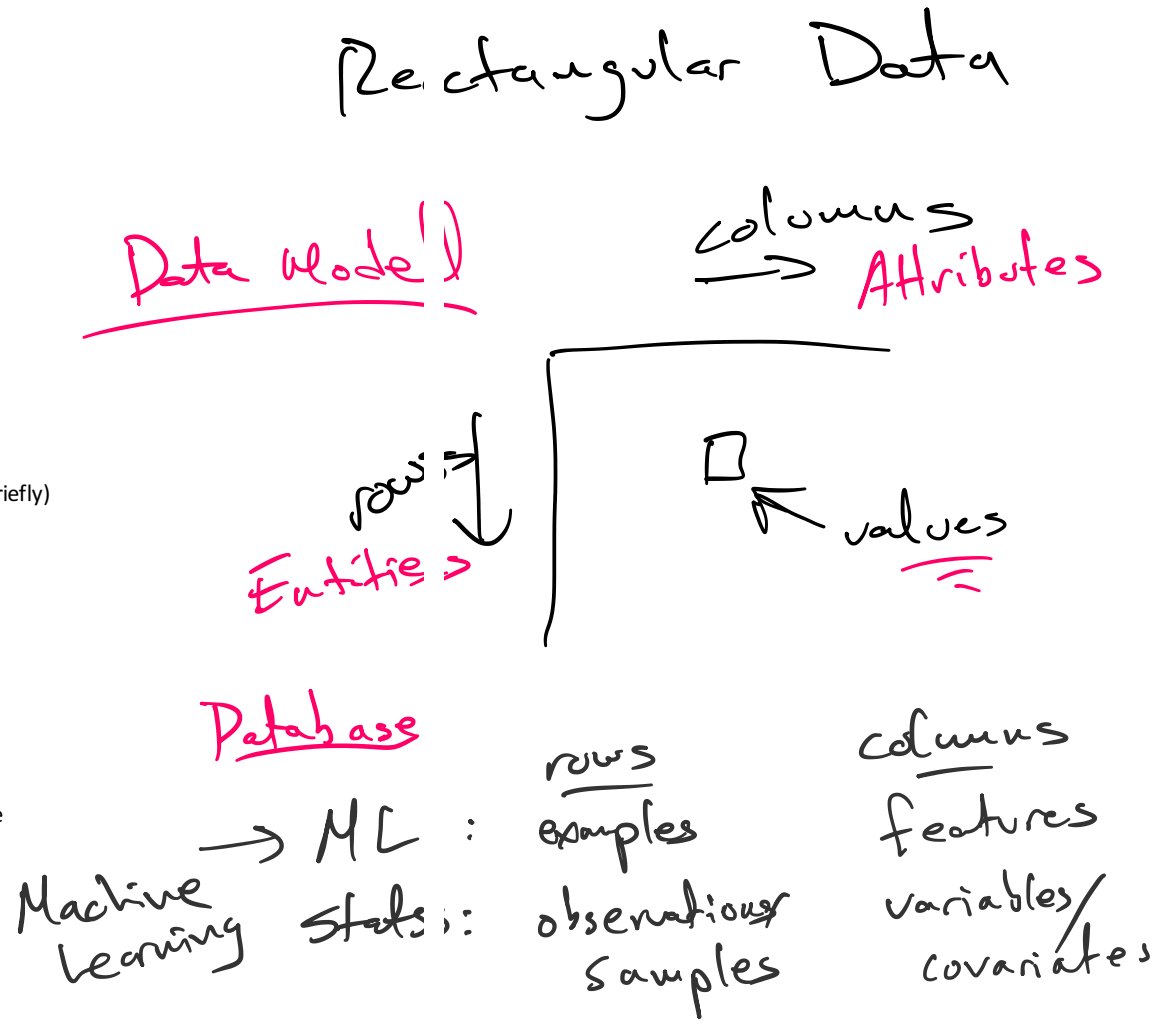
# Lesson Plan 1/30

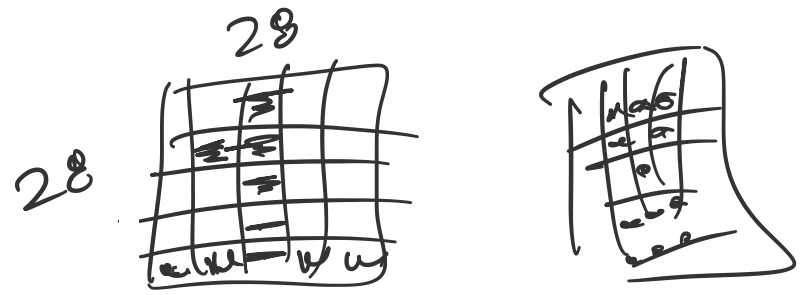
Tuesday, January 30, 2018 10:48 AM

- Admin
  - Office hours
  - Office hours sign-in sheet
  - ELMS for communication
  - Github
- Rstudio
  - Layout
  - Packages
  - Help
  - Script window
- Rectangular Data
  - Entities/Attributes
  - Translation table
  - Other kinds of data as rectangular (briefly)
- Data types
  - Categorical attributes
    - Unordered
    - Ordered
    - Factors
  - Numerical attributes
    - Discrete
    - Continuous
  - Distinctions
    - Discrete vs. continues
    - Ordered categorical vs. discrete
  - MNIST
  - Other data types
    - Datetime
    - Geolocation
  - Units
  - **EXAMPLE (flights)**

Select/rename



- Select/rename
- Slice/filter
  - How to create vectors
- Sample_frac, sample_n
- Pipeline
  - Pipe operator
  - General data -> transform -> idea
- Another pipeline example with Baltimore data




columns

| digit | pixel1 | pixel2 | ... | {x, y, intens} |
|---|---|---|---|---|
| 1 | 28 | 0 | 220 | ... |
| 8 | 0 | 250 | 220 | |

rows

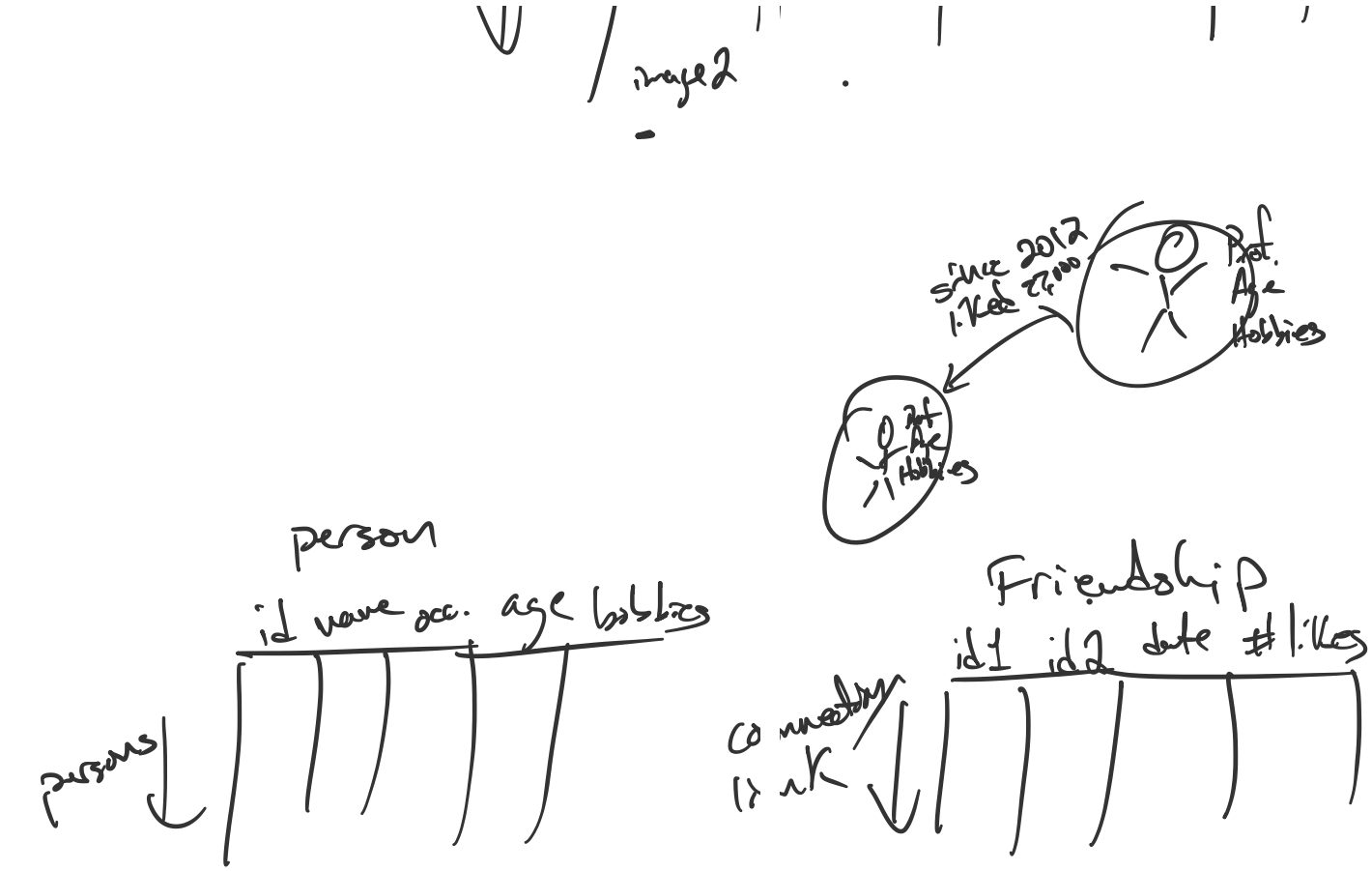
Data:
- fully described in dataset

pixels

| image | x | y | intens |
|---|---|---|---|
| image1 | 0 | 0 | 0 |
| image2 | 0 | 1 | 0 |

images

| id | digit |
|---|---|
| image1 | 1 |
| image2 | 8 |


since 2017
Prof.
Age
Hobbies
person
id name occ. age hobbies
persons
Friendship
id1 id2 date #likes
connected
is n:k

Datatypes

- Categorical
  - Take values from a finite set
  - Unordered
  - Ordered
- Numeric
  - Discrete
  - Continuous

Distinction

Ordered categorical vs

Discrete numeric

A 4
B 3
C vs 2
D 1

Operations on attributes

→ select / rename

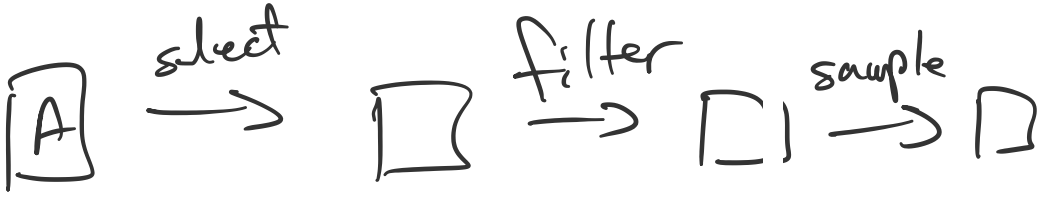
Operations on entities

slice(arrest_tab, c(1,3,10))

create vector/array

seq(start, end, by)

vectors: "c", "seq", ":"




select(A, x, y)

⇕

A %>% select(x, y) %>% filter(age < 18)