

Notes by Kosisochi Chuma-Ibe

28 January 2019

### Warm up Activity (Critique of Harvey's energy toll – Axios visualization)

- Atmospheric data, weather data, sensor data as well as a database of refinery production capacities were likely used to create this visualization
- Shows the path of the hurricane and how rainfall affects or is going to affect oil production in the displayed regions
- Overlap of the colors used is a bit confusing
- It is not easy to infer what the circle sizes mean and why there is overlap of the circles in some regions

### Files, Data, and Organization

- Text
- Binary: These are not human-readable
- Query based

### Organization

- Most data will be in array format. You can either organize by rows or by column. The way you organize your data affects your data processing and visualization

### CSV

- also called delimiter separated values
- You can easily define how your CSV file is going to work.
- You can't store nulls or arrays in a CSV file

### Binary

- Zeros and ones. When a switch is turned from zero(off) to one(on), it creates a new number
- 255 is the maximum decimal value that can be represented with 8 bits on a computer

Float[decimal points] vs string[words] vs integer[whole numbers]

### JSON

- basically XML for data storage

### HDF5

- Really popular with computational scientists.
- Flexible and very efficient
- Separated by columns

- Binary
- There are different libraries for HDF5 reading e.g. astrophysics, medical, geophysics.
- HDF5 was invented at UIUC

Nyquist limit : you want to sample the data twice as often as the highest frequency of the data