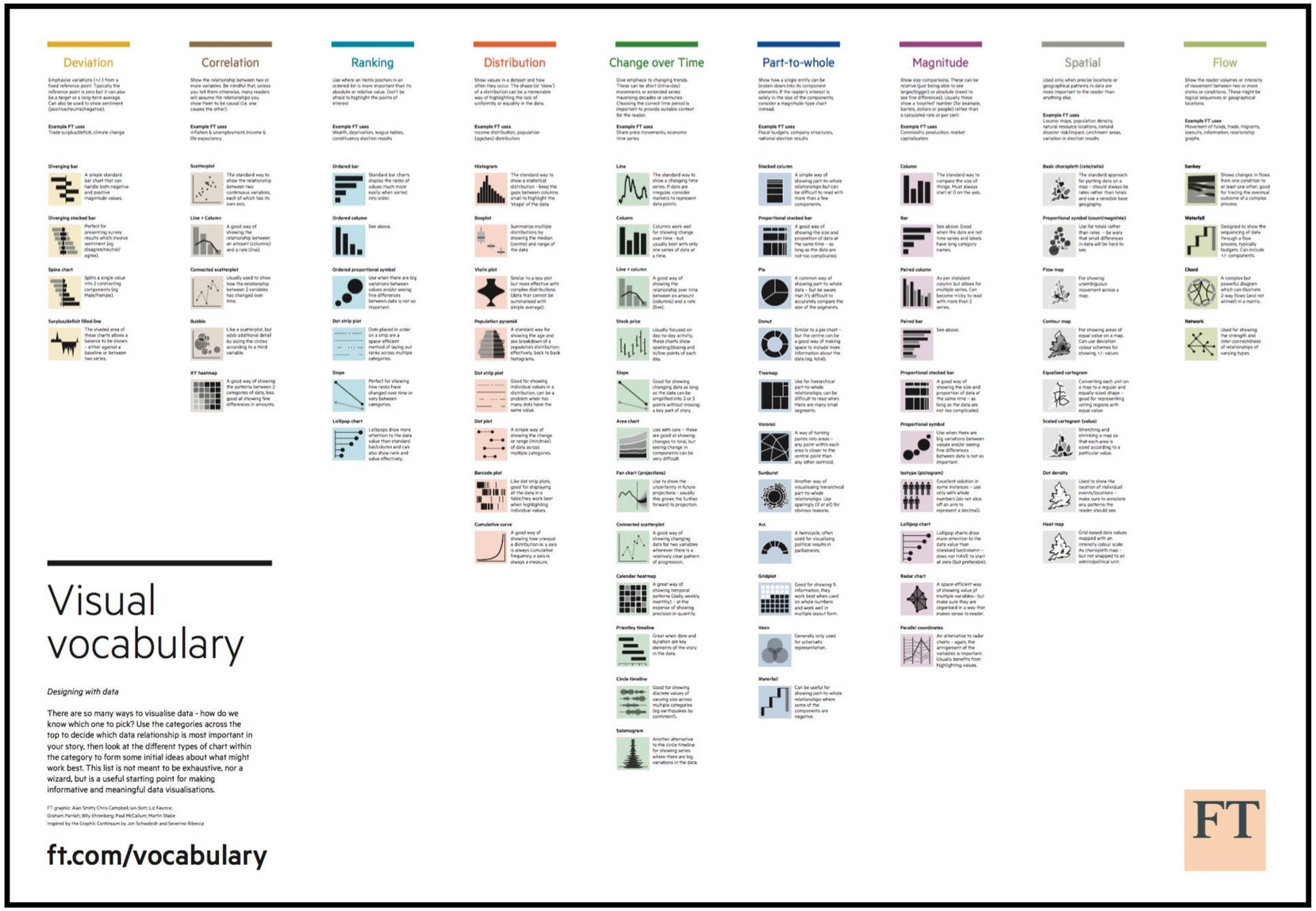






Alberto Cairo • University of Miami • www.thefunctionalart.com • Twitter: @albertocairo





http://www.datavizcatalogue.com/

https://github.com/ft-interactive/chart-doctor/blob/master/visual-vocabulary/Visual-vocabulary.pdf