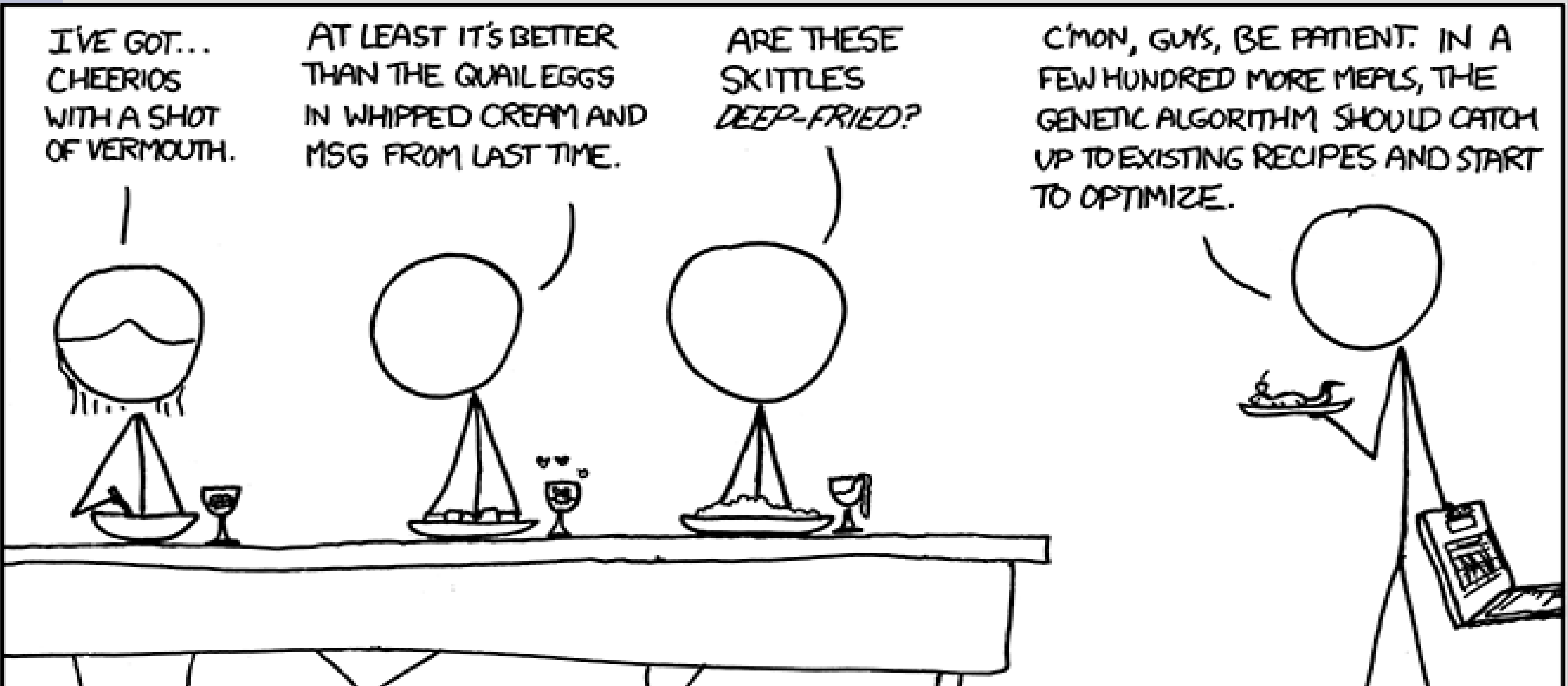
Multi-variable Optimization

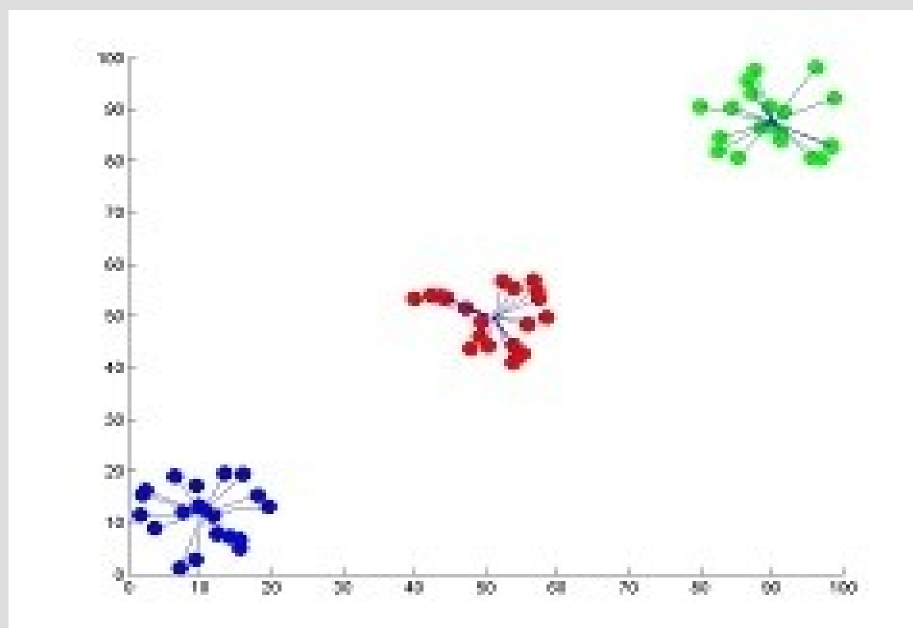


WE'VE DECIDED TO DROP THE CS DEPARTMENT FROM OUR WEEKLY DINNER PARTY HOSTING ROTATION.

K-means clustering

K-means clustering on points is finding K “central locations” that reduce the distance of each point to the nearest “central location” (summed over all points)

K=3



K-means clustering

For examples like the previous one, it is easy to find which points should be “grouped together”

Once you have a group of points, you can mathematically find the best “central location”

(“center of mass” with equally massive points)

$$center_x = \frac{1}{|G_{center}|} \sum_{i \in G_{center}} x_i$$
$$center_y = \frac{1}{|G_{center}|} \sum_{i \in G_{center}} y_i$$

K-means clustering

Suppose you wanted to find the best spot to put 5 “central locations” here:

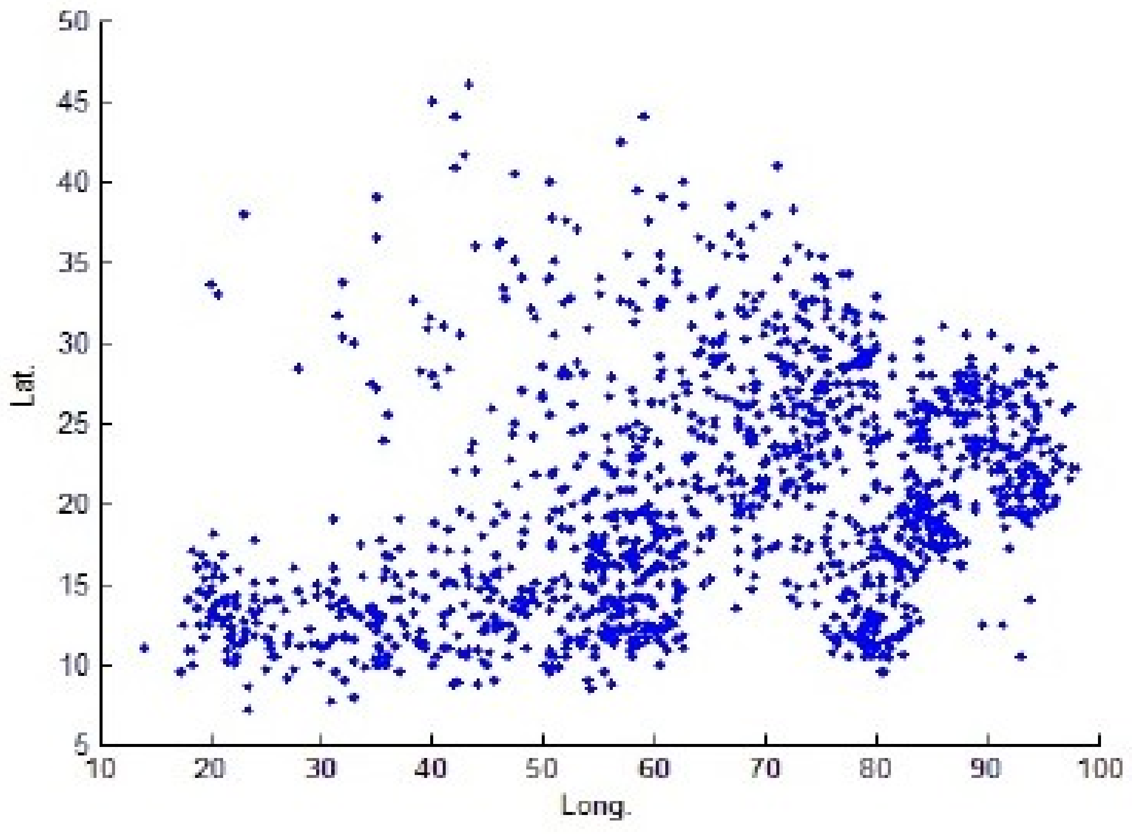


Figure 3. Starting points of all hurricane tracks

K-means clustering

Suppose you wanted to find the best spot to put 5 “central locations” here:

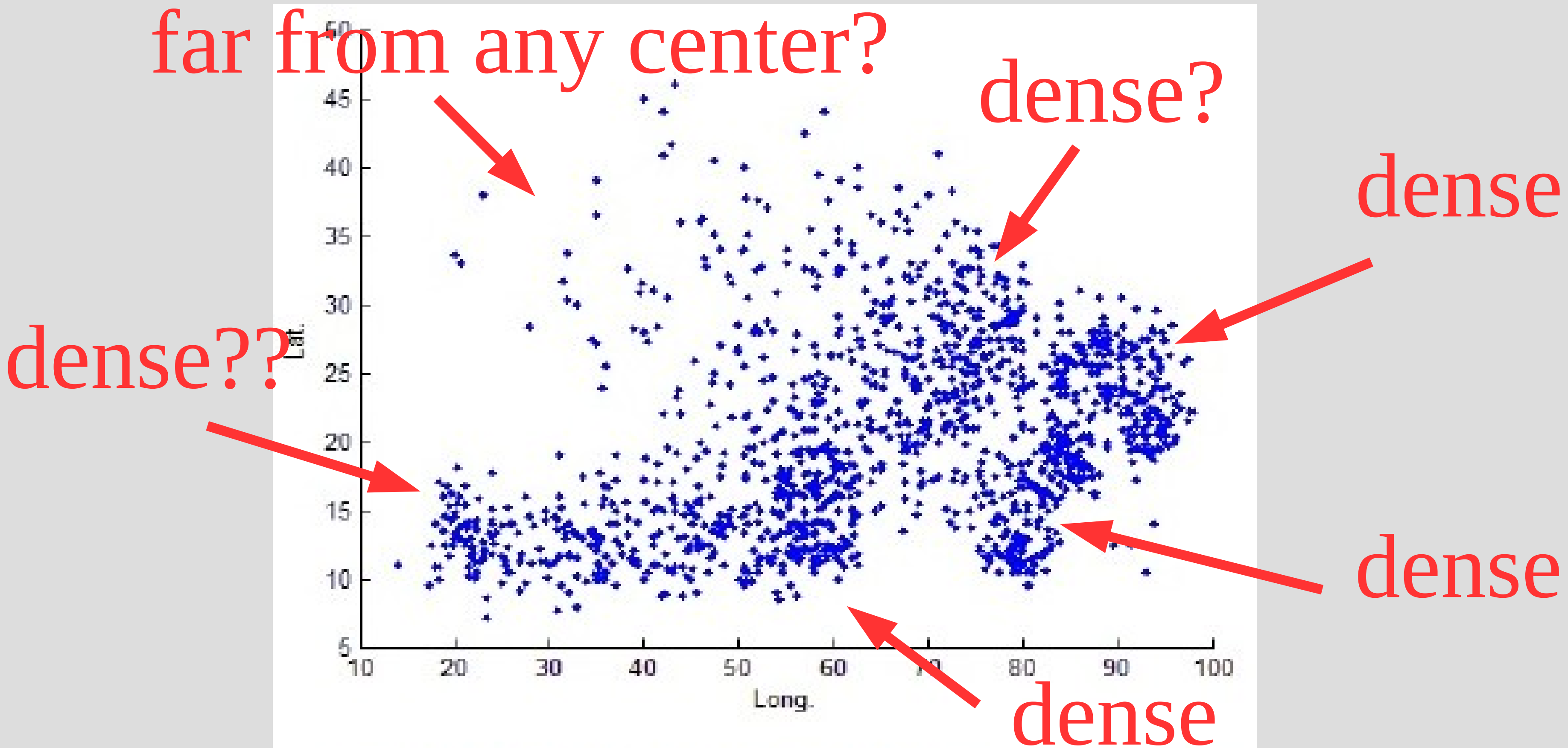


Figure 3. Starting points of all hurricane tracks

K-means clustering

Turns out you can do this the other way around as well...

If you have the “central locations” (x,y) coordinates, you can find which location all points should go to (minimum distance)

Other random topics

We have a problem:

1. If we knew point groupings, we could find the best central locations

2. If we knew central locations positions, we could find point groupings

Other random topics

One common way to solve this issue when you have multiple unknowns that depend on each other is to simply guess, then try to optimize

So, initially just make random groupings

Then find the best central locations base off of the groupings

Then find the best groupings... and repeat

Other random topics

If you set up the problem correctly (and have a “well behaved” metric), this will converge

In fact, you can do this even if you have more than two unknowns

Just make one variable while fixing all others and optimize that one
... then pick a new variable to “optimize”

Other random topics

This technique actually works in a large range of settings:

K-means clustering (this)
Bayesian networks (probabilistic reasoning)
Markov Decision Processes (policy selection)
Expectation–Maximization (parameter optimization)