

Clustering III

—

Team 6

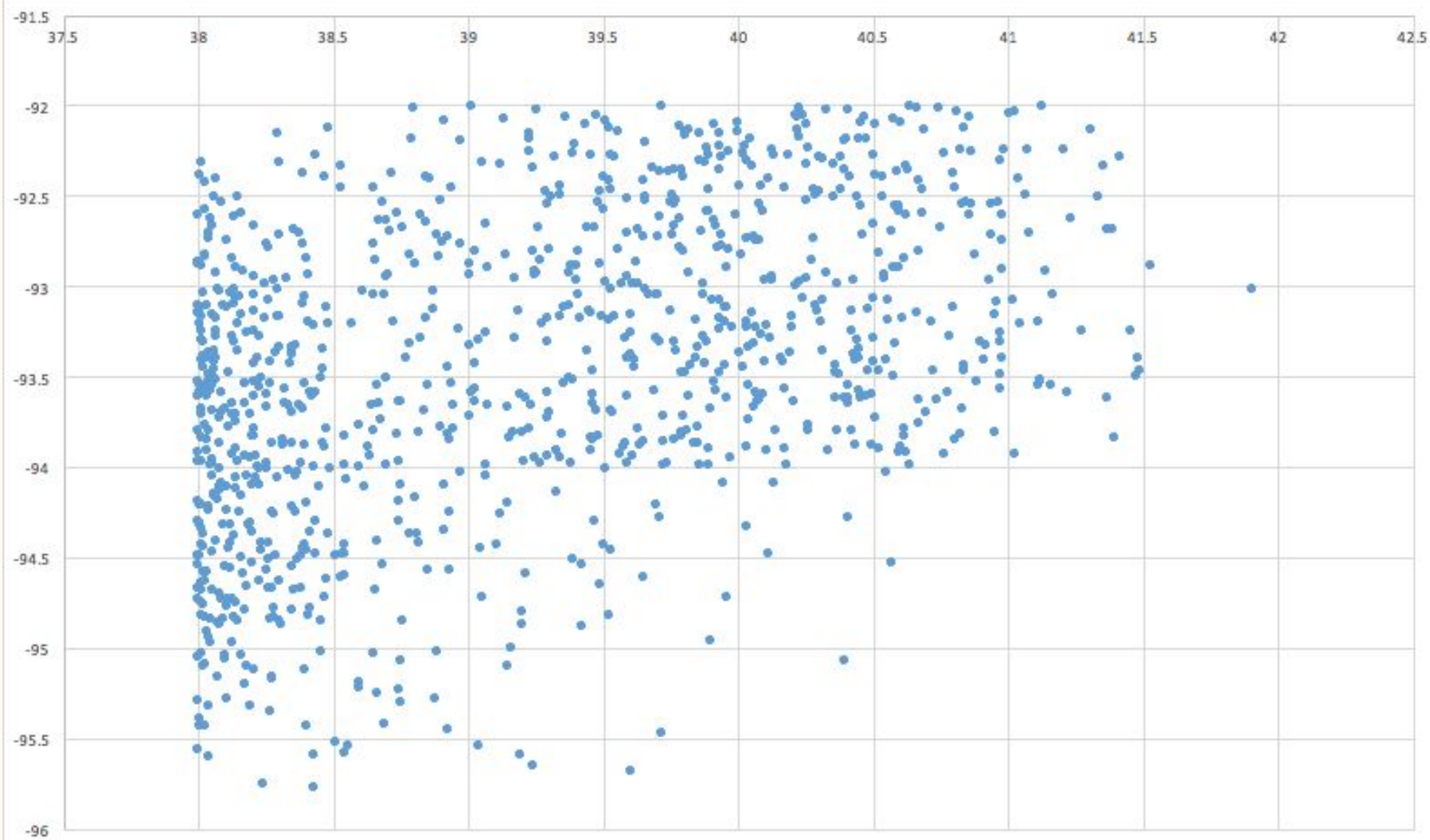
Hierarchical clustering

—

Example

For a city of 10,000 homes with uniform population density, there are 1,000 homes that have a blender. Are there specific areas of town that are more likely to have homes with blenders?

We will use hierarchical clustering to find clusters of homes that tend to have blenders.



Process

1. We will be using the agglomerative method, where we start with smaller clusters and combine them up to create larger clusters
2. Determine the appropriate metric - in this case a traditional Euclidean distance calculation would work best for calculation geographic proximity between homes: $||a-b||_2 = (\sum (a_i - b_i)^2)^{0.5}$
3. We will then link the homes and clusters of homes into larger clusters using the complete-linkage clustering method: $\max\{d(a,b): a \in A, b \in B\}$

HC Clustering outcome

