

In []:

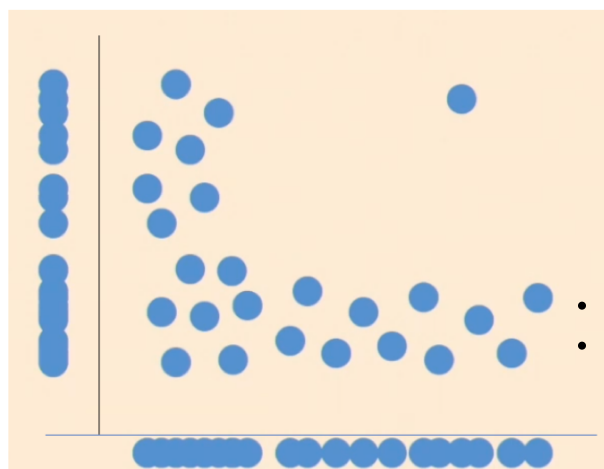
```
%run 00_basic.ipynb
```

Anomaly Detection using Isolation Forest

An **Isolation Forest** is to build trees recursively. In each iteration:

- Pick a random feature
- Split the data based randomly chosen value b/w the minimum and maximum of the chosen feature.
- Repeat until the entire data set is partitioned to form an individual tree in the forest.

Anomalies generally have shorter paths from the root than normal. An anomaly score to a point is the average path length from root to it.



A sample set of points is shown and if we see the points in one dimension, one can see that there are no anomalies.

Draw backs

- Does not work with categorical variables inherently
- Not a good one for missing values

In []:

In []: