Getting the word out:

Digital imaging, file formats, and ggsave

Jackie Wirz & Steven Bedrick & Alison Hill

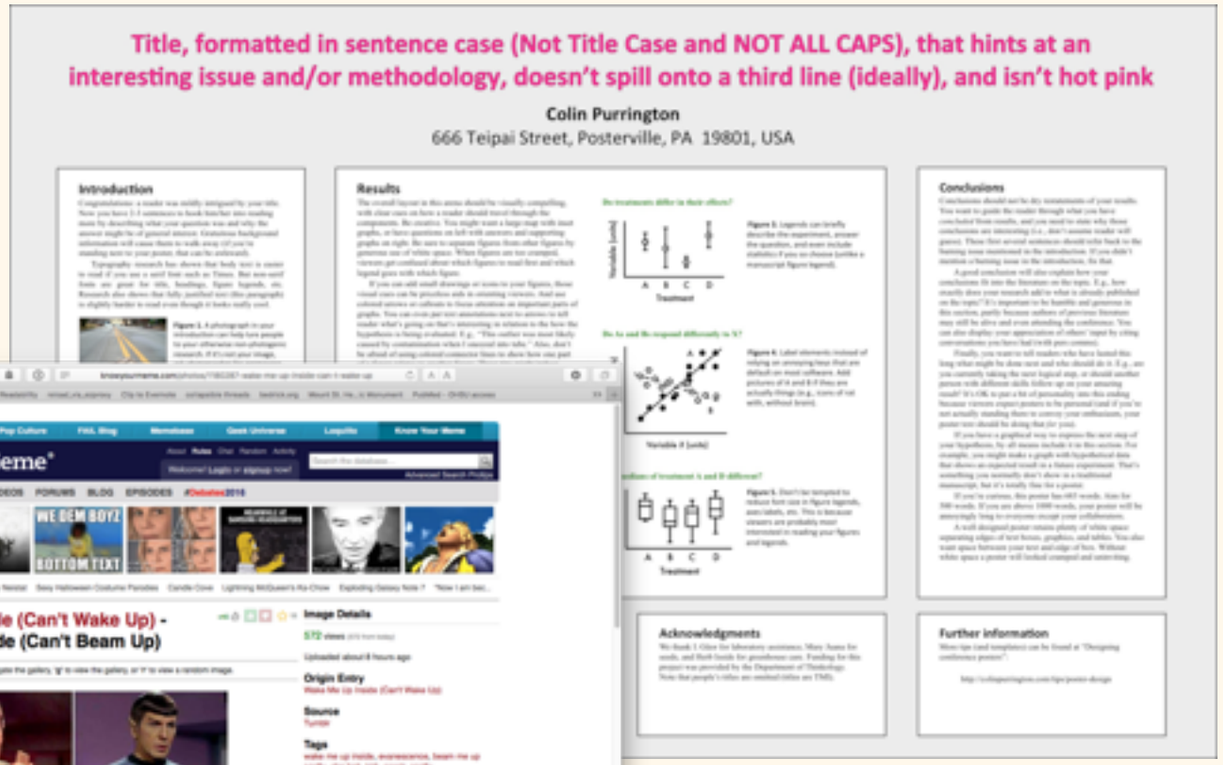
CSE 631, 10/13/16

Figure: https://medievalbooks.nl/2013/11/05/where-are-the-scriptoria/

You've made your amazing visualization...

You've made your amazing visualization...

... now you've got to get it out there!

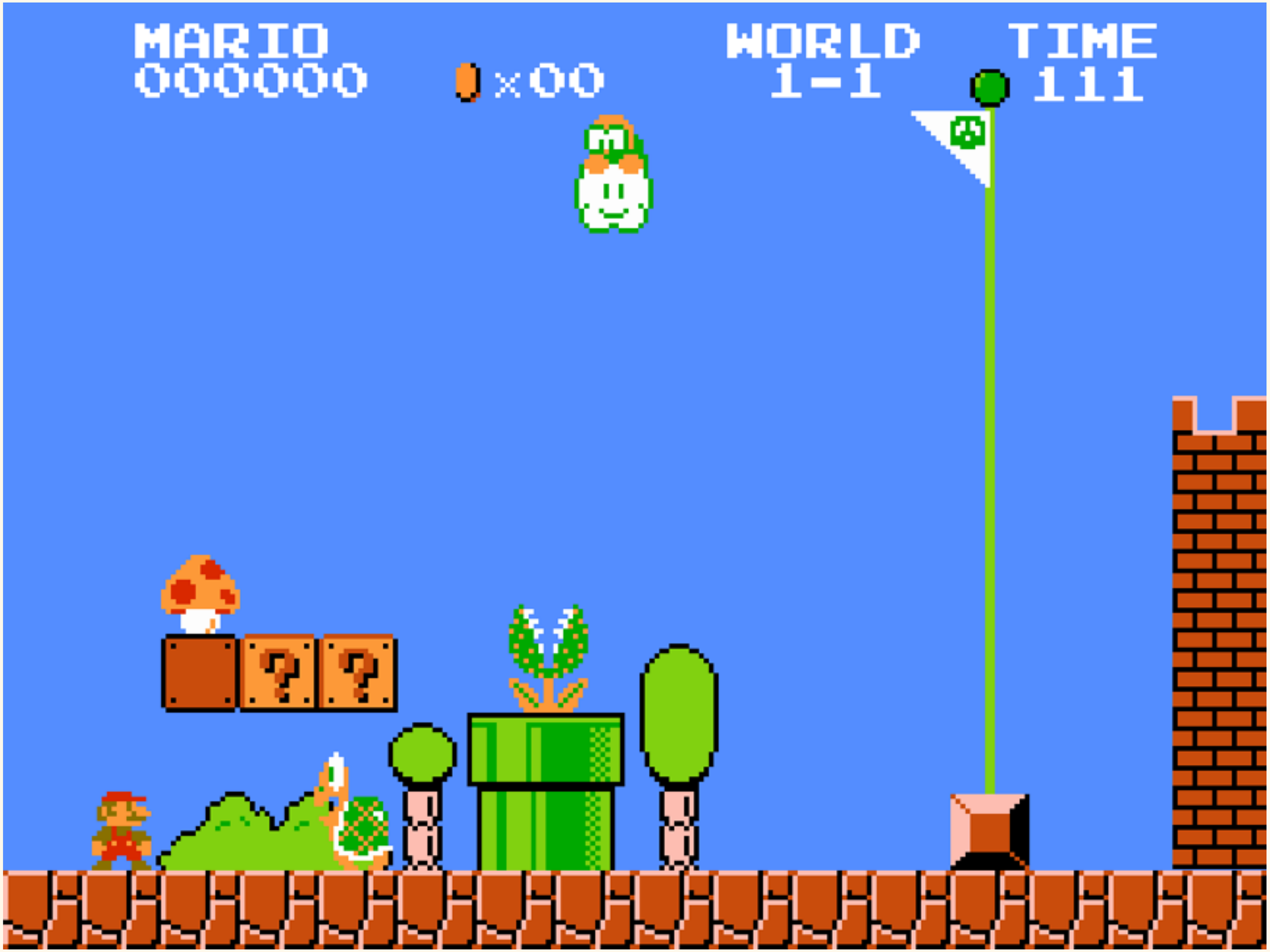


http://colinpurrington.com/wp-content/uploads/2011/09/scientific-poster-advice-purrington.jpg

A
B
B
C
CETE
DANI
GALLEA
PEPEGRI
NA
OSTRAFIORD
QVEDEFIORD
LANGE
XII.P.
HECESTHORRENDA
CARIBDIS
F
D
MIDELFIORD
REDH
HOLM
TROMS
EGGE
NALANGER
HASEVOG
LOKHELLE
B
VIKFIE.
FELBY
AVVER
GILLEFIORD
V
SANDER
GIGASVND
SKARTASVND
VANDVAD
ROLLEN
TRŌDANES
DVVANES
ANDANES
DOMVS
BIRCARORVM
SCONGEN
VNDER
VAL
K

Computers (mostly) *display* images as a grid of individual colored squares:

The fundamental unit (square) is a *pixel*.


MARIO
000000
×00
WORLD
1-1
TIME
111

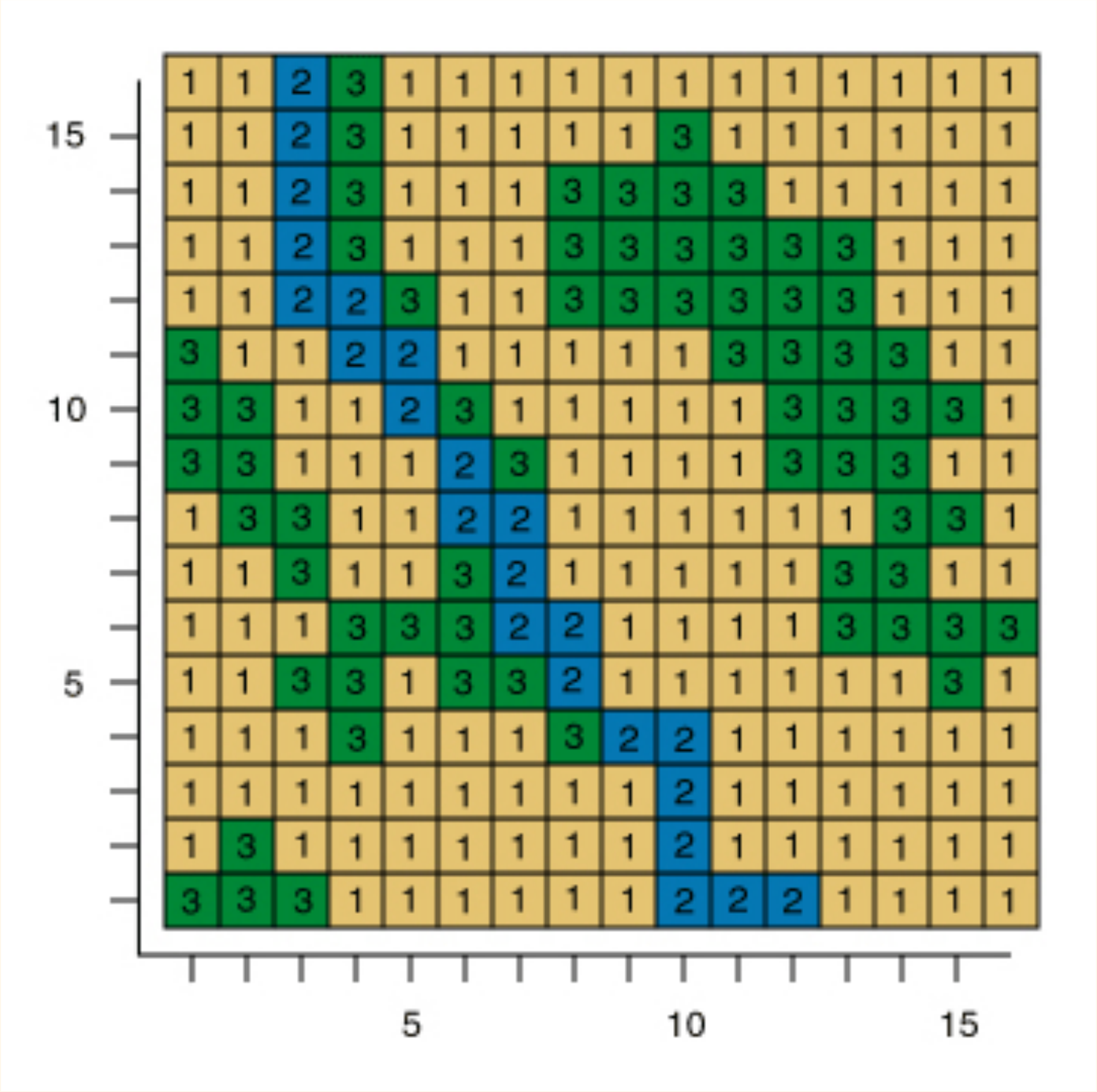
That's for *displaying* images...

What about *storing* images?

There are two main ways to store images:

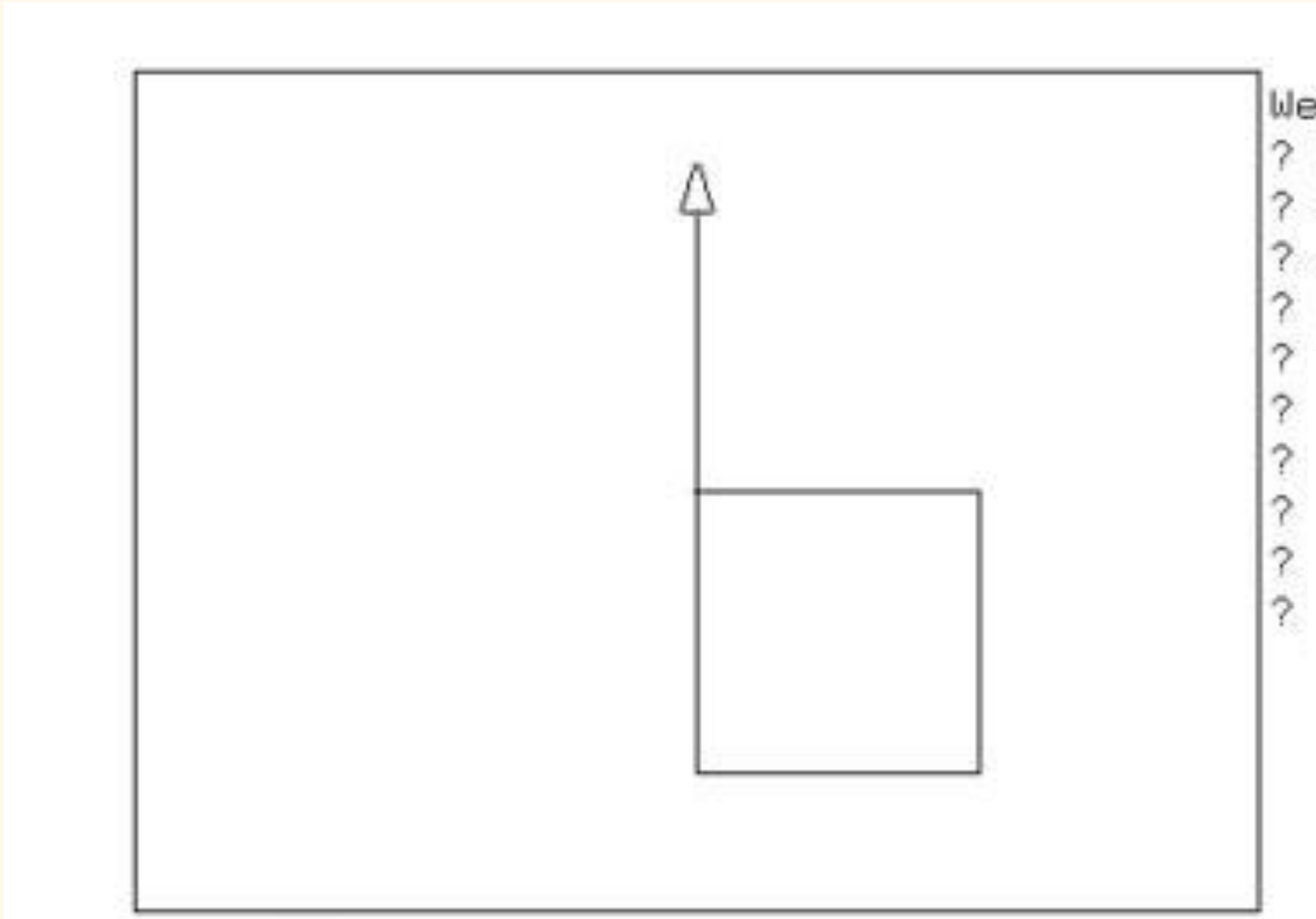
Raster files and *Vector* files.

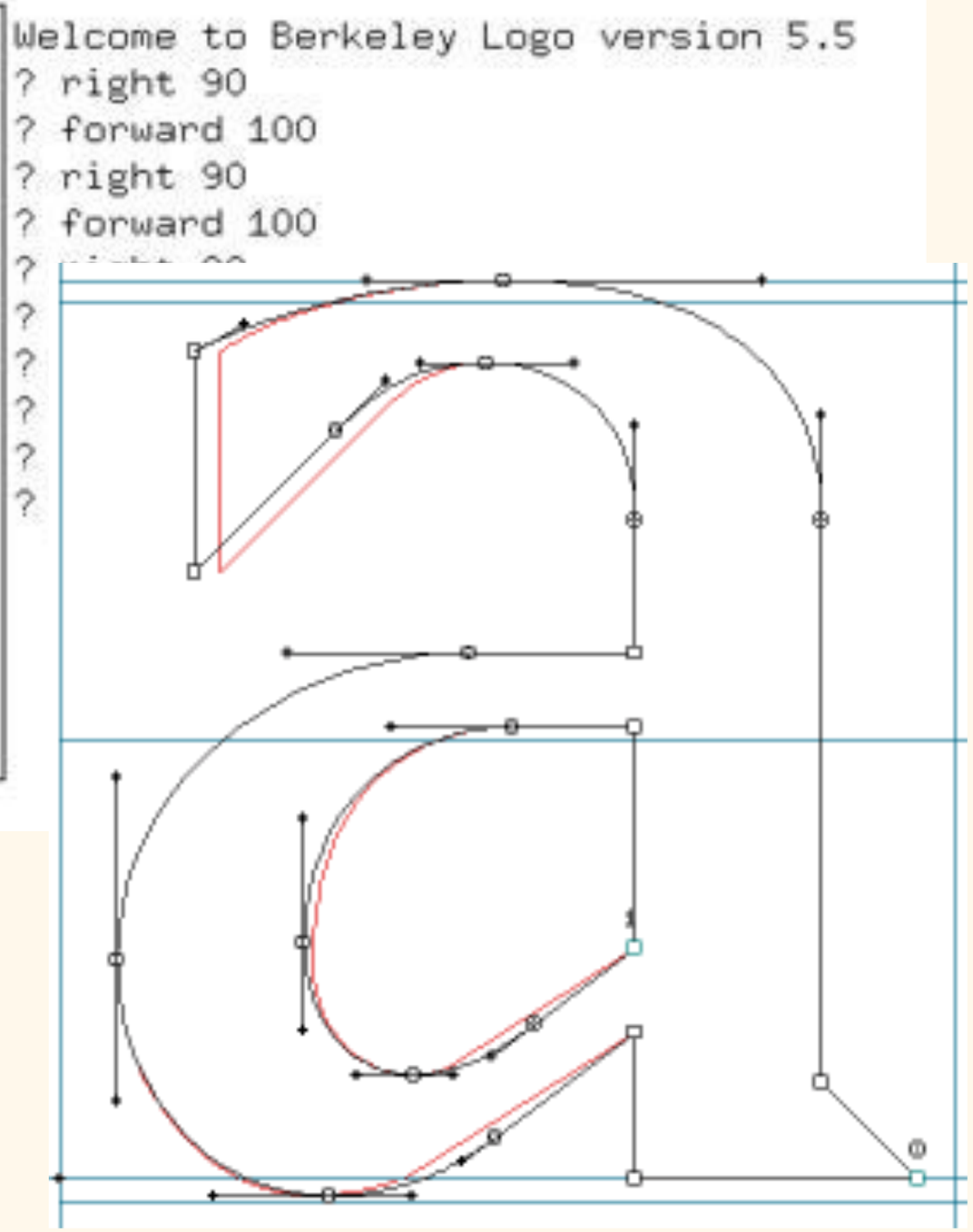
Raster files store the actual values of the pixels comprising the image:



https://www.e-education.psu.edu/geog486/node/1879

Vector files store mathematical instructions to recreate the image:

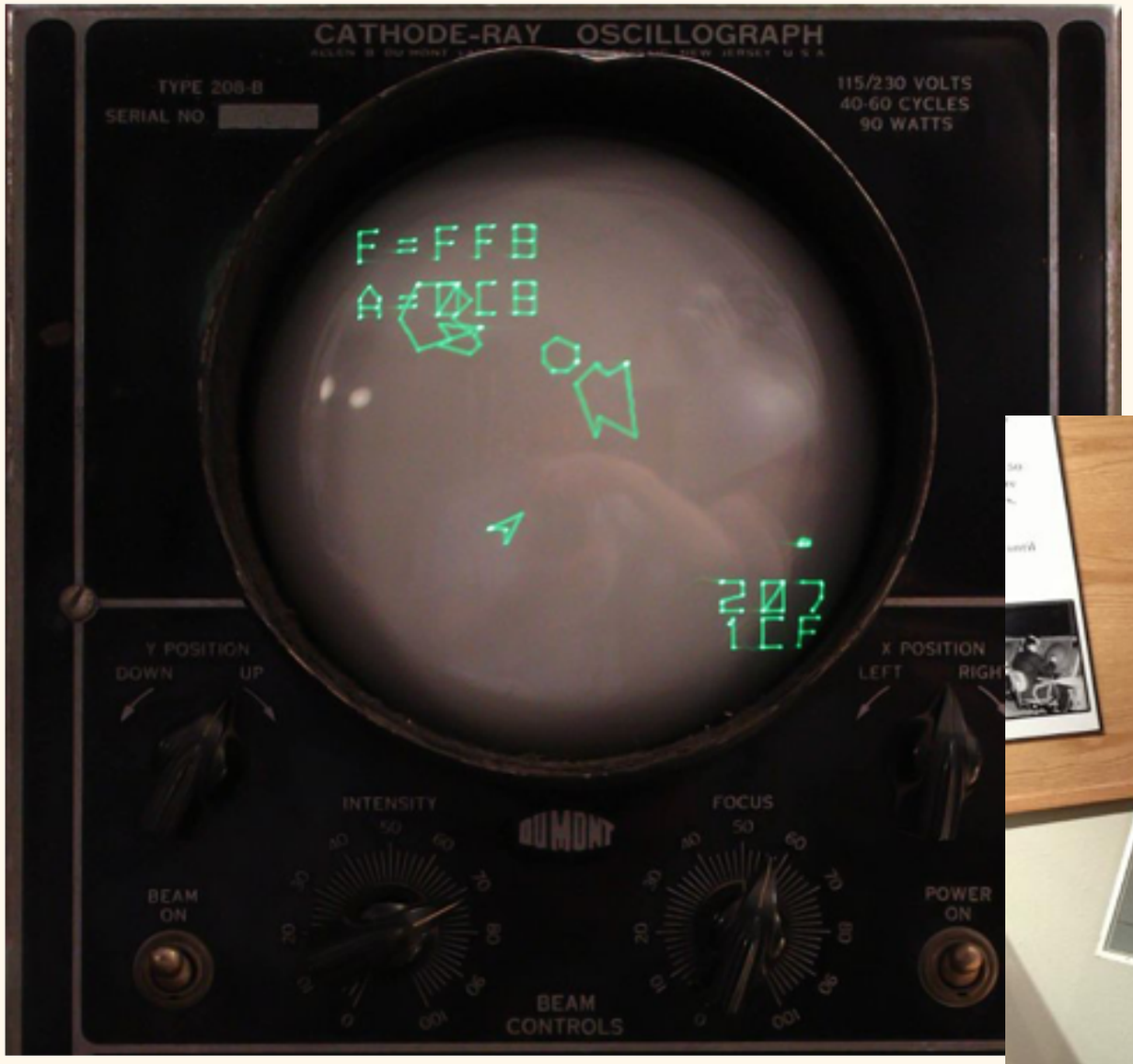


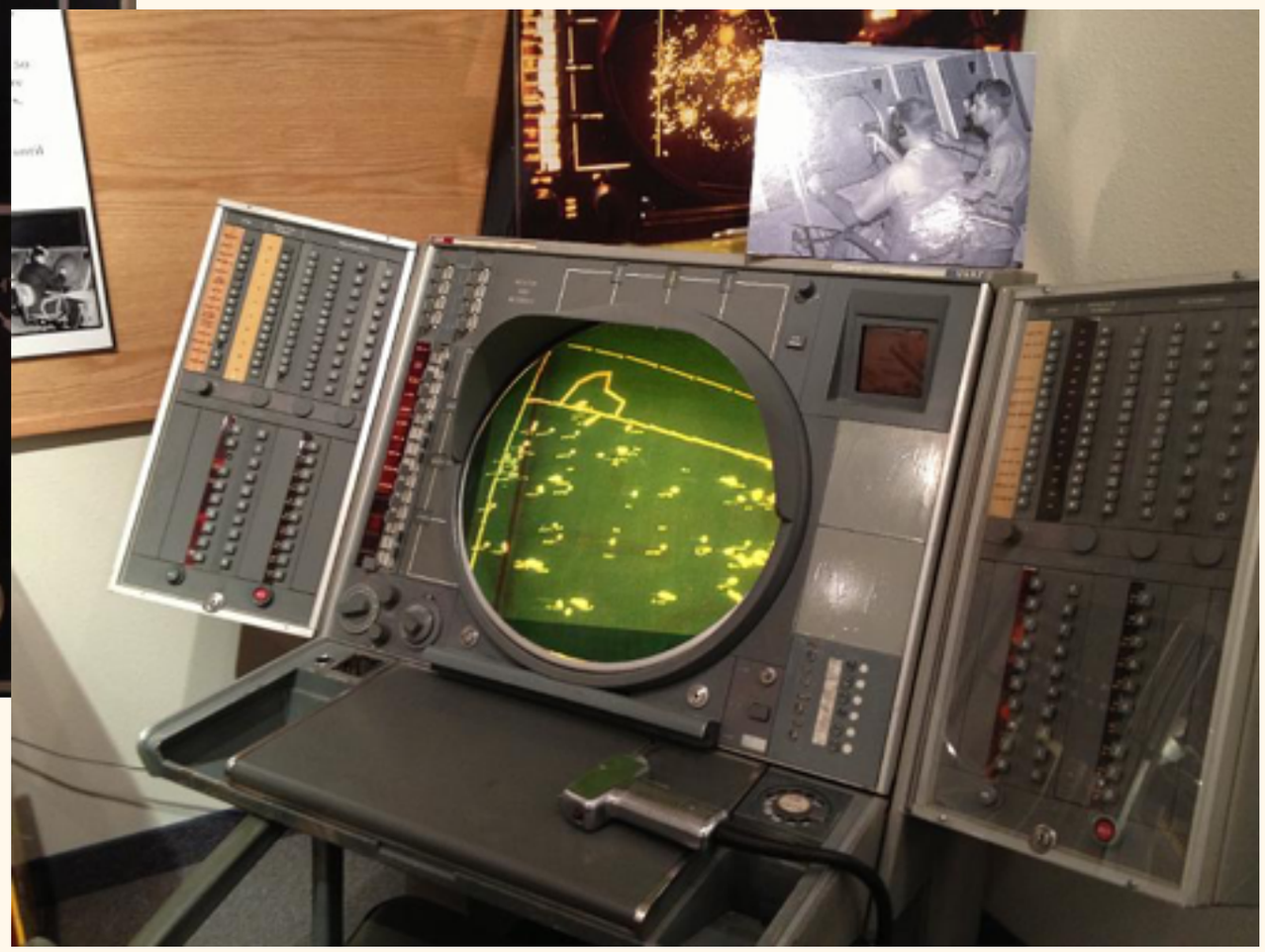


http://www.emigre.com/FF/MATRIX7.gif

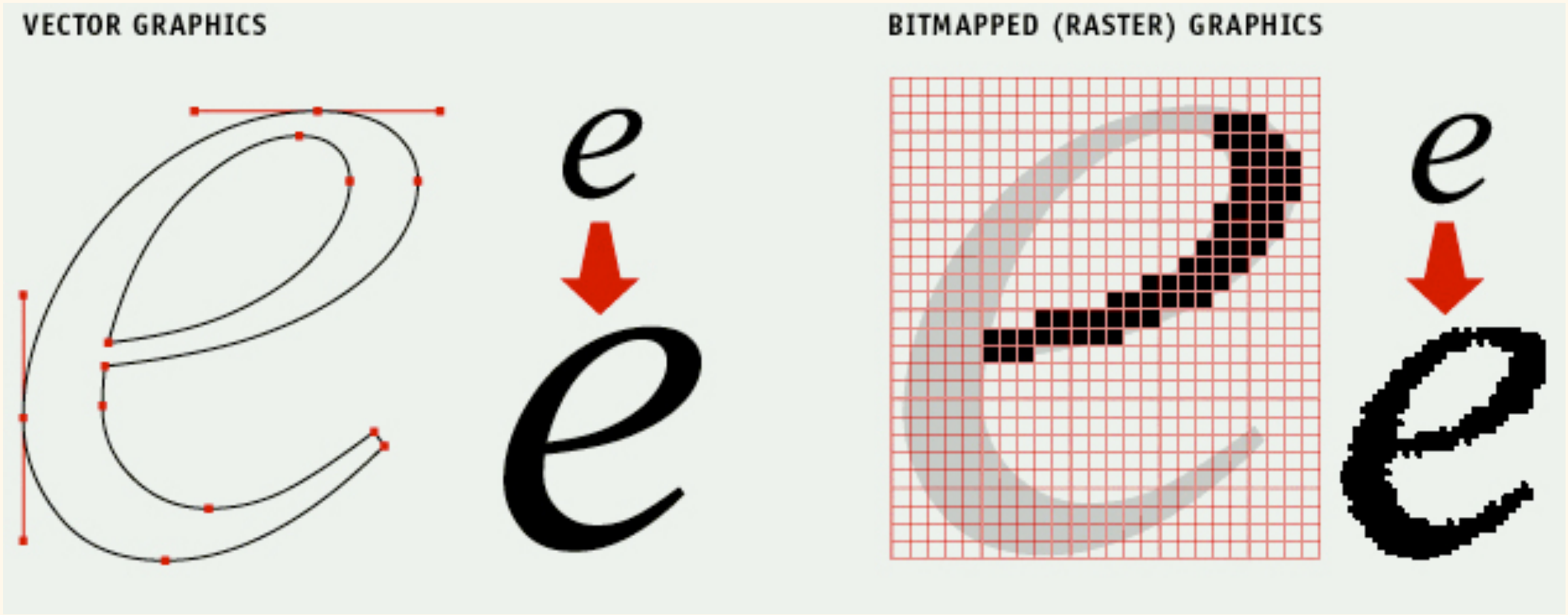
https://www.csee.umbc.edu/courses/201/spring10/assignments/proj_recursion.shtml#(6)

Fun fact: the first computer graphics systems used *vector* displays!



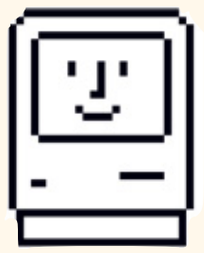


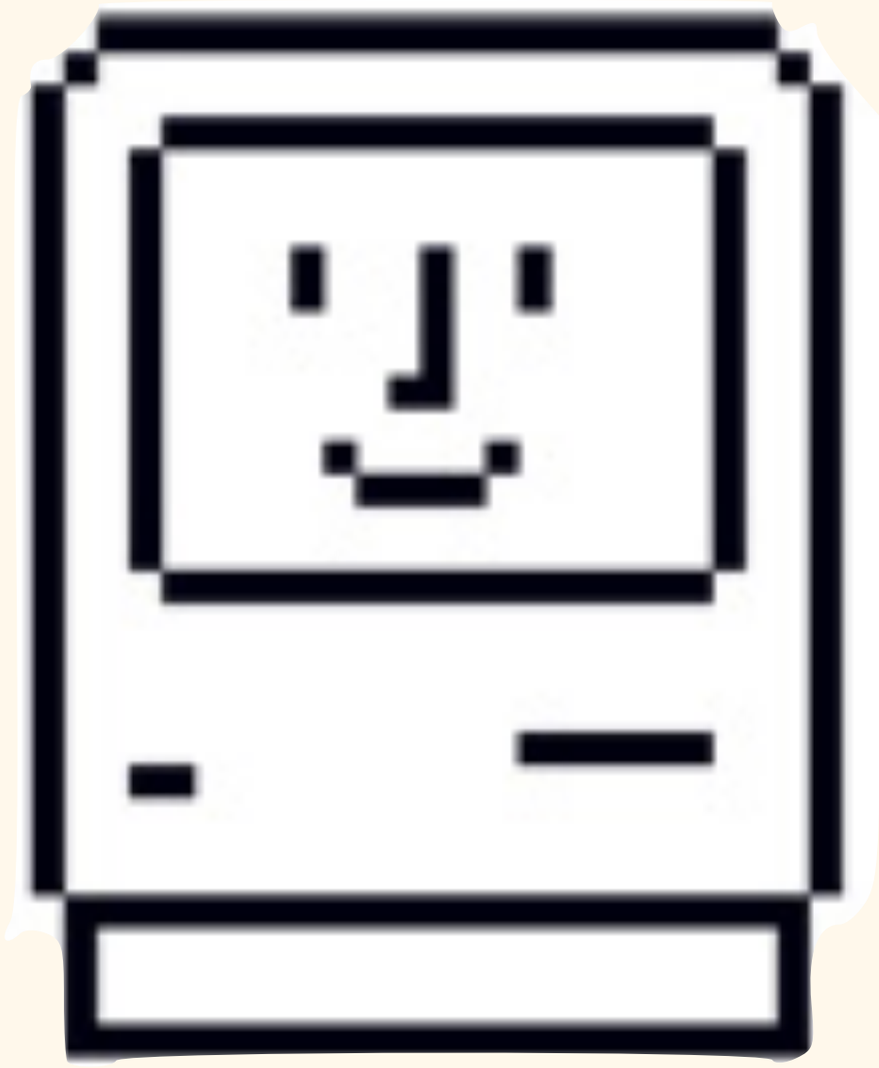
Raster images contain a finite amount of information:

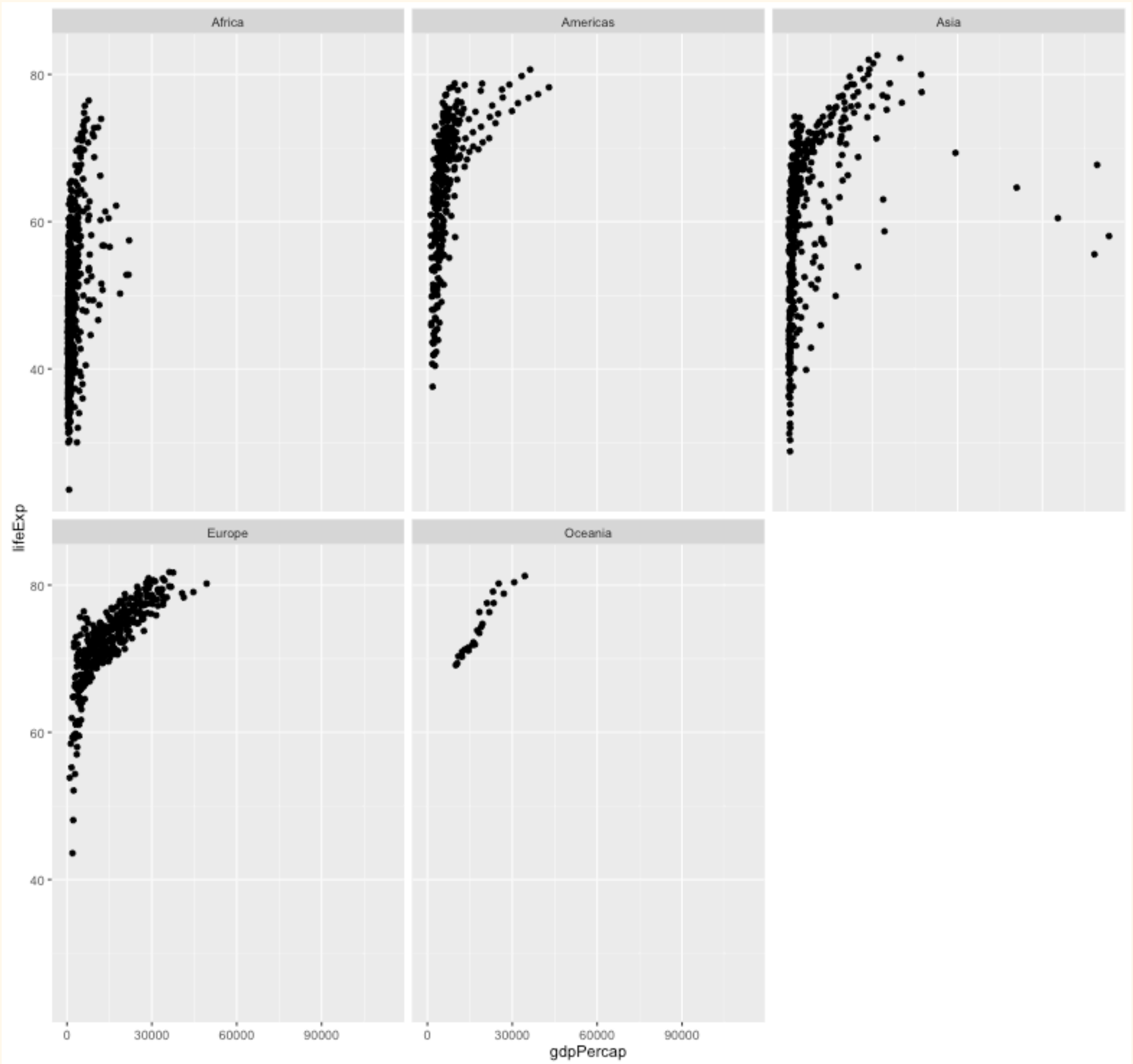


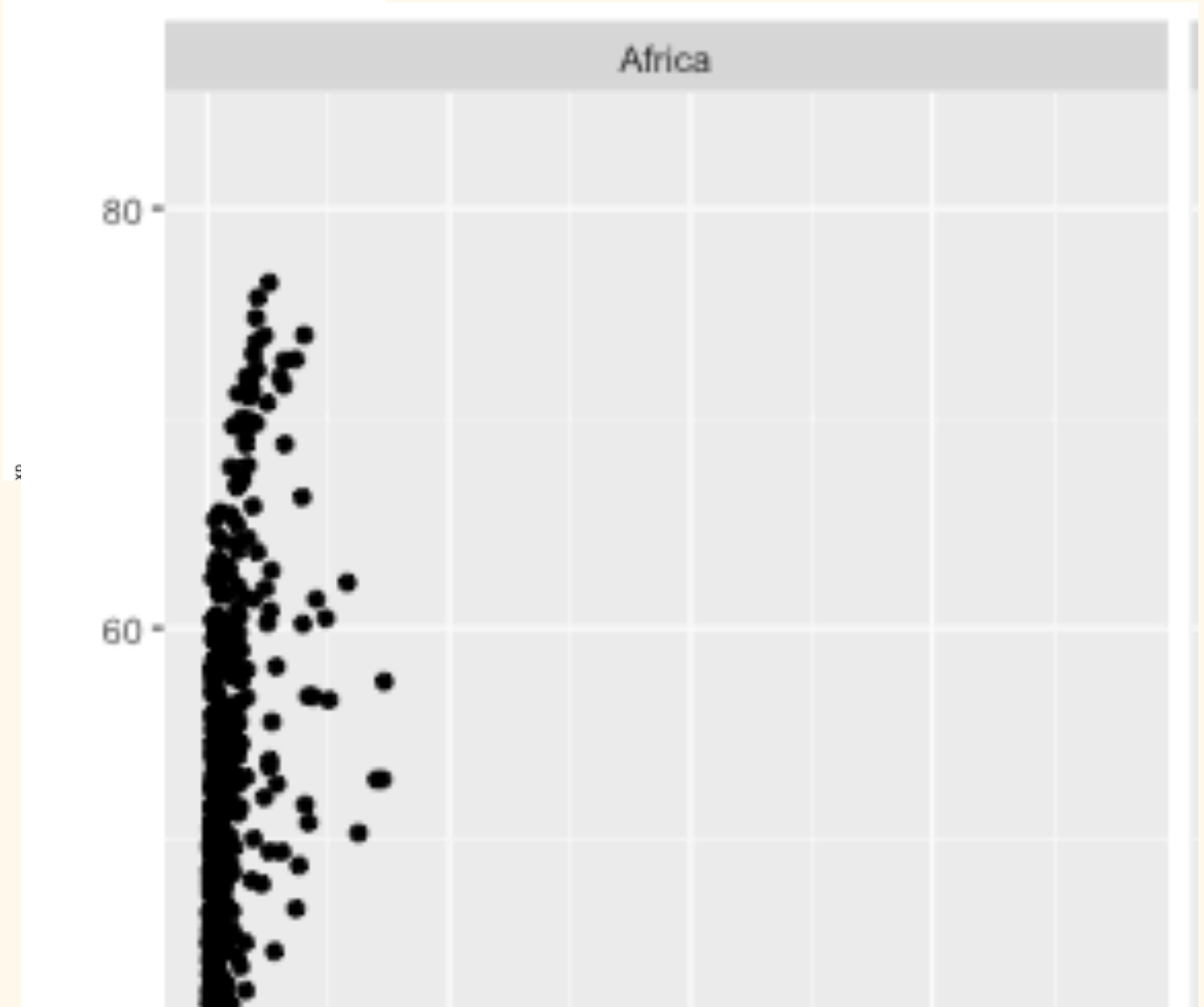
http://www.amywaggs.com/wp-content/uploads/2015/02/vector-raster.jpg

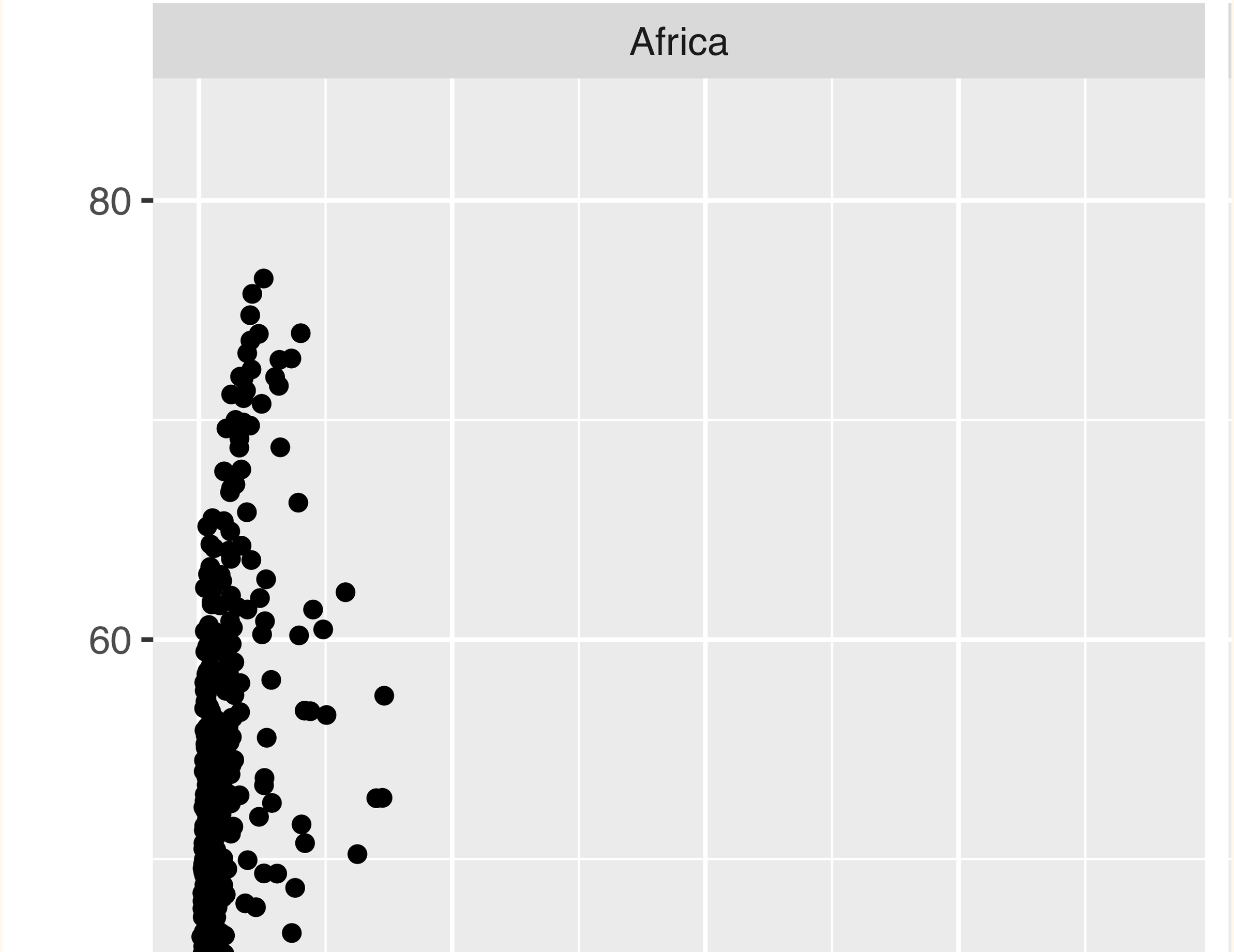
Raster images contain a finite amount of information:






Africa
Americas
Asia
Europe
Oceania
lifeExp
80
60
40
0
30000
60000
90000
gdpPercap

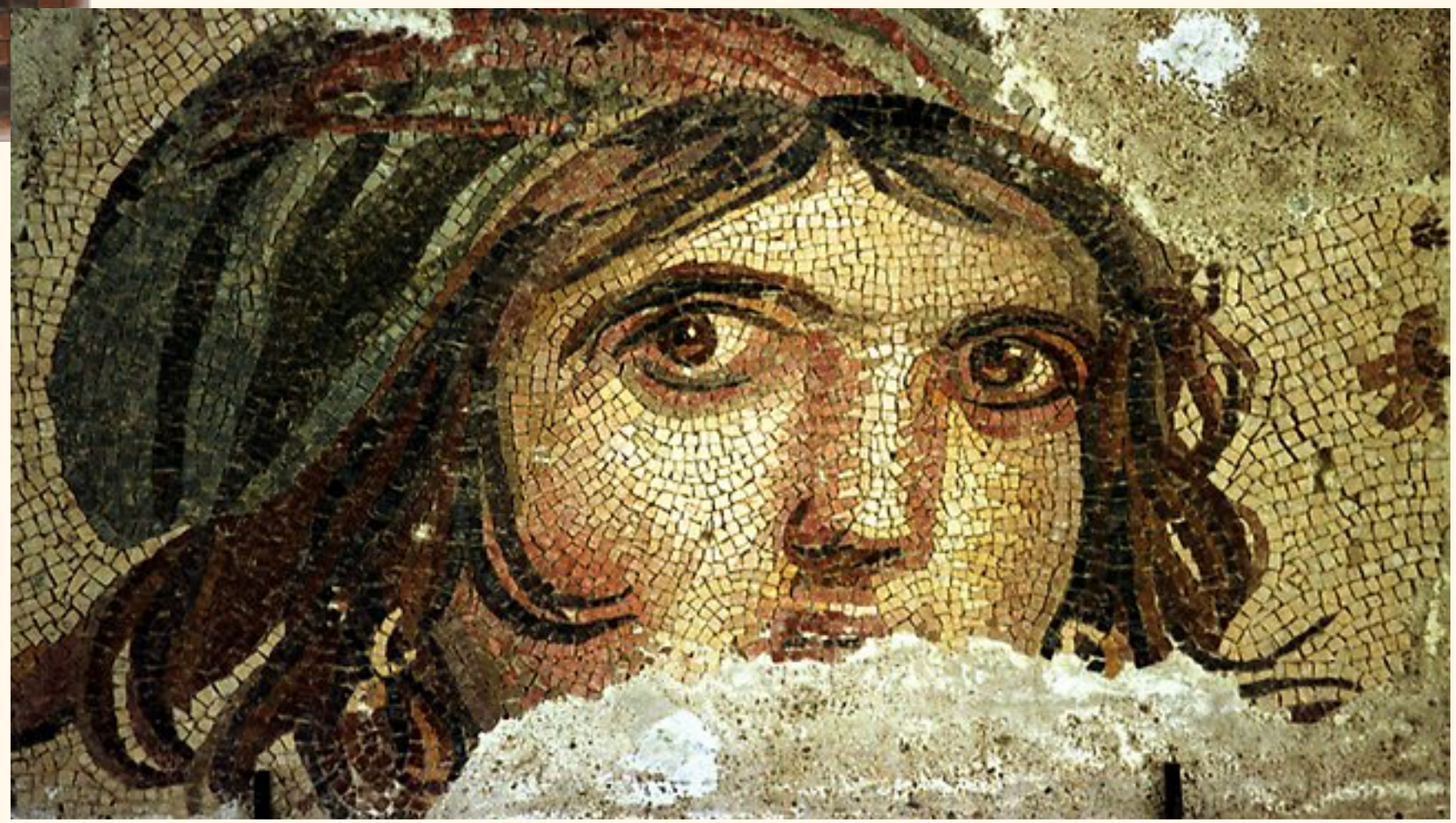

Africa
80
60

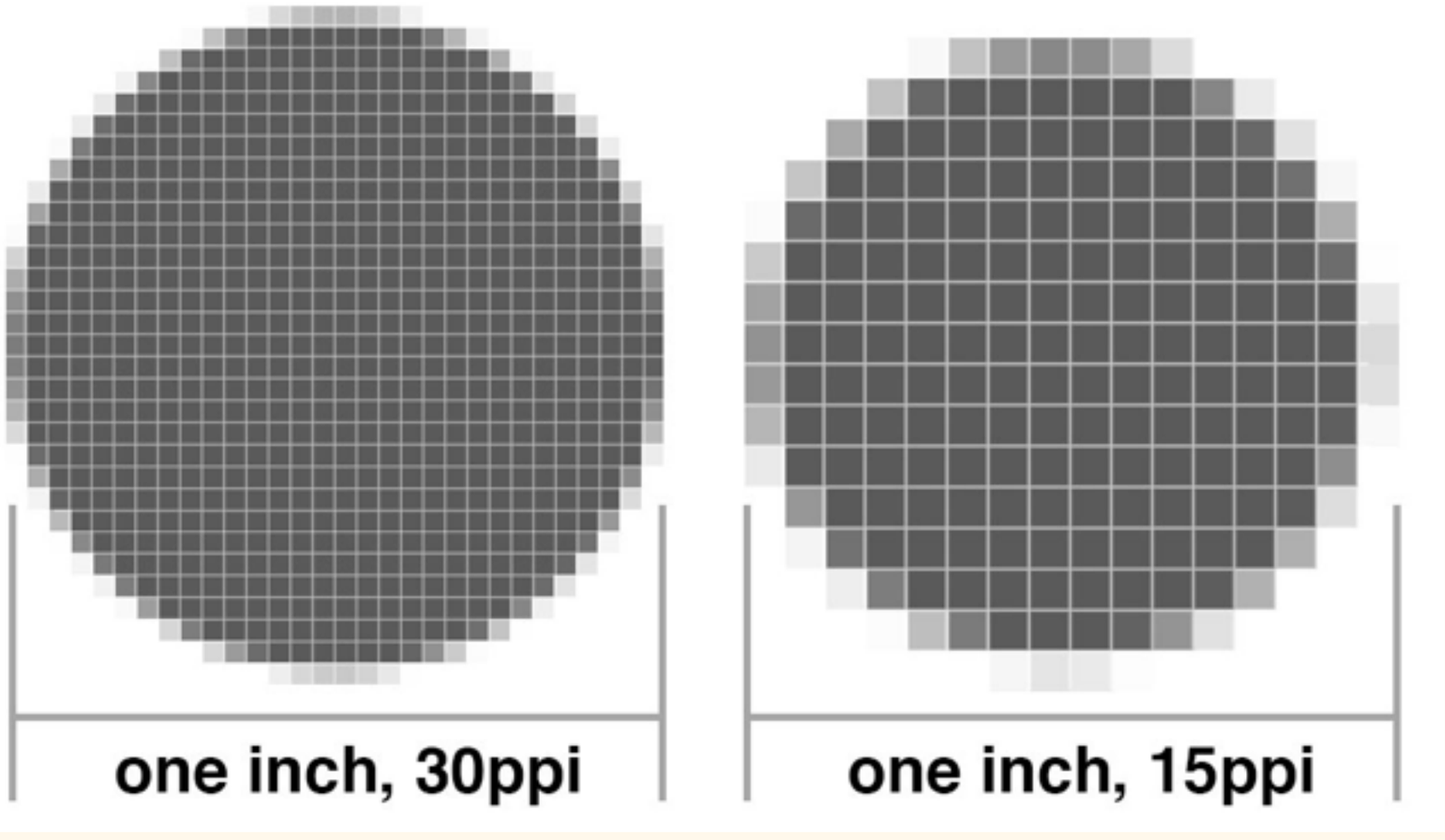

Africa
80
60

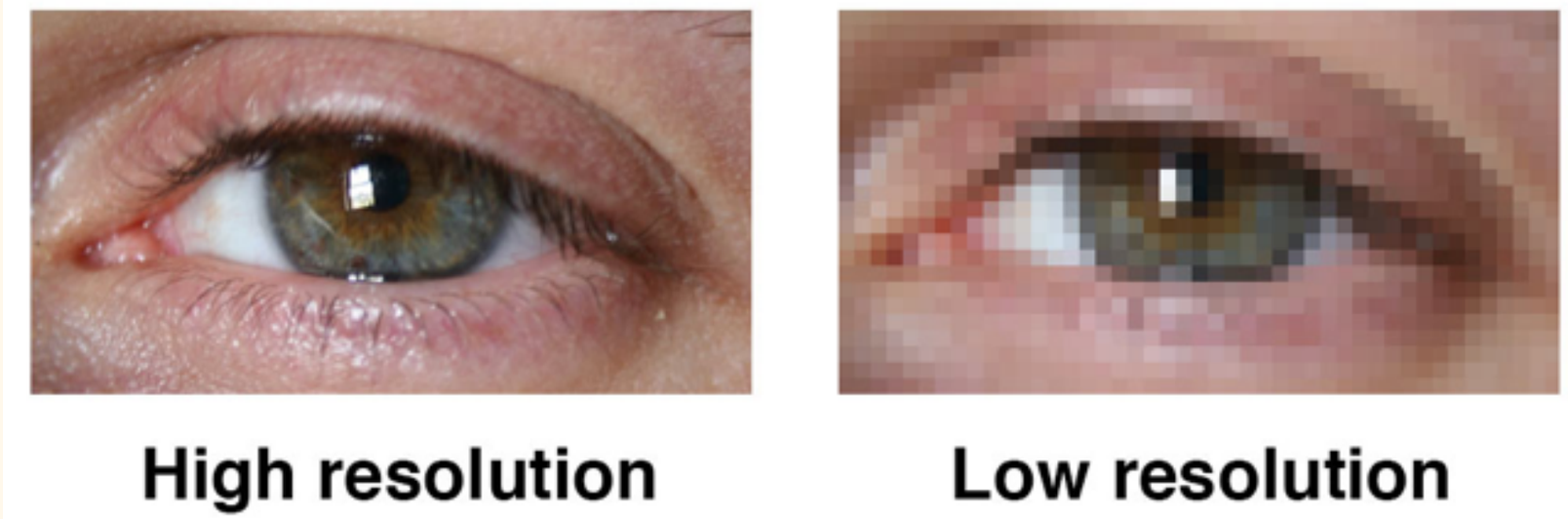
An image's *resolution* refers to the amount of space represented by each pixel.

"Higher resolution" → "Smaller" pixels

More information per unit of real space.







http://art.nmu.edu/groups/cognates/wiki/8962d/Resolution_PPIDPI.html

Print dimensions and resolution are related:

Important notes:

Many raster formats *do not* inherently specify DPI or print dimensions!

Some operating systems and layout programs *try to be clever* about figuring out DPI and size.

Others *do not*.


SMOKEY
ONLY YOU

Raster file formats:

BMP, TIFF, JPEG, GIF, PNG, etc.
Usually larger files
Good for: photographs, screen display*

Vector file formats:

PS, EPS, PDF, SVG, etc.
Usually smaller files
Good for: diagrams, graphs, print

Rules of thumb:

When possible, stick to vector images...

When you have to rasterize:

- 150 DPI or higher (300 preferred)
- For diagrams, PNG or TIFF file format

Wait as long as possible to rasterize!

Once resolution is lost, you can’t get it back.

Demo time!

Meet the ggsave() command:

```
>>> plt <- ggplot(...)

>>> ggsave(“my_file.pdf”, plt)

>>> ggsave(“my_file.png”, plt)

>>> ggsave(“my_file.pdf”, plt, width=7, height=5)

>>> ggsave(“my_file.png”, plt, width=7, height=5, dpi=300)
```