OUTLIERS

Credit: https://www.cuemath.com/data/outlier/
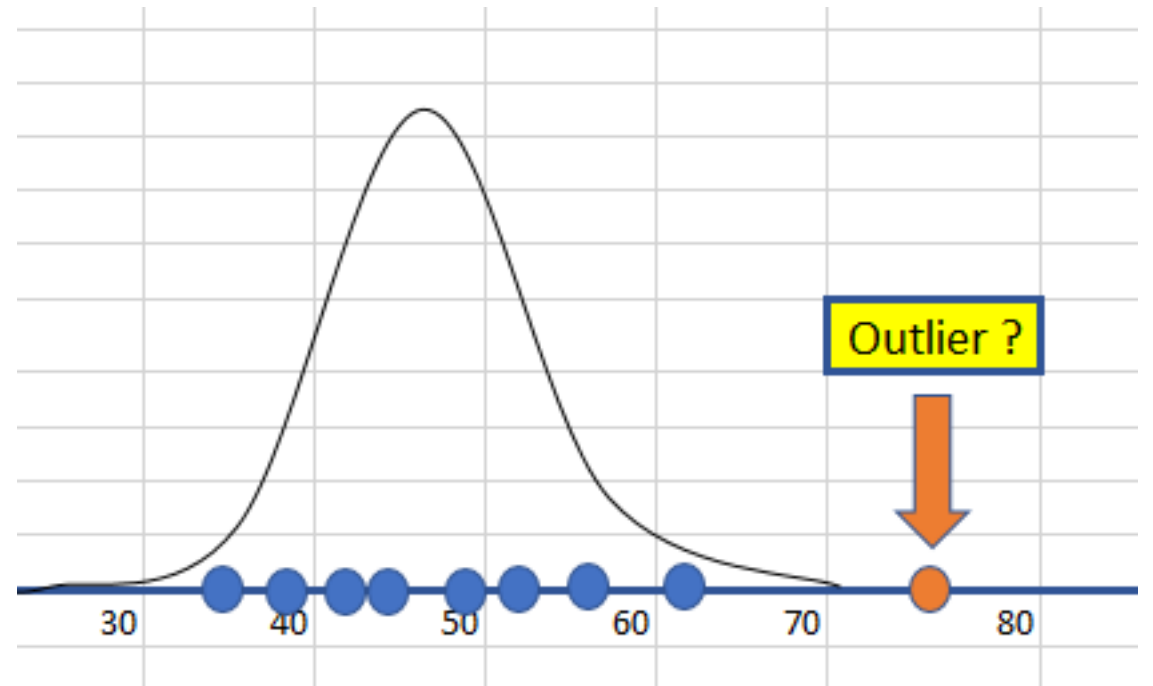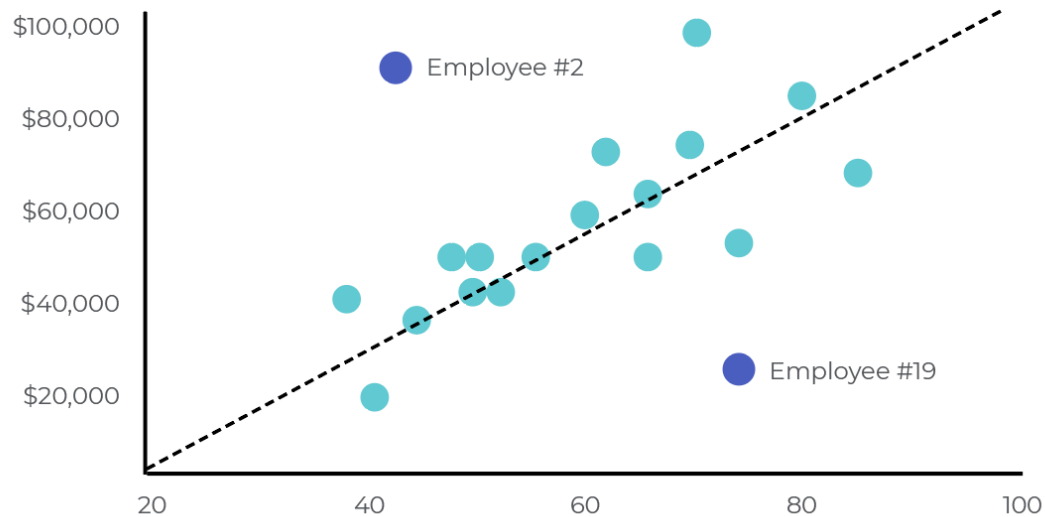
2nd ThaisCube Workshop
8-11 AUGUST 2023

Wirin Sonsrettee

What are Outliers?

Outlier: *"a value that is highly unlikely given an assumed distribution."*
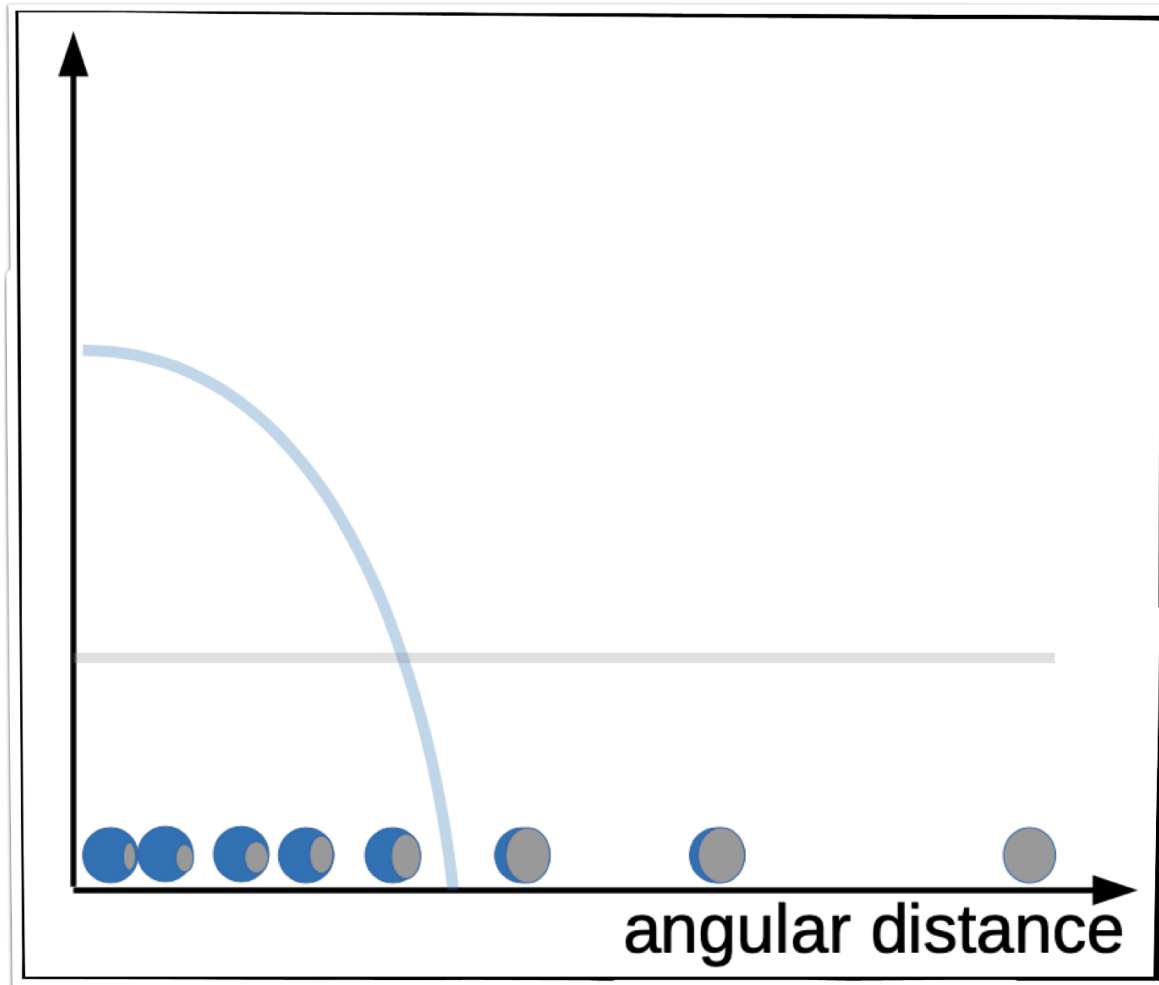

Test Scores Versus Performance Measured by Sales

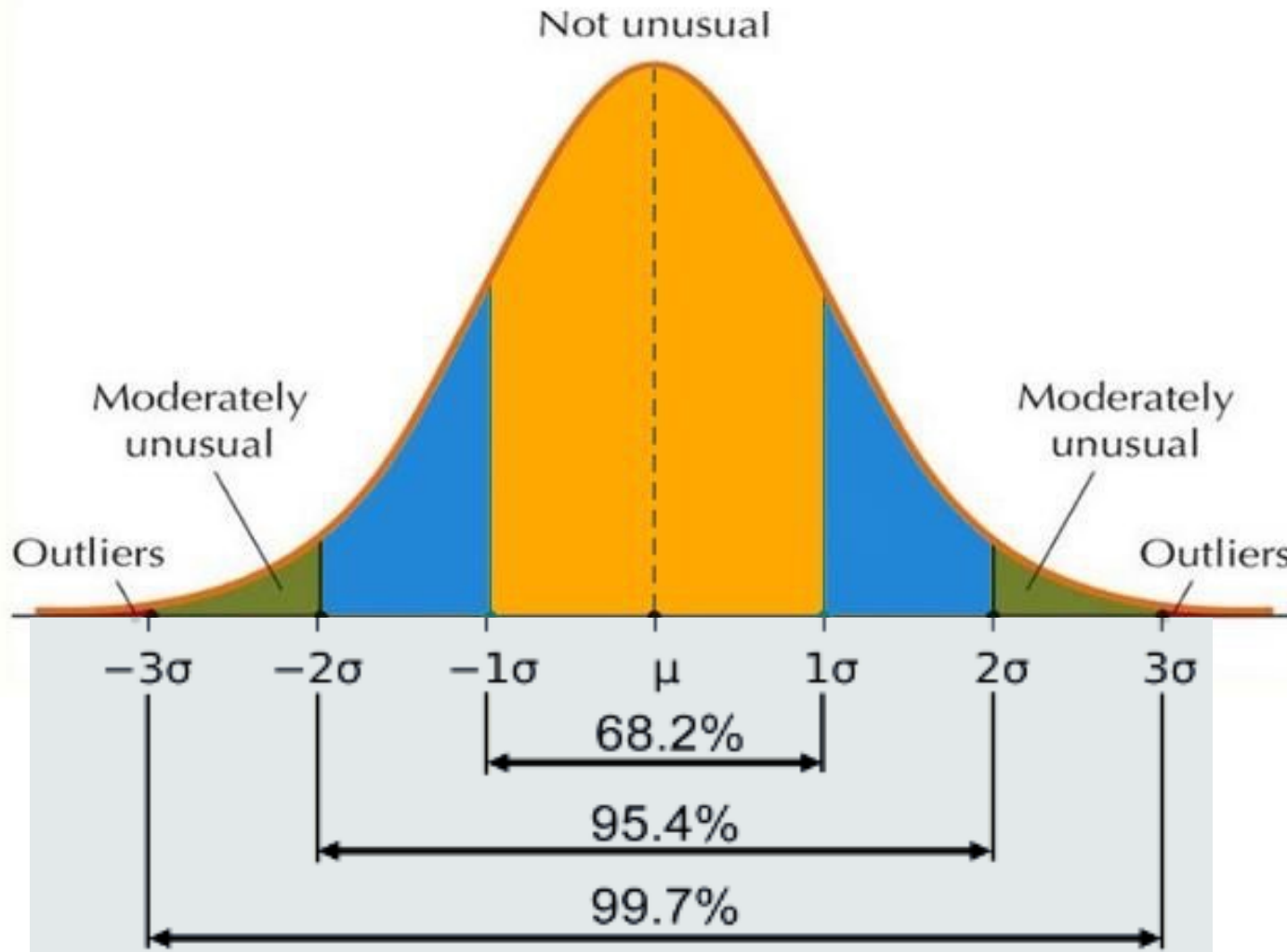# What are Outliers? (continued)

The farther a data point is from the distribution, the more probable it is to be an outlier.

Credit:A sneak peek into IceCUbe analyses, *Chiara Bellen*

Outlier Detection with Gaussian Sigma [2]

- The distribution of values in our sample follows a Gaussian or Gaussian-like pattern
- The standard deviation from the mean can effectively represent the percentage of values in the sample
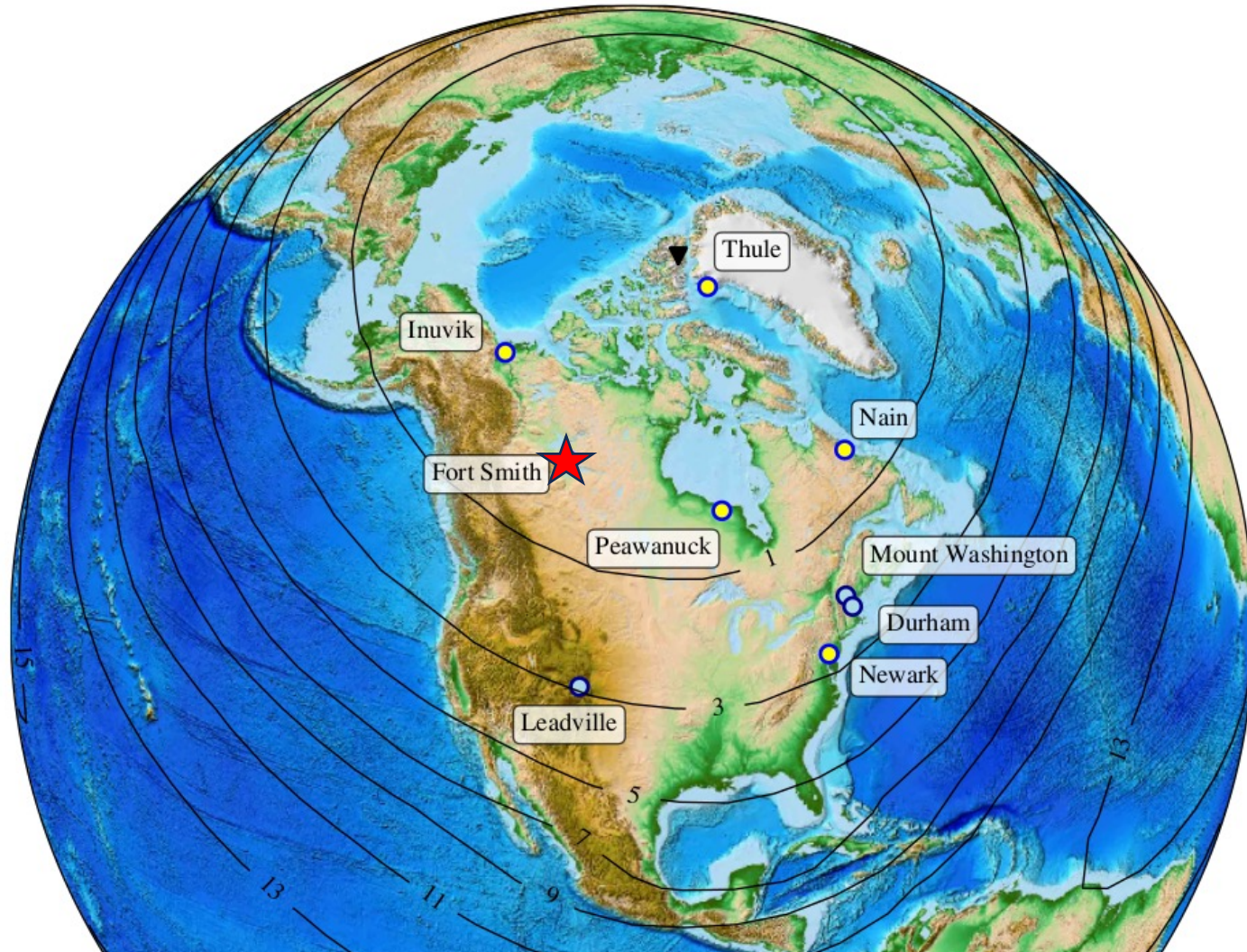
Adapted from https://www.almabetter.com/bytes/articles/outlier-detection-methods-and-techniques-in-machine-learning-with-examples

https://towardsdatascience.com/5-outlier-detection-methods-that-every-data-enthusiast-must-know-f917bf439210

Data from 'Fort Smith' neutron monitor station

Fort Smith, Canada / Bartol

Specifications

Detector	18 ^{3}He detectors in 3NM-64 modules
	with two modules in each of 3 shipping vans
Geographic latitude	60.02° N
Geographic longitude	111.93° W
Altitude	180 m asl
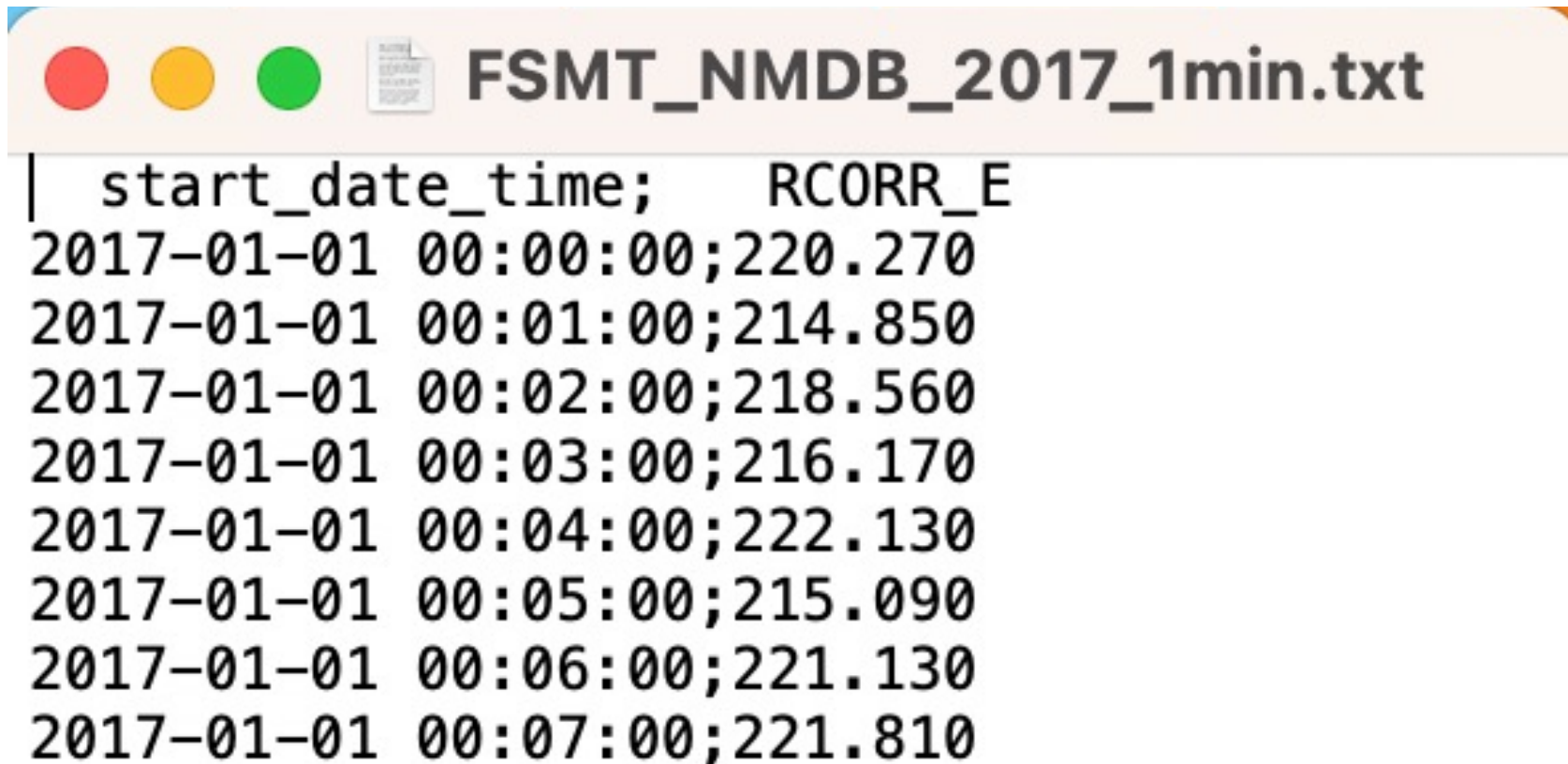
Credit: https://www.nmdb.eu/

A Data Example

- Count rates from 'Fort Smith' neutron monitor stations
- Count rates collected from 2015 to 2018.


```
     FSMT_NMDB_2017_1min.txt

| start_date_time;      RCORR_E
2017-01-01 00:00:00;220.270
2017-01-01 00:01:00;214.850
2017-01-01 00:02:00;218.560
2017-01-01 00:03:00;216.170
2017-01-01 00:04:00;222.130
2017-01-01 00:05:00;215.090
2017-01-01 00:06:00;221.130
2017-01-01 00:07:00;221.810
```