

# UITSCHIETERS

CATEGORIE 3



Sensoren veroveren de wereld. Heel veel toestellen/machines bevatten sensoren om het toestel/machine of de omgeving te monitoren. De reeksen data die deze sensoren leveren worden dan gebruikt om bijvoorbeeld een veranderd gedrag te herkennen. Het grootste probleem met de reeksen data die sensoren leveren is dat er af en toe wel eens een uitschieter kan tussen zitten. Dit zijn waarden die eruit springen ten opzichte van de omringende waarden. Dit wordt ook een anomalie genoemd. Dit kan veroorzaakt worden door een fout tijdens de meting maar kan ook de reële waarde zijn dat op iets abnormaals duidt zoals bijvoorbeeld brand. Normale waarden kunnen variëren in de reeks data. Een temperatuur van  $-5^{\circ}\text{C}$  is perfect normaal in onze winter maar iets raars in onze zomer. Daarom wordt er gewerkt met rollende gemiddelden en varianties/standaard deviaties om te bepalen wanneer een waarde een uitschieter is.

Bij deze opgave moeten jullie een programma schrijven dat deze anomalieën *on-the-fly* kan ontdekken. Daarvoor wordt er gewerkt met een glijdend venster met grootte  $aS$ . De eerste  $aS$  waarden worden gebruikt om het systeem op te starten. De anomaliedetectie start pas daarna. Een meetwaarde wordt als een anomalie beschouwd als het meer dan 3 maal de rollendeStandaardDeviatie (= vierkantswortel van de rollendeVariantie) van het rollendGemiddelde is verwijderd.

De eerste waarden voor het rollende gemiddelde ( $rolG$ ) en variantie ( $rolV$ ) zijn de gewone gemiddelde en variante van de eerste  $aS$  waarden.  $x_i$  is de  $i$ -de meetwaarde.

$$rolG_1 = \frac{\sum_{i=1}^{aS} x_i}{aS}$$
$$rolV_1 = \frac{\sum_{i=1}^{aS} (x_i - rolG_1)^2}{aS - 1}$$

Na de opstart, wordt voor elke nieuwe meetwaarde ( $x_i$ ) eerst gecontroleerd of het geen anomalie is (meer dan 3 maal de rollendeStandaardDeviatie van het rollendGemiddelde verwijderd). Als het geen anomalie is, dan worden de rollende waarden aangepast via de onderstaande formules.  $x_{i-aS}$  is de eerste waarde van het glijdende venster.

$$rolG_i = rolG_{i-1} + \frac{x_i - x_{i-aS}}{aS}$$
$$rolV_i = rolV_{i-1} + \frac{(x_i - x_{i-aS}) * ((x_i - rolG_i) + (x_{i-aS} - rolG_{i-1}))}{(aS - 1)}$$

En het glijdend venster schuift een positie op: de eerste waarde wordt eruit gehaald en de nieuwe waarde wordt achteraan toegevoegd.

Is het wel een anomalie, dan wordt dit enkel gemeld (zie uitvoer).

## Invoer

De invoer begint met een regel met daarop één getal: het aantal testgevallen. Daarna volgen per testgeval 2 regels. De eerste bevat 2 getallen:  $aS$  (het aantal startwaarden = grootte van glijdend venster) en  $aM$  (het aantal meetwaarden die moeten gecontroleerd worden). De tweede regel bevat  $aS+aM$  gehele getallen ( $[-1000, +1000]$ ).

### VOORBEELDINVOER

---

```
2
3 8
4 5 6 5 4 3 4 2 17 3 1
3 10
4 5 6 4 3 4 5 4 3 3 2 4 1
```

---

## Uitvoer

De uitvoer bestaat uit één regel per testgeval met de plaatsen van de anomalieën (gescheiden door één blanco) als er minstens één anomalie is of de tekst *geen anomalie* waarbij elke lijn start met het volgnummer van het testgeval en één blanco. De plaats van een anomalie wordt bepaald door het volgnummer in de reeks waarden (startwaarden inbegrepen) en de eerste waarde heeft nummer 1.

### VOORBEELDUITVOER

---

```
1 9
2 geen anomalie
```

---