Week 4: Concepts in Viz

Spring 2017

Matthew Turk

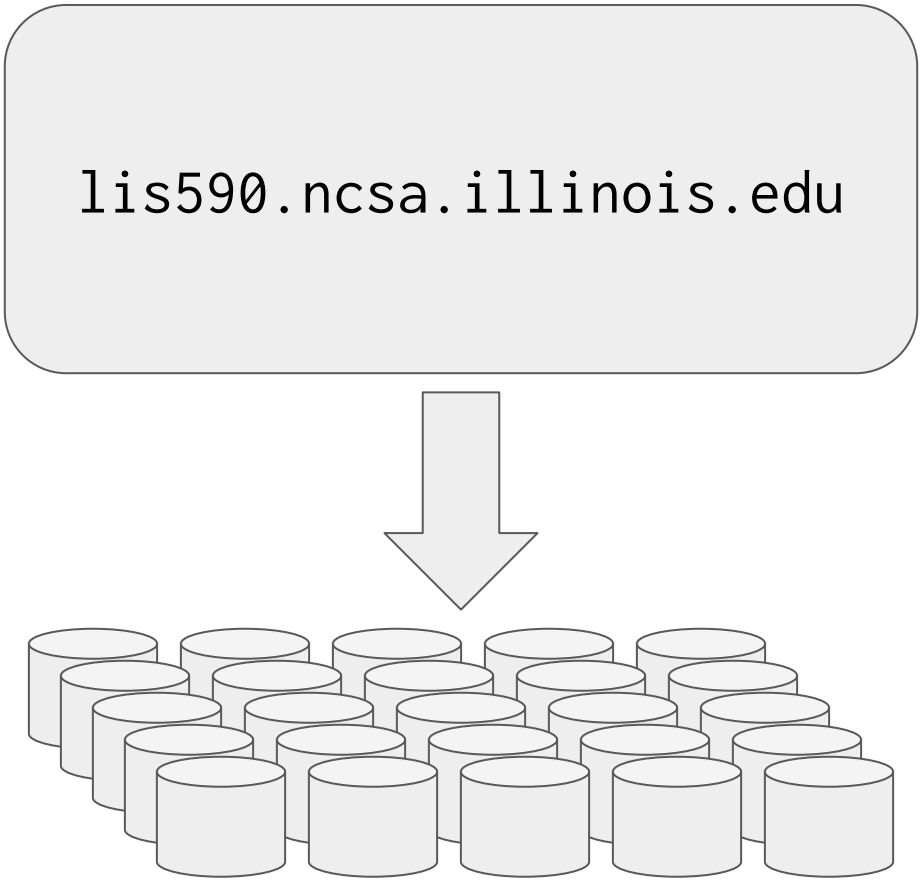
Infrastructure Review

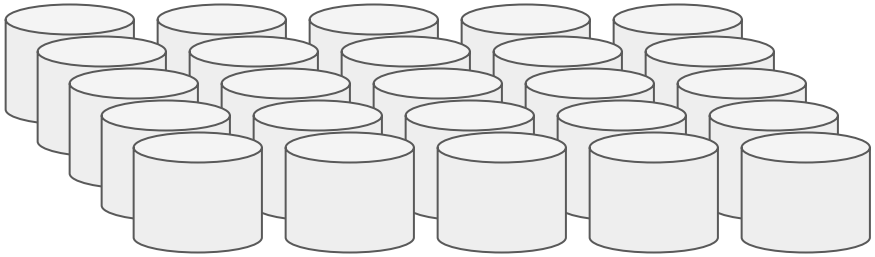
- JupyterHub & nbgrader
- Slack
- GitHub
- Broadcasting

Broadcasting

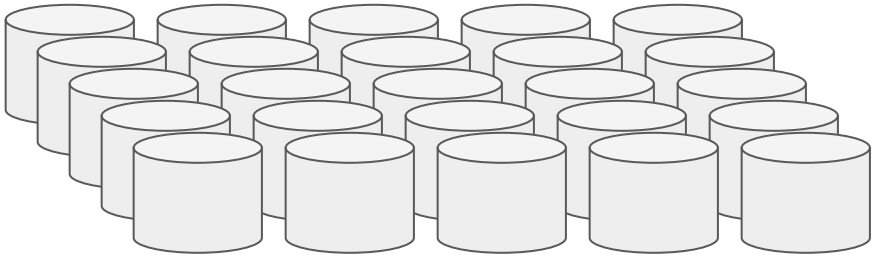
- These lectures will be simulcast, in case you can't see
- We are "Meeting Room 2" for Blackboard
- They will not be recorded

go.ischool.illinois.edu/meet2

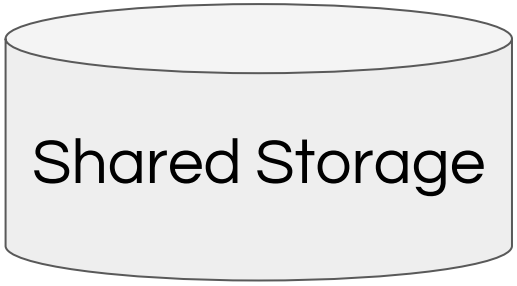

lis590.ncsa.illinois.edu
Jupyter Instances



Jupyter Images



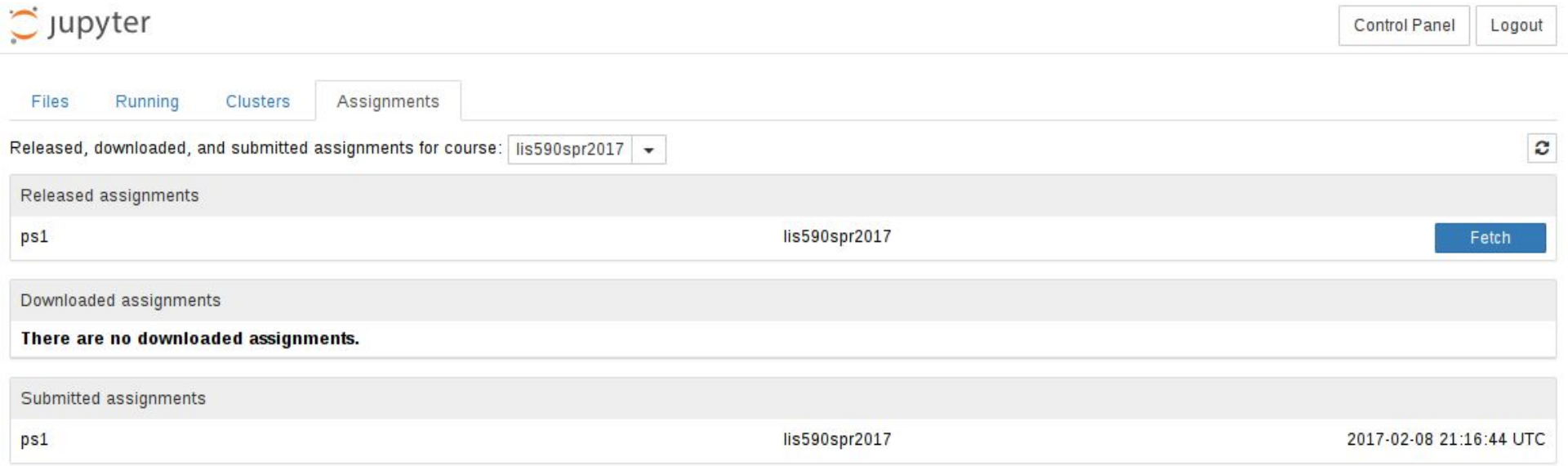
Personal Storage



- Store your notebooks on- and off-site
- Submissions will be via nbgrader
- Data will be available at `/data/` and `/home/$USER/work/data-readonly/`
- Previous lectures will be in `data-readonly/spring2017`
- You will have access to conda, etc, but I may rebuild images to add packages.

Assignment Flow

1. Instructor “releases” an assignment
2. Assignment appears in student “Assignments” tab
3. Students “fetch” assignment, which *copies* it to their work directory
4. Assignments will be notebooks, accessing shared read-only data
5. Upon completion, students “submit” assignment, which *copies* it to the instructor’s inbox
6. Some cells may be “autograded” but feedback will be provided.
7. Once feedback is available, you will be notified.



Slack

- Team is at lis590dv-spr2017.slack.com
  - #general : General announcements
  - #assignments : Help with assignments
  - #help : General help with Python, Javascript, visualization, etc
  - #lectures : During lectures, post links, comments, questions here
- Use the @ sign appropriately: @[person], @here, @channel
- Conduct will be held to same standards as any educational venue.
- Web client, standalone client and mobile devices can access this team.
- At the end of the semester, the team will be discontinued.

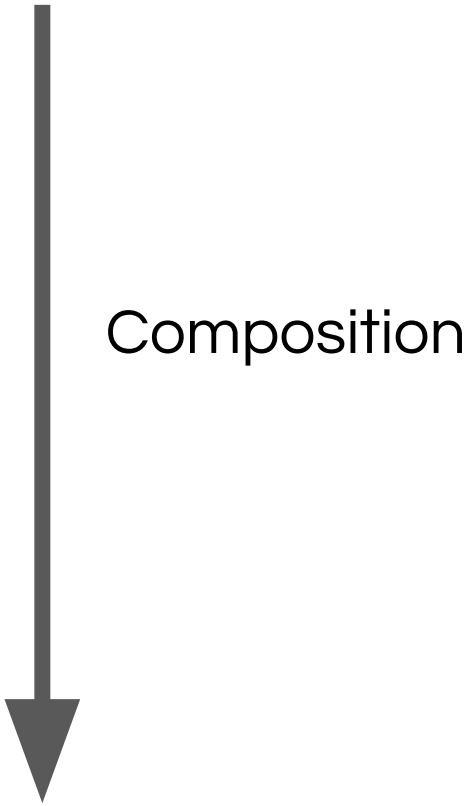
GitHub

- Reminder: http://github.com/UIUC-iSchool-DataViz/spring2017/
- Lecture notes will be placed there, and available in your JupyterHub instances in data-readonly/spring2017/weekXX .
- Copy the notebooks to your directory before using them.
- If you have not already done so, fill out the form giving me your GitHub usernames.

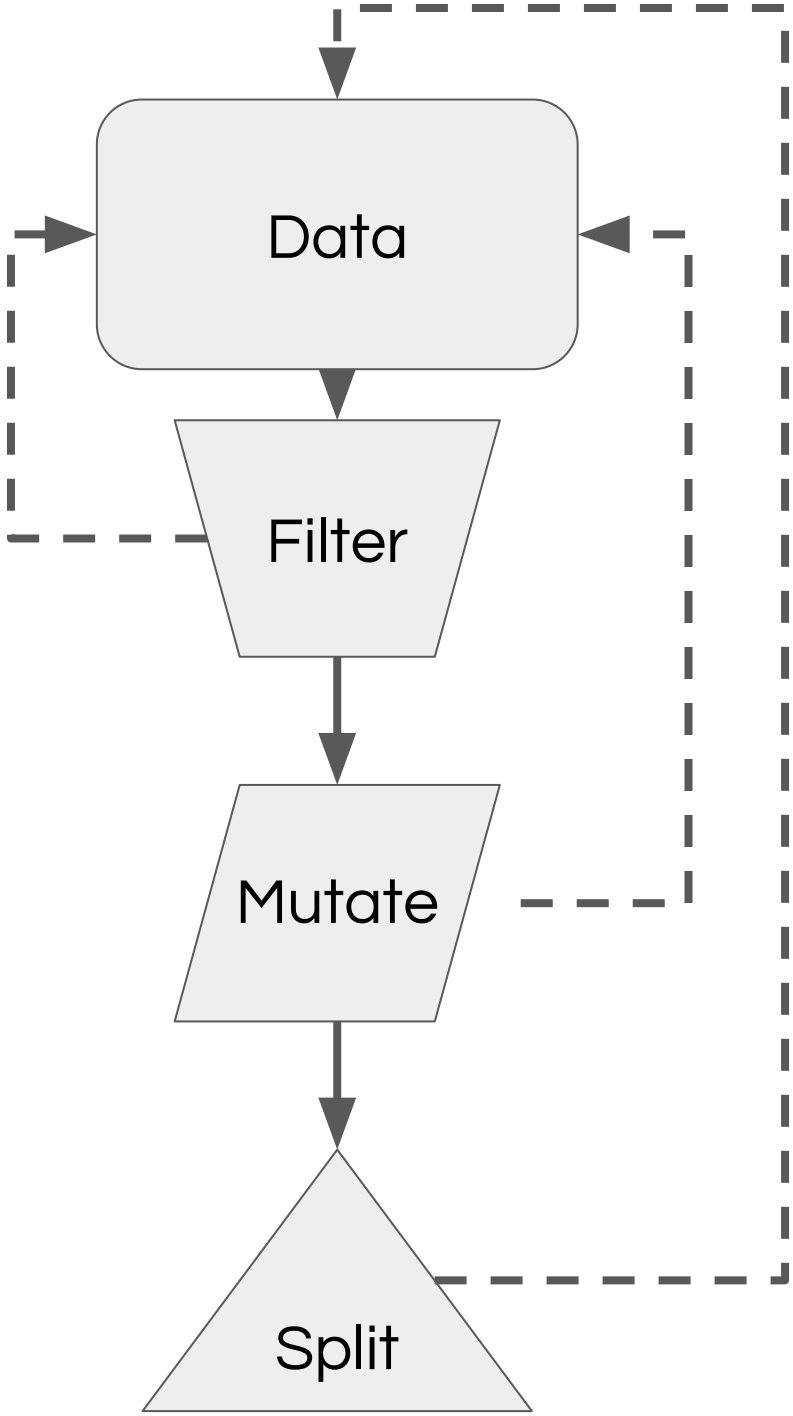
goo.gl/xCJoD9

Concepts of Visualization

Concepts of Visualization



Concepts of Visualization

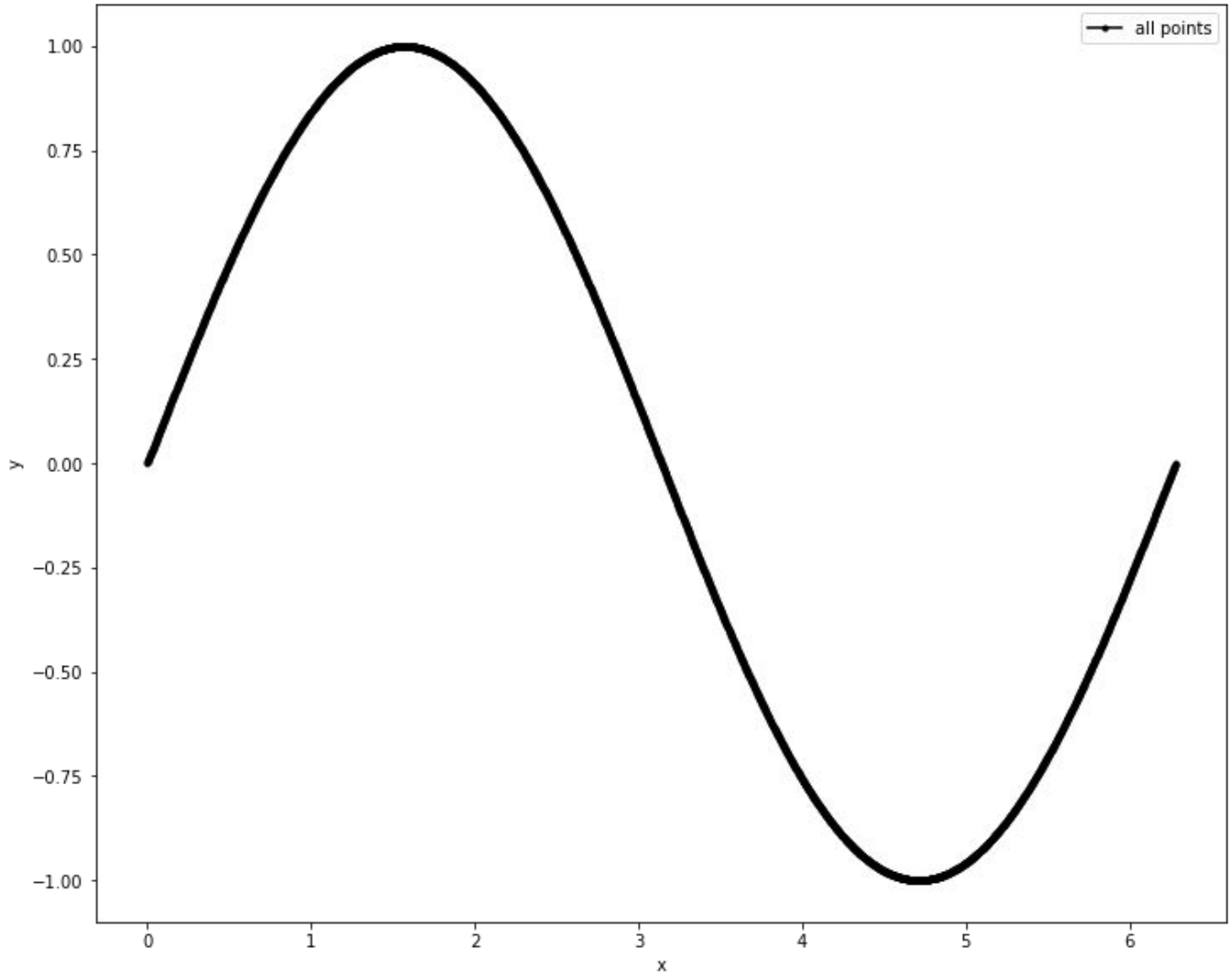


You have a palette of operations to apply.

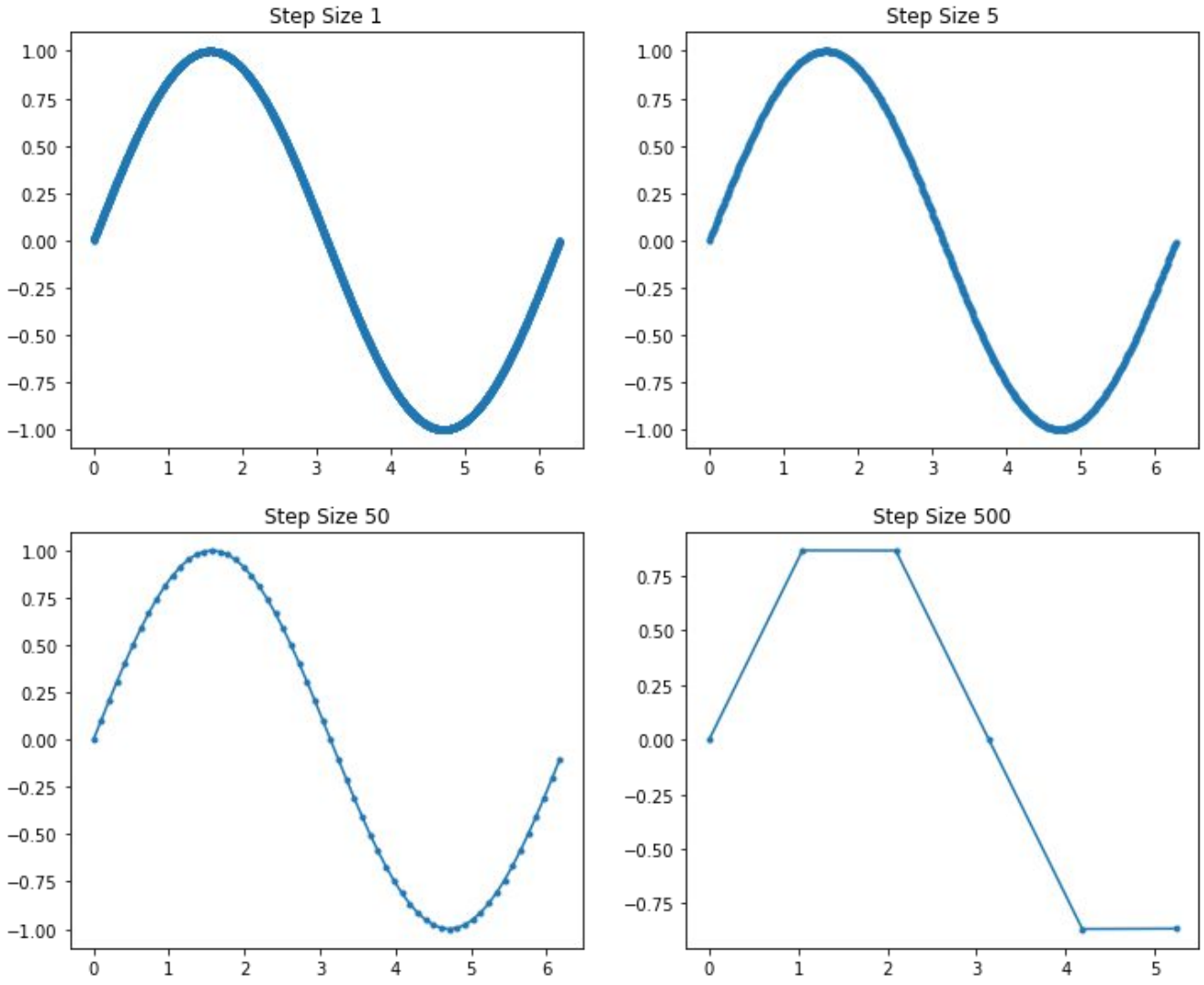
Filtering operations

- Relationships:
  - Equality, inequality
  - Quantitative value (less than, greater than)
  - Intersection, disjoint
- Subsampling
  - Regular sampling
  - Randomized sampling
  - Nyquist frequency
- Related data queries
  - Queries on other columns at fixed row location
  - External membership queries

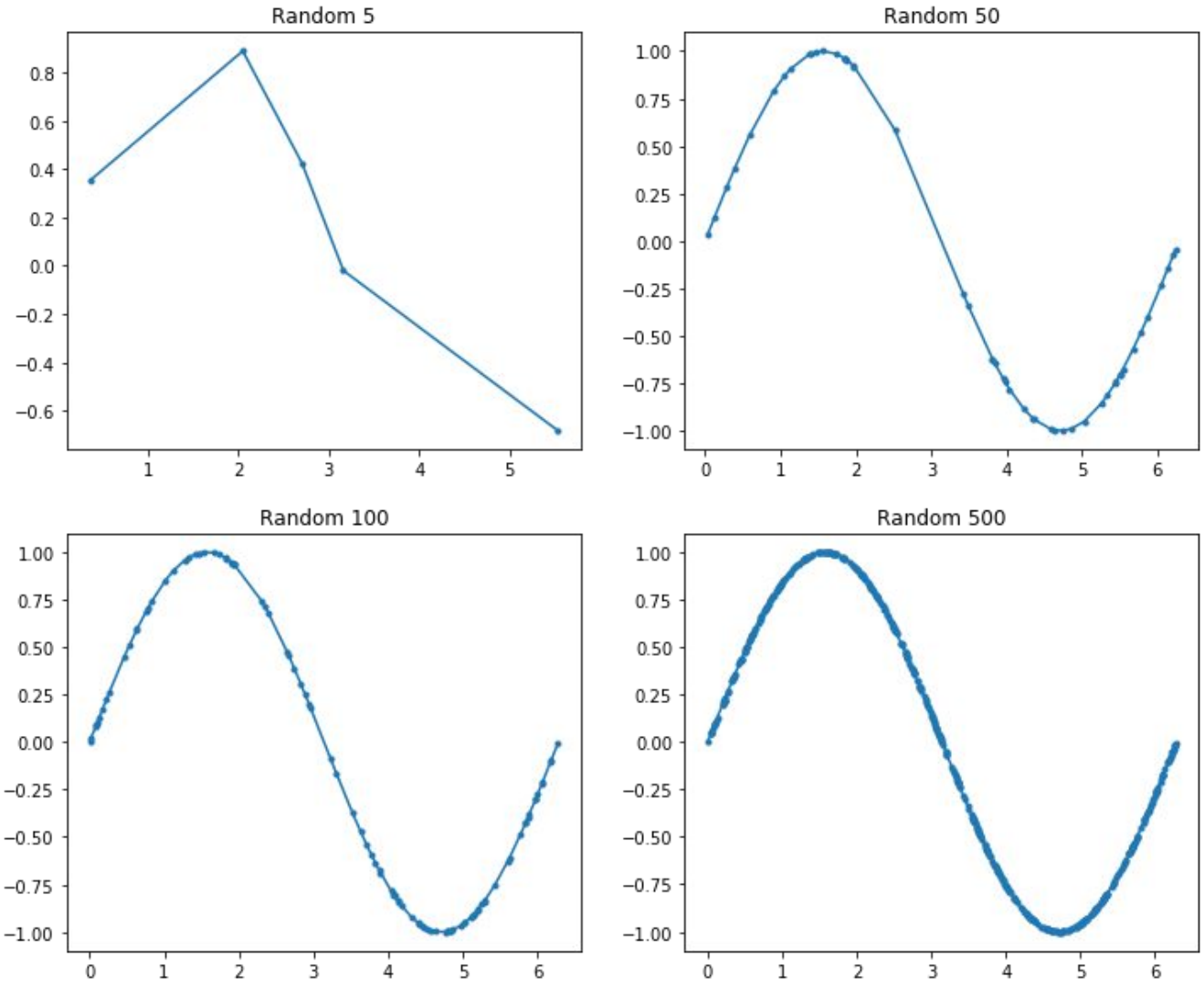
Sampling Examples



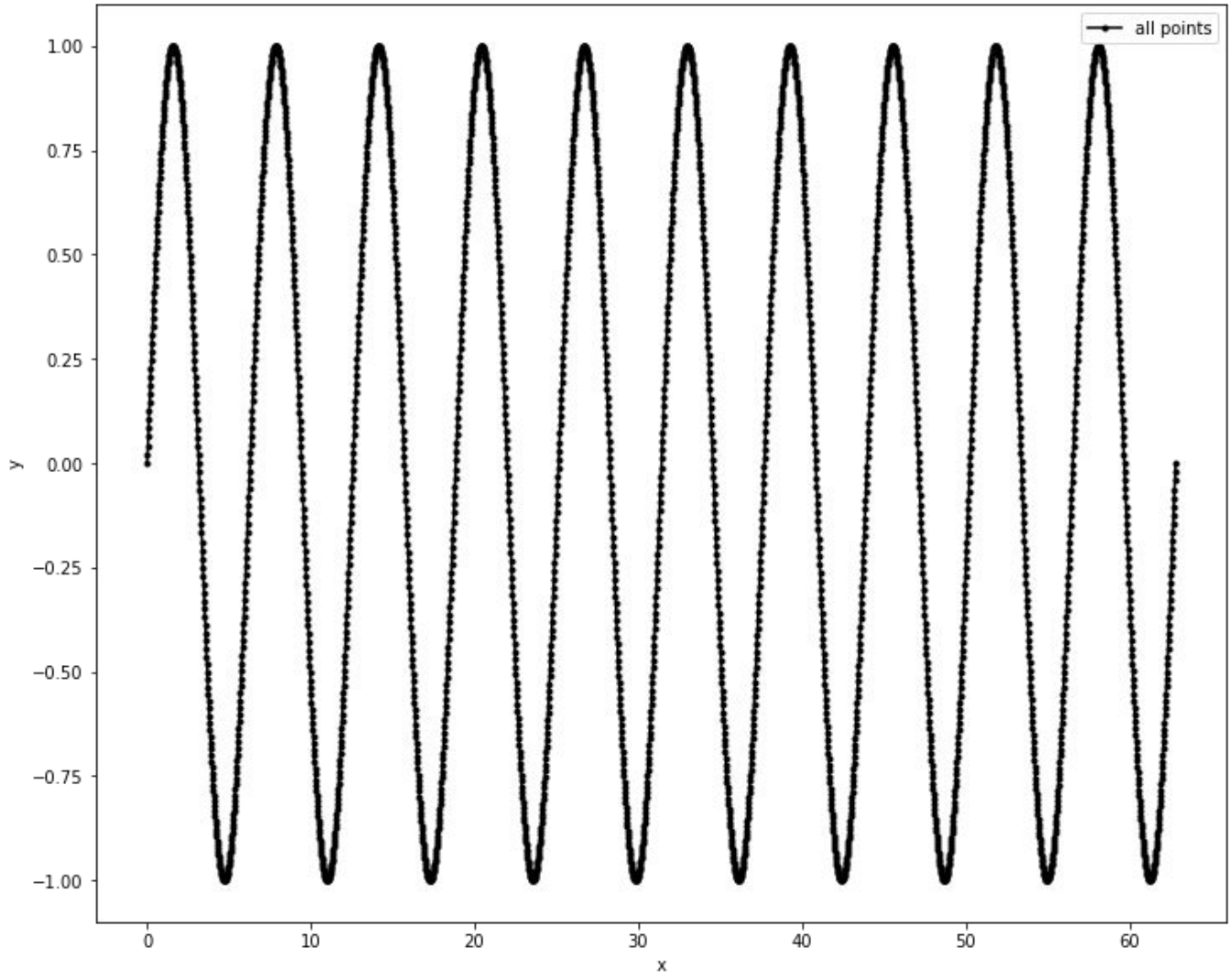
Sampling Examples



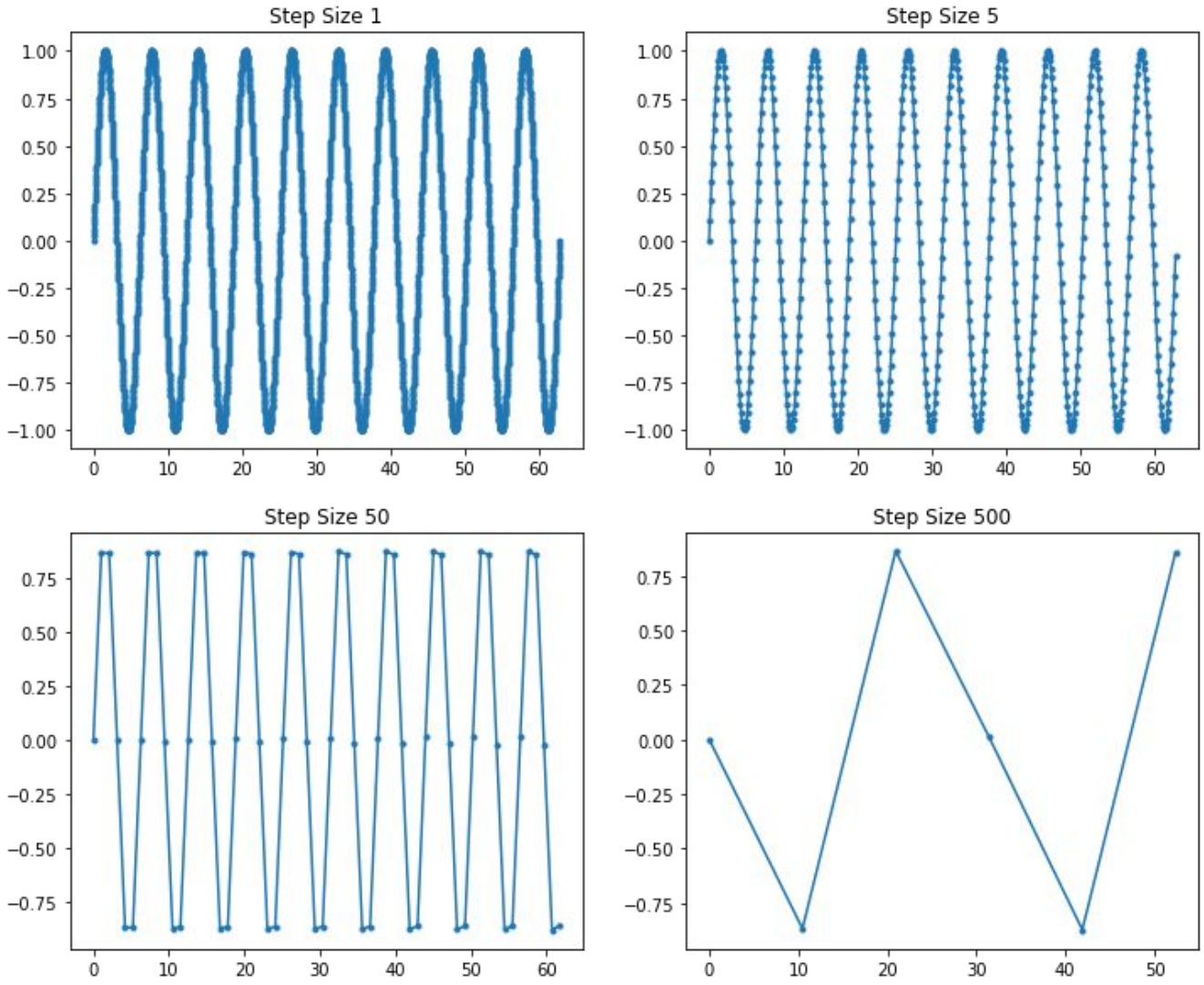
Sampling Examples



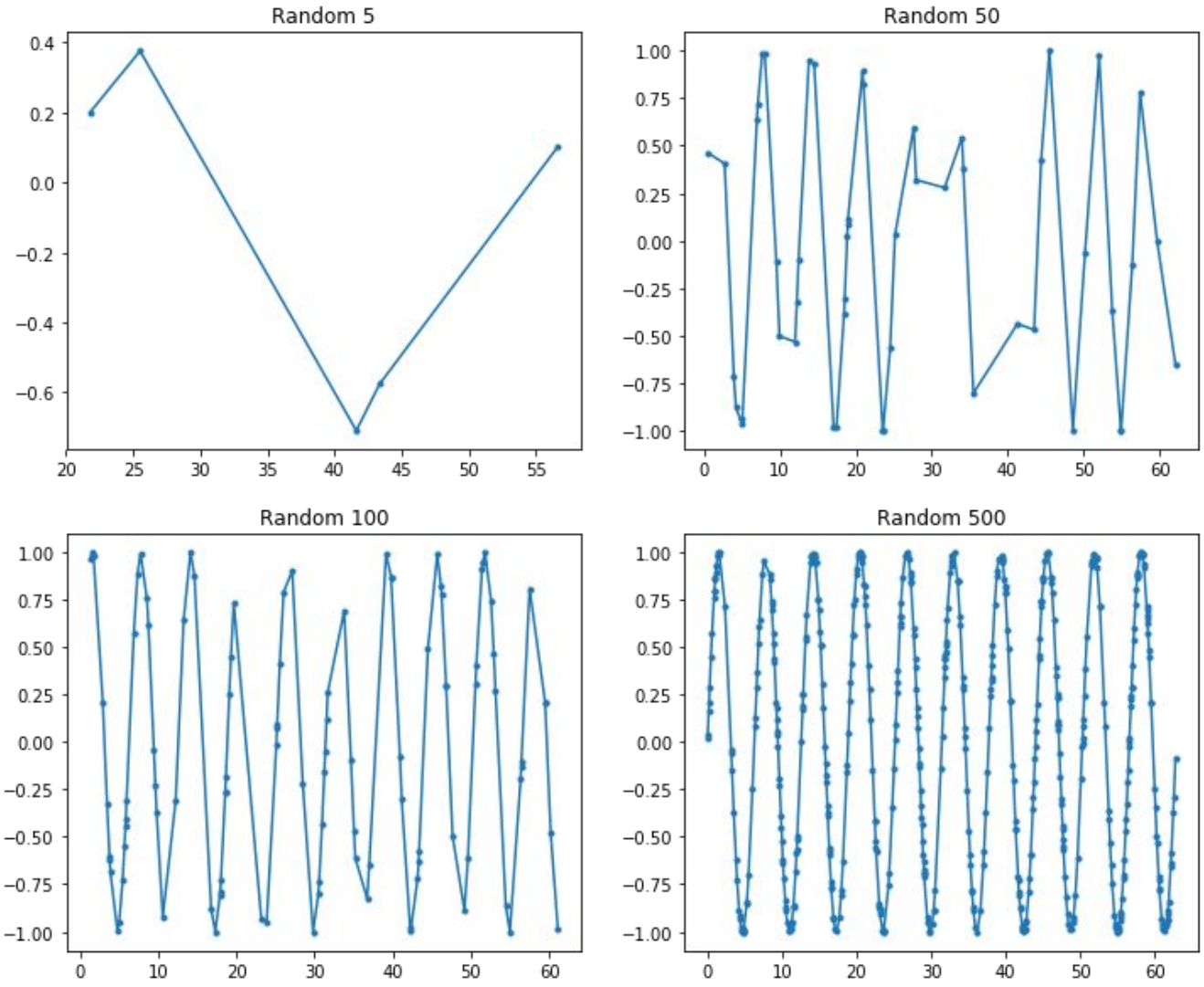
Sampling Examples



Sampling Examples



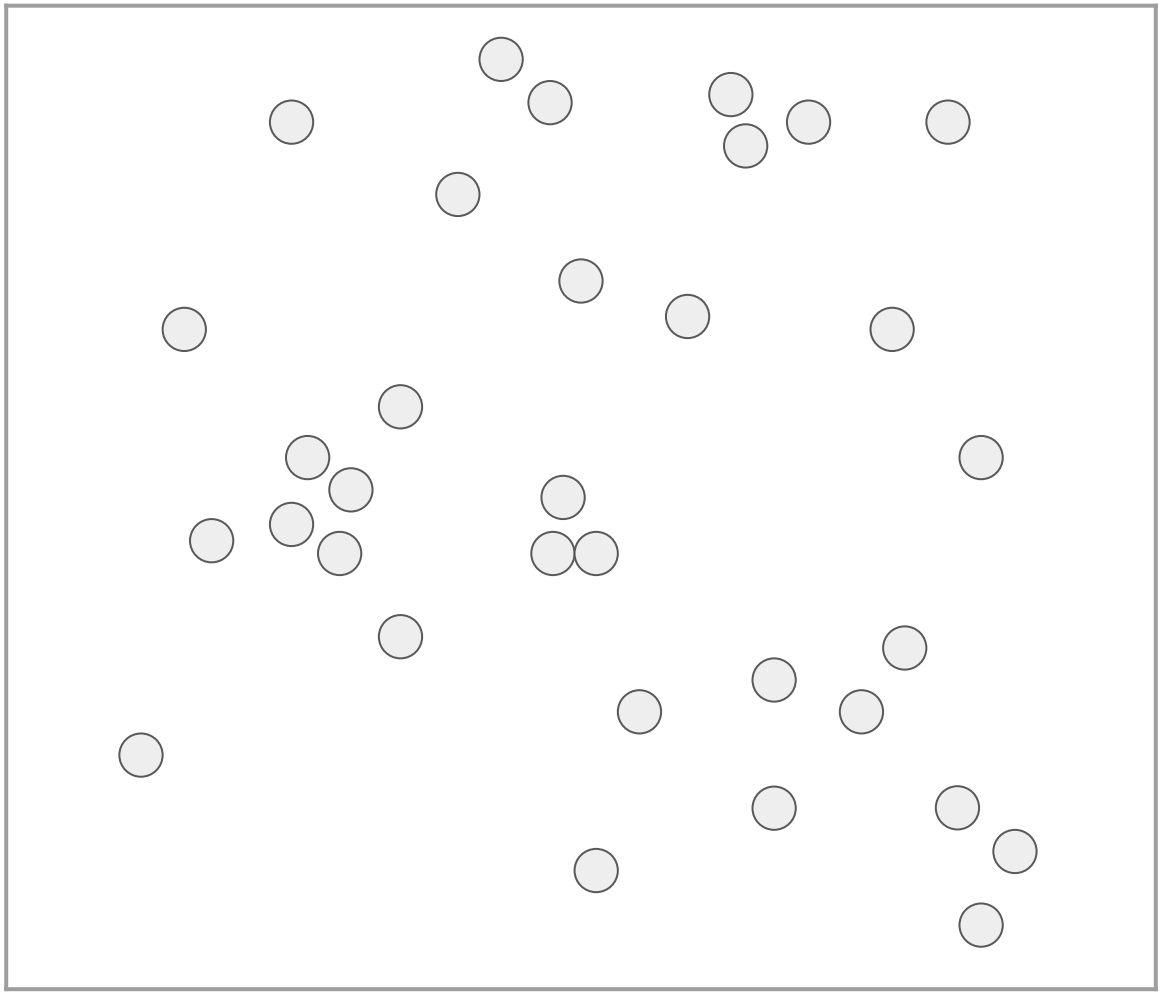
Sampling Examples



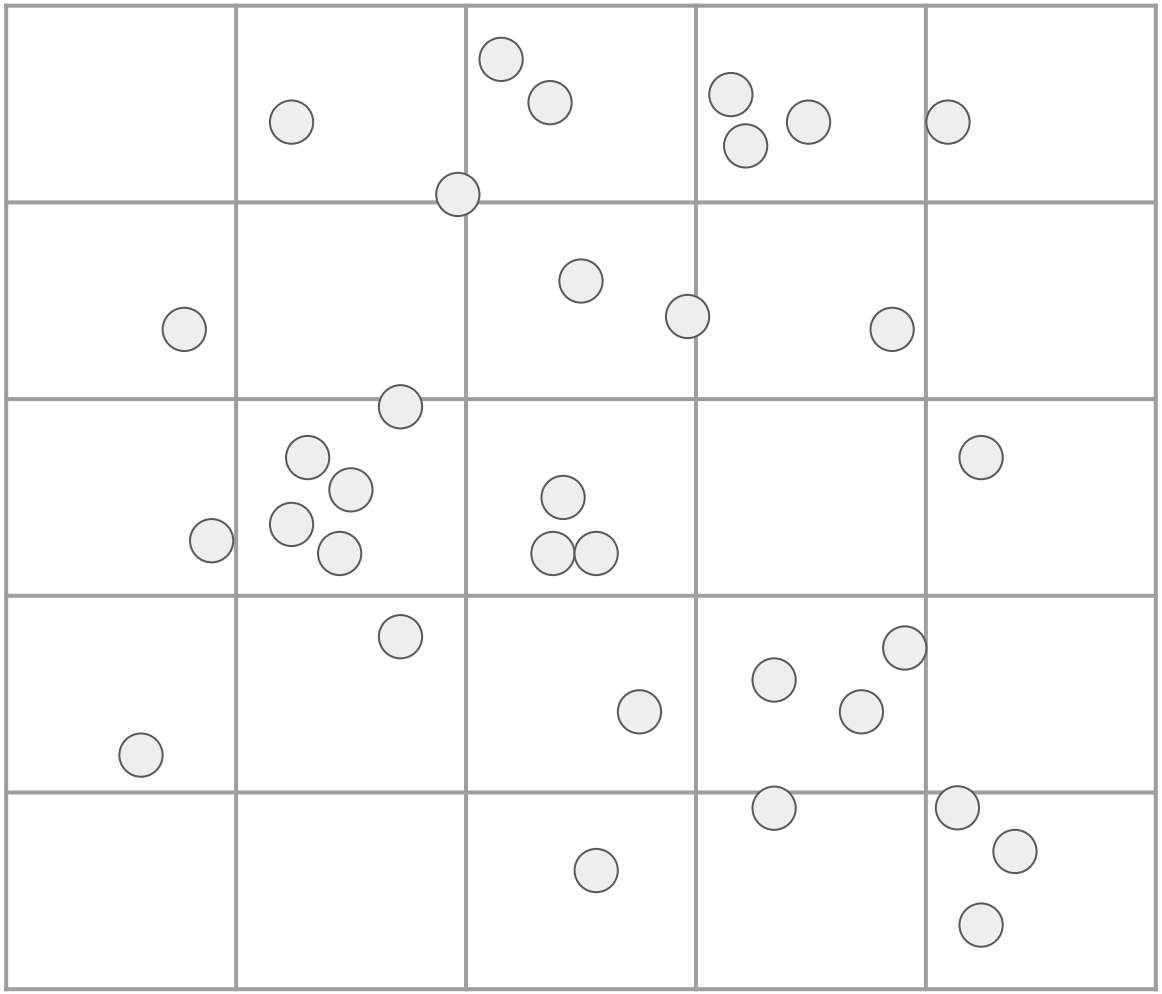
Mutation Operations

- Mathematical operations, such as injective operations.
  - Logarithmic versus linear representations
  - Arithmetic or multiplicative relationships
  - Manifold remapping
- Smoothing (reduction; not injective)
- Histograms (reduction; not injective)

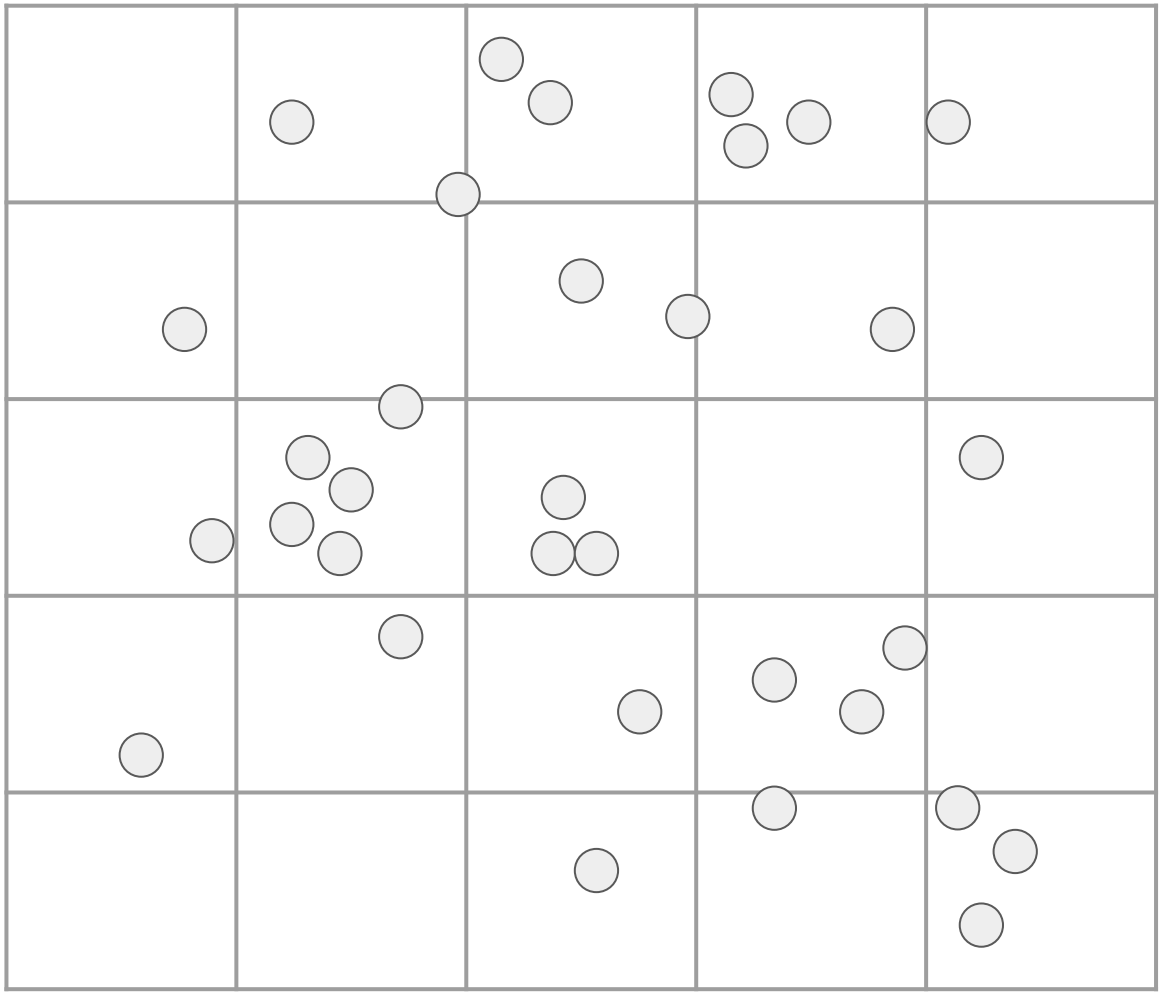
Binning and histograms



Binning and histograms



Binning and histograms



- Counts

  sum(1)
- Sum

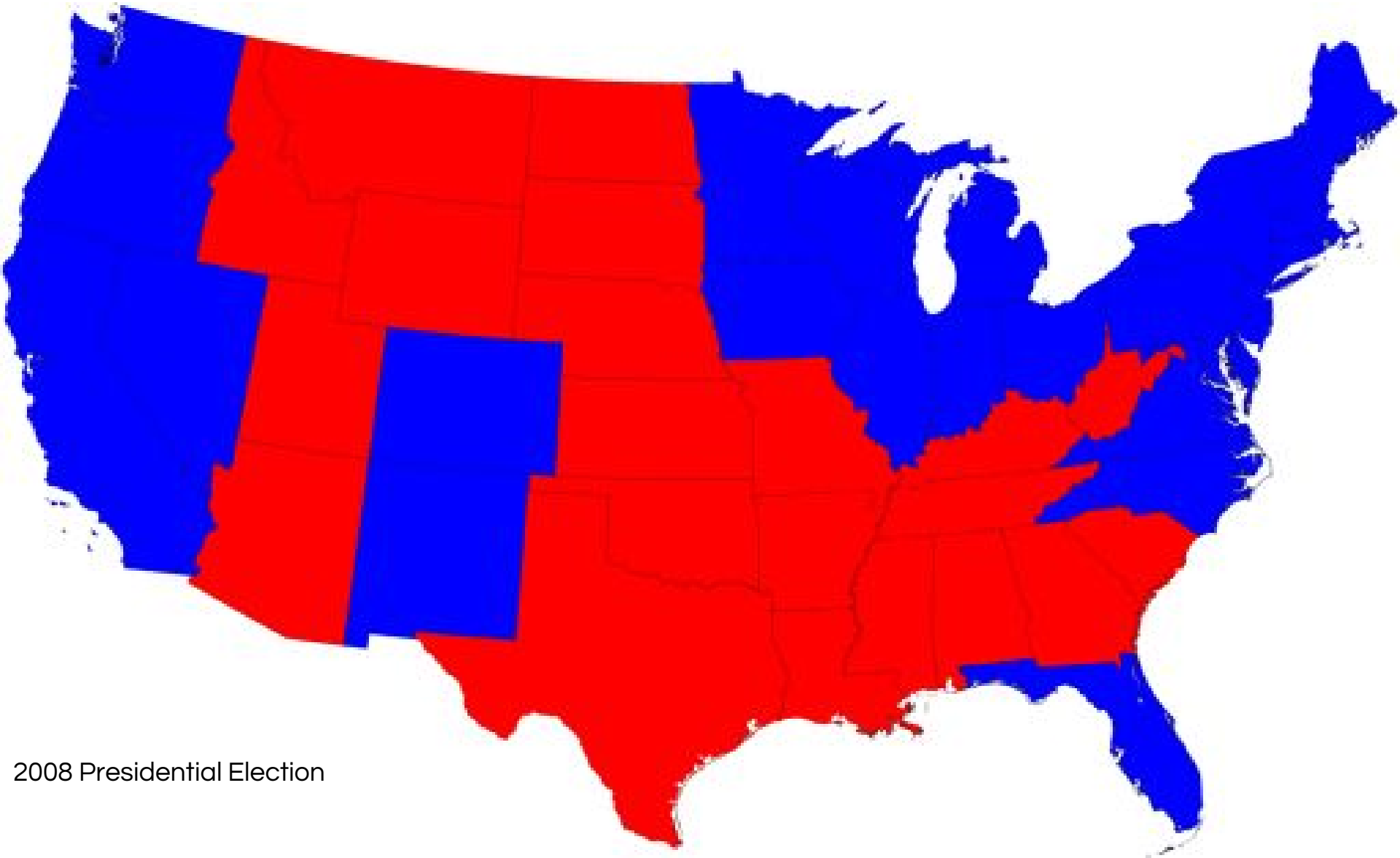
  sum(v_i)
- Average

  sum(v_i)/sum(1)
- Weighted Average

  sum(v_i * w_i) / sum(w_i)

Source: Mark Newman, University of Michigan
http://www-personal.umich.edu/~mejn/election/2008/

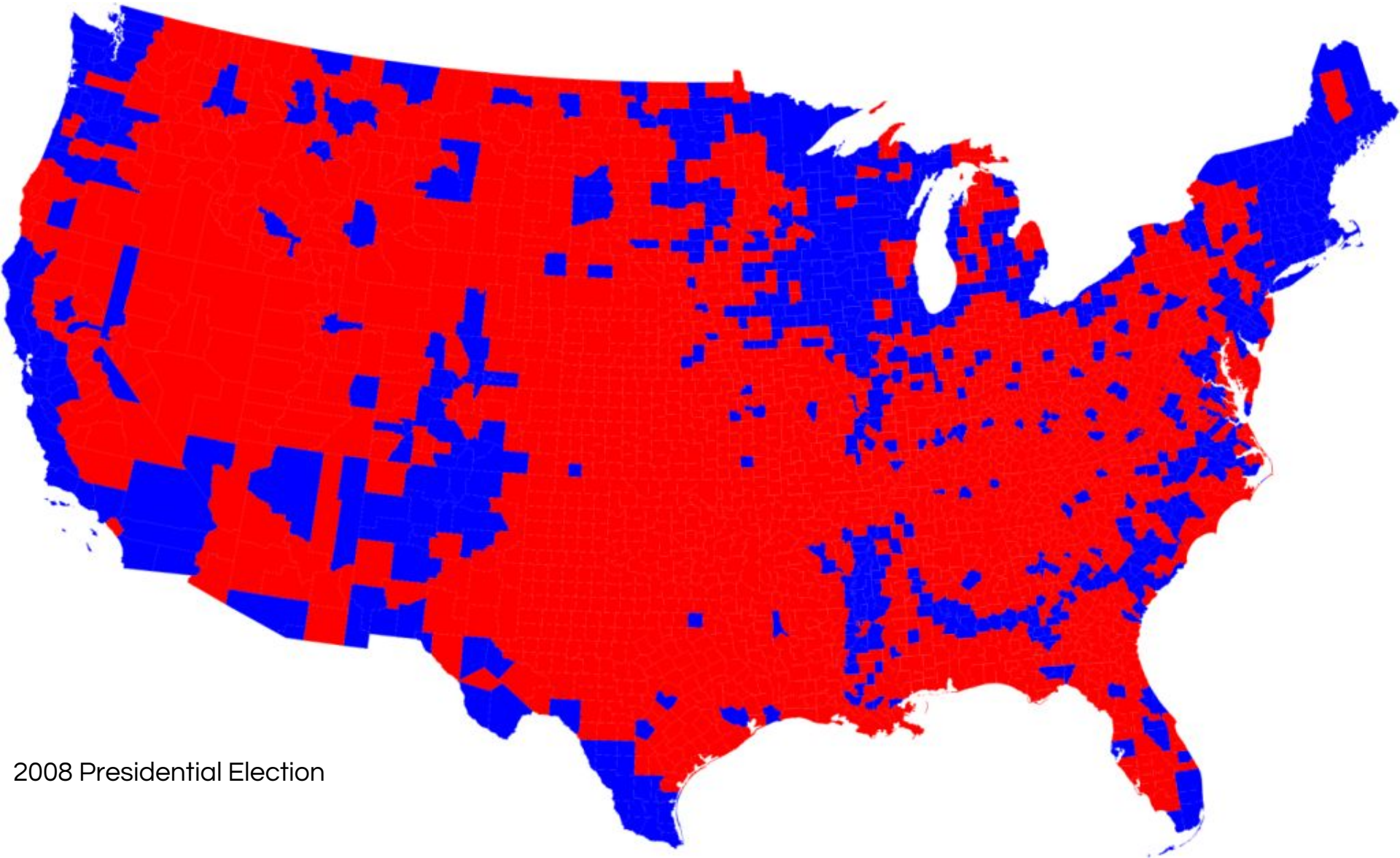
Example: Binning



2008 Presidential Election

Source: Mark Newman, University of Michigan
http://www-personal.umich.edu/~mejn/election/2008/

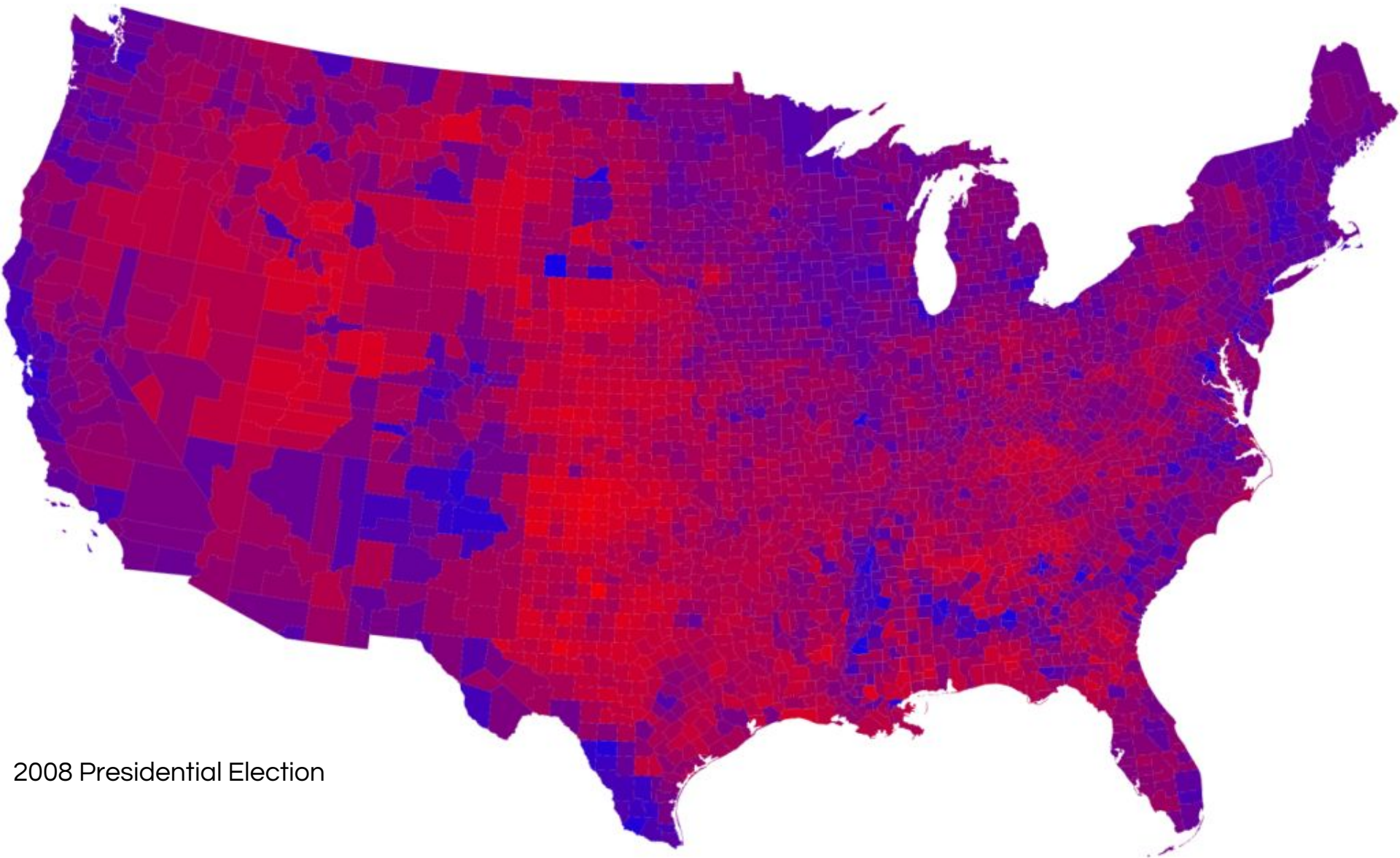
Example: Binning



2008 Presidential Election

Source: Mark Newman, University of Michigan
http://www-personal.umich.edu/~mejn/election/2008/

Example: Binning



2008 Presidential Election

Splitting Operations

Name	Category	Rating	Views	Episodes	Budget
Arrested Development	Comedy	5.0	10,000,000	30	$1,500,000
Battlestar Galactica	Sci-Fi	2.0	3,000,000	100	$2,500,000
Orphan Black	Sci-Fi	4.5	15,000,000	40	$3,000,000
Parks & Recreation	Comedy	5.0	25,000,000	90	$2,000,000

Source: Completely Made Up

Splitting Operations

Name	Category	Rating	Views	Episodes	Budget
Arrested Development	Comedy	5.0	10,000,000	30	$1,500,000
Battlestar Galactica	Sci-Fi	2.0	3,000,000	100	$2,500,000
Orphan Black	Sci-Fi	4.5	15,000,000	40	$3,000,000
Parks & Recreation	Comedy	5.0	25,000,000	90	$2,000,000

Source: Completely Made Up

Assignment will be provided Monday,
to be collected following Monday.

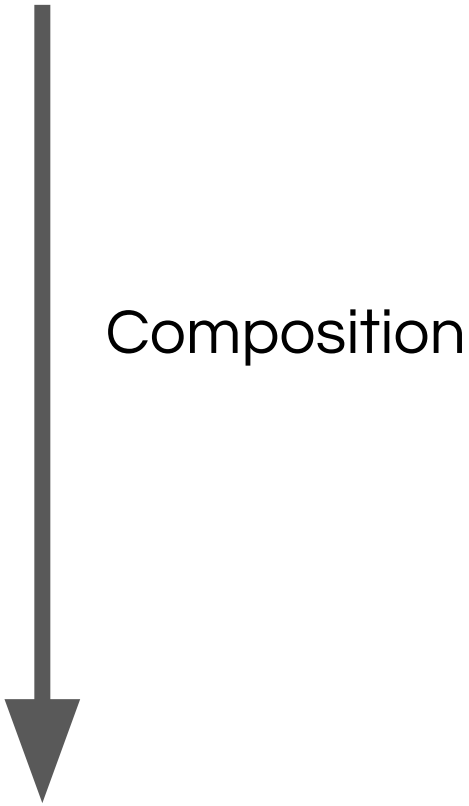
Next week, composition of visualizations

- Layers of visualizations
- Synthesis of multiple data sets: mutations and filterings
- Styling and abstraction of styles

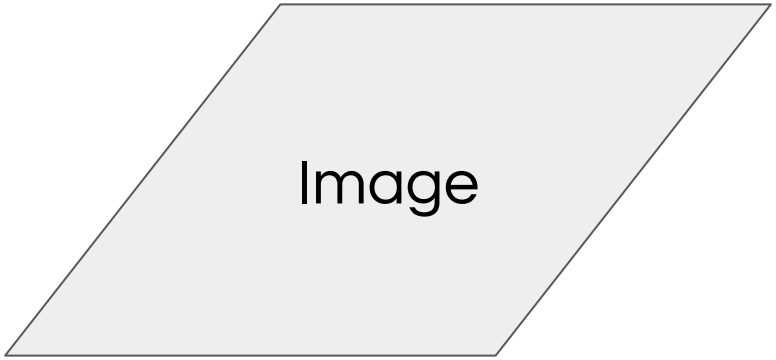
Dimensions of representation

- Position
- Color
- Size
- Shape
- Relationship

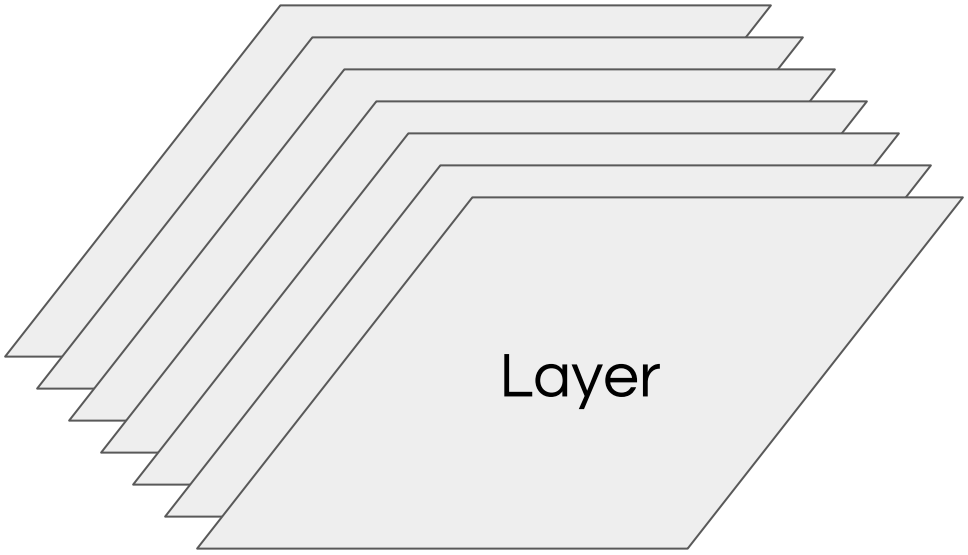
Concepts of Visualization



Concepts of Visualization



Concepts of Visualization



https://lis590.ncsa.illinois.edu/