Week 3: Elements of a Visualization

Fall 2017

Matthew Turk

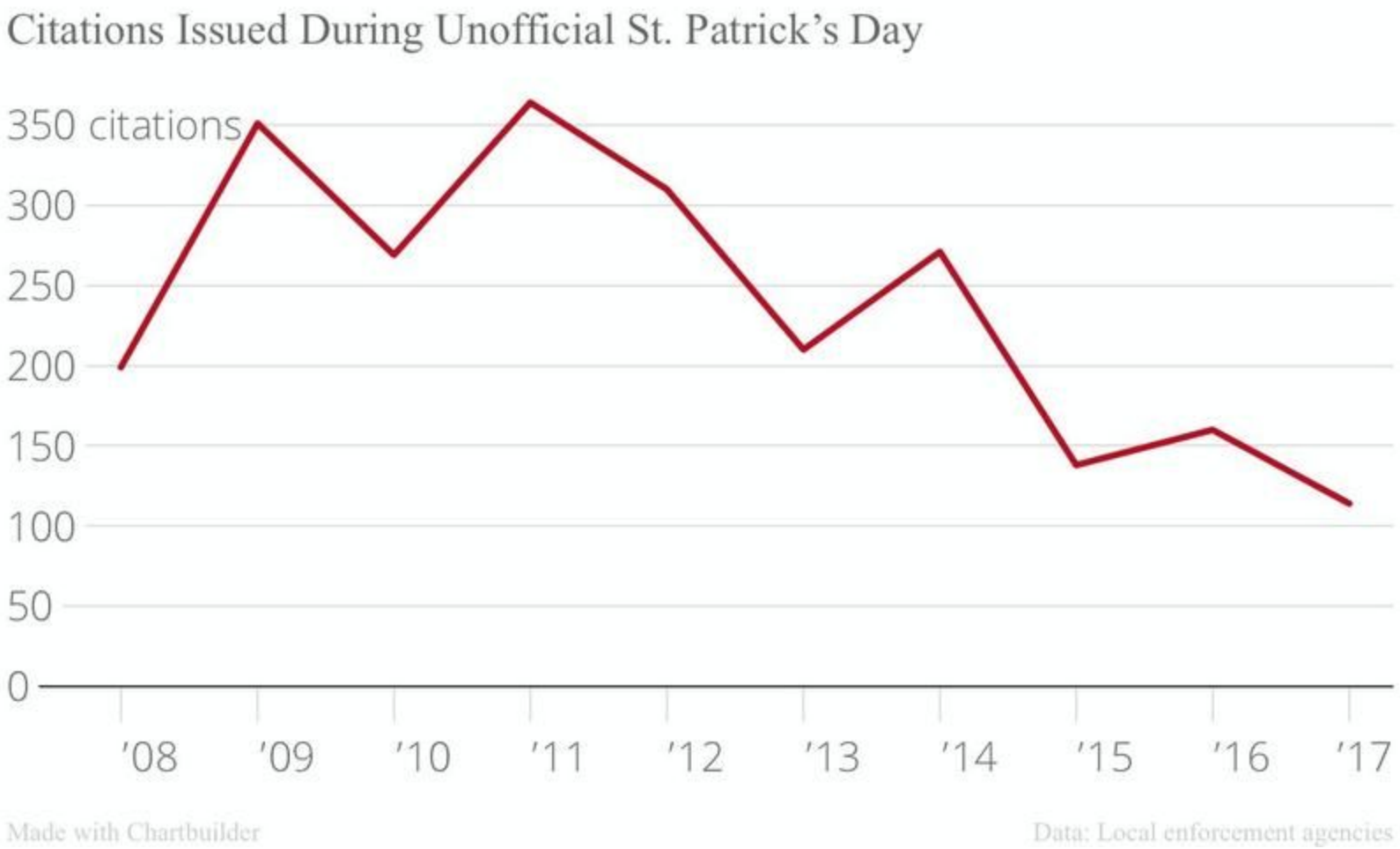
This Week

- Images and how they work
  - Raster vs vector
  - Rules vs implementation
  - Colors
- Components of a Visualization
  - Abstract components
  - Data representation
  - Visual guidance
  - Styles
- Basic Quantitative Visualization in Matplotlib

Warm-Up Activity

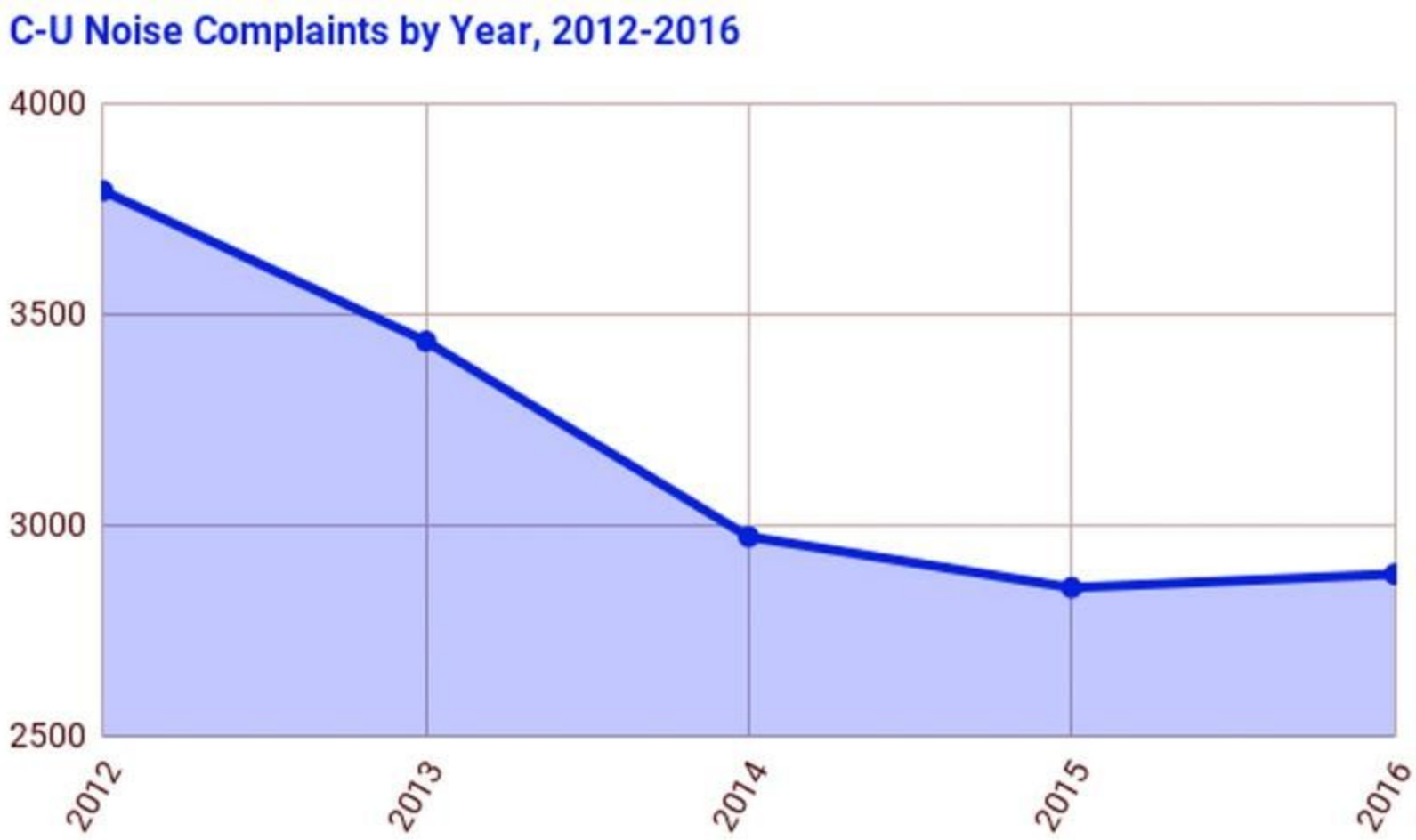
1. What is the visualization trying to show?
2. What are its methods?
3. What are the strengths / weaknesses?
4. (Bonus) How was the data collected?

Warm-Up Activity



http://www.news-gazette.com/news/local/2017-04-12/official-numbers-are-unofficial-2017.html

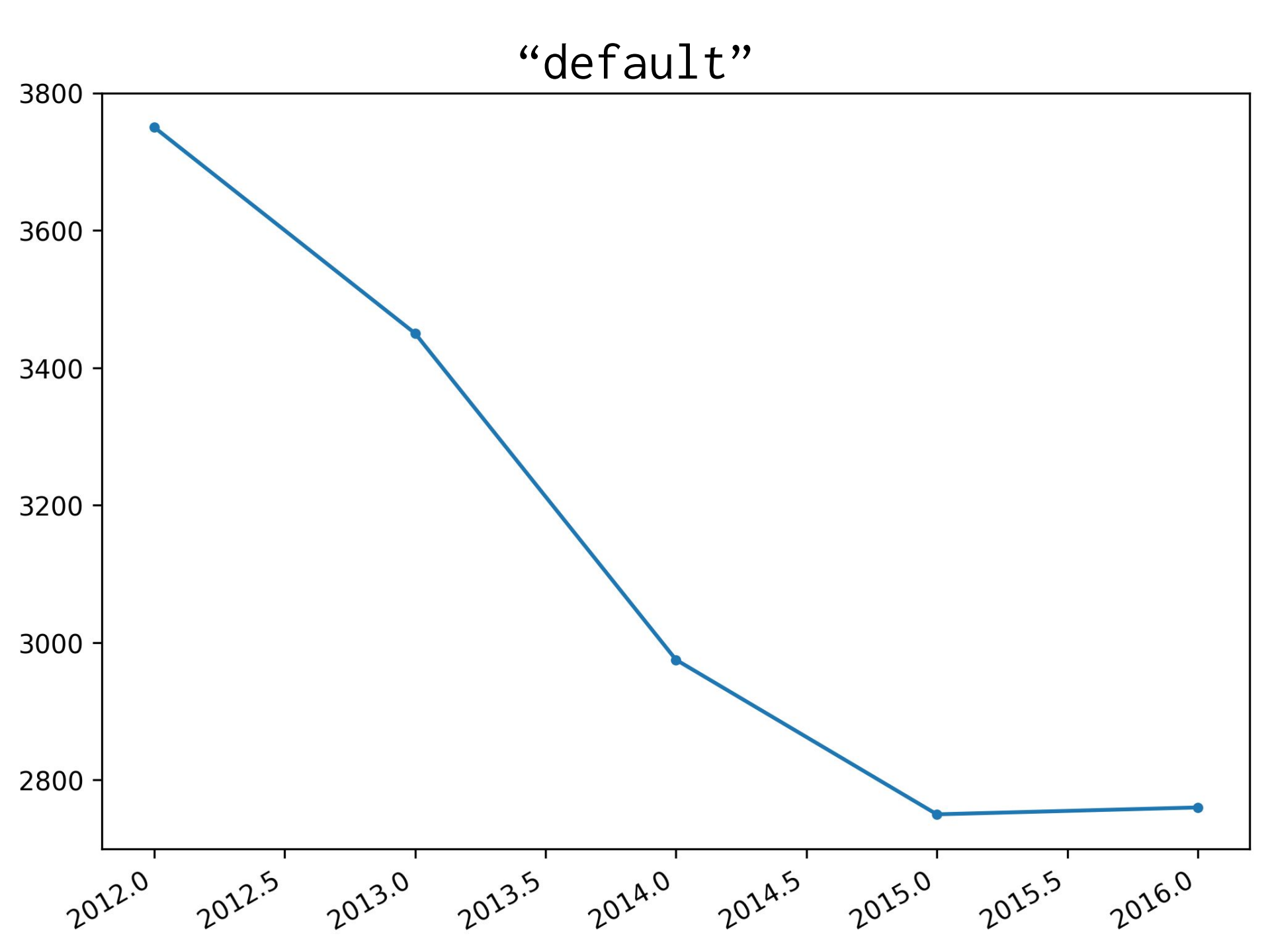
Warm-Up Activity



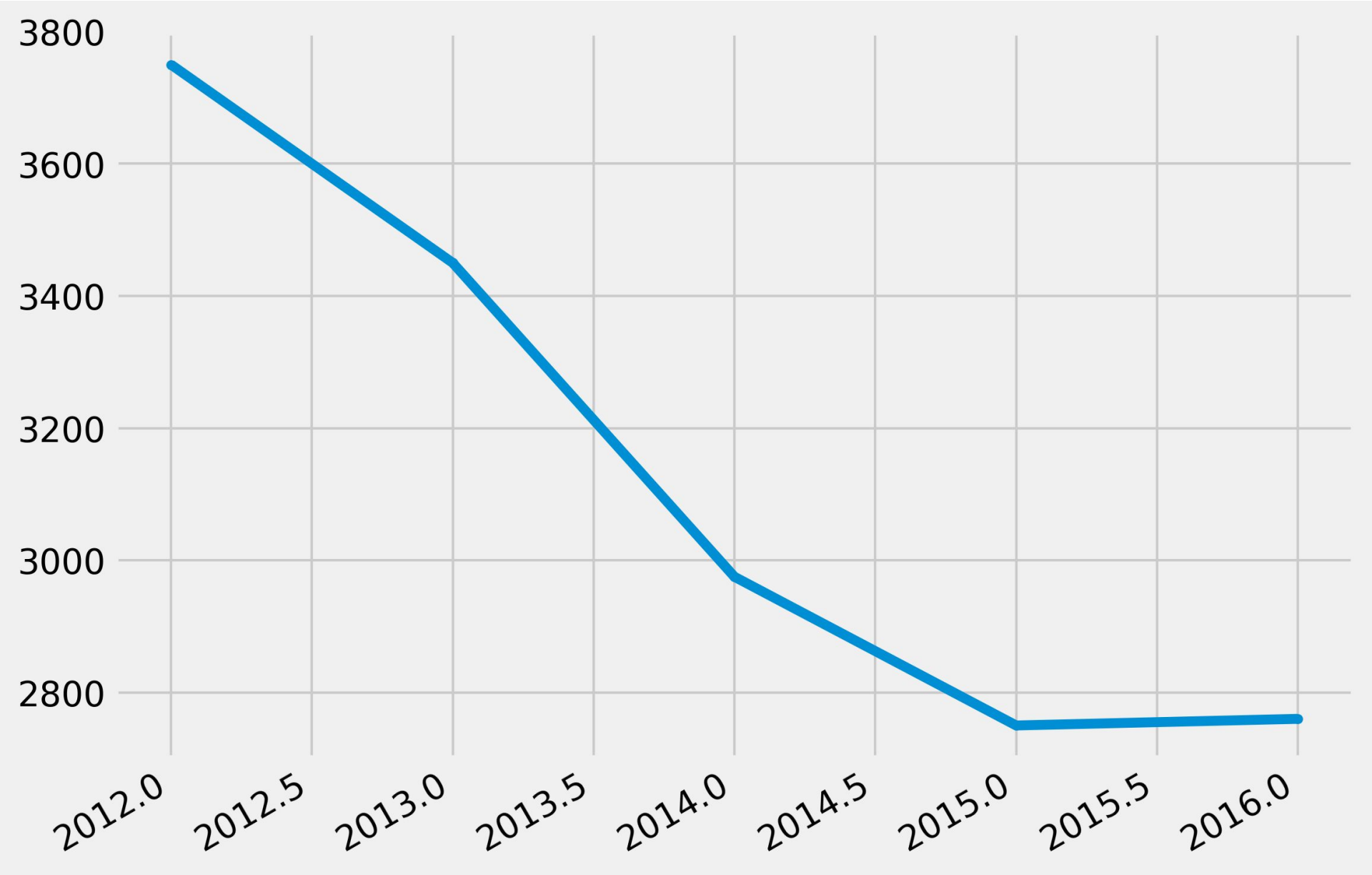
http://www.news-gazette.com/news/local/2017-09-04/top-morning-sept-4-2017.html

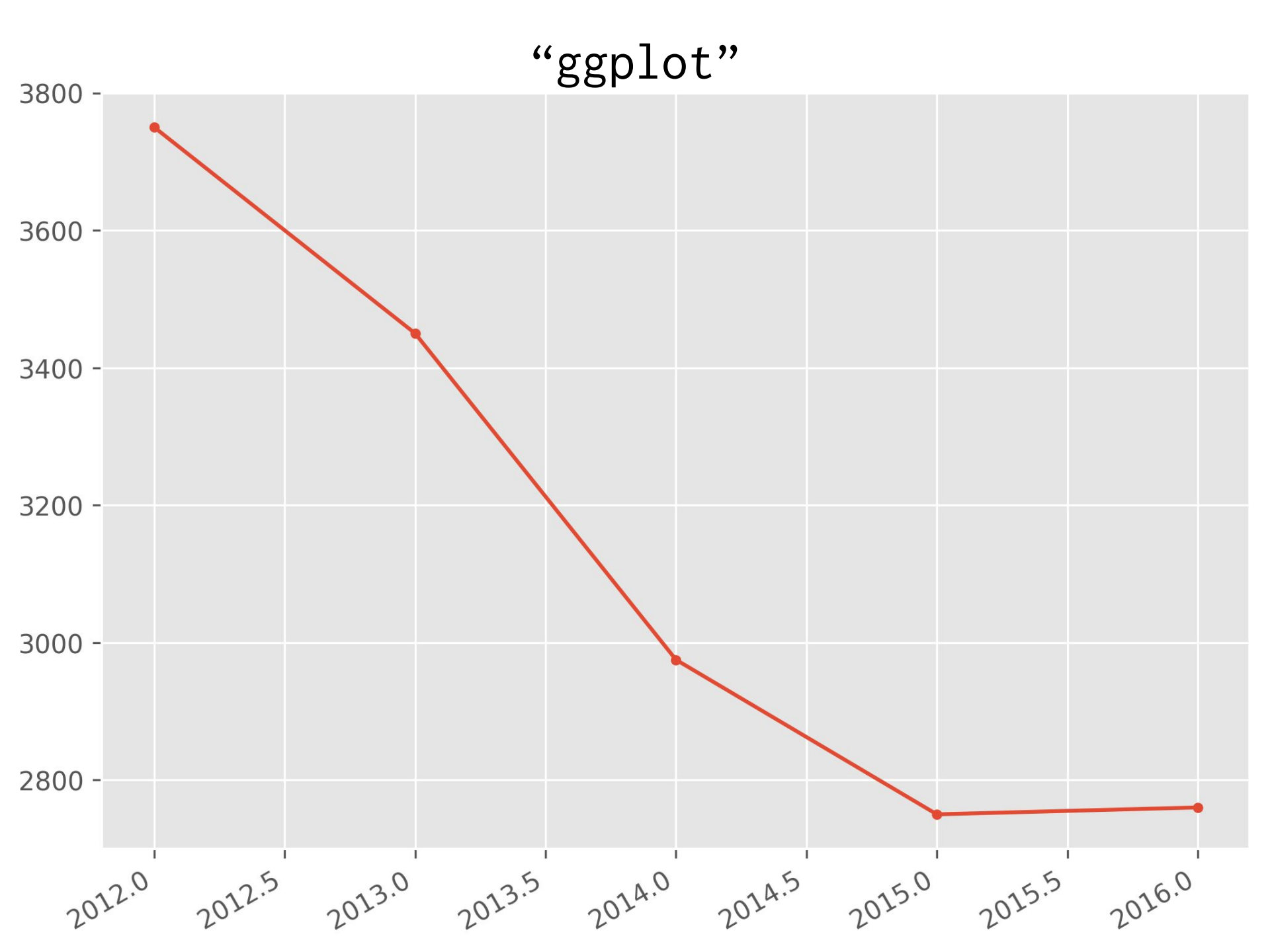
```
fig = plt.figure(figsize = (8, 6), dpi = 300)
ax = fig.add_subplot(111)
ax.plot(years, complaints, '.-')
fig.autofmt_xdate()
```

(the results aren't that great.)


“default”
3800
3600
3400
3200
3000
2800
2012.0
2012.5
2013.0
2013.5
2014.0
2014.5
2015.0
2015.5
2016.0

“fivethirtyeight”

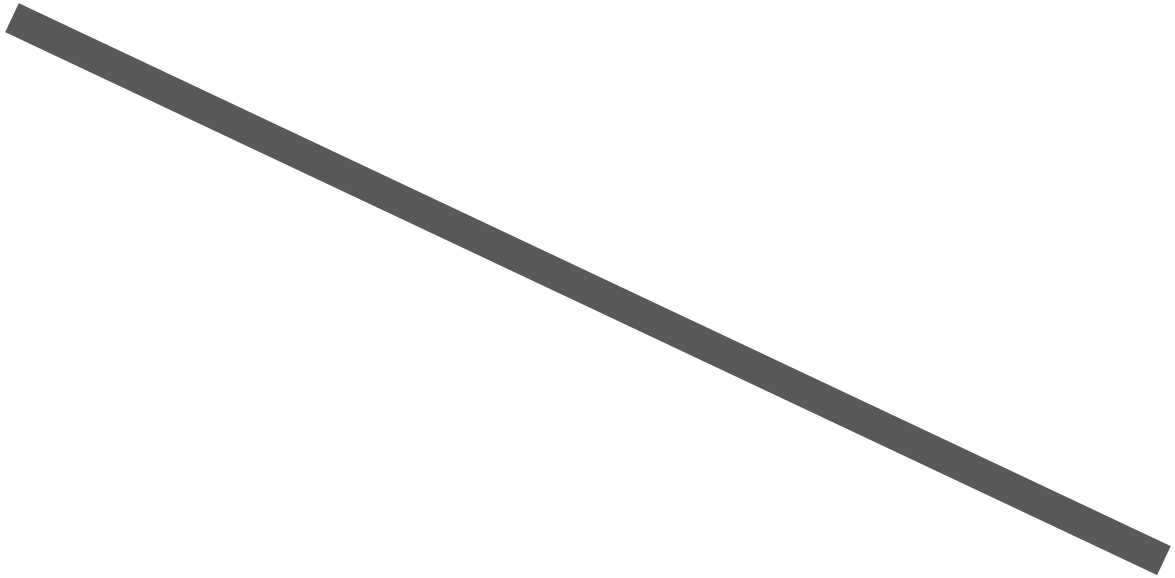


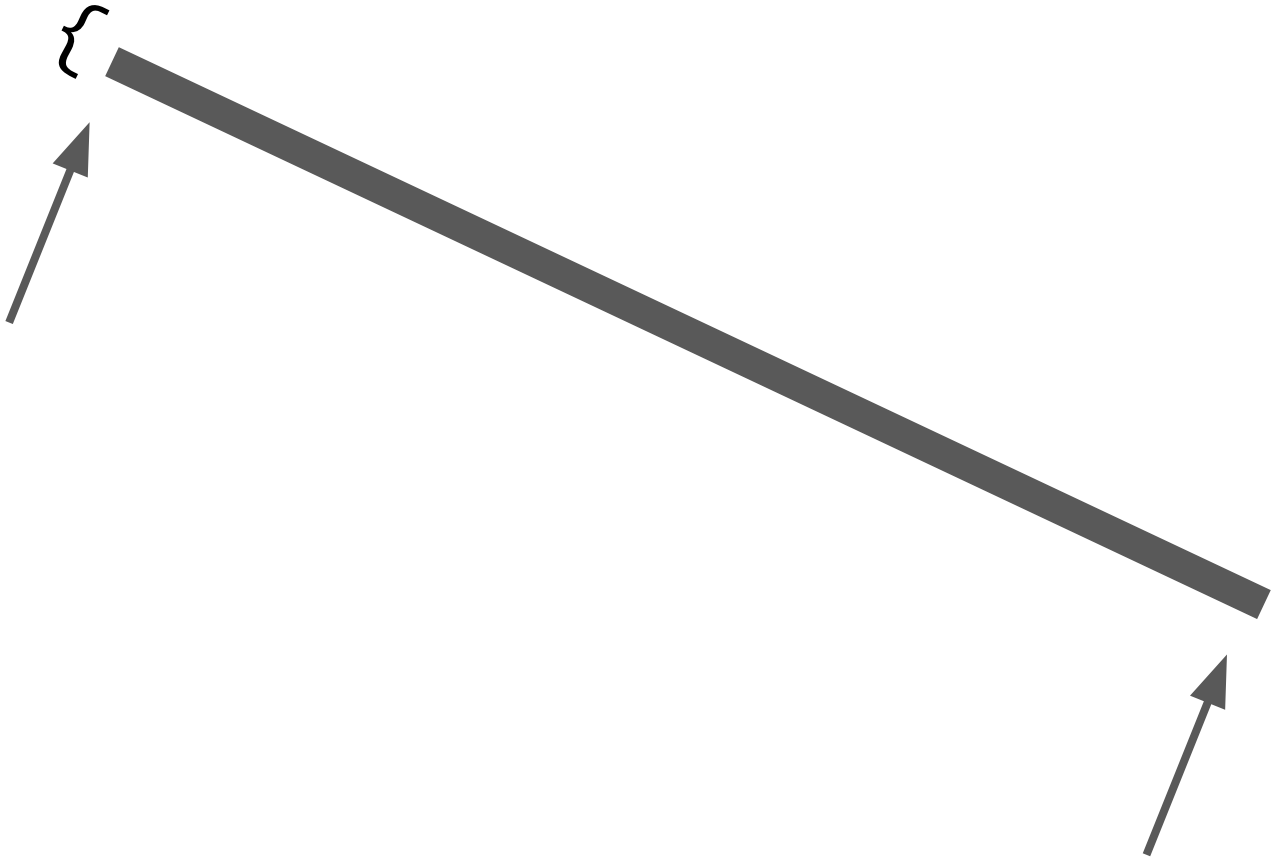

“ggplot”
3800
3600
3400
3200
3000
2800
2012.0
2012.5
2013.0
2013.5
2014.0
2014.5
2015.0
2015.5
2016.0

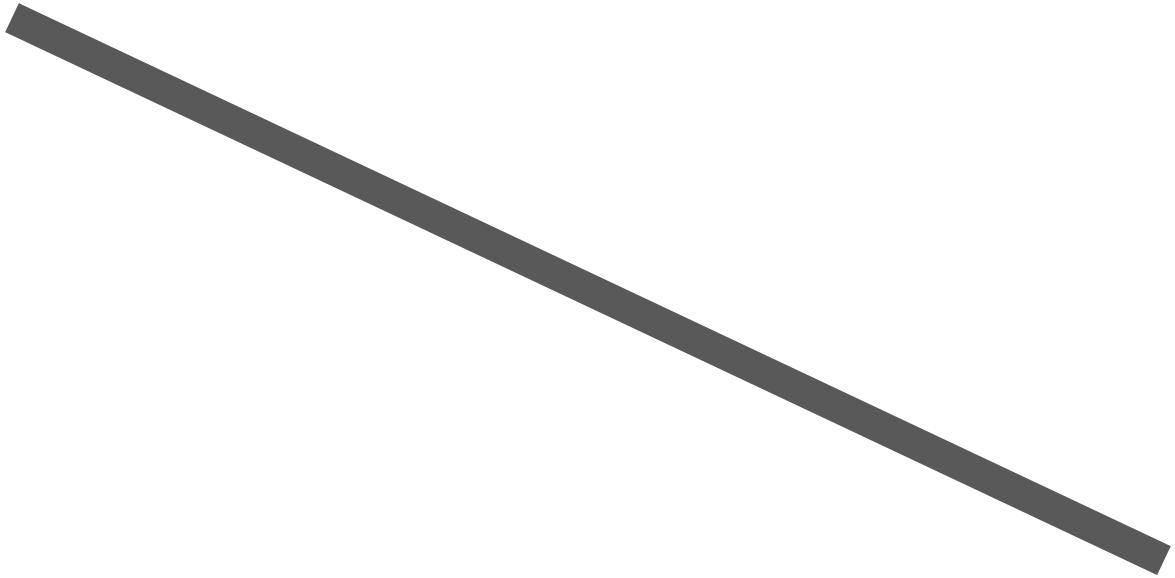
How do we draw an image?


SCORE
11,591
2
SHIELD
3 HAVE
EXHAUST PORT AHEAD

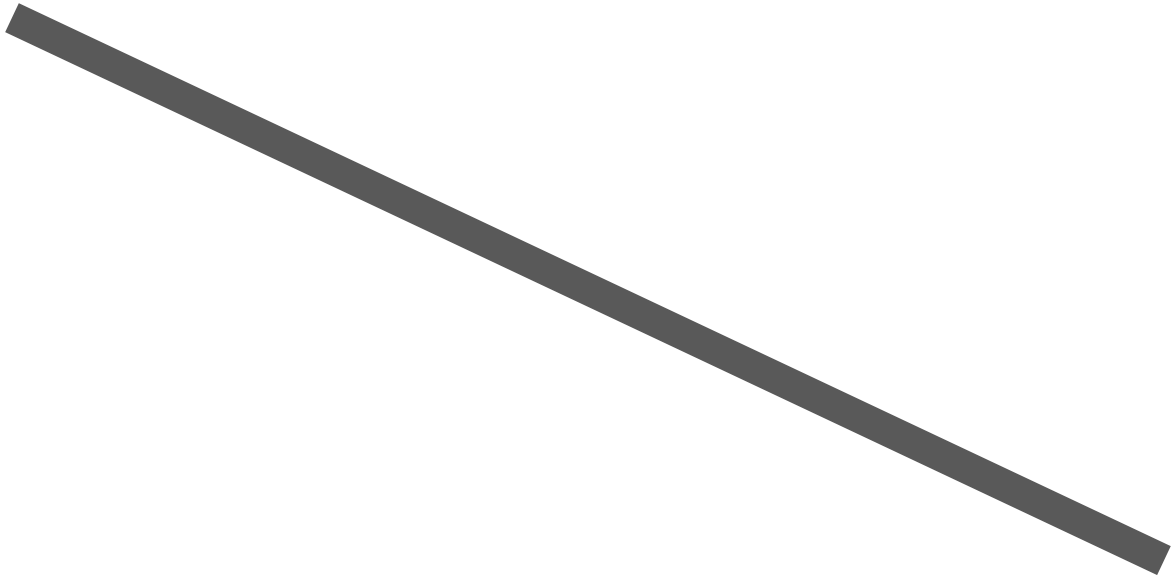
Let’s say we wanted to draw a line.





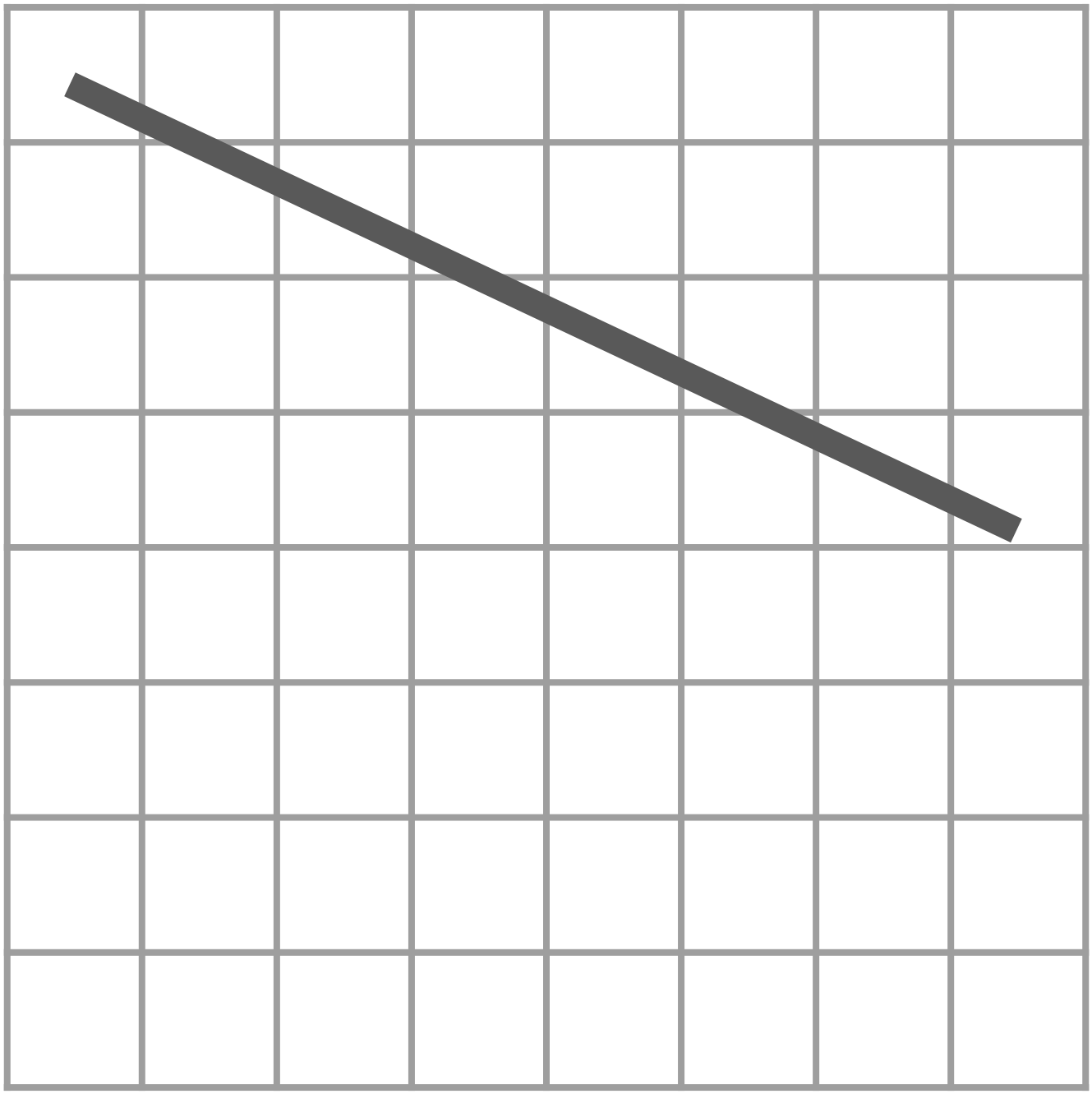


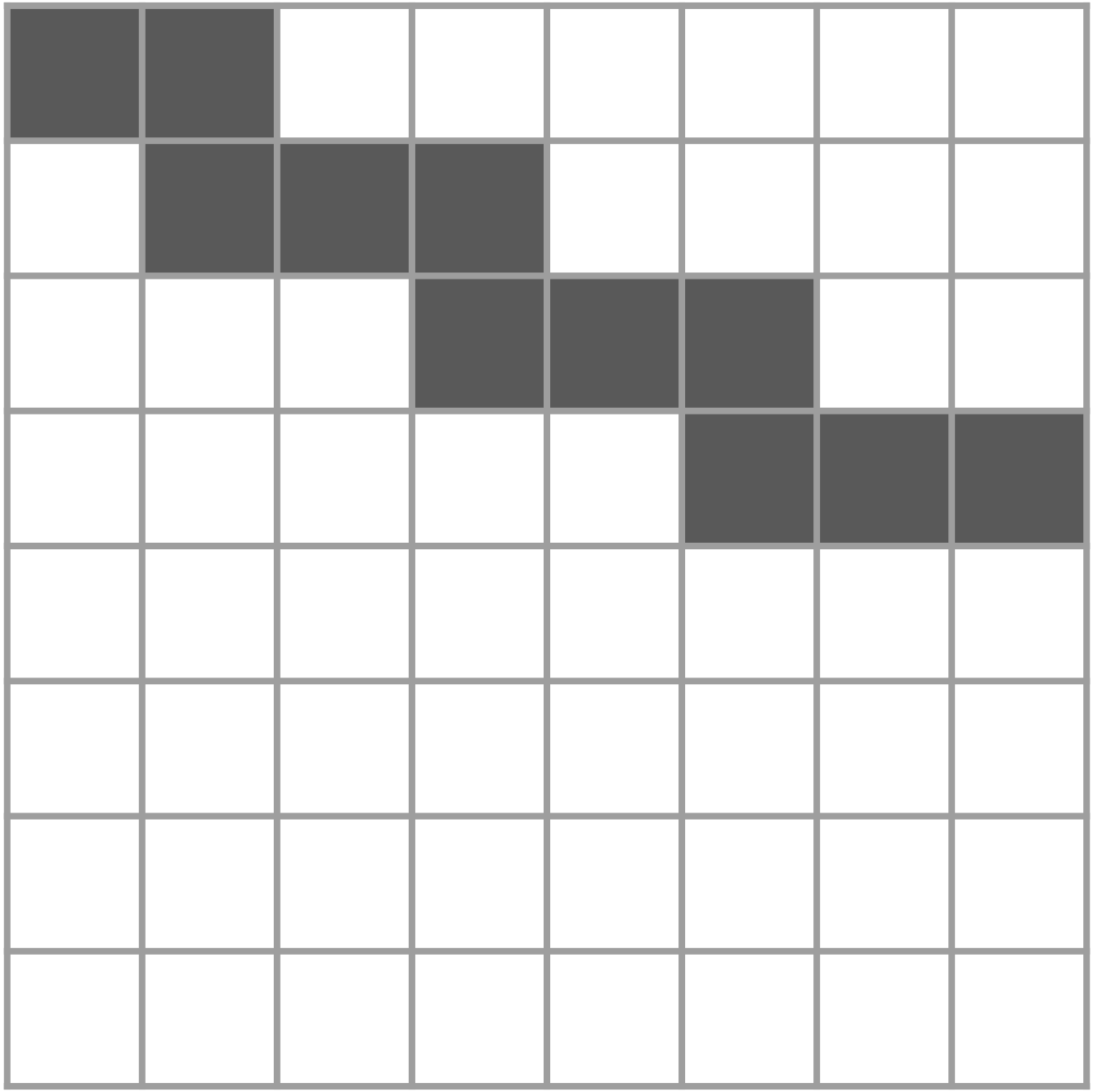
```
starting_point = ( x0, y0 )
ending_point   = ( x1, y1 )
width          = 1.0
```

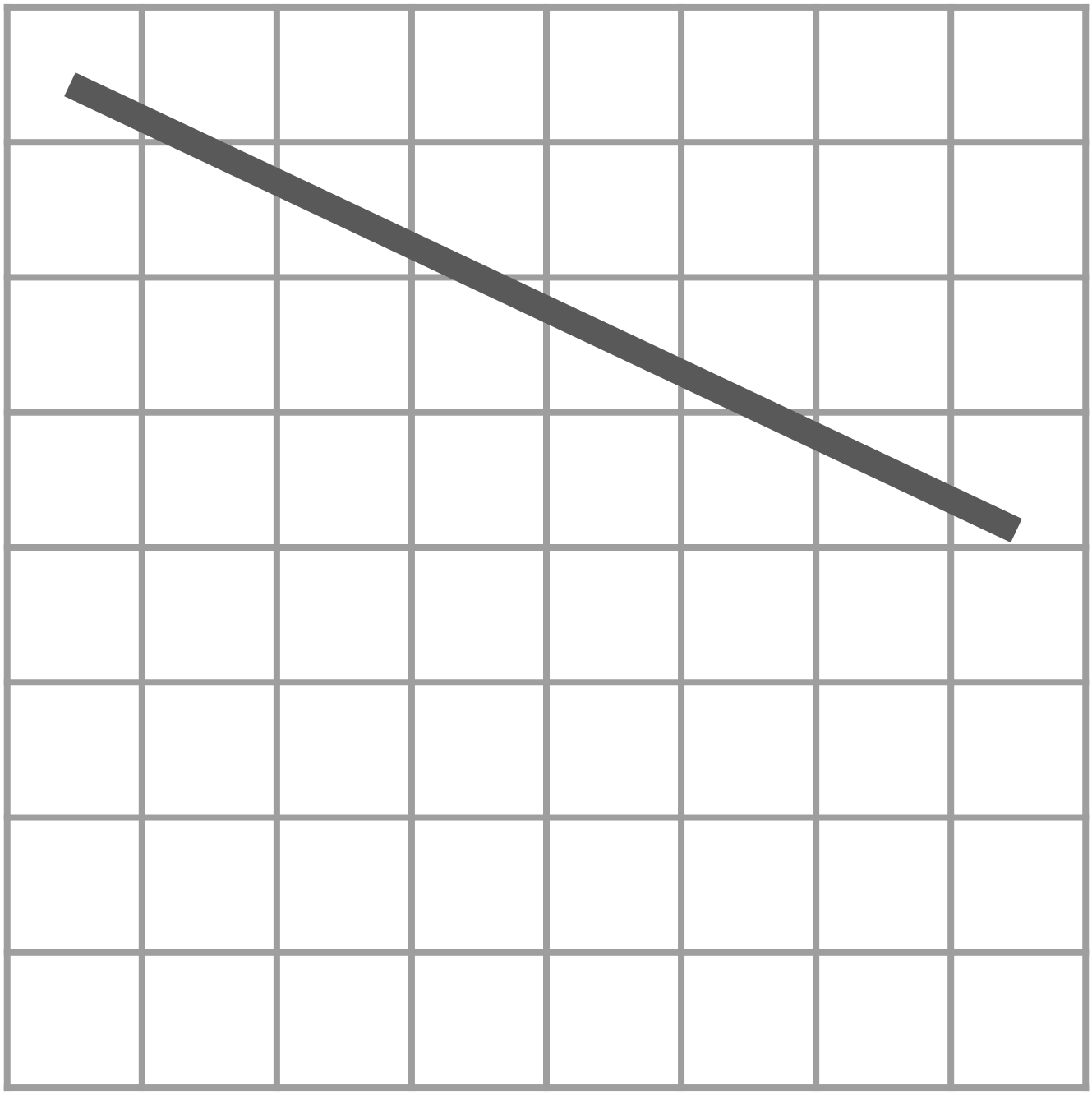


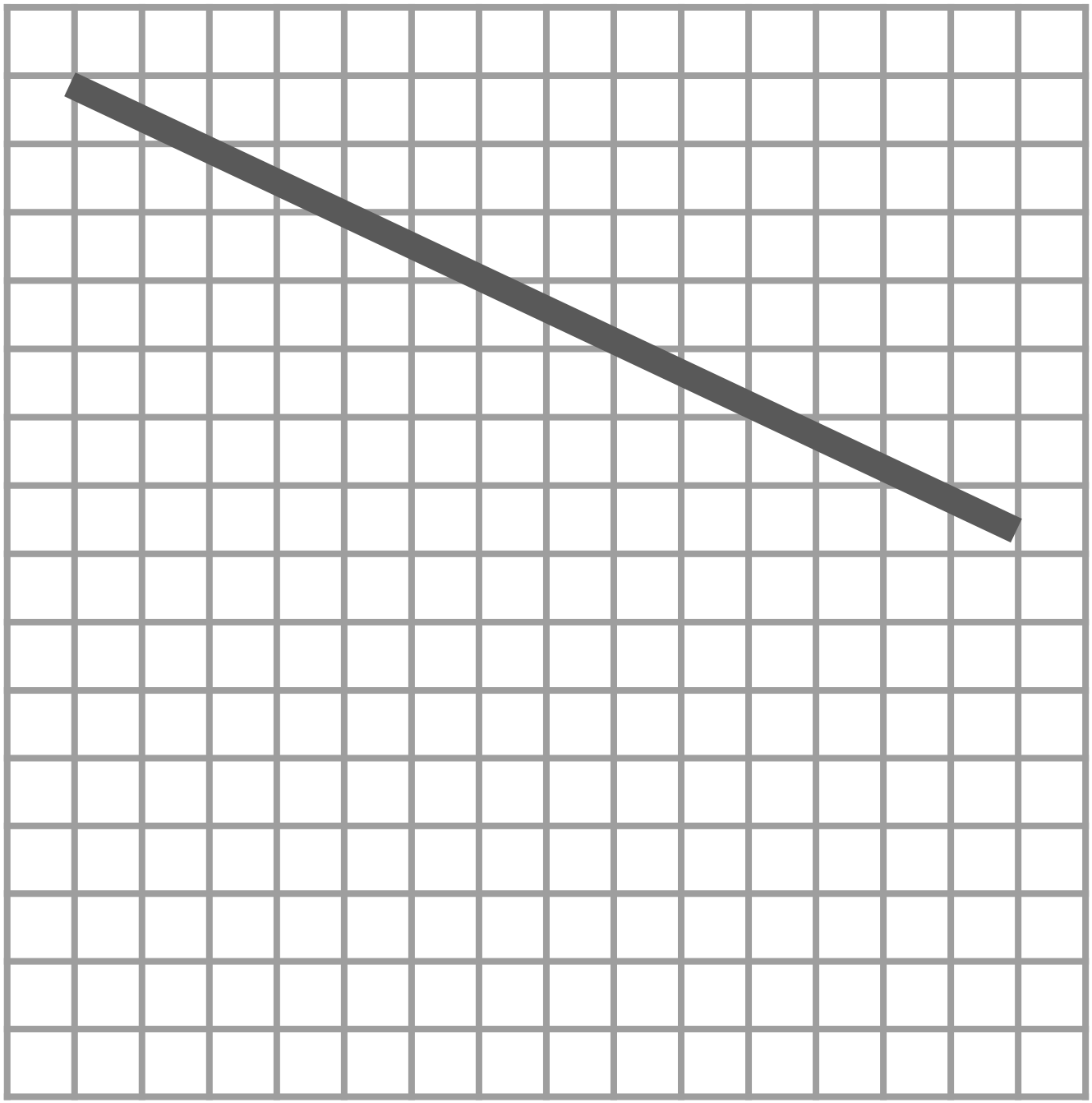
5 floating point numbers

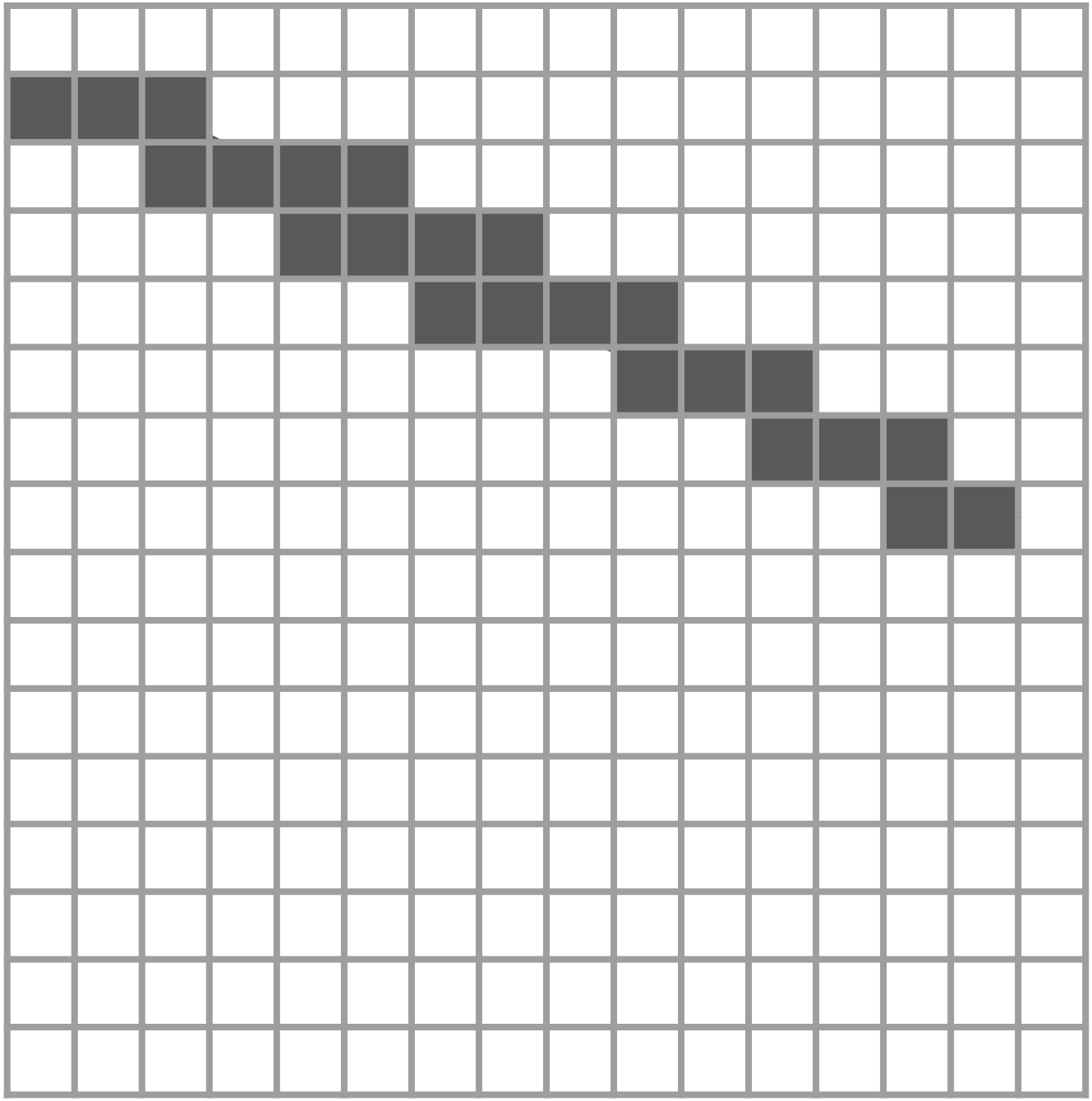
```
starting_point = ( x0, y0 )
ending_point   = ( x1, y1 )
width          = 1.0
```











Raster Representation

	1 Line	2 Lines	30 Lines	1000 Lines
600x600	45 kb	45 kb	45 kb	45 kb
1200x1200	180 kb	180 kb	180 kb	180 kb
2400x2400	720 kb	720 kb	720 kb	720 kb

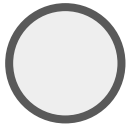
(uncompressed, 1 bit images)

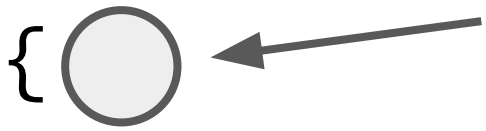
Vector Representation

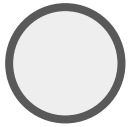
	1 Line	2 Lines	30 Lines	1000 Lines
600x600	5 bytes	10 bytes	150 bytes	5000 bytes
1200x1200	5 bytes	10 bytes	150 bytes	5000 bytes
2400x2400	5 bytes	10 bytes	150 bytes	5000 bytes

(uncompressed, single precision)

Let’s say we wanted to draw a dot.

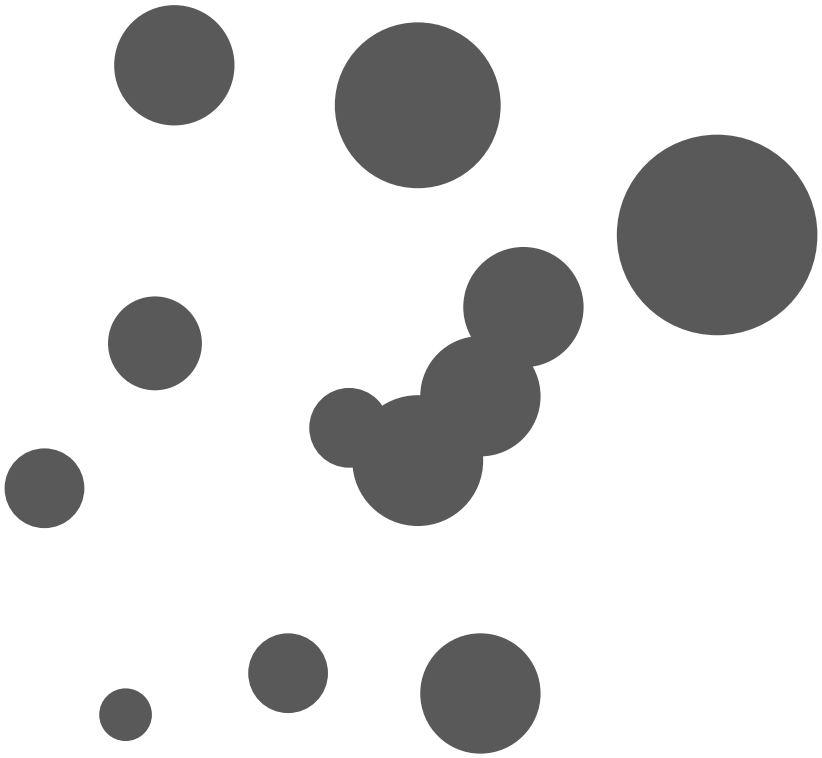


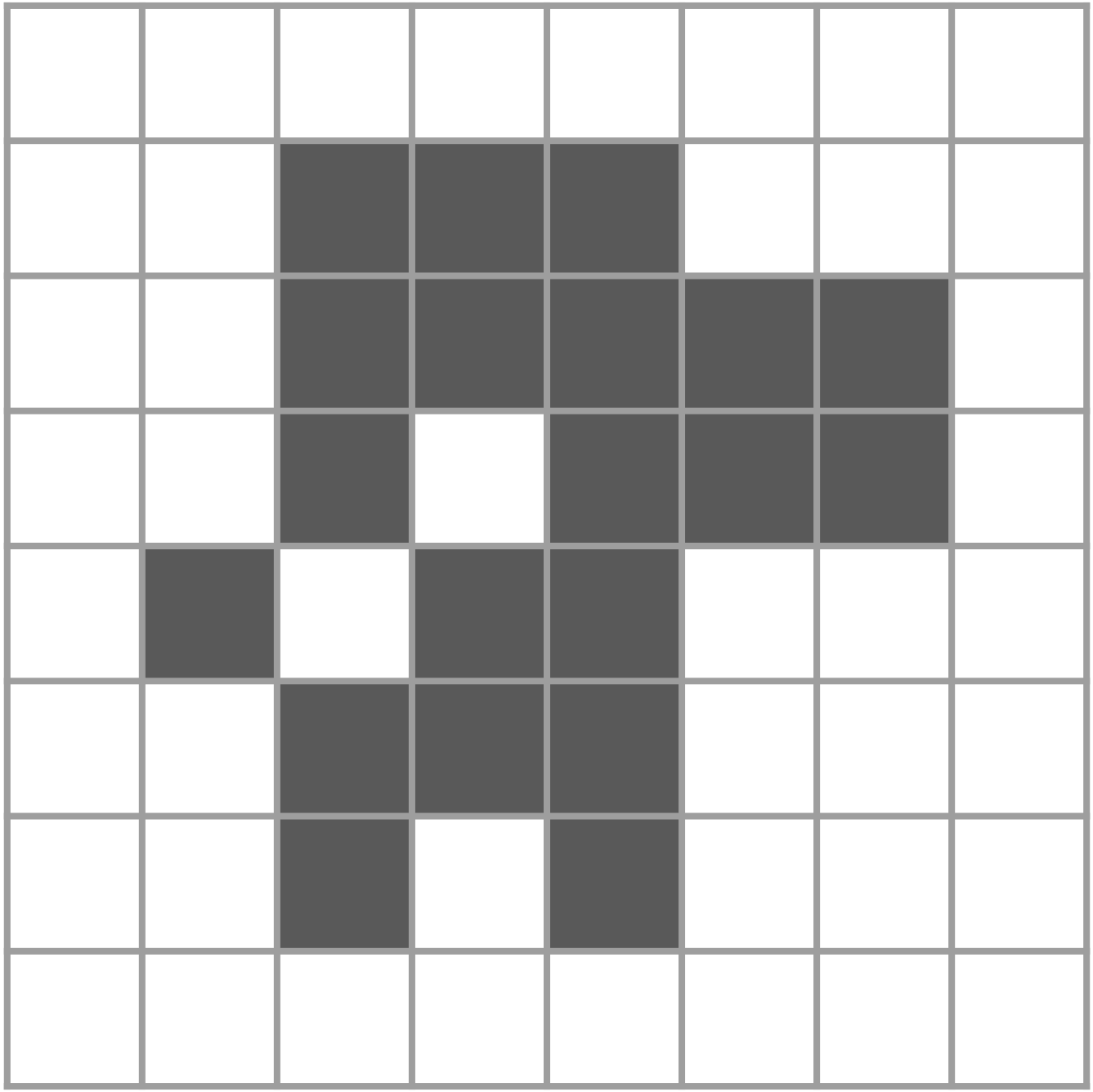


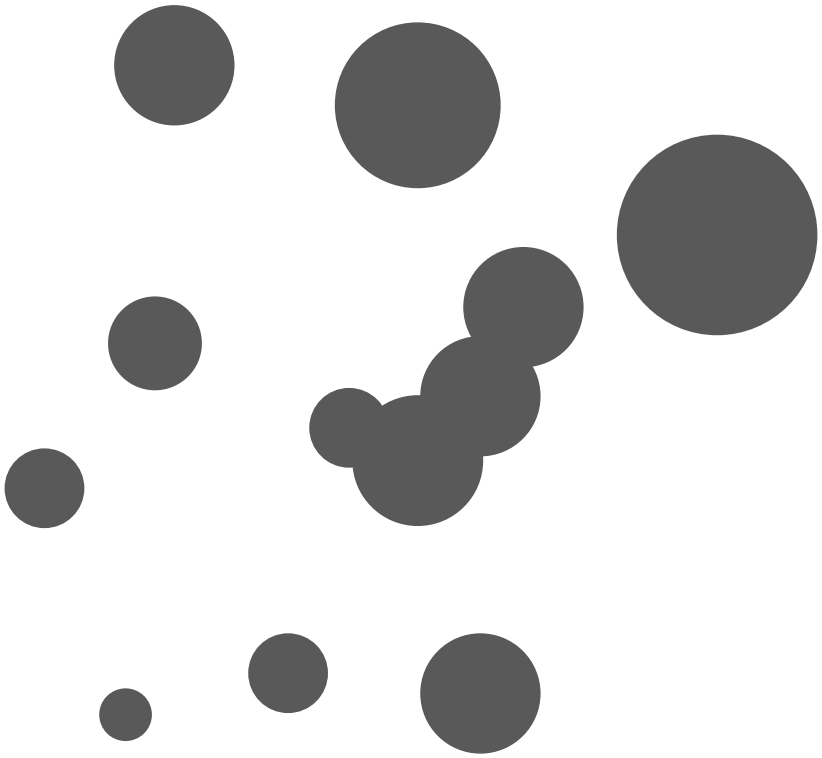


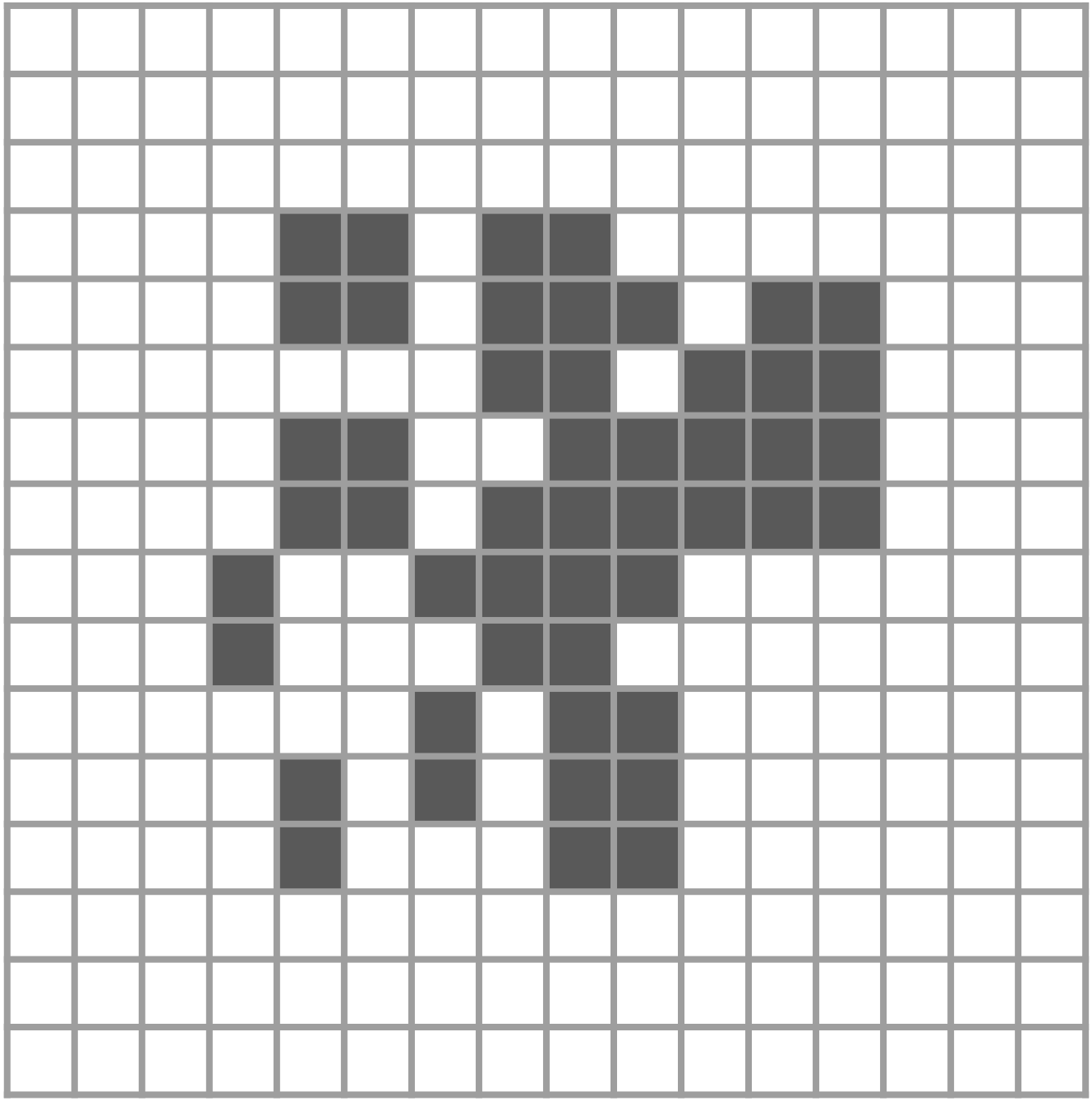
3 floating point numbers

```
center = ( x0, y0 )
radius = 1.0
```









Raster Representation

	1 Line	2 Lines	30 Lines	1000 Lines
600x600	45 kb	45 kb	45 kb	45 kb
1200x1200	180 kb	180 kb	180 kb	180 kb
2400x2400	720 kb	720 kb	720 kb	720 kb

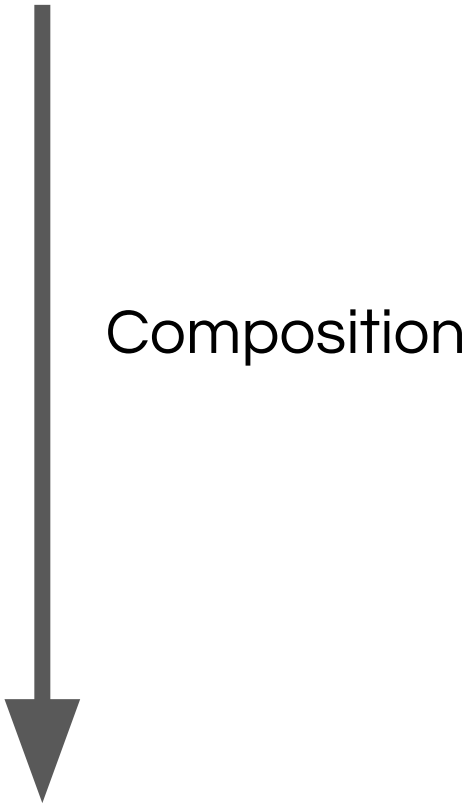
(uncompressed, 1 bit images)

Vector Representation

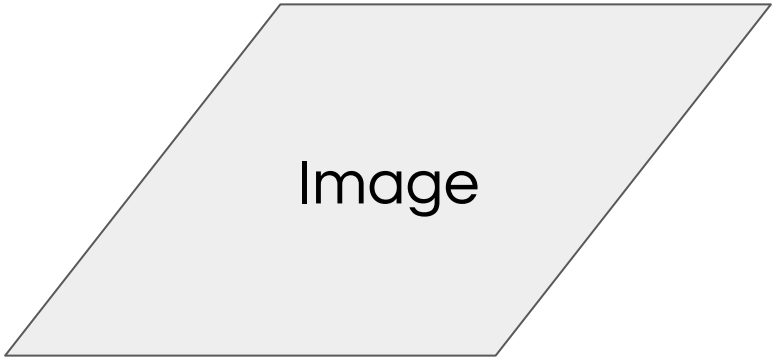
	1 Circle	2 Circles	1000 Circles	1e6 Circles
600x600	3 bytes	6 bytes	3000 bytes	3 Mb
1200x1200	3 bytes	6 bytes	3000 bytes	3 Mb
2400x2400	3 bytes	6 bytes	3000 bytes	3 Mb

(uncompressed, single precision)

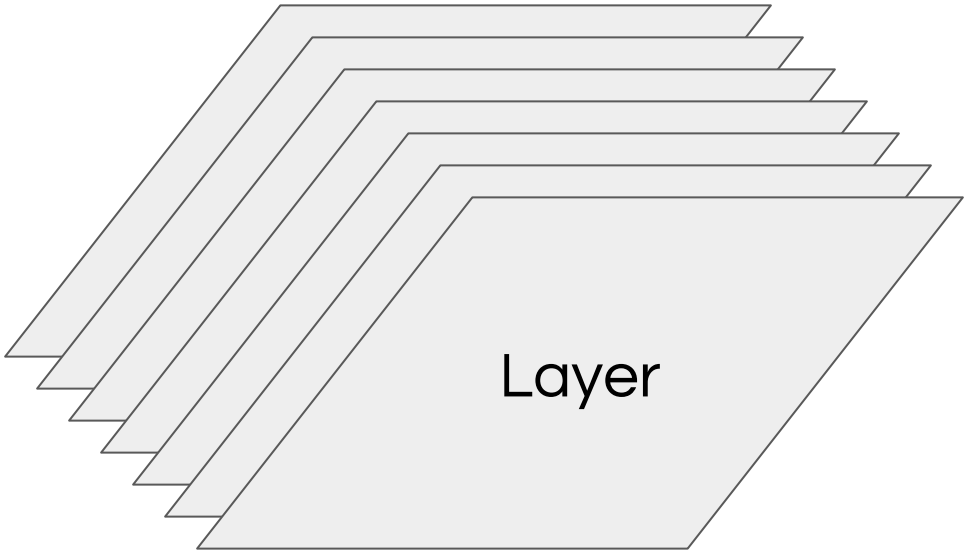
Concepts of Visualization



Concepts of Visualization



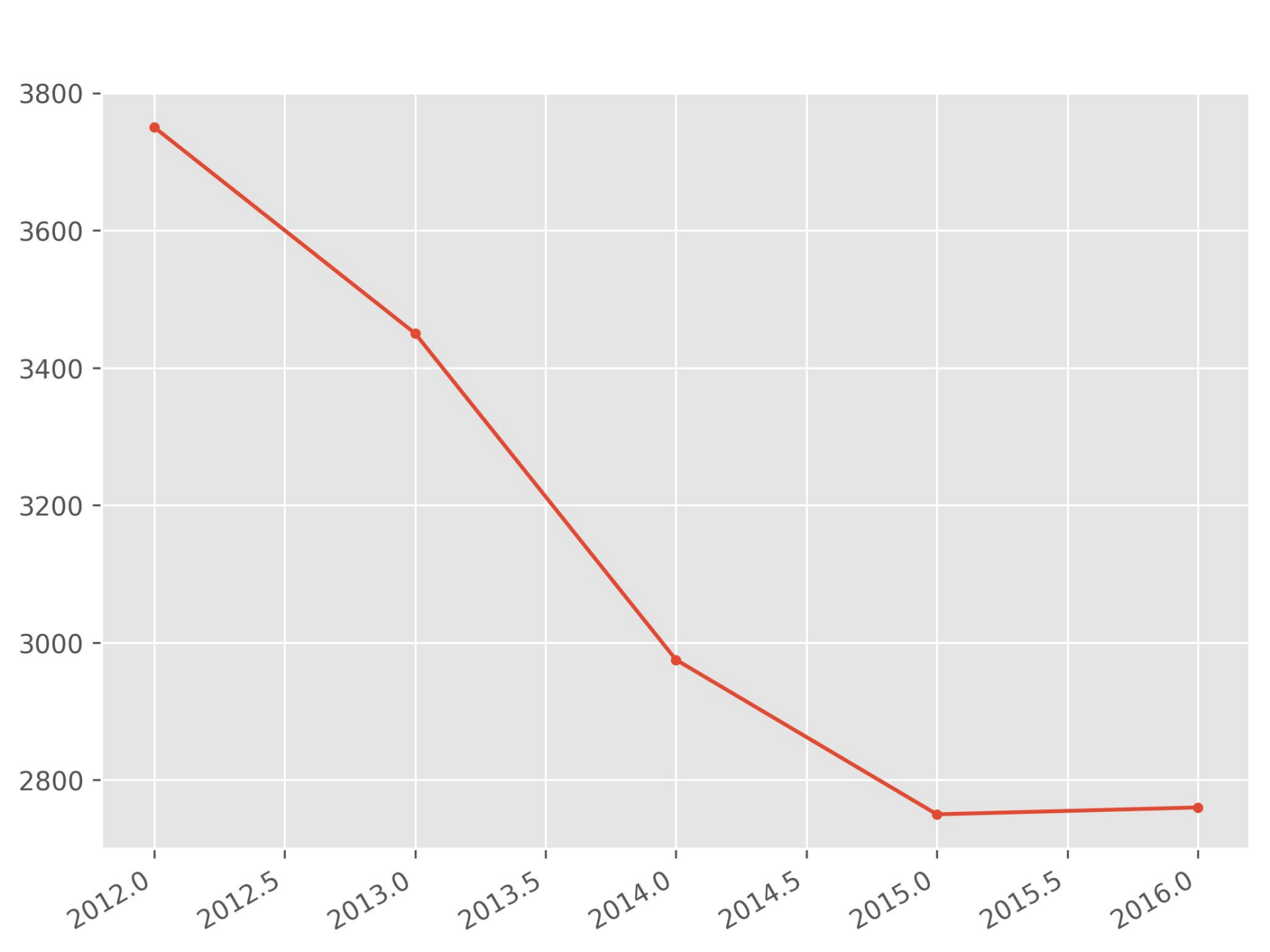
Concepts of Visualization

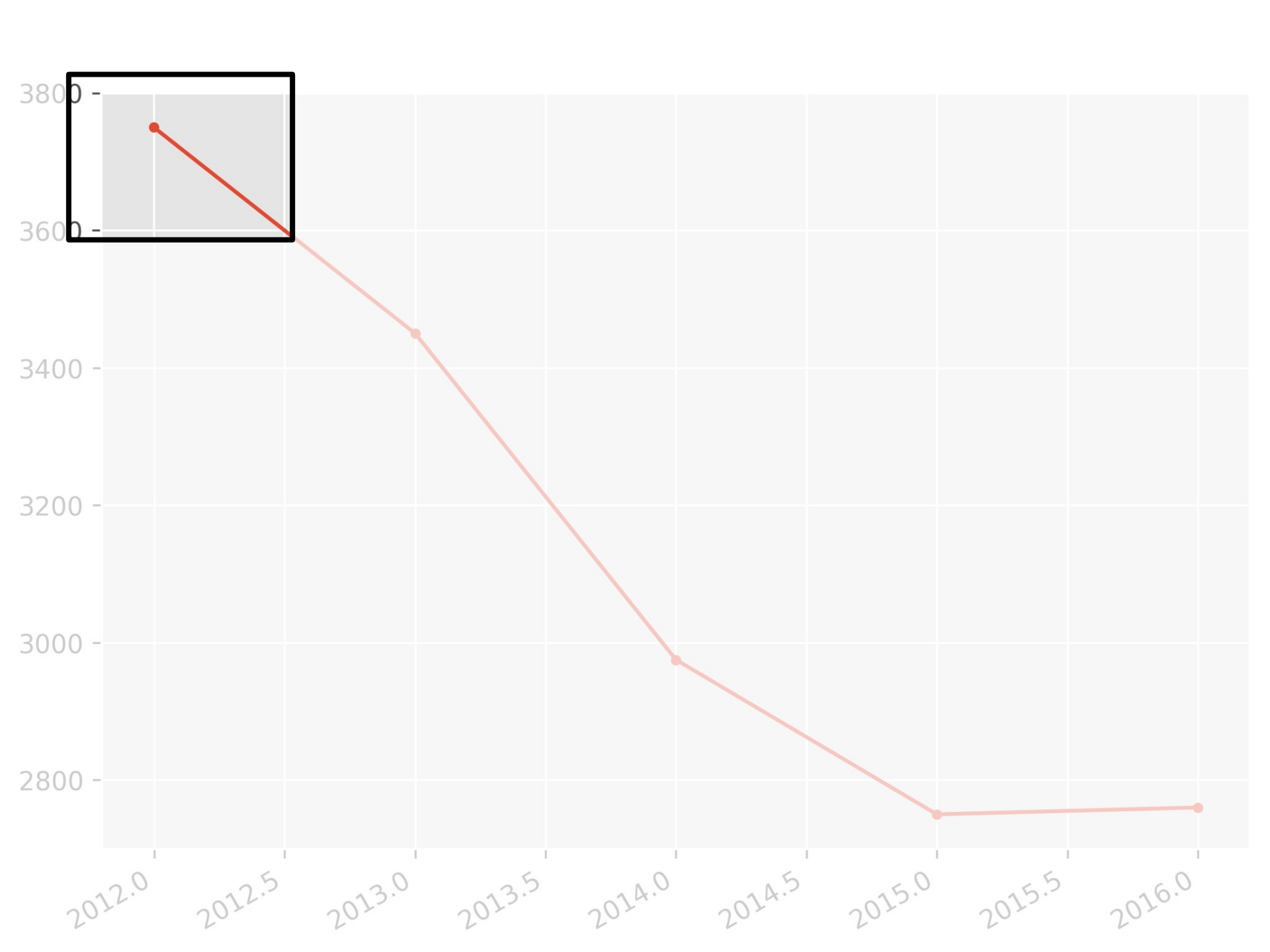


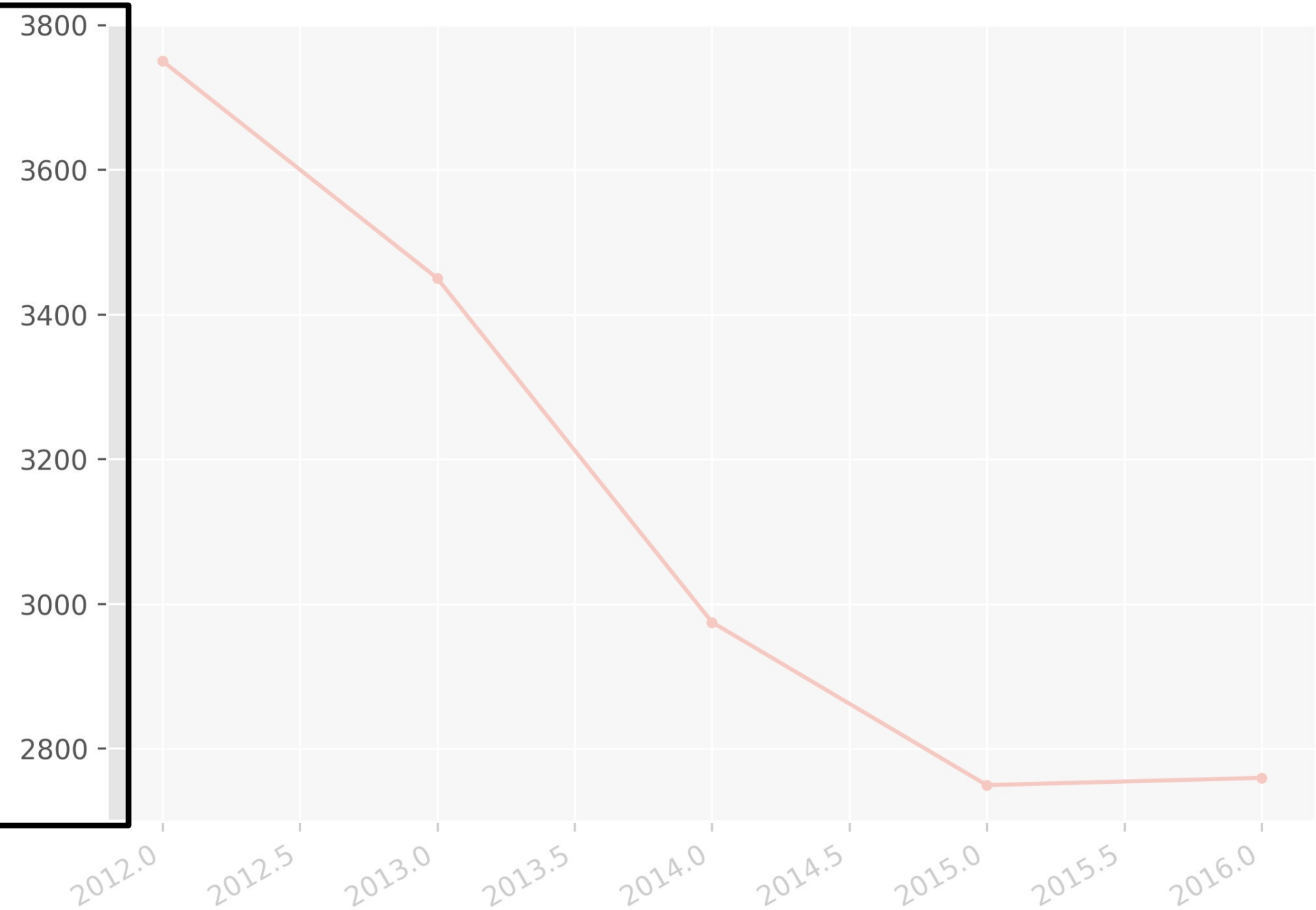
Tufte proposes we think about the ratio between essential ink and inessential ink. What can be erased without removing the ability to reconstitute the meaning?

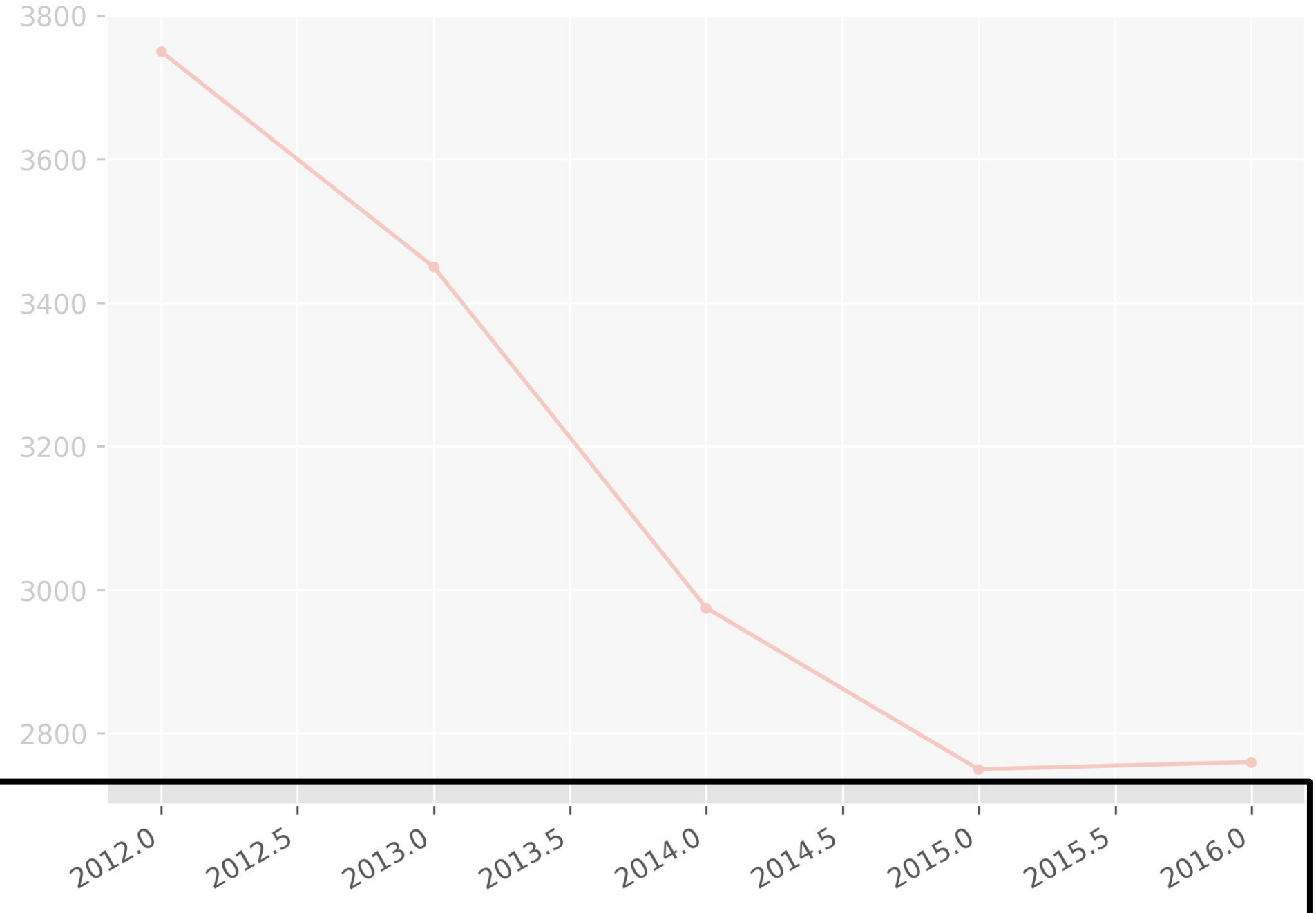
$$\frac{\texttt{data\_ink}}{\texttt{ink}}$$

Components of a Visualization: Case Study


3800
3600
3400
3200
3000
2800
2012.0
2012.5
2013.0
2013.5
2014.0
2014.5
2015.0
2015.5
2016.0


3800
3600
3400
3200
3000
2800
2012.0
2012.5
2013.0
2013.5
2014.0
2014.5
2015.0
2015.5
2016.0


3800
3600
3400
3200
3000
2800
2012.0
2012.5
2013.0
2013.5
2014.0
2014.5
2015.0
2015.5
2016.0

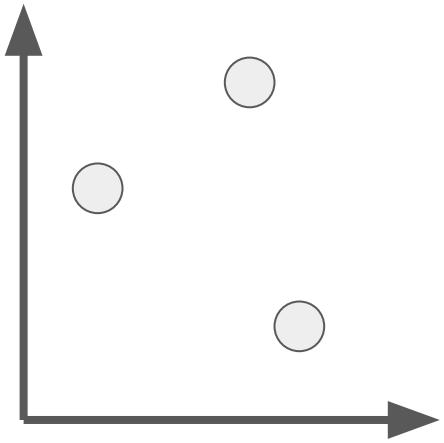

3800
3600
3400
3200
3000
2800
2012.0
2012.5
2013.0
2013.5
2014.0
2014.5
2015.0
2015.5
2016.0

Dimensions of representation

- Position
- Color
- Size
- Shape
- Relationship

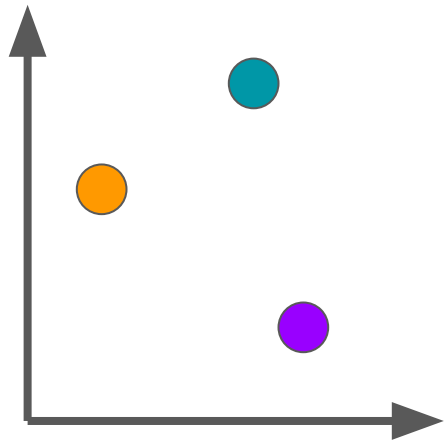
Dimensions of representation

- Position
- Color
- Size
- Shape
- Relationship



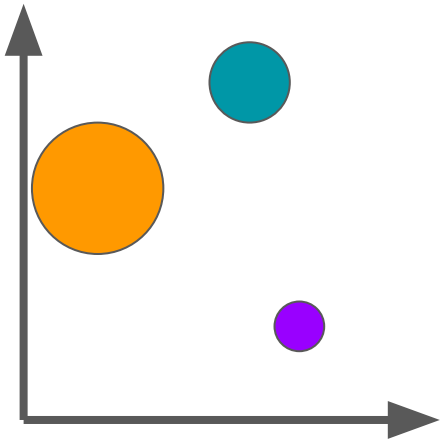
Dimensions of representation

- Position
- Color
- Size
- Shape
- Relationship



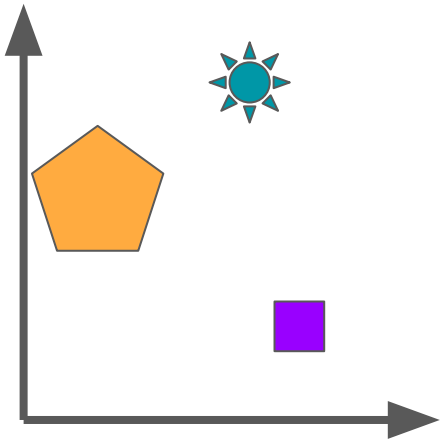
Dimensions of representation

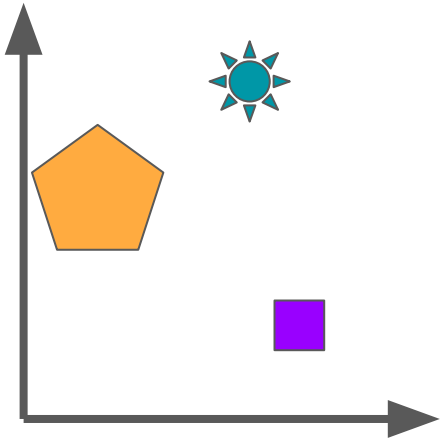
- Position
- Color
- Size
- Shape
- Relationship



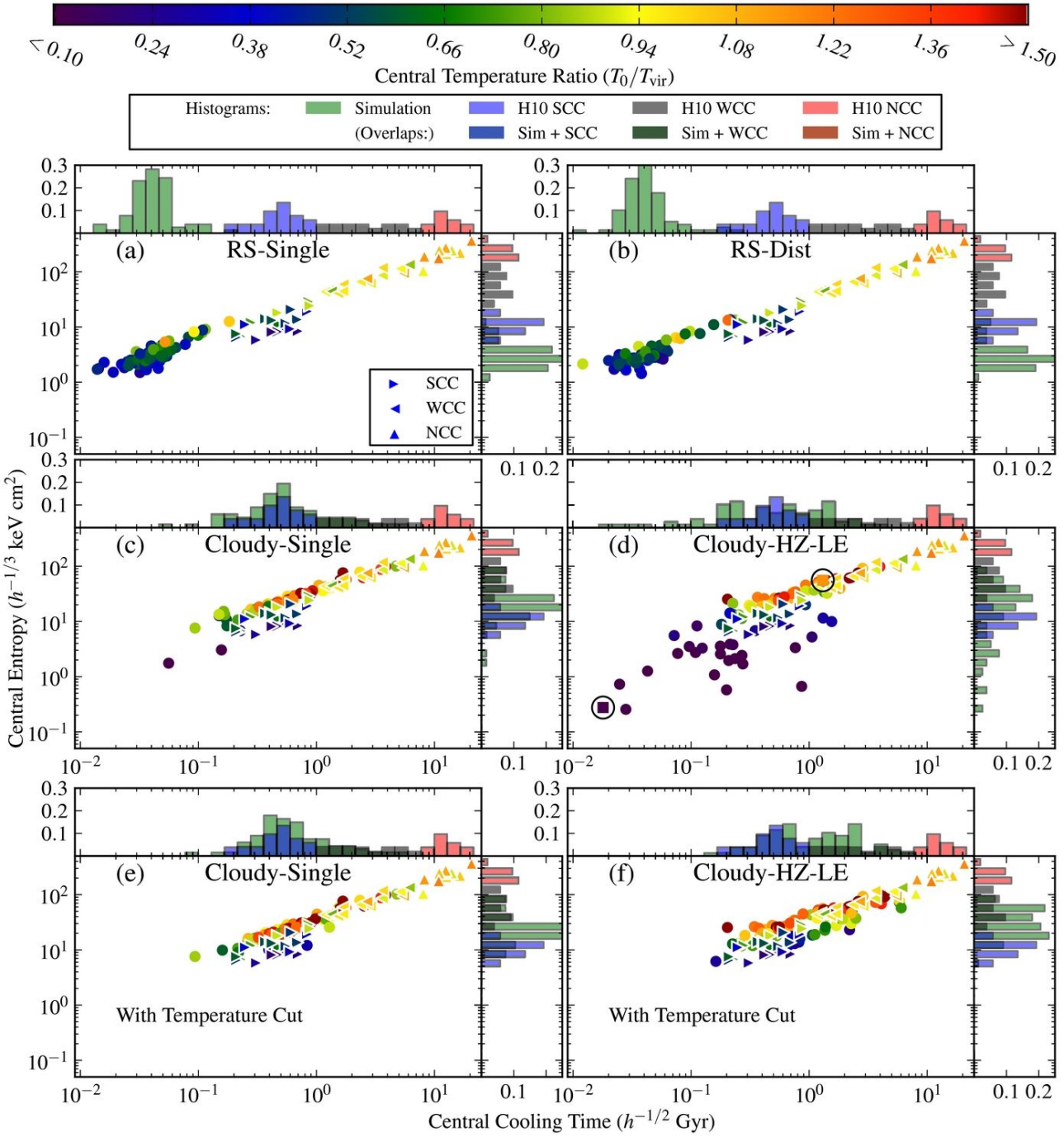
Dimensions of representation

- Position
- Color
- Size
- Shape
- Relationship





Components of a Visualization: Case Study



Skory et al, 2013

http://adsabs.harvard.edu/abs/2013ApJ...763...38S

Components of a Visualization: Case Study

http://vis.sciencemag.org/space-graveyard/

Topics

- Very simple visualization
  - Bivariate
  - Trivariate
  - Simple Binning
- Adjusting axes and properties
- Binning
  - Manual in 1D and 2D
  - Automated in ND
- Interaction with Jupyter widgets

Next Up

- Histograms
- Distributions

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