**Your involvement, phase by phase:**

**Beat development:**

* Encourage reporters to find their data “sources,” add them to their Rolodex, take them out for “coffee”, etc.
* Give them time to explore a dataset that is crucial to their beat; perhaps generate a minimum story from it

**Idea generation:**

* Questions you should be asking/Things to keep in mind:
  + Story idea should be framed as a question, not a noun. I.E. “I want to do a story about how many bridges are deficient” is better than “I want to do a story about unsafe bridges” because the question helps point you to the data that will help tell your story. It’s okay if there’s more than one question, especially in the beginning.
  + Tipoffs for you that maybe the reporter should consider their own data analysis:
    - The reporter has some summary numbers but there are many questions that can’t be answered with the summary numbers. For example: let’s say the idea is to look at whether some neighborhoods have had an uptick in crime recently, but the police department is only able to provide the change over time for the city as a whole. Or they can only show the change from last month to this month and you think there’s a bigger trend (a better story would be to compare a recent time period to the same months in previous years)
  + There’s a balancing act you need to consider in deciding whether to pursue data analysis for a story.
    - One factor is time – how quickly do you have to turn this around
    - Another is whether somebody else has already done the analysis that you want (i.e. a government agency or non-profit).
    - Another is how complicated the question is. For example, if the question is whether the percentage of students on free or reduced price lunch has gone up in recent years and where it’s gone up the most – that data is very easy to obtain and answering that question would only take a couple hours, at most. But if you want to prove that county workers are “spiking” their pensions by working excessive overtime in their last five years on the job, that data is harder to obtain and this question would require a more complicated analysis.
    - Another is how important the story is.
  + If you’re going to launch into an analysis that is even a little bit complex, always, always make sure the reporter spends time looking for examples of similar analyses that could serve as a template – i.e. other news organizations, non-profits, academics, etc. There’s a very good chance somebody has done something in the same vein and you could learn from their work, rather than reinventing the wheel. A good place to start is IRE’s Extra! Extra! Website (publicly accessible) or the story database in the Resource Center (requires IRE membership). Also check the IRE tipsheets for guidance on the topic (also requires IRE membership)

**Requesting data:**

* Ask how readily accessible the data is. If the data will require a public records request, be prepared for it to take longer than you might expect.
* Best practices for making a request (i.e. wording of request letter, getting record layouts, documentation, doing some “reporting” about the data in advance)
* Dealing with denials and excessive cost

**Reporting/Data analysis:**

* Ask about the quality of the data. Does it need to be “cleaned.” If so, build in extra time for that cleaning.
* As the reporter starts the analysis, ask some questions about the data to get a feel for how well the reporter understands the data. If he/she can’t answer your questions, encourage them to seek out someone who knows the data and can answer those questions. Some potential questions to ask: what does each record represent? What’s the time frame of the data – is it just one year or multiple years? What are the key questions you’re planning to ask the data? What’s your working theory about what you’re going to find? Are there any fields that you don’t understand?
* Encourage the reporter to keep a “trail” of their data cleaning and analysis. This trail should be a form of documentation that you would be able to understand and would enable an outsider to recreate the analysis on their own
* Sit down with the reporter and have them show you the data and the analysis they’ve done. You may want to repeat this multiple times during the process (if it’s a larger project)
* Be sure the reporter repeats the analysis, start to finish, to ensure accuracy
* If something seems too good to be true, it probably is
* The reporter should attempt to prove the key finding wrong before going ahead with it
* Don’t wait until the last minute to show key sources the findings; if fact, that should be the first step in the reporting process. Often the sources know the data better than you do and they might point out something that you missed or didn’t understand.
* Get digital and graphics people into the loop to discuss data visualization and/or other digital options as soon as you have the analysis defined. It doesn’t need to be perfected, but just so that you don’t have any major lingering questions or problems.

**Writing/Editing:**

* Numeracy—“star number”; use graphics to share the details; avoid too many numbers in same sentence or paragraph; you truly don’t need decimal points (almost never)
* Your goal should be to have a story that is supported by data but is not solely focused on the data (with some very rare exceptions). Data will either be the “spine” of the story (if that’s what’s driving the content) or it will be used in smaller bits to support key parts of the story.
* Do the math
* Don’t be afraid to consult outside experts
* As assignment editor, you should review any digital content, just like you’d review the story. Make sure that it makes sense, that it complements the story and consider how friendly it is for the reader (usability is particularly important if it’s an interactive element)
* Make sure that it also gets reviewed by a copy editor just before publication.

**What options do we have for data-driven digital?**

* Visualizations – graphics, maps
* News apps – allows readers to explore data, play a game, etc.

**What skills do reporters need?**

Minimum:

* Have a “data state of mind”
* Comfortable in spreadsheets--basic formulas & functions, Pivot Tables, filtering, importing
* Numeracy—be comfortable using numbers in stories, doing basic math, understand simple statistical concepts

Better:

* Be able to make some simple graphics for internal use and possibly to go online
* Be able to pull data out of a PDF
* Higher level of skill in spreadsheets

More advanced:

* Make data visualizations or simple interactive maps
* Analyze data in a database manager (such as Access or SQLite)
* Be able to scrape data from a web page

**What software should they learn?**

Minimum:

* Excel or Google spreadsheets
* Simple online charting software like Google Charts, Google Fusion Tables or DataWrapper
* Tabula, CometDocs or other software for pulling data out of PDFs

Better:

* Visualization software--Tableau Public (free)
* Database managers --Microsoft Access ($) or SQLite (free)
* Online mapping -- Google Fusion Tables or CartoDB or Tableau Public (all free)
* Cleaning data -- OpenRefine (free)

Advanced:

* Web scraping (there are various tools for this, usually free)
* GIS software, such as ArcMap ($) or QGIS (free)
* Statistical software, such as SPSS ($) or R (free)