

How to choose metrics

Tressa L. Fowler





Overview

Asking the right questions

Disparate observation sources

Statistical and Practical significance

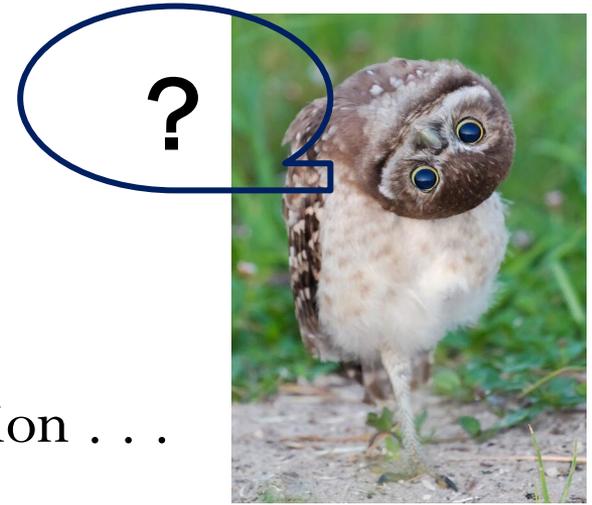
Sample size

Data types

Types of Statistics



Avoiding *Type III* Error

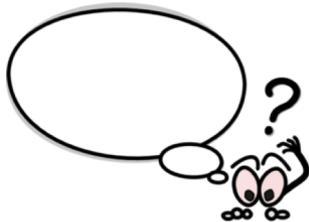


Finding the right answer for the wrong question . . .

What exactly are the **research questions** for different users?

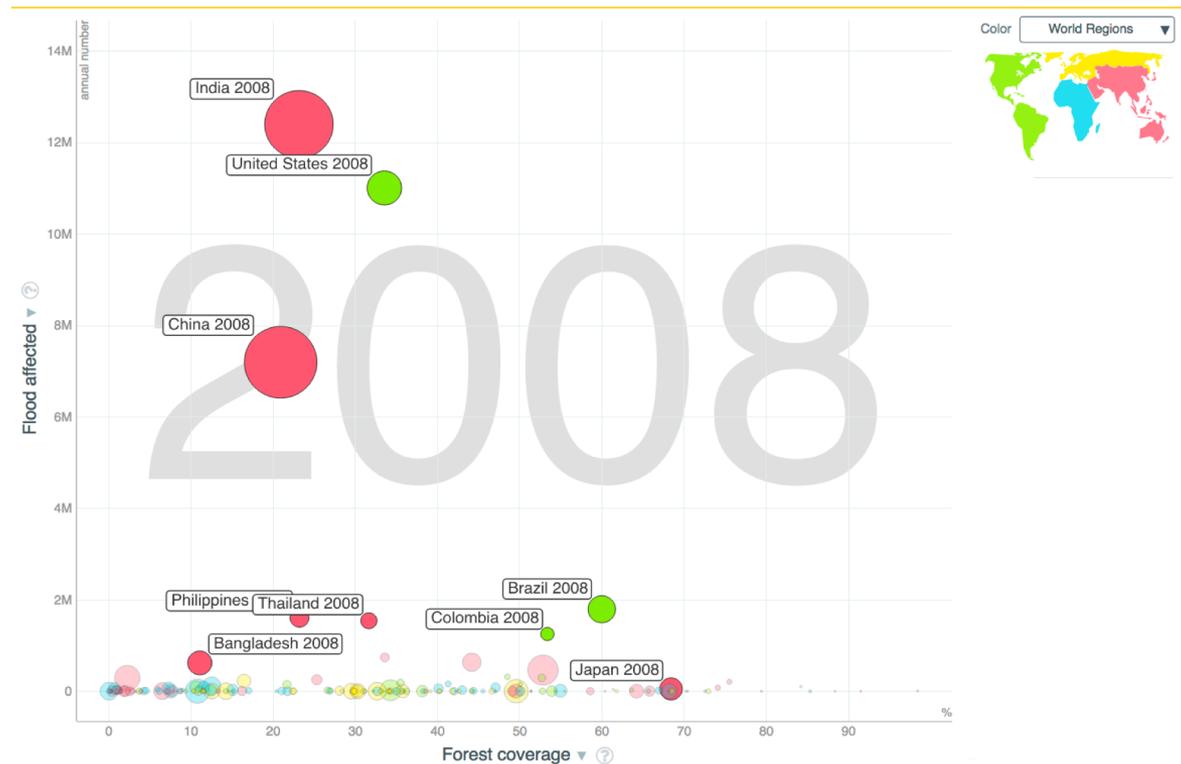
More precise questions are easier to answer:

- Which forecast is better?
- Which has the closest temperature forecast?
- Which identifies extreme temperature hazards earliest?



Inconsistent (world) Observations

- All observations are wrong, most are still useful.
- Possible Differences
 - Precision
 - Accuracy
 - Frequency
 - Timeliness
 - Drift
 - Quality Control
- Homogenous sample
 - Increased Power
 - More precise answer



Significance – Statistical and Practical

- **Statistical significance** is mathematical and sample-size centric.
- **Practical significance** arises out of the applicability of the result in decision making.
 - More subjective
 - Depends upon external factors
 - Cost
 - Time
 - Measurement error
 - Objectives
- Actionable levels are best defined ahead of time.



Samples



- Need an appropriately selected sample, with sufficient size.
- Appropriately selected
 - Representative
 - ~Homogenous

May be from specific regions, regimes, seasons, times of day, etc.
- Forecasts and observations are not independent in space or time, so our count of cases is an over-estimate.
- Go big – within considerations of time and cost.
 - Lots of lead times and / or lots of grid points \neq a large sample.
 - Many “cases” is a large sample.

Data Types



- Categorical Data
 - Nominal (mutually exclusive, not ordered)
 - Ordinal (order matters but not the difference between values)
- Numerical Data
 - Discrete
 - Continuous
 - Interval (differences are meaningful, e.g. Temp F or C)
 - Ratio (interval plus meaningful 0, e.g. Temp K).

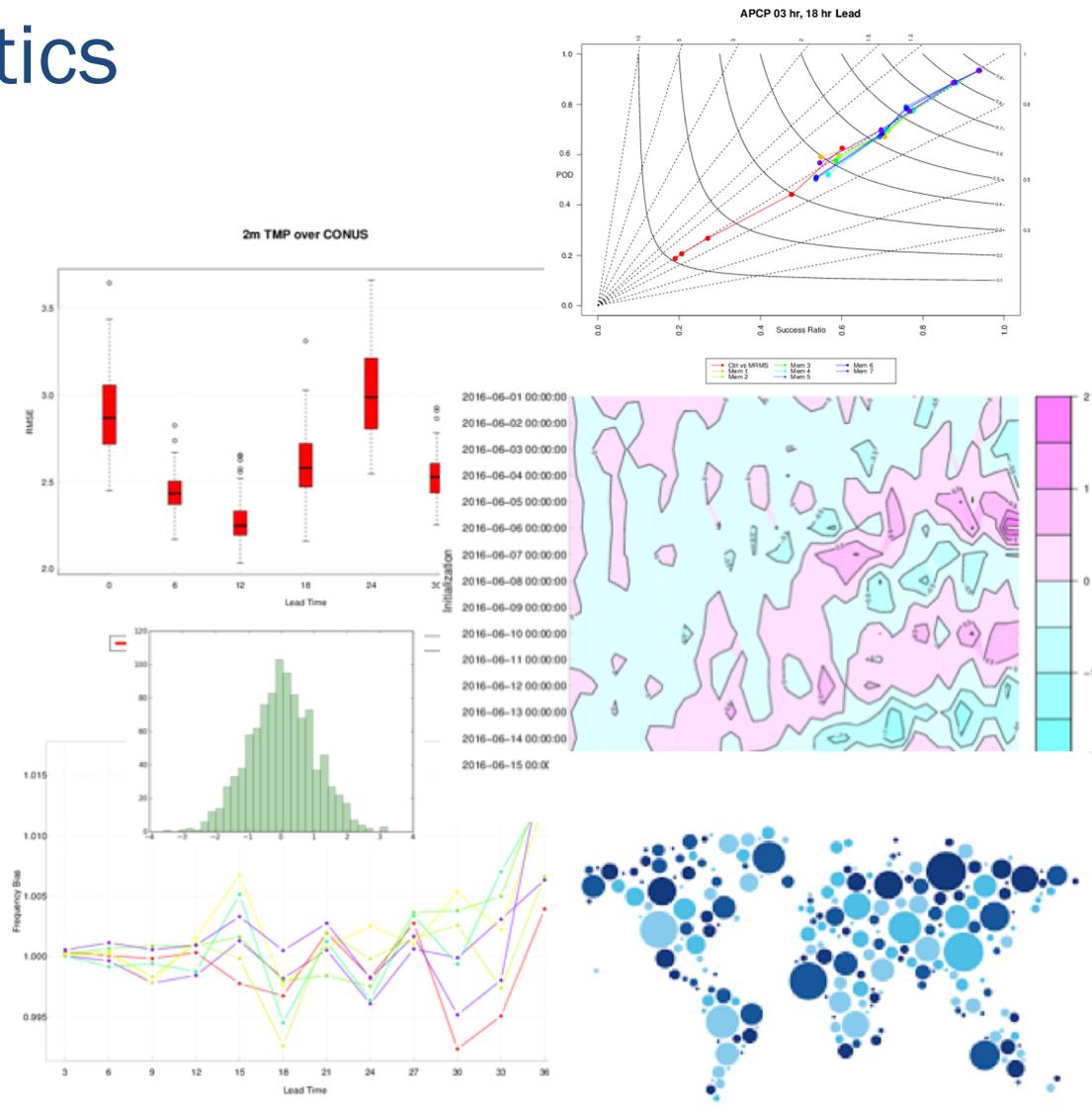
Having a good understanding of the different data types is crucial since you can use certain statistical measurements only for specific data types.

We have all learned to be careful with precip and reflectivity.



Choosing statistics

- Descriptive statistics
- Patterns
- Relationships
- Differences
- Classification
- Compare distributions
- Hypothesis test
- Point estimate
- Confidence interval
- Measure extremes
- Temporal changes



Extra Questions (dotting your i's and crossing your t's)

1. What **type of variables** do we have? Do forecasts and observations match?
2. What is the approximate **observation error**?
3. Should a **conditional** evaluation be considered?
4. How many **groups** (regions, seasons) are being studied or compared?
5. What **assumptions** can / should we have?
6. Do we have **power** to detect real differences?
7. Do we need error adjustments for **multiple testing**?



Guidelines

- Define questions and actionable outcomes.
- Identify, then check, data types.
- Consider extremes.
- Identify a set of statistics (and graphics).
- Define sample (representativeness and size).

