Sources of News Value

Proximity

Prominence (celebrities, public officials)

Magnitude (a car wreck in a small town is news; in NYC, it’s not)

Impact (how does it affect me?)

Process in data journalism:

1. Acquire the data
2. Clean the data
3. Analyze the data
4. Present it

Newsworthy data shows the unexpected (ex: the health inspection grades in which a disproportionate number of restaurants get a 90)

Numbers only have meaning in relation to other numbers

Scatterplots – usually only newsworthy if there’s a pattern (exception: birthdate-draft number chart) or an outlier

Ideas

Where Rhodes Scholars (or other interesting category of people) come from (map)

Interview the data:

* Ask what’s normal, then look for what’s unexpected

EXCEL

Functions:

Adding a column: Go to bottom of column, type =SUM(M2:M2354), enter

Sorting:

Click on the top left 1; click funnel, drop down arrows now appear in each column

Click on a column header, click A-Z button at top, descending, expand

Finding averages:

Go to bottom of column, =AVERAGE(M2:M2354)

Note: use medians, not averages when there is a wide range of numbers

Sort/add combos (ex: how much McCrory got from Durham residents)

=sumif(range, criteria, [sum\_range])

example” =sumif(E:E, “Durham”, [M:M])

=COUNTIF(M:M,”50”) …count how many donations in column M were $50

=countif(M:M, “>50) … count how many donations were greater than $50

Add $ before numbers if you don’t want the value in particular function cell to change when you add new data

Use MySQL, SQL lite or Overview to process data

RStudio for stat analyis