

DATA VISUALIZATION

MPA 630: Data Science for Public Management

September 13, 2018

*Fill out your reading report
on Learning Suite*

PLAN FOR TODAY

Addressing some concerns

Data visualization

The grammar of graphics

5NG

Birds and airports
(for real this time)

ADDRESSING SOME CONCERNS

SOME ANSWERS

Learning curve

RStudio.cloud vs. RStudio Desktop

DataCamp vs. RStudio

Websites

Problem sets

R in the real world

DATA VISUALIZATION

ALL WE NEED IS RAW DATA

x	y
55.3846	97.1795
51.5385	96.0256
46.1538	94.4872
42.8205	91.4103
40.7692	88.3333
38.7179	84.8718
35.641	79.8718
33.0769	77.5641
28.9744	74.4872
26.1538	71.4103
23.0769	66.4103
22.3077	61.7949
22.3077	57.1795
23.3333	52.9487
25.8974	51.0256
29.4872	51.0256
32.8205	51.0256
35.3846	51.4103
40.2564	51.4103
44.1026	52.9487
46.6667	54.1026
50.0	55.2564
53.0769	55.641
56.6667	56.0256
59.2308	57.9487
61.2821	62.1795
61.5385	66.4103
61.7949	69.1026

```
> mean(new_data$x)
[1] 54.26327
> mean(new_data$y)
[1] 47.83225
> sd(new_data$x)
[1] 16.76514
> sd(new_data$y)
[1] 26.9354
> cor(new_data$x, new_data$y)
[1] -0.06447185
```

No correlation!

J K L O L



HUMANS LOVE PATTERNS

frontiers in
NEUROSCIENCE

REVIEW ARTICLE
published: 22 August 2014
doi: 10.3389/fnins.2014.00265



Superior pattern processing is the essence of the evolved human brain

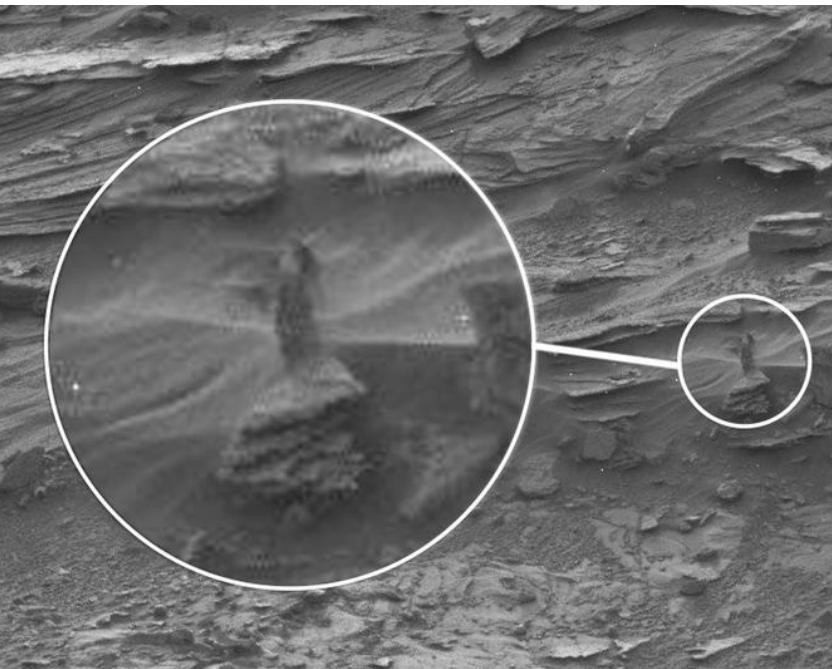
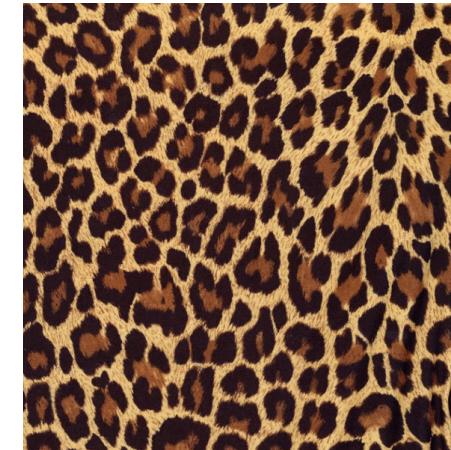
Mark P. Mattson^{1,2*}

¹ Laboratory of Neurosciences, National Institute on Aging Intramural Research Program, Baltimore, MD, USA

² Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Edited by:
J. Michael Williams, Drexel
University, USA

Humans have long pondered the nature of their mind/brain and, particularly why its capacities for reasoning, communication and abstract thought are far superior to other species, including closely related anthropoids. This article considers superior pattern



WHY VISUALIZE DATA?

Graphs let us see patterns in our data

Sometimes graphs alone are sufficient for telling a story and drawing inference from data



THIS IS DIFFICULT

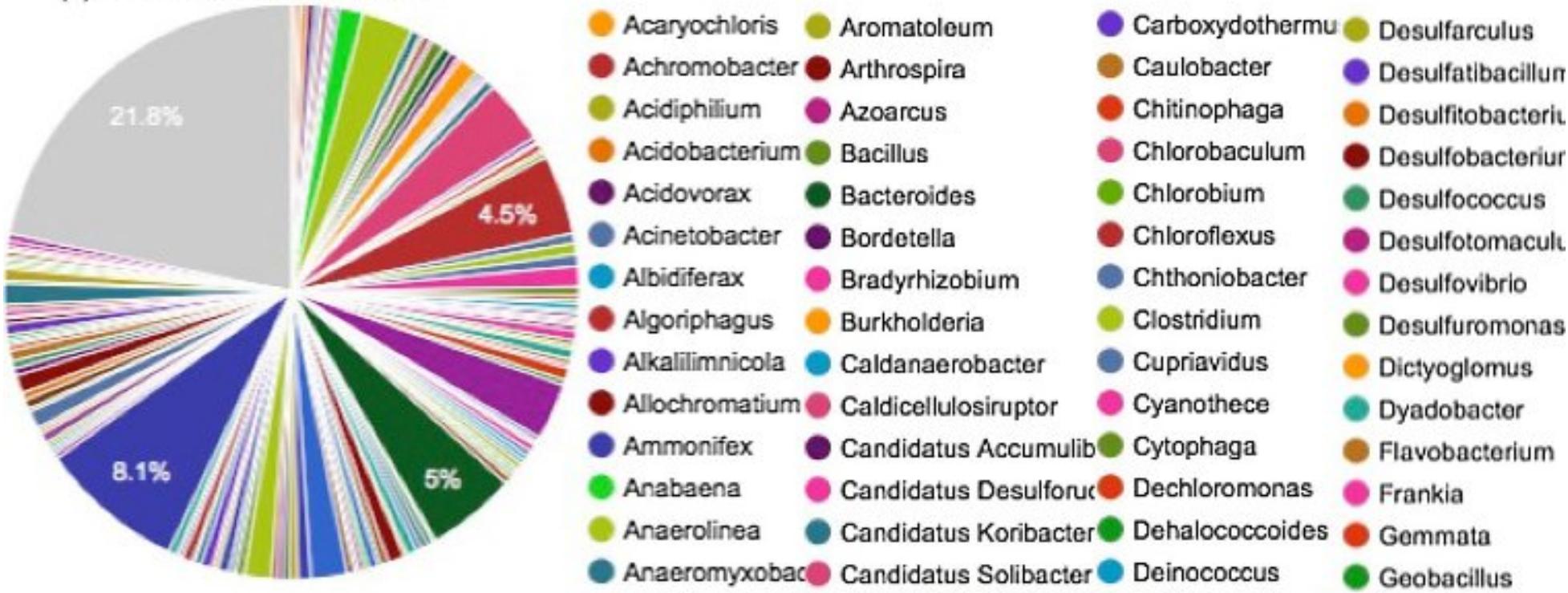


Incompetence

Deceit

Complexity

(f) Distribution of Genus



Geobacter	Methylacidiphilum	Nodularia	Pseudomonas	Sphingomonas	Thermodesulfovibrio
Gloeobacter	Methylibium	Nostoc	Ralstonia	Spirochaeta	Thermomicrobium
Haliangium	Methylobacterium	Opitutus	Rhizobium	Spirosoma	Thermosinus
Heliobacterium	Methylococcus	Oscillatoria	Rhodopirellula	Stigmatella	Thermosynechococcus
Herpetosiphon	Microcoleus	Oscillochloris	Rhodopseudomonas	Streptomyces	Thermotoga
Isosphaera	Microcystis	Paenibacillus	Rhodospirillum	Symbiobacterium	Xanthomonas
Ktedonobacter	Microscilla	Pedobacter	Rhodothermus	Synechococcus	Thermus
Leadbetterella	Moorella	Pedosphaera	Roseiflexus	Synechocystis	Thioalkalivibrio
Leptospira	Mycobacterium	Pelobacter	Roseomonas	Syntrophobacter	Thiobacillus
Leptothrix	Myxococcus	Pelotomaculum	Rubrobacter	Syntrophus	Trichodesmium
Lyngbya	Nitrosococcus	Pirellula	Shewanella	Thermaerobacter	Truepera
Magnetospirillum	Nitrosomonas	Planctomyces	Sideroxydans	Variovorax	Other
					unclassified from Bacteria
					unclassified Deltaproteobacteria
					Verminephrobacter

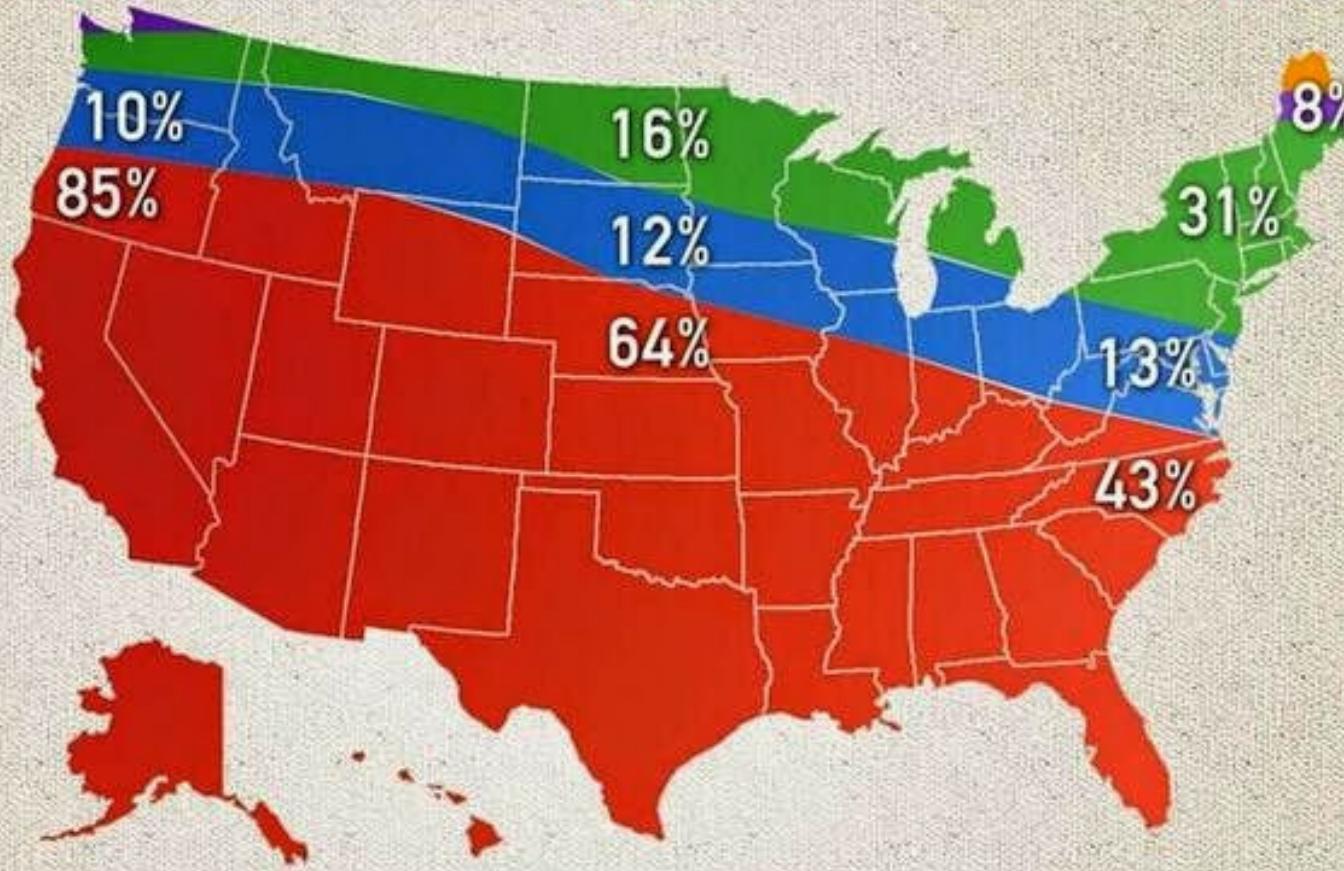
CHANGING FACE OF AMERICA

Percent of total U.S. population by race and ethnicity, 1960–2060

1960

2010

2060



OTHER

ASIAN

HISPANIC

BLACK

WHITE

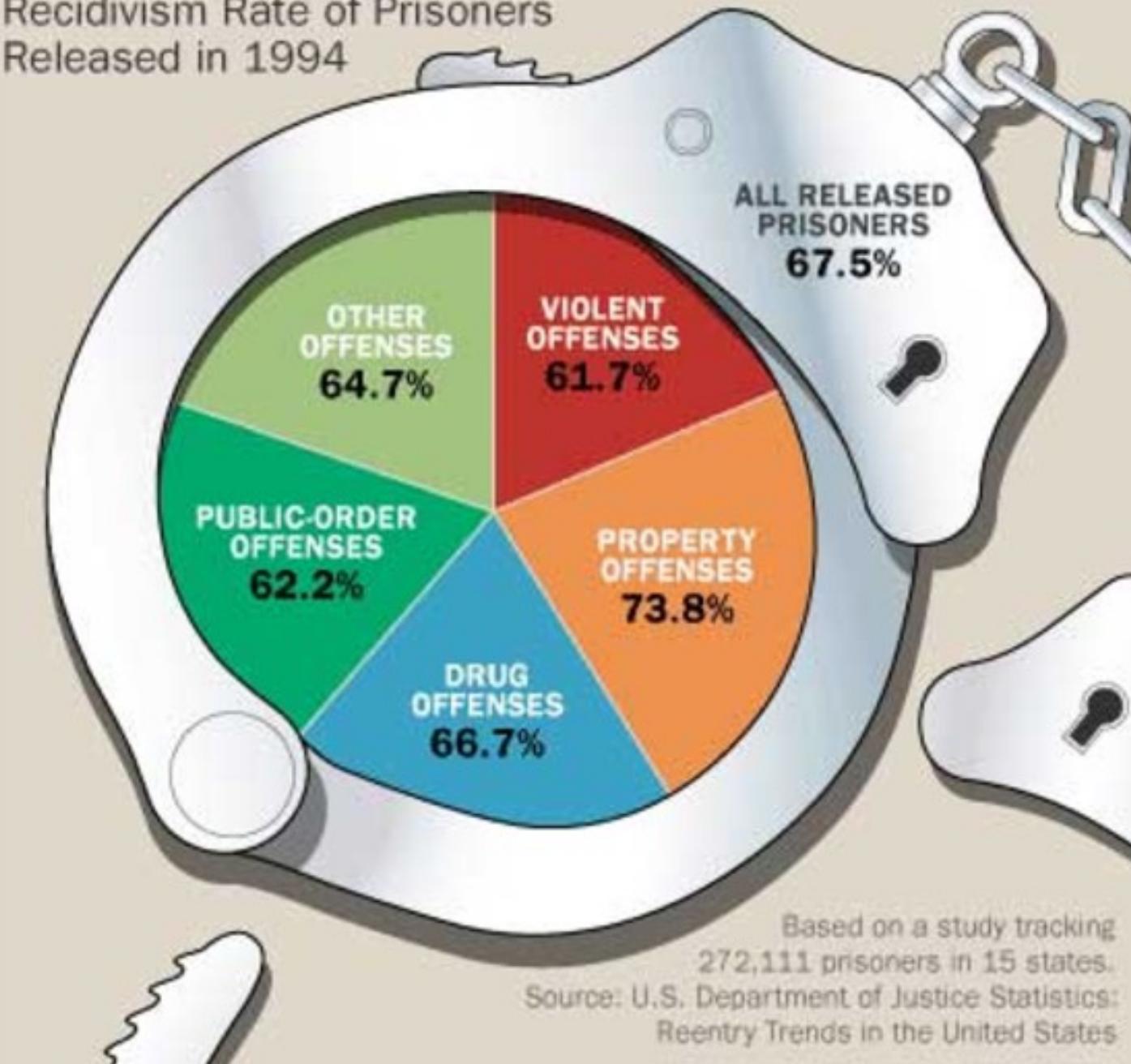
SOURCE: PEW RESEARCH CENTER

NIGHTLY NEWS
with BRIAN WILLIAMS

How Prison Works

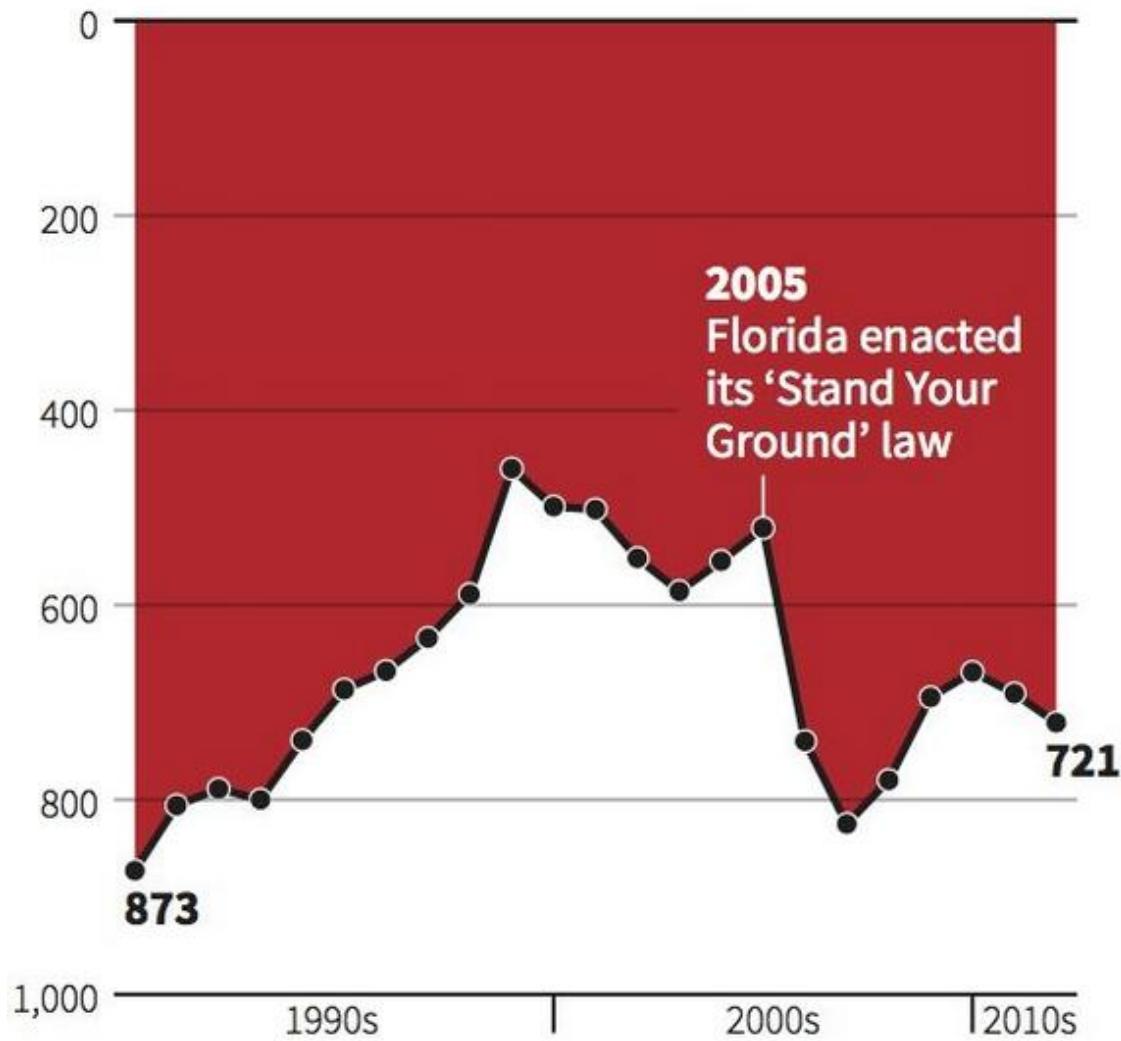
©2007 HowStuffWorks

Recidivism Rate of Prisoners
Released in 1994



Gun deaths in Florida

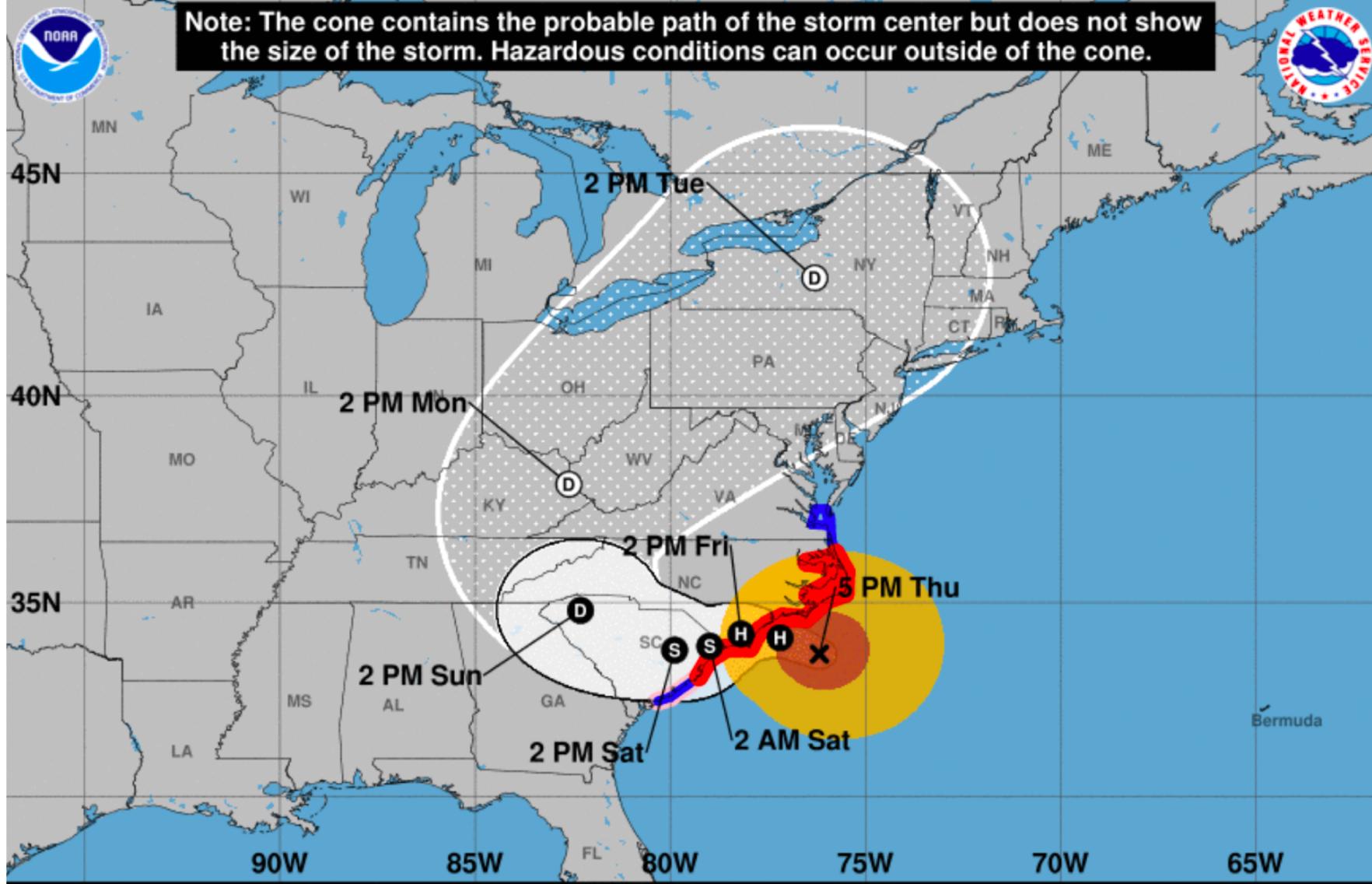
Number of murders committed using firearms



Source: Florida Department of Law Enforcement



Note: The cone contains the probable path of the storm center but does not show the size of the storm. Hazardous conditions can occur outside of the cone.



Hurricane Florence

Thursday September 13, 2018

5 PM EDT Advisory 58

NWS National Hurricane Center

Current information: **X**

Center location 33.7 N 76.2 W

Maximum sustained wind 100 mph

Movement WNW at 5 mph

Forecast positions:

● Tropical Cyclone ○ Post/Potential TC

Sustained winds: D < 39 mph

S 39-73 mph H 74-110 mph M > 110 mph

Potential track area:



Day 1-3



Day 4-5

Watches:

Hurricane Trop Stm

Warnings:

Hurricane Trop Stm

Current wind extent:

Hurricane Trop Stm

Labels

Greenville

NORTH
CAROLINA

Kinston

New Bern

Pamlico Sound

Havelock

Jacksonville

30km

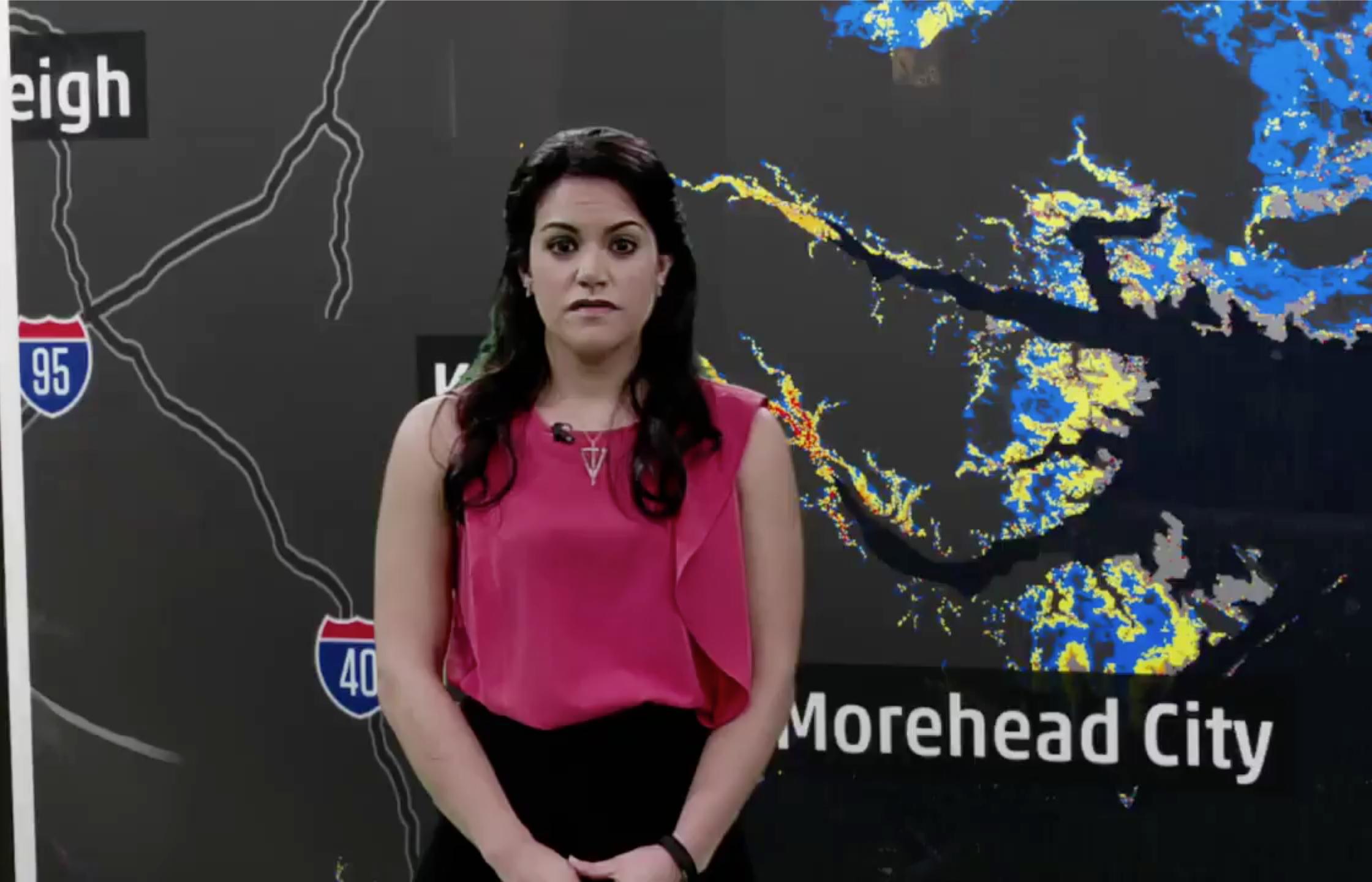
20mi



9 FT
The Weather Channel

6 FT

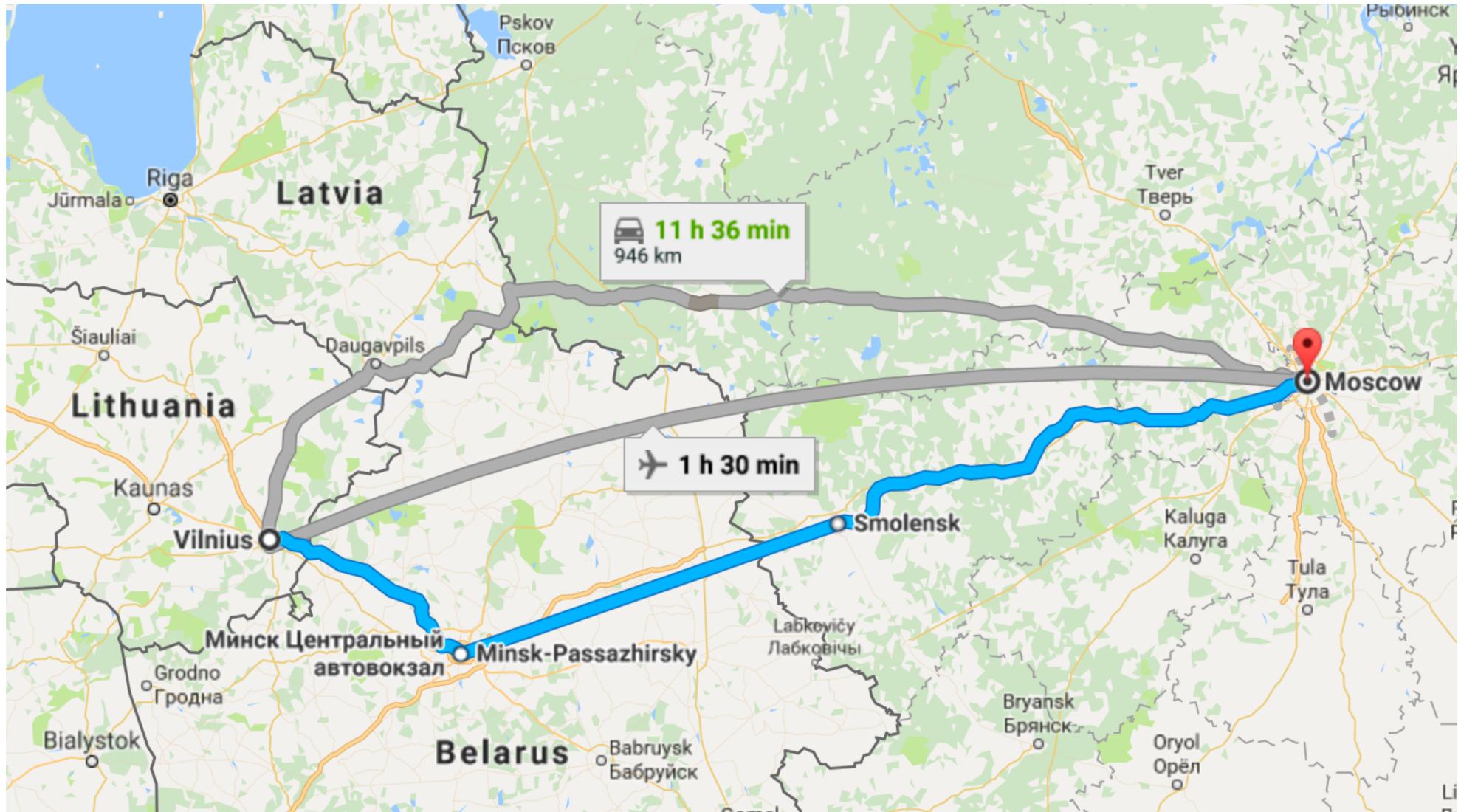
3 FT



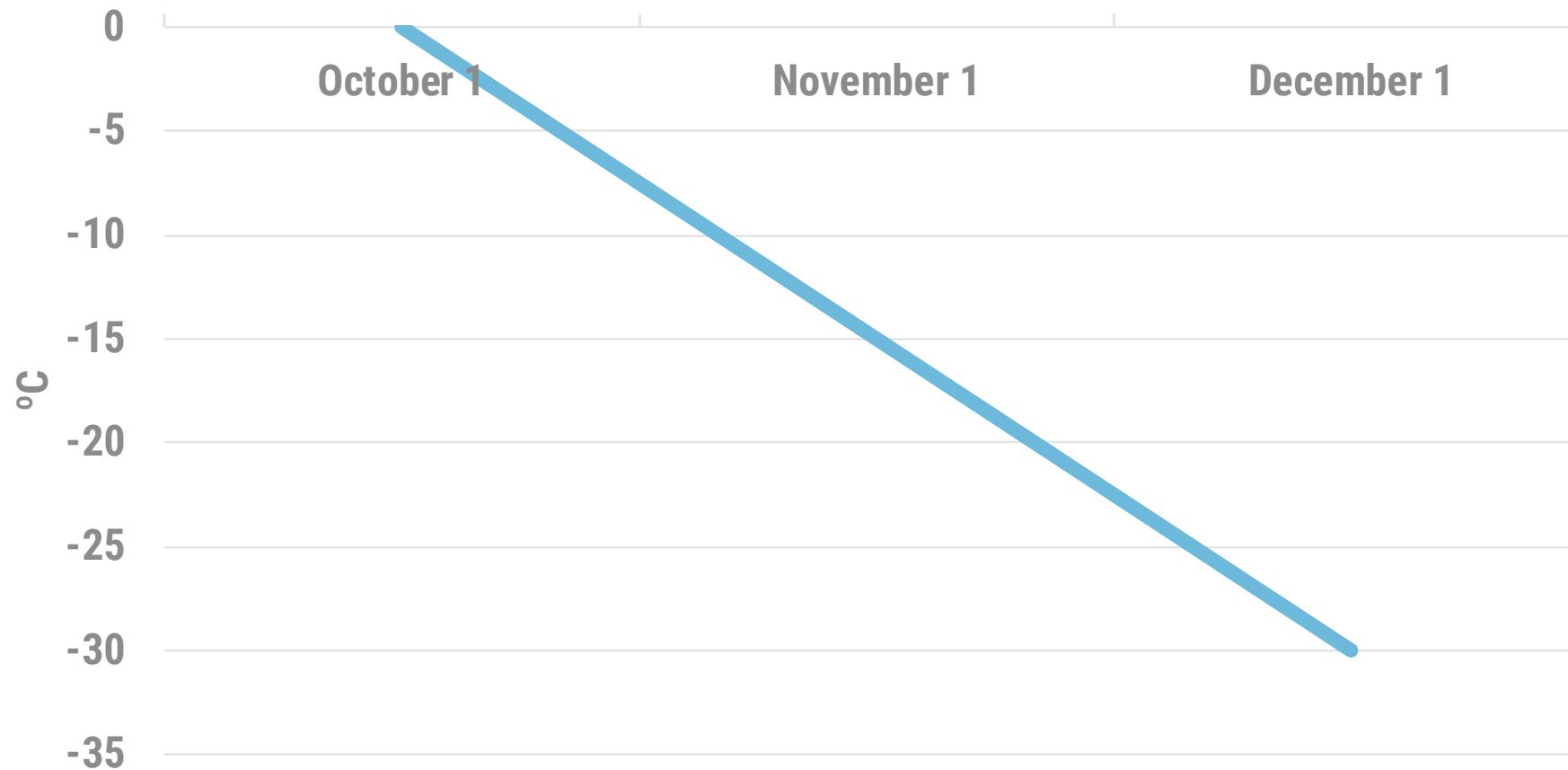
THE GRAMMAR OF GRAPHICS



THE WAR OF 1812

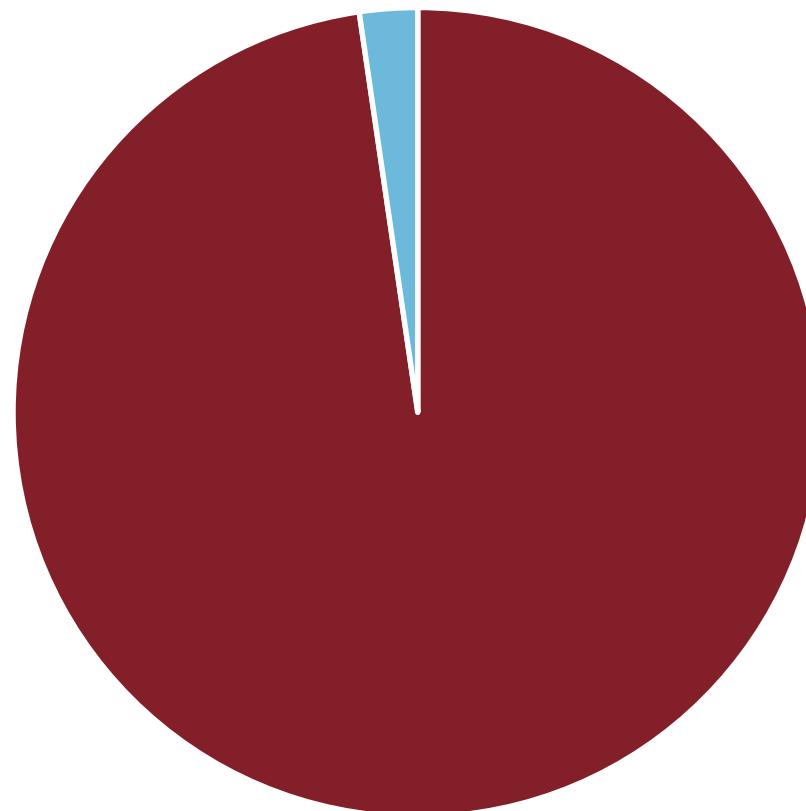


THE WAR OF 1812



THE WAR OF 1812

Napoleon's Grande Armée



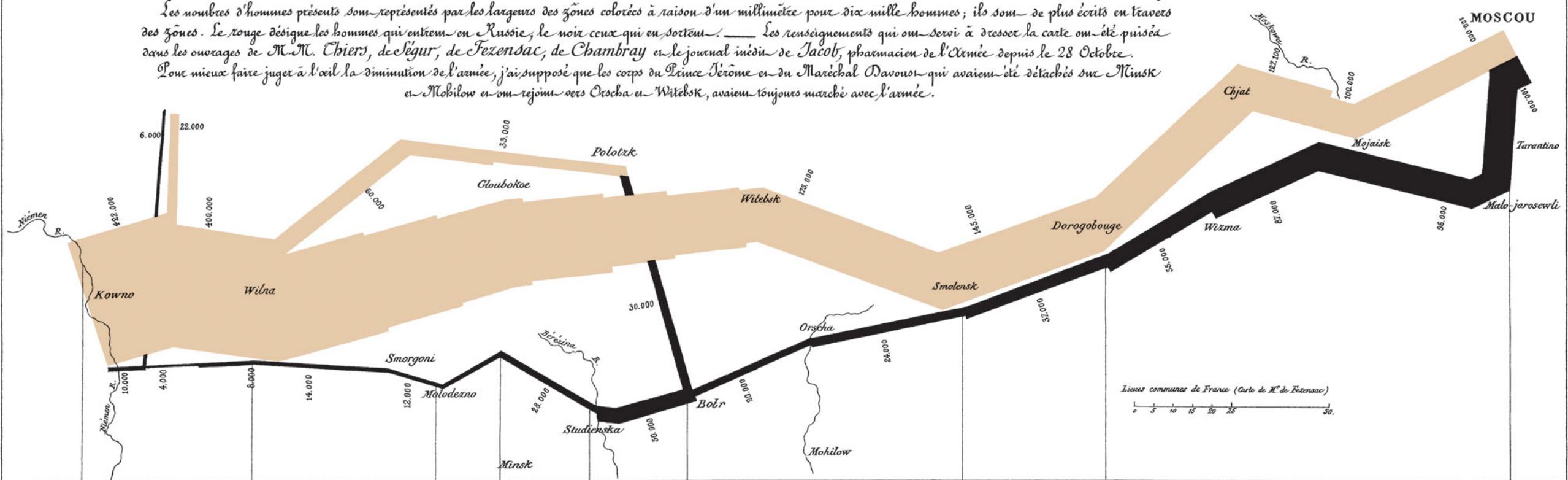
■ Died ■ Survived

Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussees en retraite
Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie; le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Séguir, de Fézensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk en Mohilow et se rejoignaient vers Orscha en Wilebsk, avaient toujours marché avec l'armée.



Lieux communs de France (Carte de M. de Fézensac)

0 5 10 15 20 25 30

TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les Cosaques passent au galop
le Niemen gelé.



AESTHETICS AND DATA

Data	Aesthetic	Graphic
Longitude	x	point
Latitude	y	point
Army size	size	path
Army direction	color	path
Date	x	line and text
Temperature	y	line and text

AESTHETICS AND DATA

data

Longitude

Latitude

Army size

Army direction

Date

Temperature

aes()

x

y

size

color

x

y

geom_

point

point

path

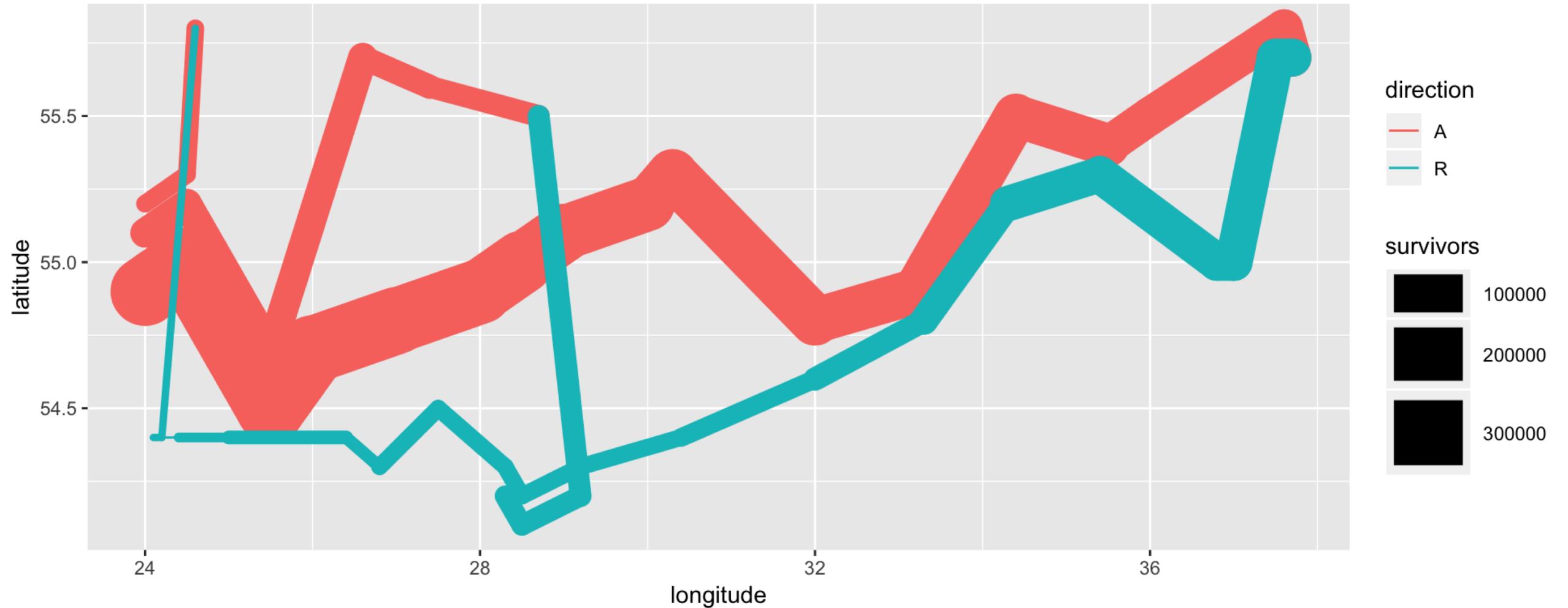
path

line and text

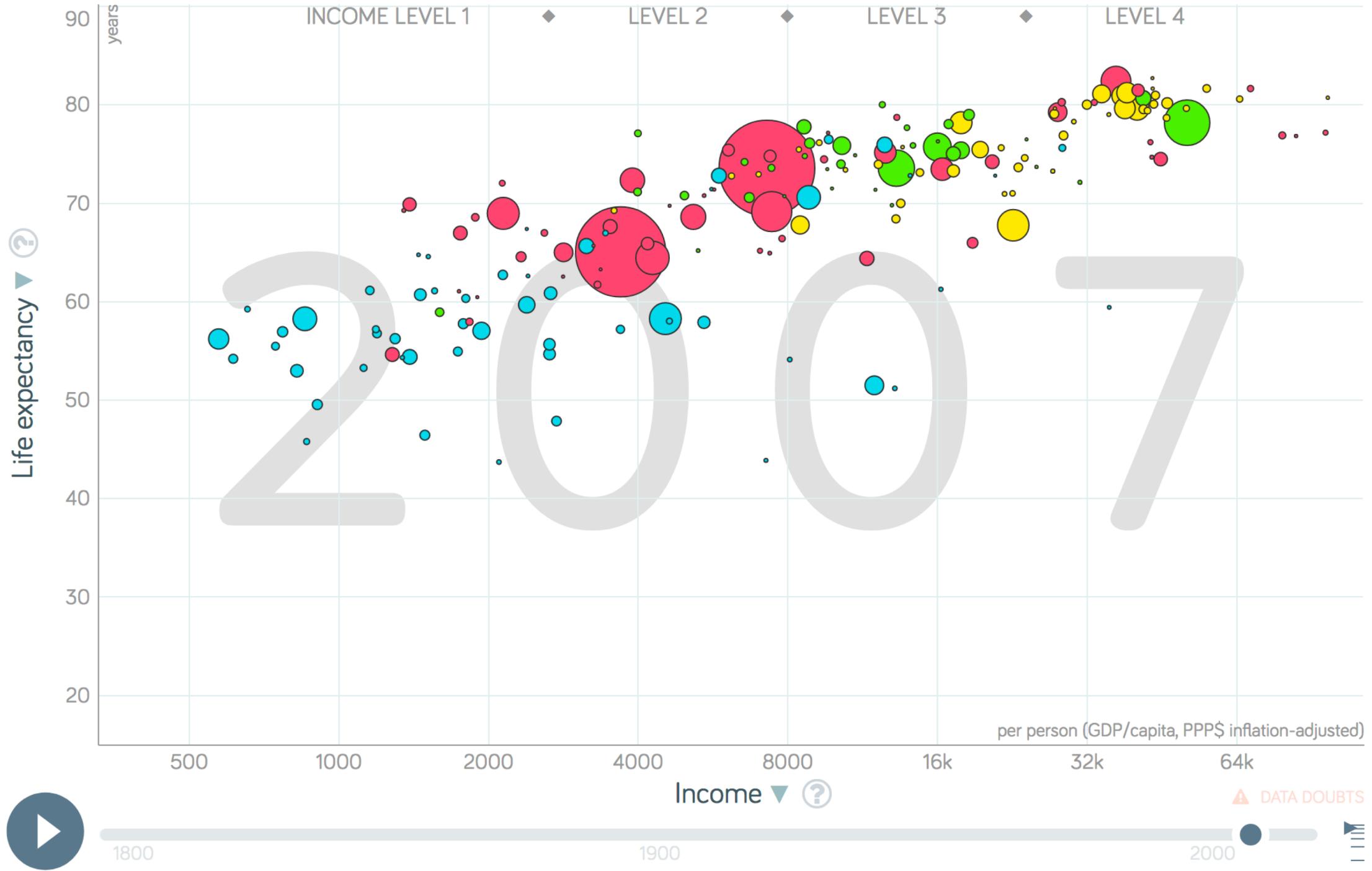
line and text

ANATOMY OF `ggplot()`

ANATOMY OF `ggplot()`







AESTHETICS AND DATA

data

Wealth (GDP/capita)

Health (Life expectancy)

Continent

Population

aes()

x

y

color

size

geom_

point

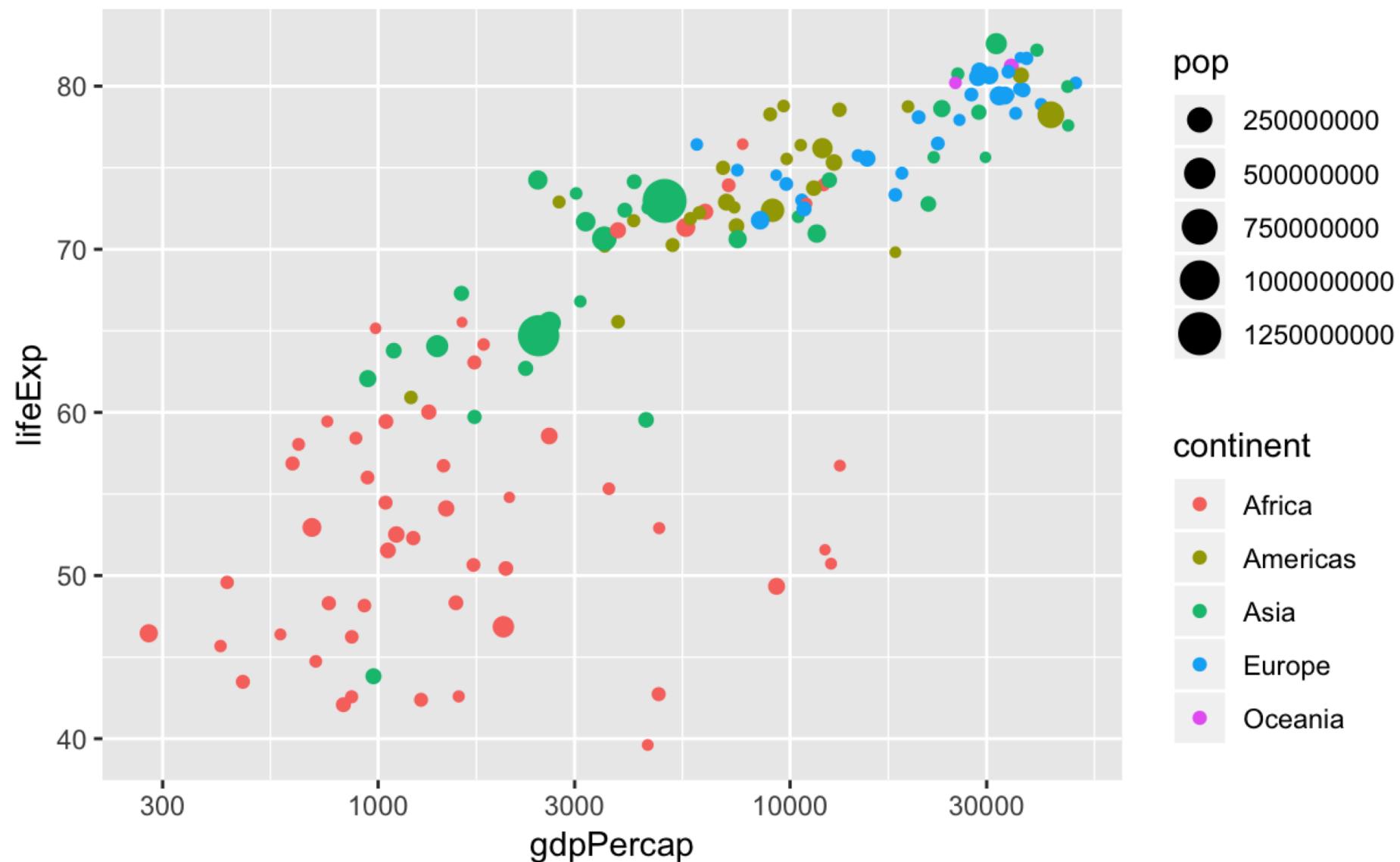
point

point

point

HEALTH AND WEALTH

HEALTH AND WEALTH



5 N G

THE BIG 5

Scatterplots

Linegraphs

Histograms

Boxplots

Barplots

LIVE EXAMPLES

B R E A K

BIRDS AND AIRPORTS

On your own!