4 designs and 5 iterations and categorty changes across 2 granularities (g3 categories const)

Sayani Gupta

1 *Simulation design*

```
t = 3000
n <- seq(0, t, 1)
g1 <- n %%2
g2 <- n %%3
g3 <- n %%5
mu1= c(0, 2) # mean of cat of g1
mu2 = c(7, 5, 1)# mean of cat of g2
mu3 = c(0, 0, 0, 0, 0) # mean of cat of g3
```

Three circular granularities g1, g2 and g3 are considered with levels 2, 3 and 5 respectively. Many time series with 3000 observations are created using the five designs below, each of which is iterated five times. We anticipate to have 3 clusters, each with five time series conforming to the same design, once we execute the clustering.

2 *Distance computation*

1) wpd for three granularities computed and then distance between designs are computed by computing the euclidean distances between them - gran as variables, designs across rows and cell values as wpd.

2) Robust scaling done for each customer and then JSD is computed between same categories for a granularity.

3 *Granularities design*

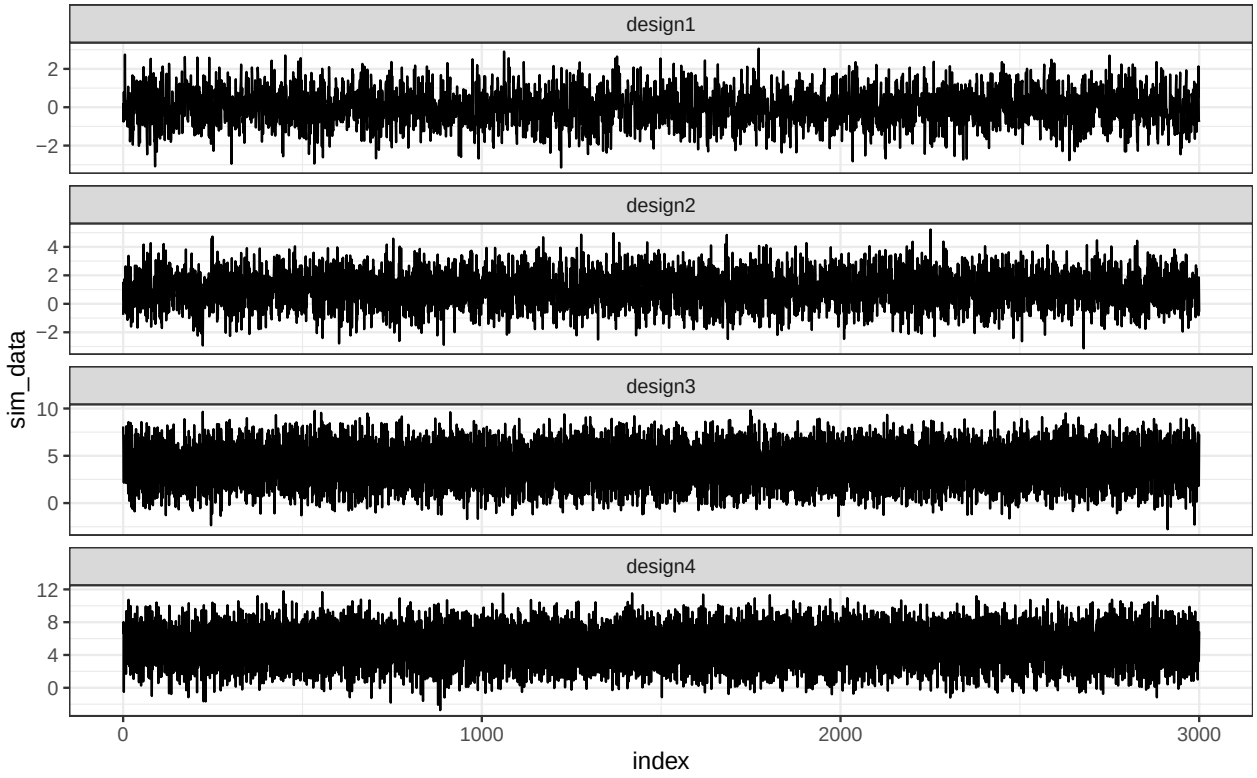
design	g1	g2	g3
design-1	no	no	no
design-2	no	yes	no
design-3	yes	no	no
design4	yes	yes	no

4 *Simulate data*

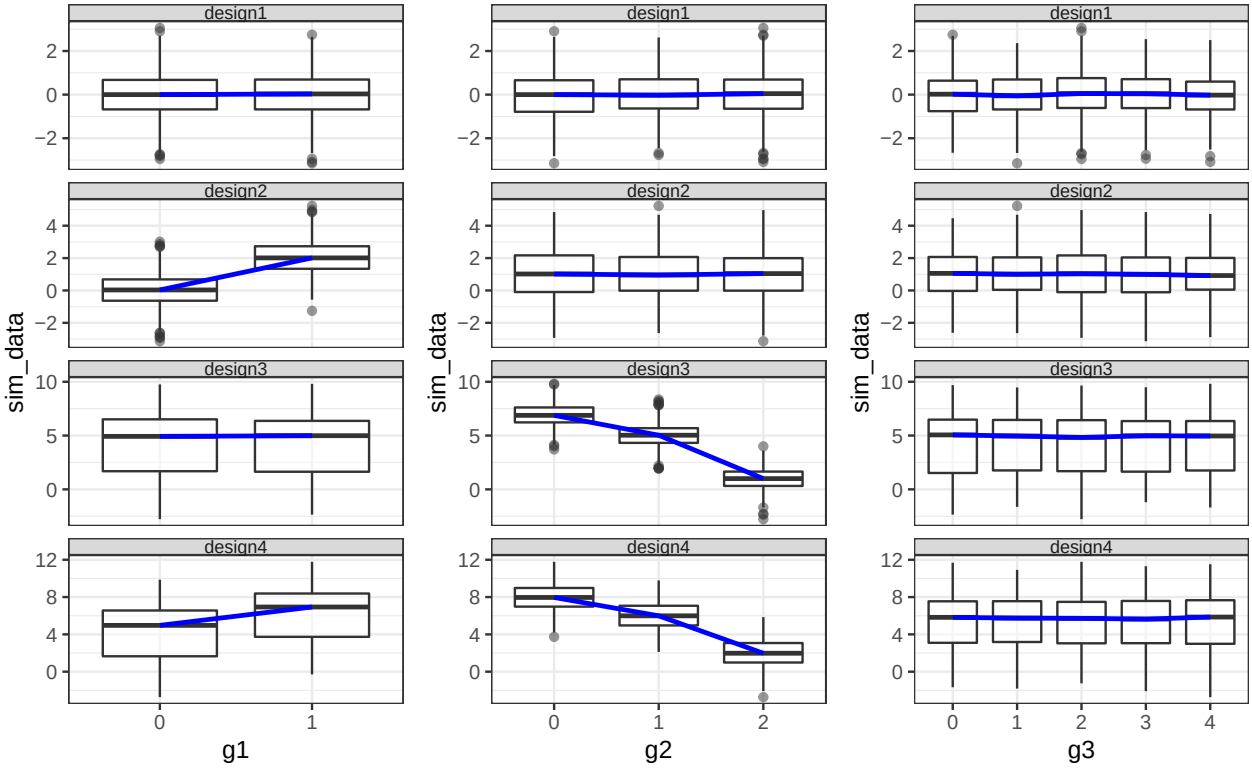
```
## # A tibble: 10 x 6
##    index    g1    g2    g3 design  sim_data
##    <dbl> <dbl> <dbl> <dbl> <chr>      <dbl>
##  1     0     0     0     0 design1    0.217
##  2     0     0     0     0 design2   -0.778
##  3     0     0     0     0 design3    8.10
```

```
##  4     0     0     0     0 design4    6.61
##  5     1     1     1     1 design1   -0.747
##  6     1     1     1     1 design2    1.51
##  7     1     1     1     1 design3    5.47
##  8     1     1     1     1 design4    8.02
##  9     2     0     2     2 design1    0.786
## 10     2     0     2     2 design2    0.244
```

5 *Plot raw data*



6 *Plot distribution across granularities*



7 *Iterate designs*

```
## # A tsibble: 10 x 8 [1]
## # Key:       customer_id [1]
##    seed_id index    g1    g2    g3 design  sim_data customer_id
##    <chr>   <dbl> <dbl> <dbl> <dbl> <chr>      <dbl> <chr>
##  1 1           0     0     0     0 design1    0.586 design1-1
##  2 1           1     1     1     1 design1    0.709 design1-1
##  3 1           2     0     2     2 design1   -0.109 design1-1
##  4 1           3     1     0     3 design1   -0.453 design1-1
##  5 1           4     0     1     4 design1    0.606 design1-1
##  6 1           5     1     2     0 design1   -1.82  design1-1
##  7 1           6     0     0     1 design1    0.630 design1-1
##  8 1           7     1     1     2 design1   -0.276 design1-1
##  9 1           8     0     2     3 design1   -0.284 design1-1
## 10 1           9     1     0     4 design1   -0.919 design1-1
```

8 *Cluster designs: distance using wpd*

```
##            Reference
## Prediction design1 design2 design3 design4
##    design1       5       0       0       0
##    design2       0       5       0       0
##    design3       0       0       5       0
##    design4       0       0       0       5
```

9 *Cluster designs: distance using js distance between categories*

```
##           Reference
## Prediction design1 design2 design3 design4
##    design1       5       0       0       0
##    design2       0       5       0       0
##    design3       0       0       5       0
##    design4       0       0       0       5
```

10 *Cluster designs: distance using js distance between categories (nqt)*

```
##           Reference
## Prediction design1 design2 design3 design4
##    design1       5       0       0       0
##    design2       0       5       0       0
##    design3       0       0       5       0
##    design4       0       0       0       5
```