

## Rapid Analysis : Clustering of Toxicity for Vaccine Lots

By Craig Paardekoooper

### Data Source:

2021 data VAERS USA : <https://howbad.info/tox-data.csv>

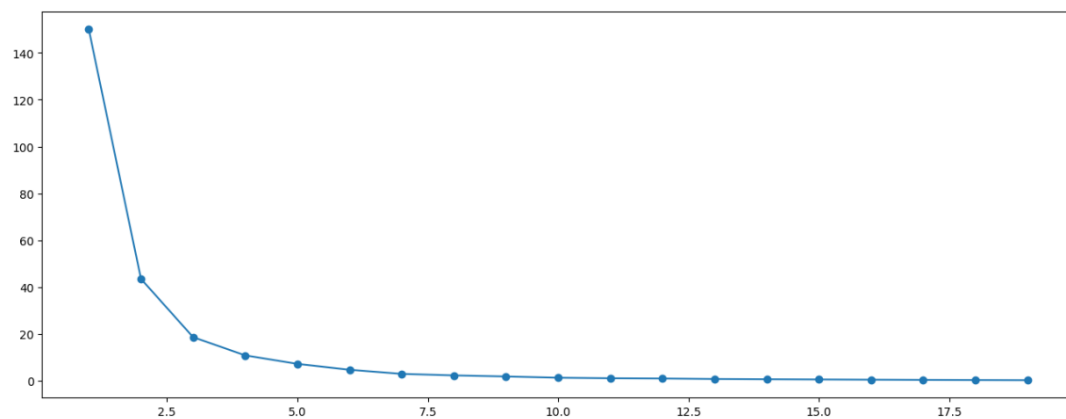
I had lot sizes shipped for 150 different lots, and the number of adverse reaction reports for each lot

Toxicity was defined as number of adverse reactions per 100,000 doses shipped for each lot.

### Method :

K-means clustering was used to see if there was any grouping of toxicities.

The “Elbow method” showed 3 clusters of toxicity –



k-means clustering was then applied based on 3 clusters

### Results :

#### Cluster 1

low toxicity - averaging 24.5 adverse reactions per 100,000 doses shipped

88 lots = 59% of the total number of lots analysed (88 out of 150)

Pfizer F series (FA, FC, FD, FE, FF, FG, FH, FJ)

#### Cluster 2

highest toxicity - averaging 231.7 adverse reactions per 100,000 doses shipped

15 lots = 10% of the total number of lots analysed (15 out of 150)

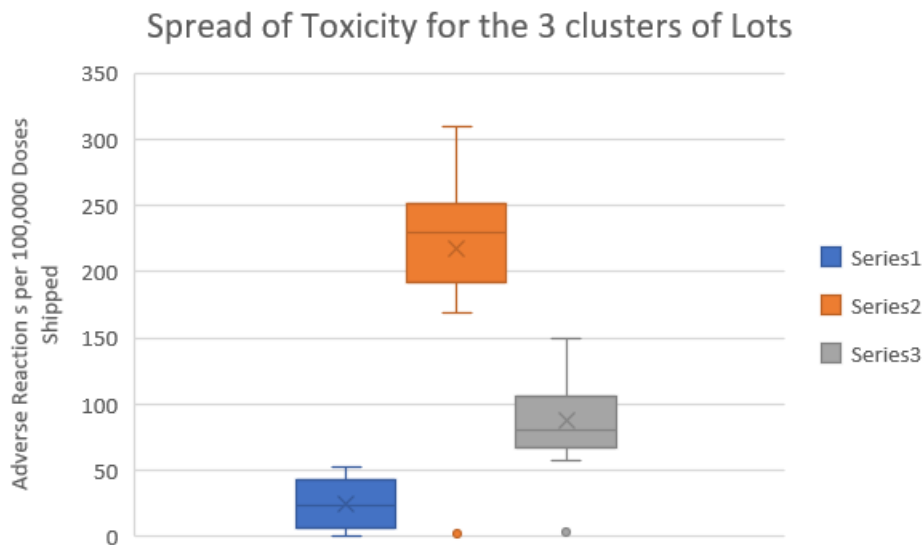
Pfizer E series (EH, EJ, EK, EL)

### Cluster 3

medium toxicity - averaging 89 adverse reactions per 100,000 doses shipped

47 lots = 31% of the total number of lots analysed (47 out of 150)

Pfizer E series (EL, EM, EN, ER, EW)



### Discussion :

For US data we see that almost two thirds of the lots were of low toxicity, almost one third of medium toxicity and 10% of high toxicity.

USA data shows –

High toxicity	10%	Pfizer E series (EH, EJ, EK, EL)
Medium toxicity	31%	Pfizer E series (EL, EM, EN, ER, EW)
Low toxicity	59%	Pfizer F series (FA, FC, FD, FE, FF, FG, FH, FJ)

We can compare this to the Denmark study

Denmark data showed –

High toxicity	4%	Pfizer EJ EK EL EM
Medium toxicity	64%	Pfizer EP ER ET EW EX EY FA FC FD FE FF and FG
Low toxicity	32%	Pfizer FG FH FJ FK FL FM

Ref : [Danish Study](#)

[Placebo Batch Numbers](#)

The Pfizer lot number series that correspond to high, medium and low toxicity are similar in both the USA data and in the Danish study.

In both cases, as the alphabet ascends the toxicity appears to decrease.

The main difference between the USA and Danish data is that in the USA there appear to be a higher proportion of high toxicity and low toxicity batches. In other words there is more of a polarisation – with less medium toxicity.

So in the USA 10% of the batches were highly toxic, and 60% were low toxicity. The larger number of low toxicity batches would generate more support for the vaccines, which would help maintain the vaccine rollout despite a larger % of high toxicity batches.

These findings are compatible with the V-Safe findings where 7.7% of vaccinated sought medical treatment after vaccination for COVID-19. See [V-Safe Data](#)

The average adverse reactions per 100,000 doses shipped provides a relative idea of toxicity variation between batches. However this does not take into account the under-reporting factor.