General Knowledge

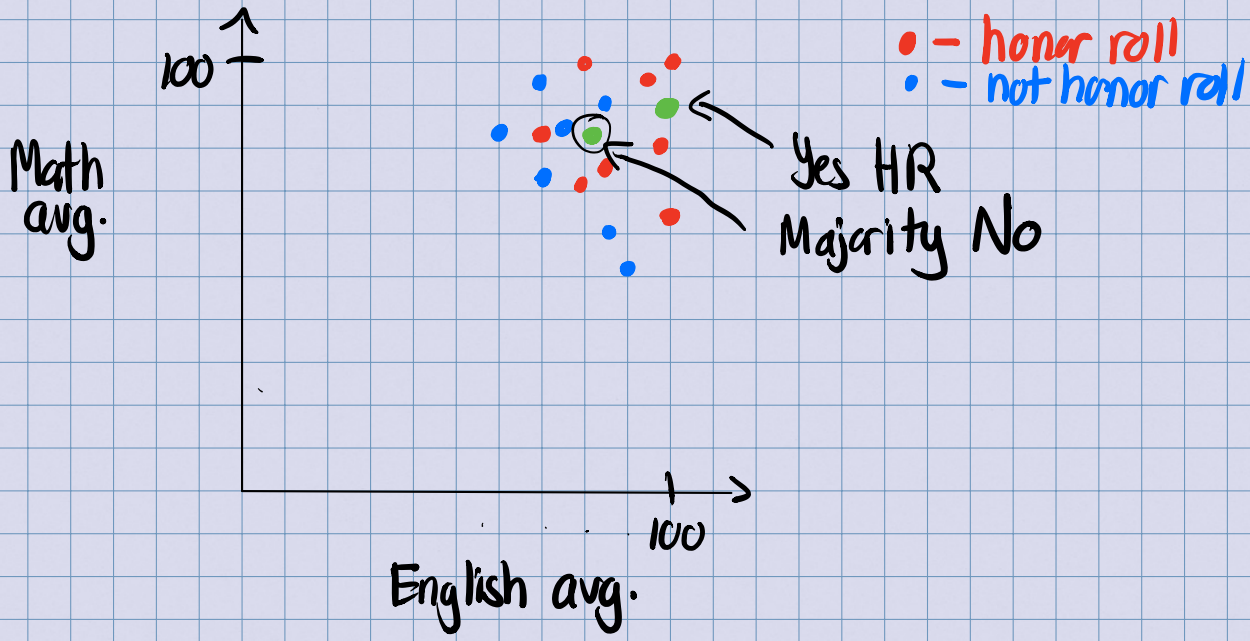
- "Models" (NN, predicting programs) require two things:
  - Data to classify/predict
  - Previous data to base our predictions off of
- All this data should be numeric; no words and such
  - To replace words/descriptions such as colors or yes/no

$$\left(\frac{*}{\text{color}}, \frac{36}{\text{length}}, \frac{8}{\text{fastiness}}\right)$$

red = 1
blue = 2
green = 3

k-Nearest Neighbors

- Example:



What was reasoning/intuition based off of data?

Close people aren't
Closer to average

The idea of k-Nearest Neighbors is to literally look at the k (# nearest pieces of data

Theoretical Implementation:

I) How to get "distance" from one to another?
   → Euclidean distance
II) How to get k-nearest?
   → Store in list, take first k; make sure to keep track of class

Potential Downsides

- Does not "weigh" things differently; if we were trying to classify something like fish with red = 1 and blue = 2, colors should have a big influence, but it might not
- Does not learn from mistakes/incorrect classifications