Distributions & uncertainty

Nate Cohn ✔
@Nate_Cohn
Following

The GA-6 live model is live.
nytimes.com/elections/resu …

Median forecast

ESTIMATED MARGIN, in pct. points
Even
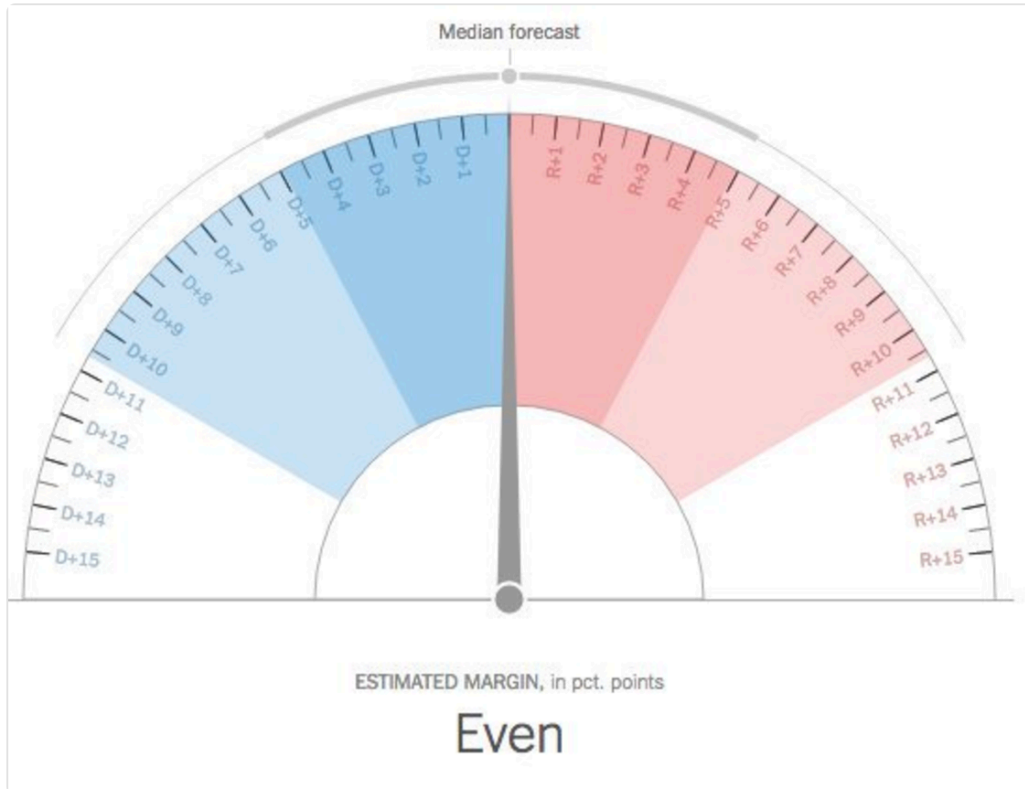
4:47 PM - 20 Jun 2017 from Manhattan, NY

481 Retweets 660 Likes

Virgil Texas ✔ @virgiltexas · Jun 20
Replying to @Nate_Cohn
Nate

Alp Ozcelik ✔ @alplicable · Jun 20
Replying to @Nate_Cohn
DO NOT DO THIS TO ME AGAIN

Sarcasmorator @Sarcasmorator · Jun 20
Replying to @Nate_Cohn @jacquicollins_
ah, yes, the election stress-o-meter

2   36

Rubric draft

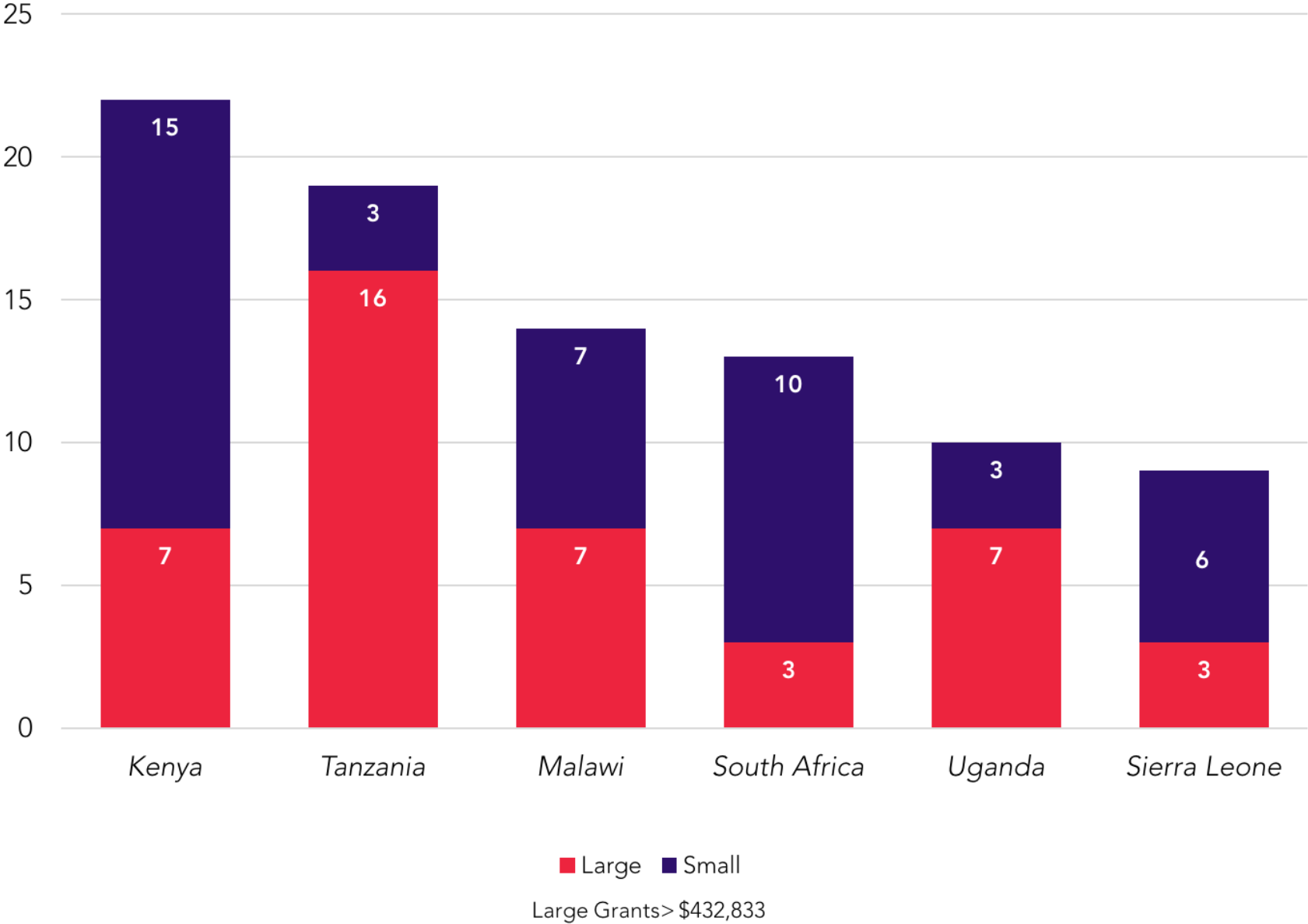# Comic Relief redesign

Uncertainty

Sooooo much R

Rubric draft

VISUALIZATION RUBRIC RUBRIC

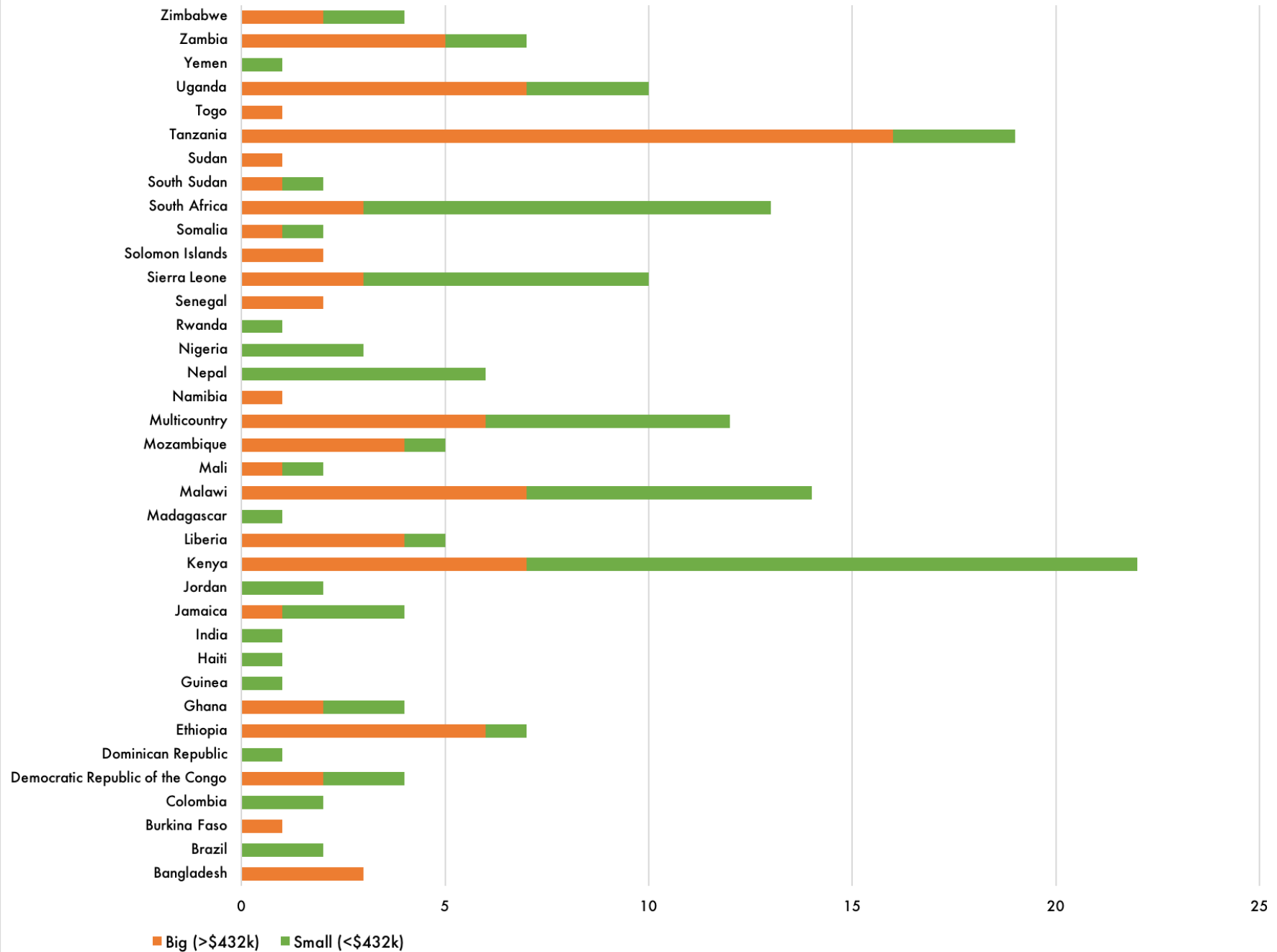|	STRONG	MEDIUM	WEAK
COVERS THE RIGHT CONTENT	The content of the rubric represents the best thinking in the field of data visualization about what it means to create meaningful and truthful graphics.	Much of the content of the rubric represent the best thinking in data visualization, but there are a few places that are questionable.	It is difficult to tell what the rubric is intended to assess and the content is far removed from the best thinking in data visualization.
	The content of the rubric aligns with the objectives of the course and covers issues of credibility, ethics, aesthetics, and accessibility.	Some features don't align with the objectives of the course and omit features related to credibility, ethics, aesthetics, or accessibility.	The rubric does not alight with the objectives of the course and does not cover features related to credibility, ethics, aesthetics, or accessibility.
CRITERIA ARE WELL ORGANIZED	The rubric is divided into easily understandable criteria. The number of criteria reflects the complexity of assessing data visualizations.	The number of criteria needs to be adjusted a little—either a single criterion should be made into two criteria, or two criteria should be combined.	The rubric is a long list of everything; there is no organization. The rubric looks like a brainstormed list.
	The details that are used to describe a criterion go together; you can see how they are facets of the same criterion.	Some details that are used to describe a criterion are in the wrong criterion, but most are placed correctly.	The rubric seems mixed up. Descriptors that go together don't seem to be placed together; things that are different are put together.
	The relative emphasis on various features of performance is right—things that are more important are stressed more; things that are less important are stressed less.	The emphasis on some criteria or descriptors is either too small or too great; others are okay.	The rubric is out of balance—features of more importance are emphasized the same as features of less importance.
	The criteria are independent. Each important feature appears in only one place in the rubric.	Some features are included in more than one criterion.	Features are so redundant that the criteria don't really cover different things.
NUMBER OF LEVELS	The number of levels of quality used in the rating scale makes sense. There are enough levels to accurately rate the visualization, but not so many levels that it is impossible to distinguish among them.	More levels would be helpful to make finer distinctions in the assessment of the visualization.	The number of levels is not appropriate—there are either so many levels that it is impossible to reliably distinguish between them, or too few to make important distinctions.
LEVELS DEFINED WELL	Each score point is defined with indicators or descriptions.	Only the top level is defined; the other levels are not defined.	No levels are defined; the rubric is little more than a list of categories to rate followed by a rating scale.
	There is enough descriptive detail to allow you to match the visualization's performance to the right score.	There is some attempt to define terms and include descriptors, but some key ideas are fuzzy in meaning.	Wording of levels (if present) is vague or confusing.
	Two independent users raters assign the same rating most of the time.	It is uncertain if two independent testers, even with practice, could assign the same rating most of the time.	It is unlikely that two independent raters could assign the same rating.
	Wording is descriptive, not evaluative.	Wording is mostly descriptive, but there are a few instances of evaluative labels.	Wording is evaluative and not descriptive (e.g. the visualization is "interesting" or "above average")
PARALLEL LEVELS	The levels of the rubric are parallel in content—if a feature is discussed in one level, it is discussed in all levels. If the levels are not parallel, there is a good explanation why.	The levels are mostly parallel in content, but there are some places where there is a feature at one level that is not present at the other levels.	Levels are not parallel in content and there is no explanation of why.

Comic Relief redesign

Countries Receiving the Highest Number of Grants

Source: Comic Relief

	Kenya	Tanzania	Malawi	South Africa	Uganda	Sierra Leone
Small	15	3	7	10	3	6
Large	7	16	7	3	7	3

■ Large   ■ Small

Large Grants> $432,833

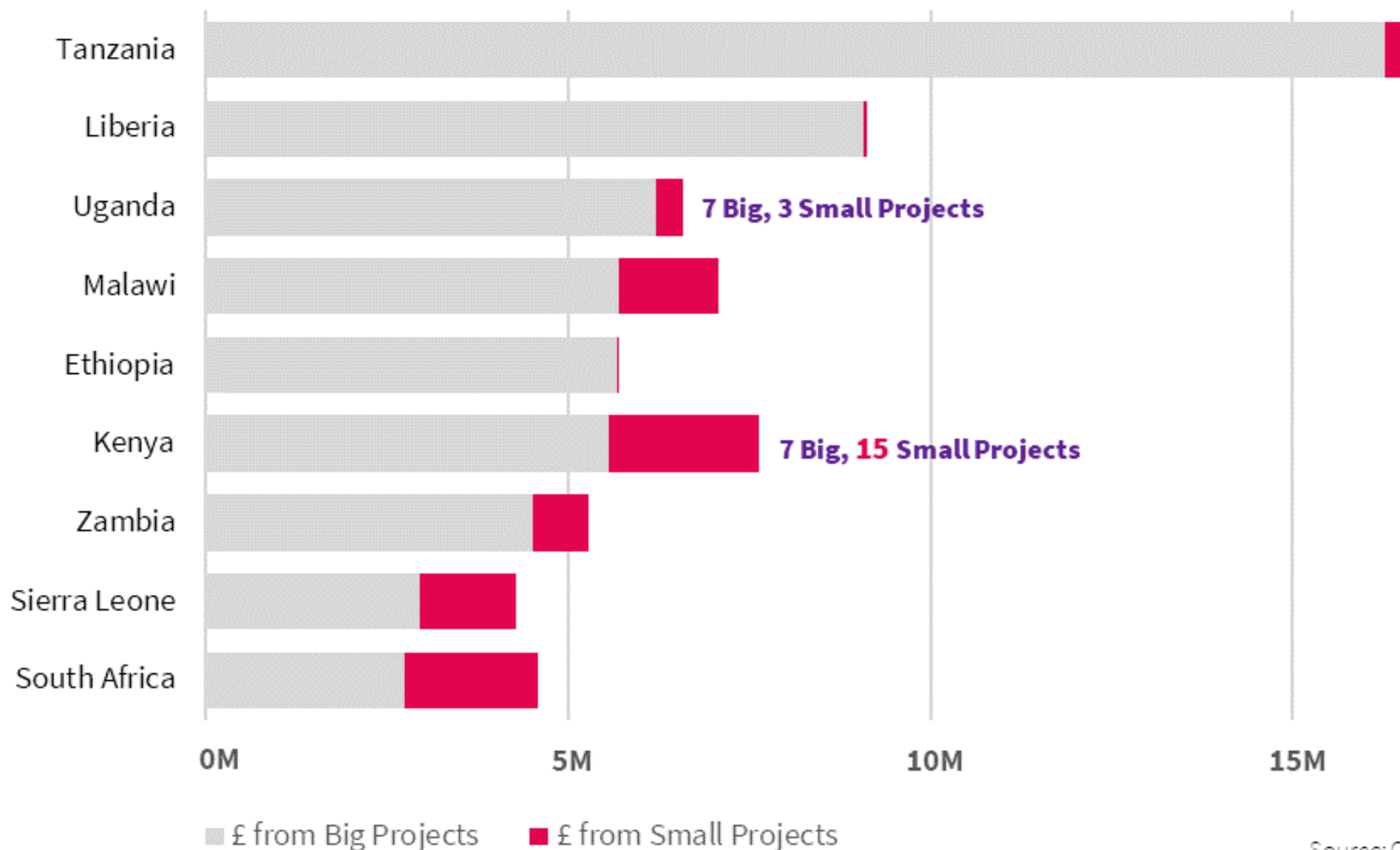Number of Grants Awarded by County
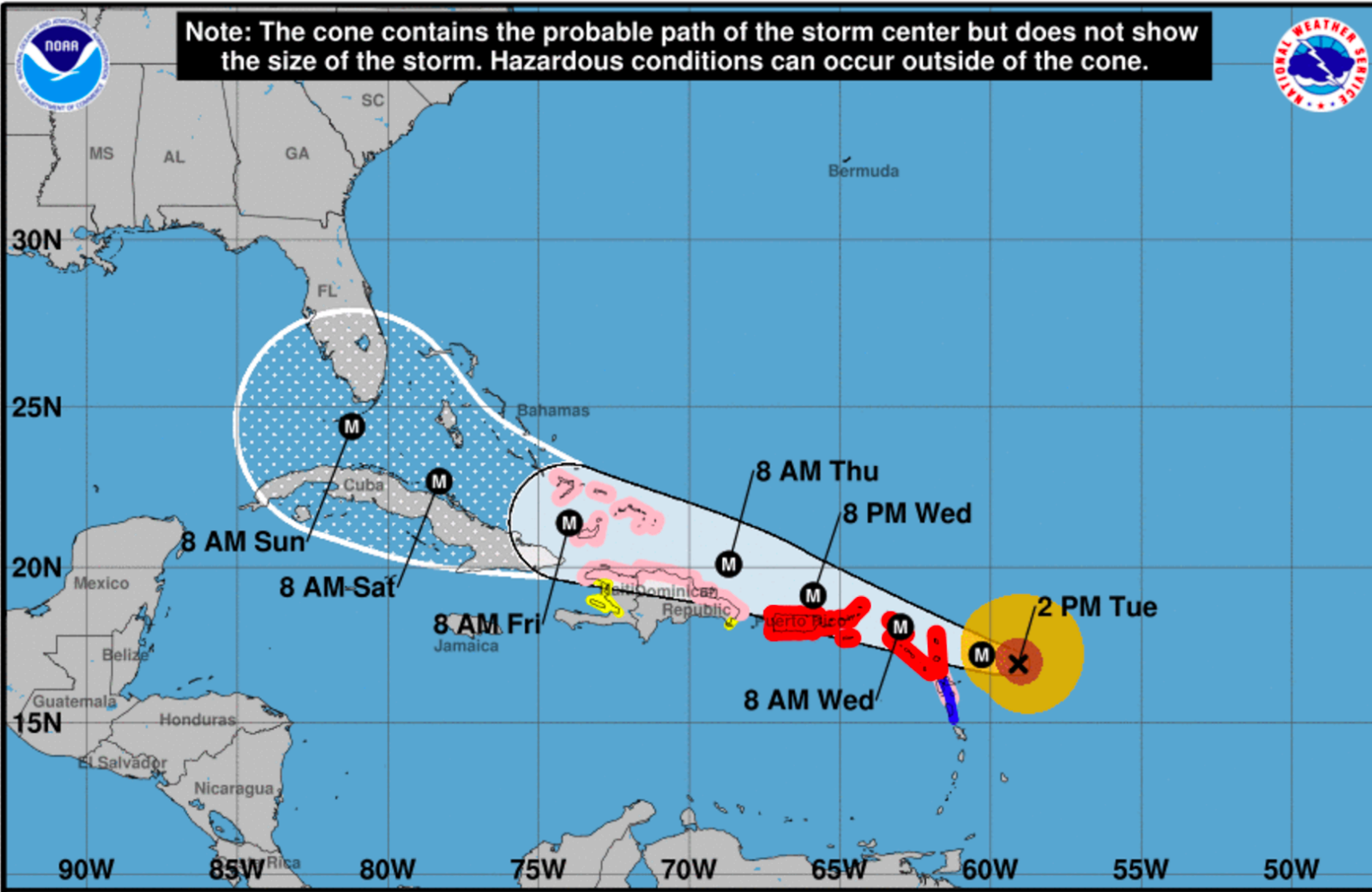
Saving PDFs from Excel in Windows

whyyyyy

Uncertainty

Why is it important?

Why is it so difficult?

Some Countries Receiving Funding for Big Projects also Receive Many Small Projects

Country	£ from Big Projects	£ from Small Projects
Tanzania		
Liberia		
Uganda	7 Big, 3 Small Projects	
Malawi		
Ethiopia		
Kenya	7 Big, 15 Small Projects	
Zambia		
Sierra Leone		
South Africa		

0M   5M   10M   15M

■ £ from Big Projects   ■ £ from Small Projects

Source: Comic Relief

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.

8 AM Thu

8 PM Wed

8 AM Sun

8 AM Sat

2 PM Tue

8 AM Fri

8 AM Wed

Hurricane Irma
Tuesday September 05, 2017
2 PM AST Intermediate Advisory 26A
NWS National Hurricane Center

Current information: ✕
Center location 16.9 N 59.1 W
Maximum sustained wind 185 mph
Movement W at 14 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
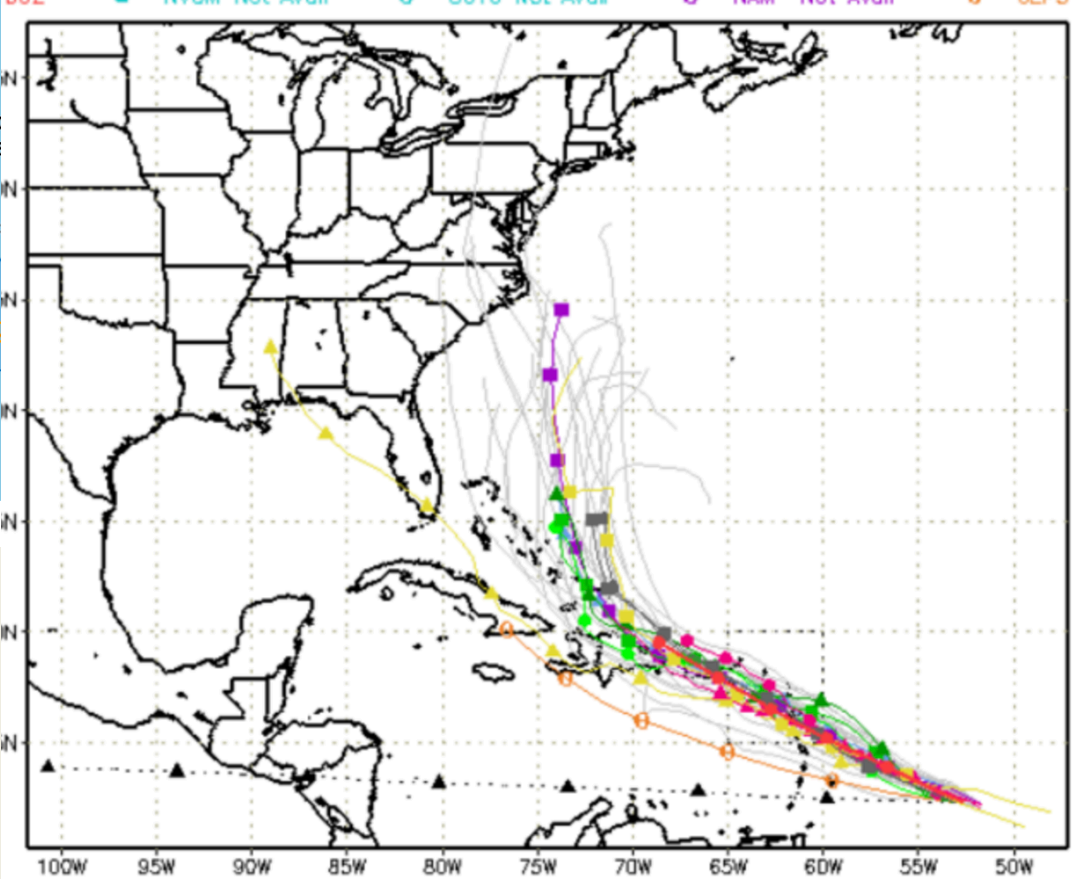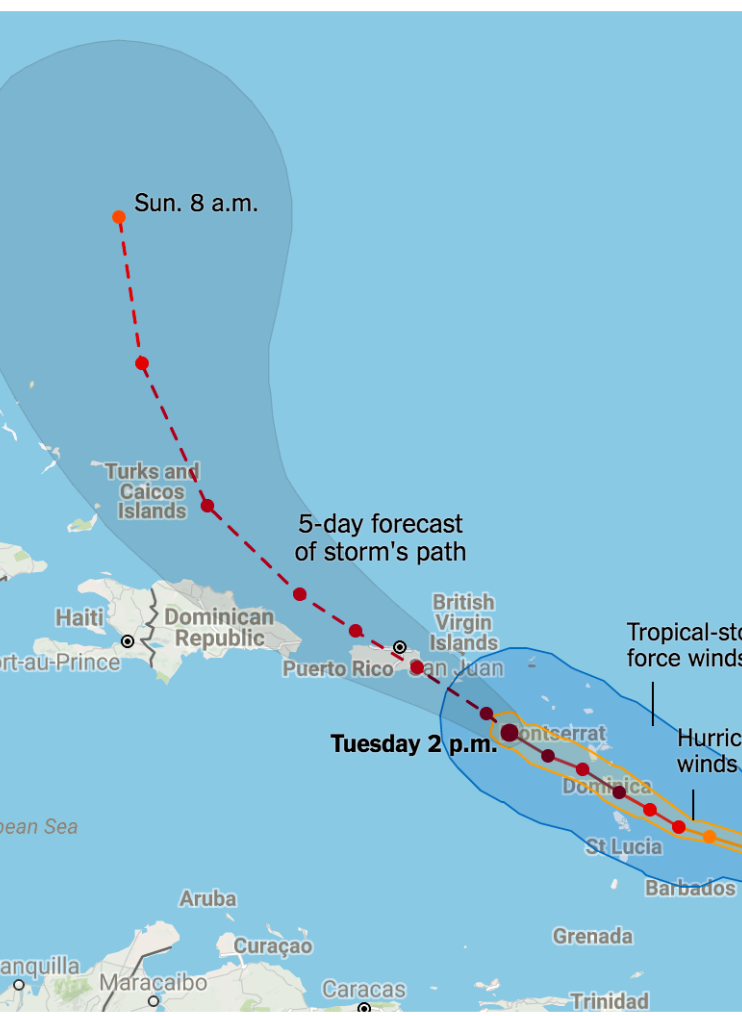
Inches

20

15

10

6

4

2

1
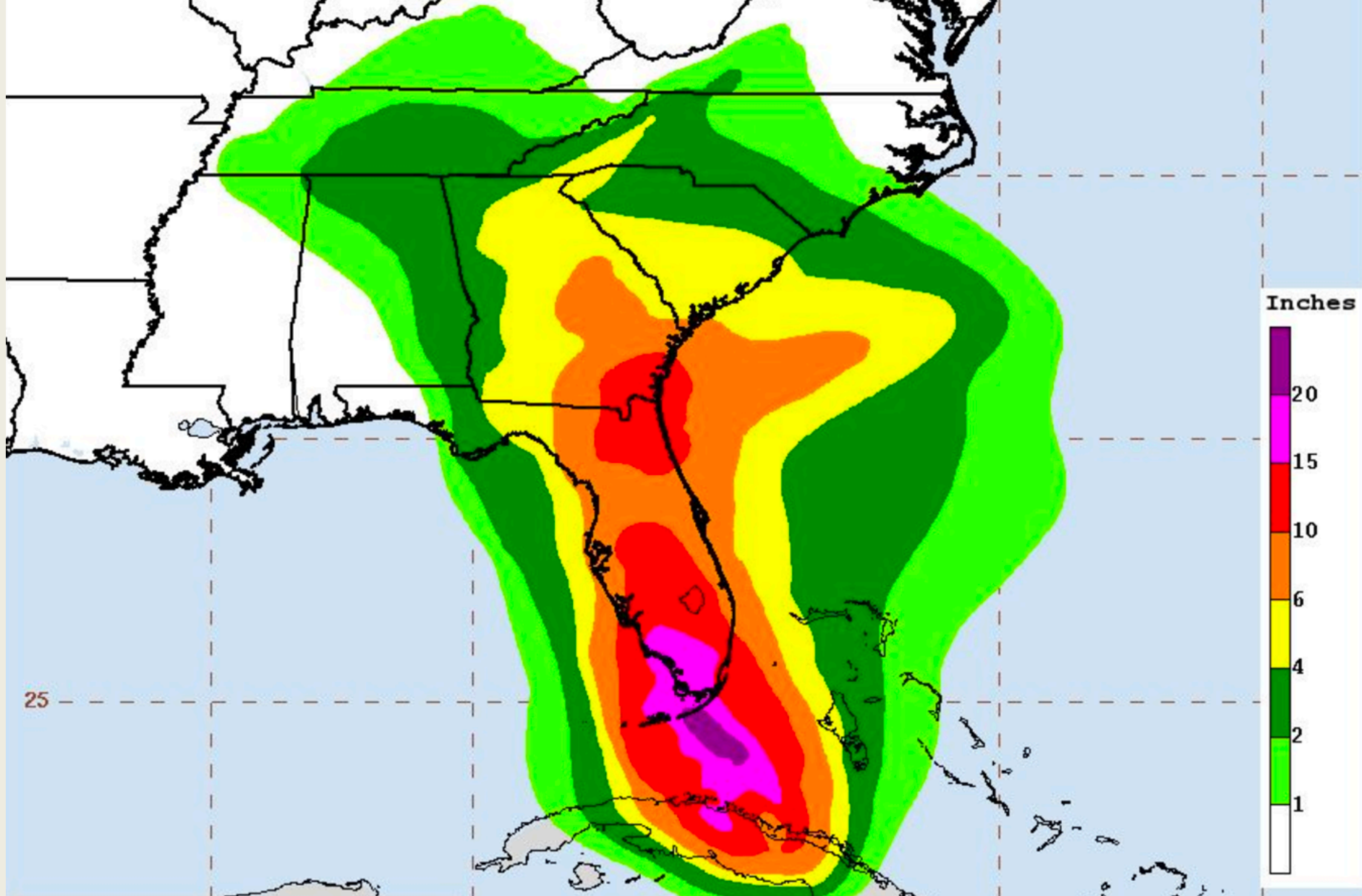
Hurricane Irma
120-hour Day 1-5 Rainfall Forecast (inches)
Created 5:34 AM EDT Sat Sep 9 2017
Valid 8:00 AM EDT Sat Sep 9 2017
through 8:00 AM EDT Thu Sep 14 2017
NOAA/NWS/NCEP/WPC

Local point maximum rainfall may be higher than shown.
See the NHC public advisories for the latest tropical cyclone information.

SEVERITY Category ■5 ■4 ■3 ■2 ■1 ■Tropical storm

Sun. 8 a.m.

Turks and Caicos Islands

5-day forecast of storm's path

Haiti

Dominican Republic

Port-au-Prince

Puerto Rico

British Virgin Islands

San Juan

Tropical-storm force winds

Tuesday 2 p.m.

Montserrat

Hurricane winds

Dominica

St Lucia

Barbados

Caribbean Sea

Aruba

Curaçao

Maracaibo

Barranquilla

Caracas

Grenada

Trinidad

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.

NOAA

NC
SC
MS AL GA
FL

Bermuda

33
30N
25N

8 AM Sun
8 AM Sat
8 AM Thu
8 AM Wed
8 AM Fri
8 PM Wed
2 PM Tue

H
M
M
M
M
M
M

Cuba

Haiti
Dominican Republic
Puerto Rico

75W 70W 65W 60W 55W 50W

Current information: ✗
Center location 16.6 N 63.6 W
Maximum sustained wind 160 mph
Movement WNW at 10 mph

Forecast positions:
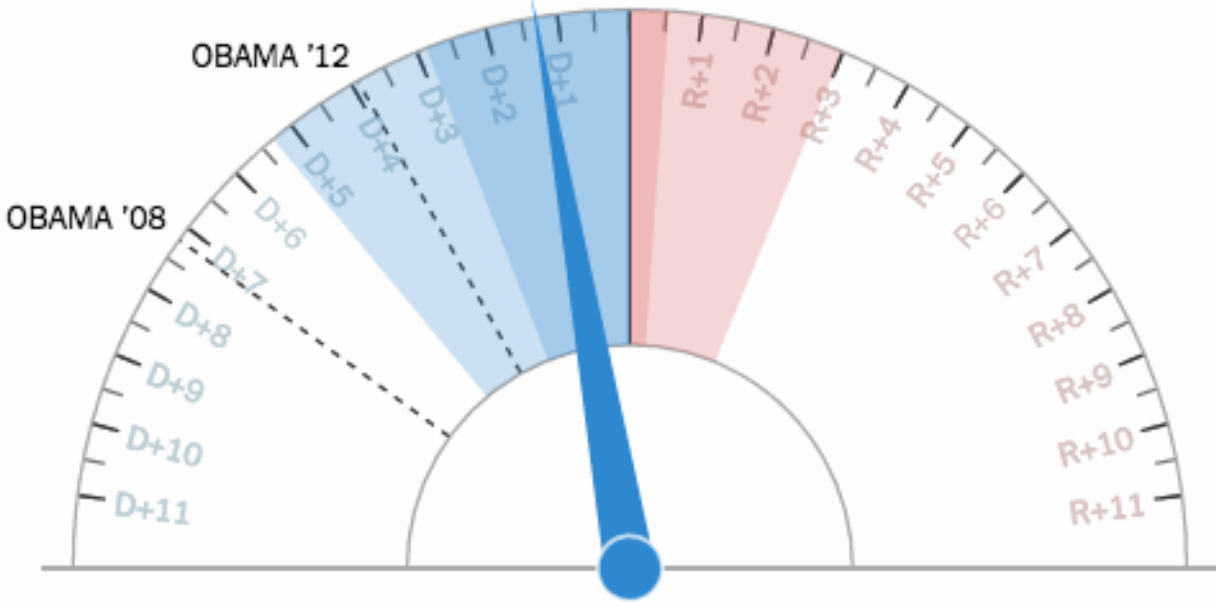● Tropical Cyclone ○ Post/Po
Sustained winds: D < 39 mp
S 39-73 mph H 74-110 mph M >

Warnings:
Trop Stm (yellow) Hurricane (red) Trop Stm (blue) Hurricane (brown) Trop (orange)

Current wind ext

NVGM Not Avail   GOTC Not Avail   NAM Not Avail   CLP5

100W 95W 90W 85W 80W 75W 70W 65W 6DW 55W 50W

Popular vote margin

OBAMA '12

OBAMA '08

D+11 D+10 D+9 D+8 D+7 D+6 D+5 D+4 D+3 D+2 D+1 R+1 R+2 R+3 R+4 R+5 R+6 R+7 R+8 R+9 R+10 R+11

Clinton +1.4

FORECAST, in pct. points

Predicted CSRE in following year

Political stability of neighbors

- No coups in neighboring countries
- Coup activity in neigboring countries

Relative protest activity

Less (1) Normal (3) More (5)

- Nonviolent
- Violent

Average number of absences

Weeks before/after truancy intervention

Lines
- Actual
- Predicted

Colors
- 80% Confidence
- 95% Confidence
- Truancy intervention

⚡Stats 101⚡

Variable types

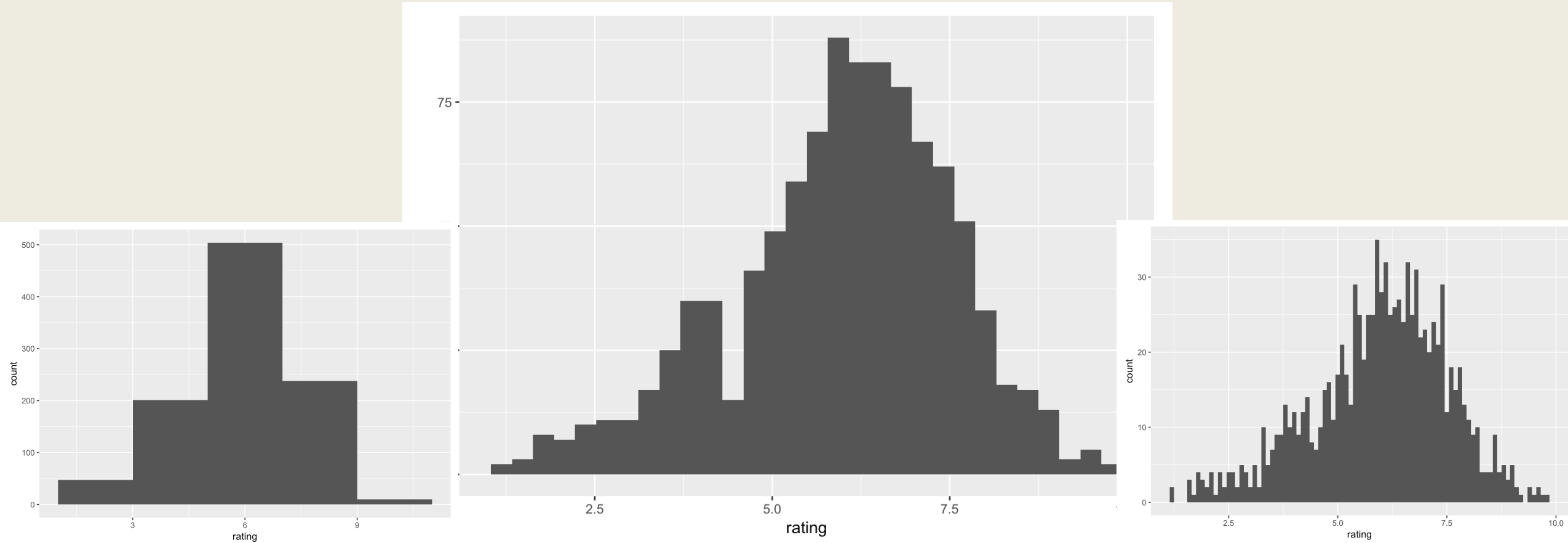| Nominal	Ordinal	Continuous

Gender

Income

Stock prices
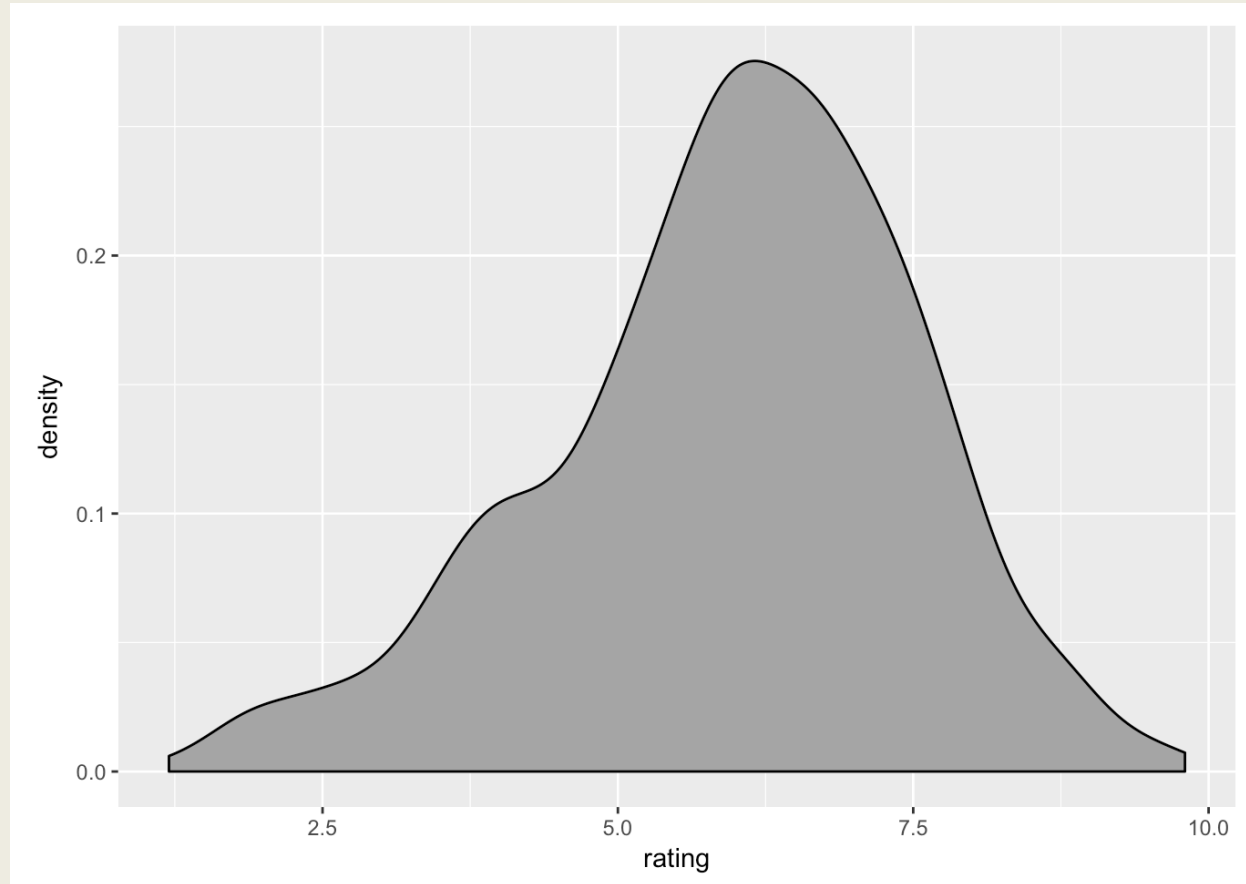
Fertility

Nationality

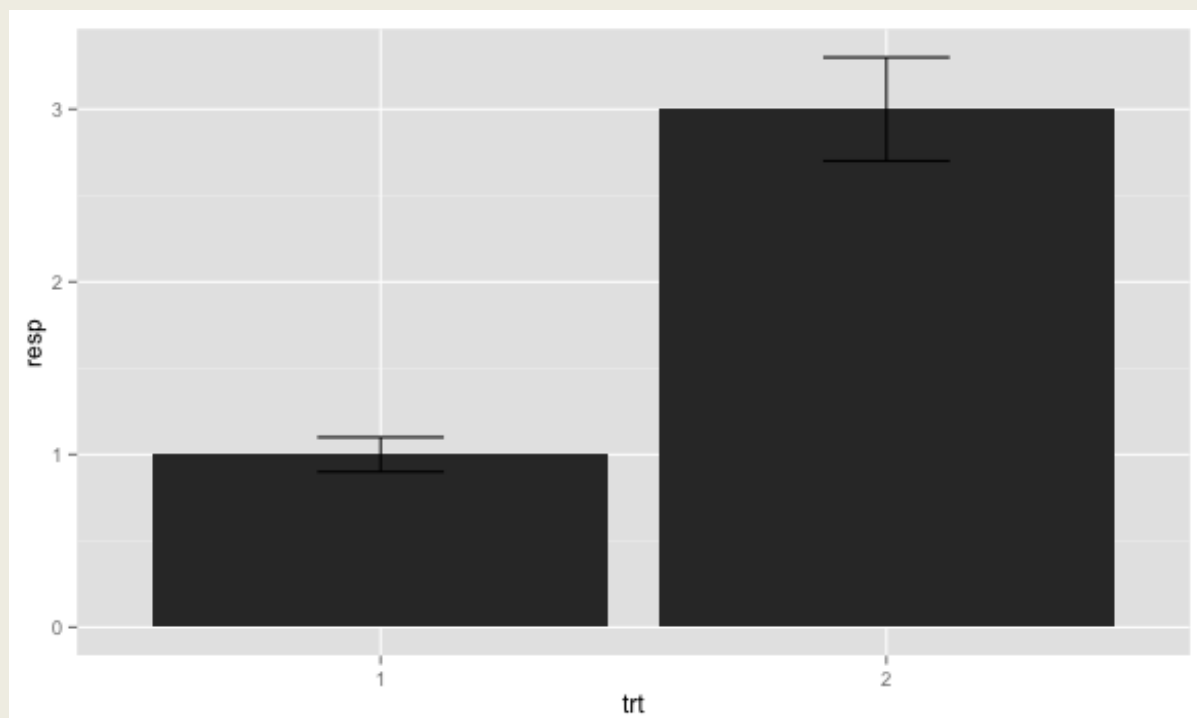Job title

Date

Murders per capita

Population

Likert scale

Histograms
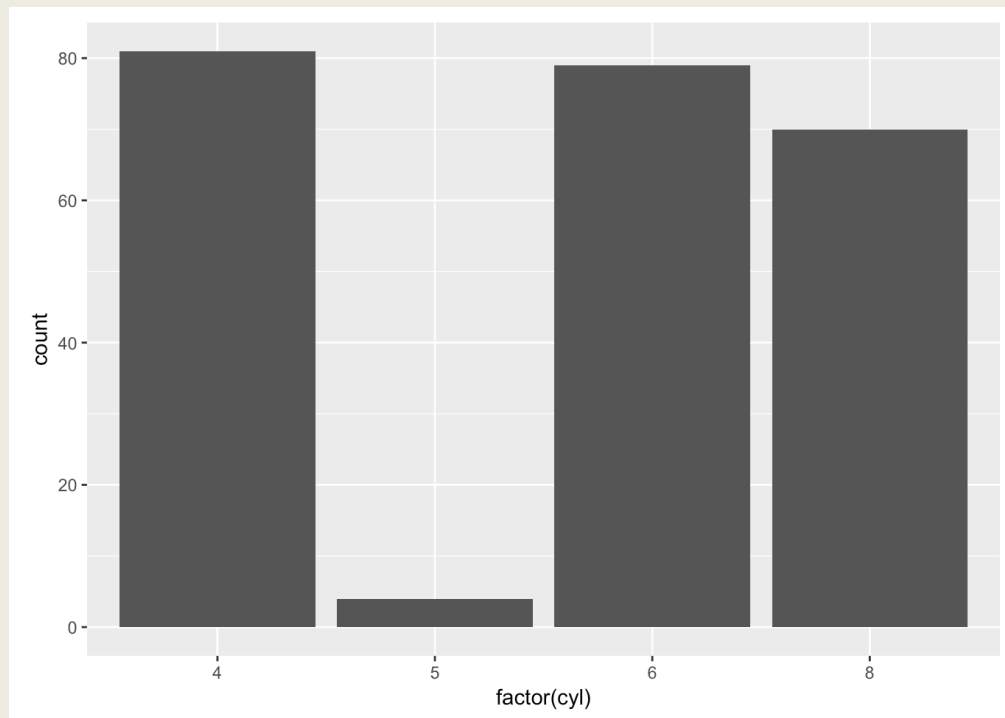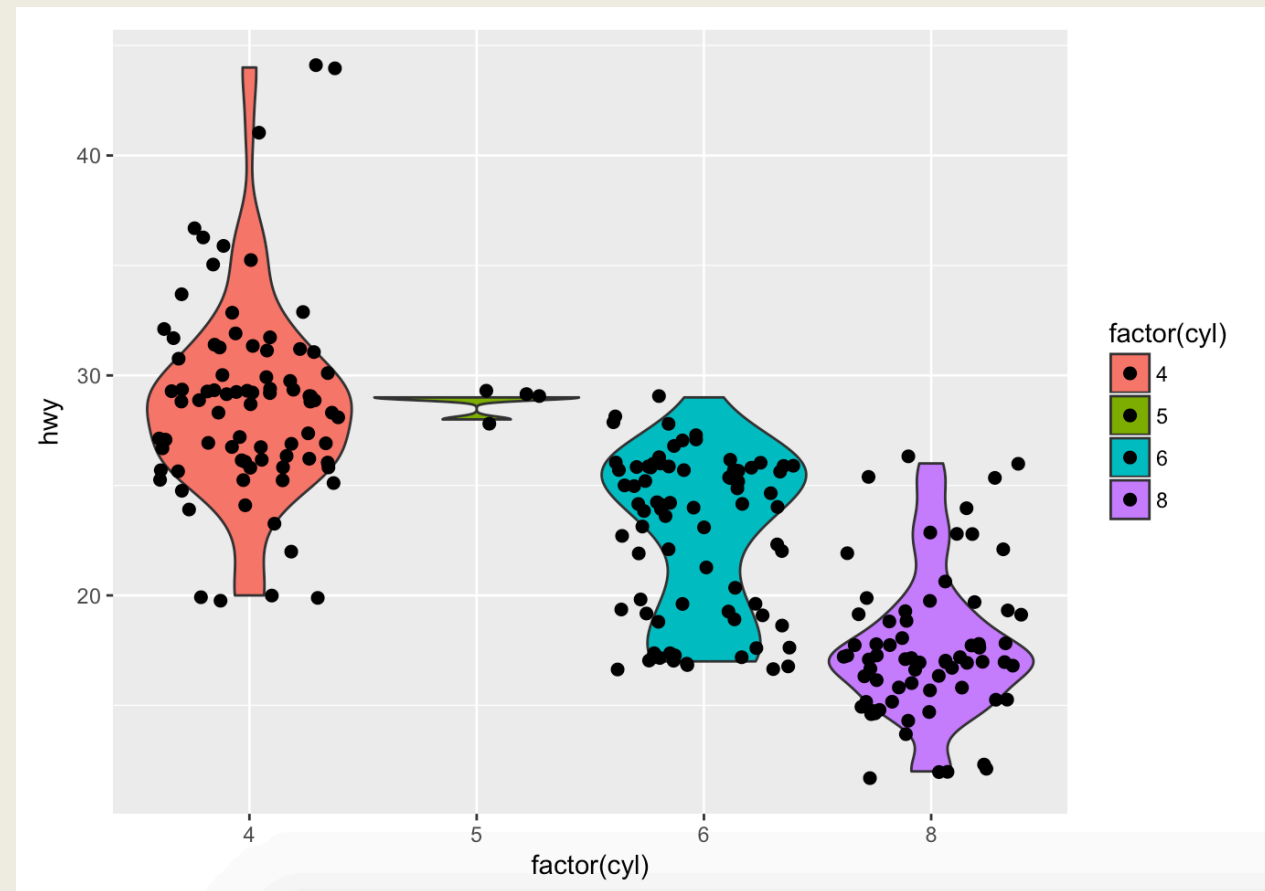
Density plots

Box plots
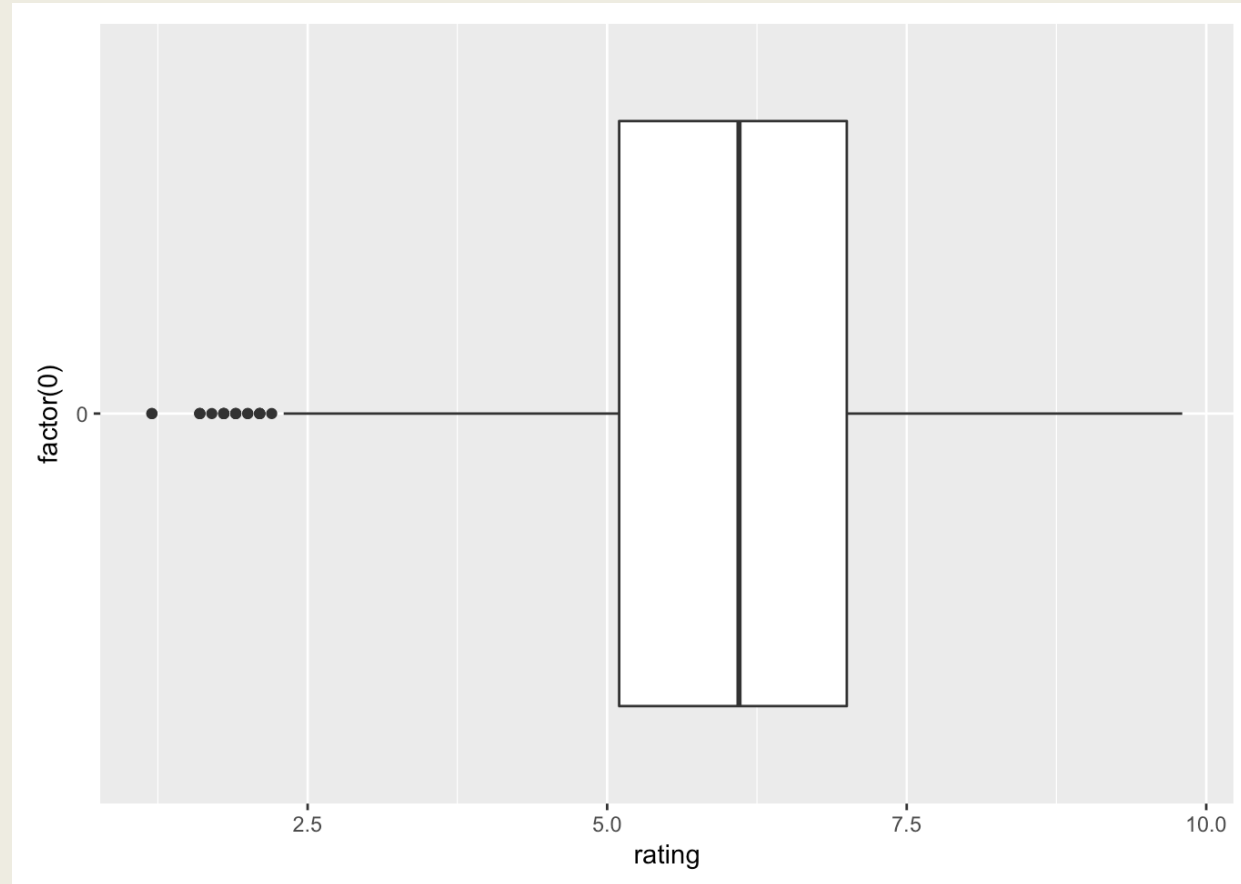
Bar charts

Violin charts

Ridge plots

DW-NOMINATE by party of U.S. House: 1963-2013

Distribution of 1st Dimension DW-NOMINATE by party
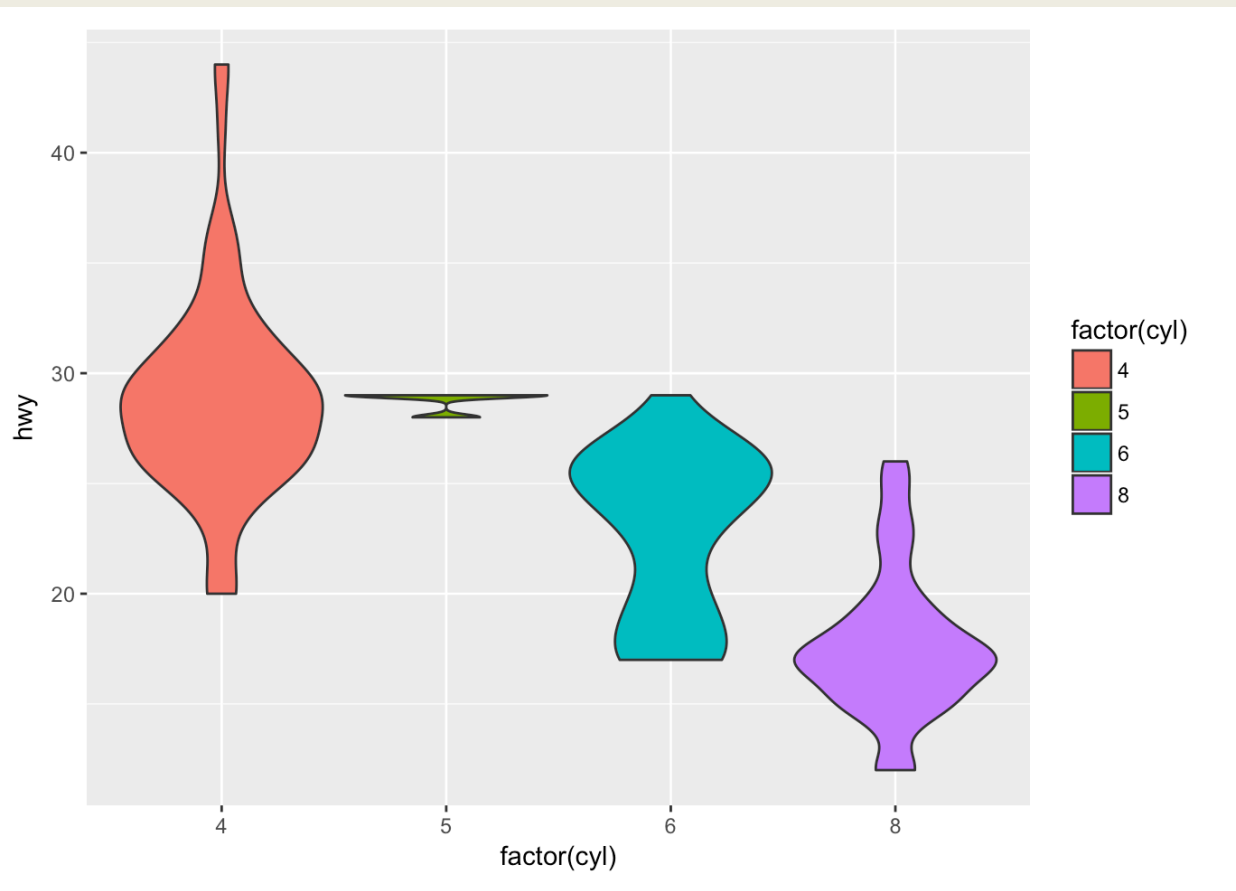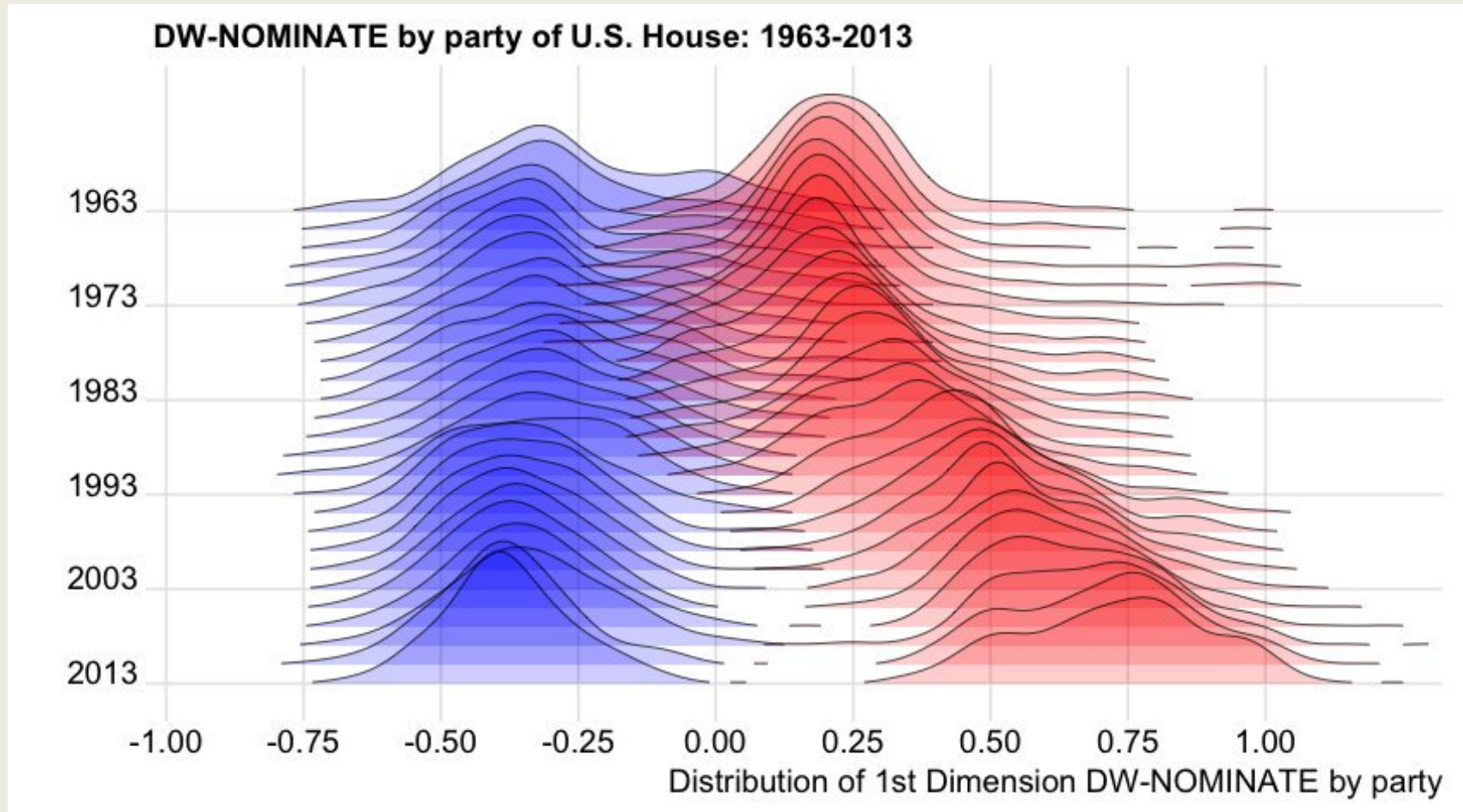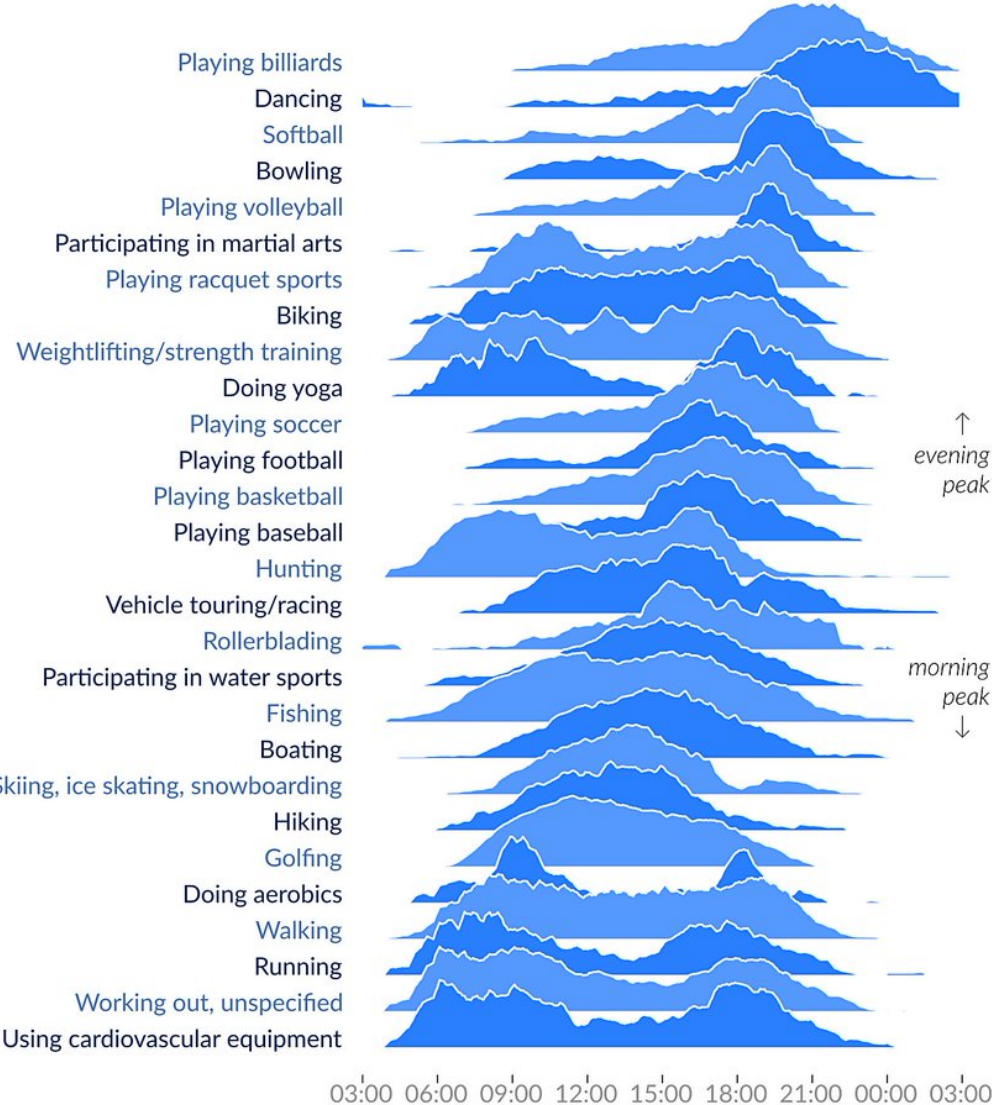
Peak time of day for sports and leisure

Number of participants throughout the day compared to peak popularity.
Note the morning-and-evening everyday workouts, the midday hobbies,
and the evenings/late nights out.

Playing billiards
Dancing
Softball
Bowling
Playing volleyball
Participating in martial arts
Playing racquet sports
Biking
Weightlifting/strength training
Doing yoga
Playing soccer
Playing football
Playing basketball
Playing baseball
Hunting
Vehicle touring/racing
Rollerblading
Participating in water sports
Fishing
Boating
Skiing, ice skating, snowboarding
Hiking
Golfing
Doing aerobics
Walking
Running
Working out, unspecified
Using cardiovascular equipment

↑
*evening
peak*

*morning
peak*
↓

03:00 06:00 09:00 12:00 15:00 18:00 21:00 00:00 03:00

Perceptions of Probability

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

Almost Certainly

Highly Likely

Very Good Chance

Probable

Likely

We Believe

Probably

Better Than Even

About Even

We Doubt

Improbable

Unlikely

Probably Not

Little Chance

Almost No Chance

Highly Unlikely

Chances Are Slight

Assigned Probability

created by /u/zonination

🎉 R time! 🎉