Anamoly Detection

is also known as outlier detection, is the process of identifying unusual patterns, data points or events that deviate significantly from the majority of the data.

Data Points : 10, 12, 11, 102, 05, 14

Anomalous : 102

Example: Anamoly Detection can be done by using Z-Score stats method & isolation forest ML method.

Z-Score Method

Datapoints : [10, 12, 13, 12, 14, 11, 990, 1000]

Saturday 28

(118-247) Wk 17

Step 1 : Mean (μ)

$$\mu = \frac{10+12+13+12+14+11+990+1000}{8}$$

$$= \frac{2062}{8} = 257.75$$

Step 2 : Standard Deviation (σ)

$$\sigma = \sqrt{\frac{1}{n}\Sigma(x_i-\mu)^2}$$

Value	$x_i - \mu$	$(x_i-\mu)^2$
10	$10-257.75 = -247.75$	61387.06
12	-245.75	60396.06
13	-244.75	59900.56
12	-245.75	60396.06

April / May 2018

29 Sunday
(119-246) Wk 17

value	$x_i - \mu$	$(x_i - \mu)^2$
14	-243.75	59408.06
11	-246.75	60891.06
990	732.25	536193.06
1000	742.25	550942.06

$\Sigma (x_i - \mu)^2 = 889514.98$

$\sigma = \sqrt{\frac{889514.98}{8}} \approx \sqrt{111189.37} \approx 333.42$

Step 3: Z-score Calculation

$$Z = \frac{x - \mu}{\sigma}$$

30 Monday
(120-245) Wk 18 ☺

Now Calculate for each value

value	Z-score
10	$\frac{10 - 257.75}{333.43} \approx -0.743$
12	-0.737
13	-0.734
12	-0.737
14	-0.731
11	-0.740
990	2.196
1000	2.242

Tuesday 1

(121-244) Wk 18

Step 4: Anomaly flag ($|z| > 1.5$)

You've threshold to 1.5, so:

- Values with $|z| > 1.5$ → True (red)
- Rest → False (blue)

Anomalies Detected.

Wednesday 2

(122-243) Wk 1