VKT analysis

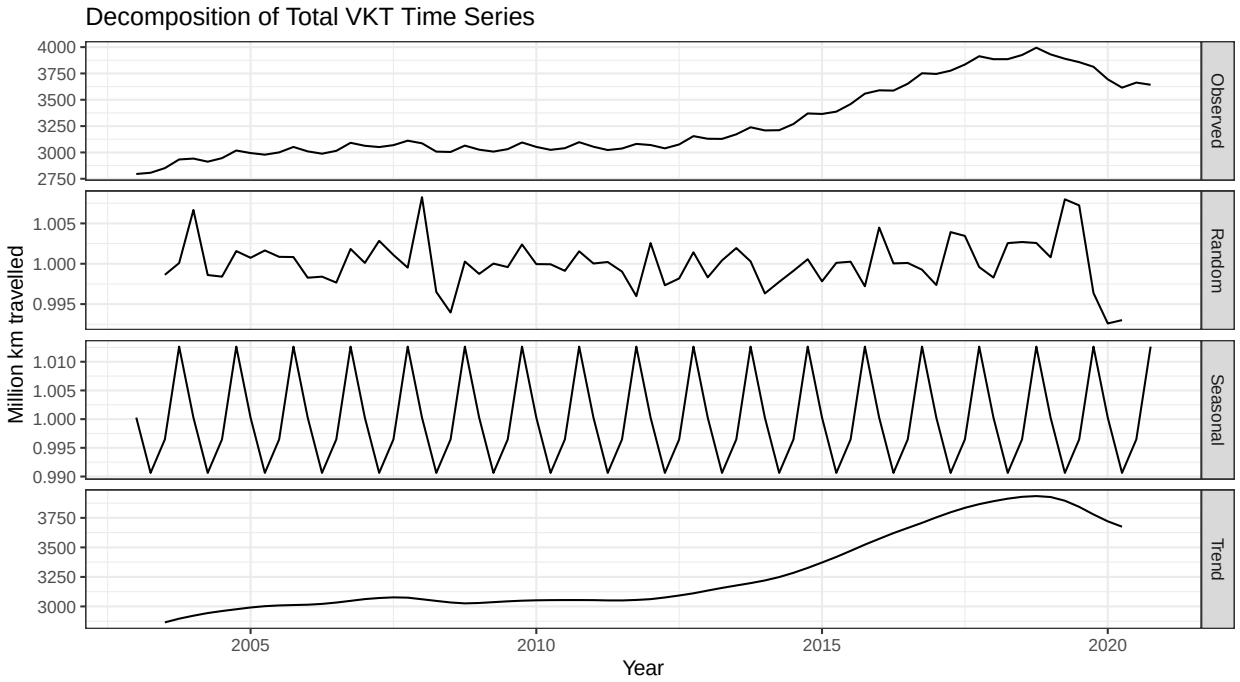
```
vkt = read.csv("downloaded_stats/VKT_main_quart.csv")
vkt_other_yearly = read.csv("downloaded_stats/VKT_other_yearly.csv")
vkt_fuel_yearly = read.csv("downloaded_stats/VKT_fuel_vehicle_type_yearly.csv")
```

```
decomp_vkt = vkt[vkt$year >= 2003,]$Auckland %>%
  ts(frequency = 4, start = min(2003)) %>%
  decompose("multiplicative")

plot_decomp(decomp_vkt, ylab = "Million km travelled", title = "Decomposition of Total VKT Time Series")
```

```
## Don't know how to automatically pick scale for object of type ts. Defaulting to continuous.

## Warning: Removed 2 row(s) containing missing values (geom_path).
```



if we make the assumption of the seasonal trend of VKT following a sin wave we can use the equation

$$y = a \sin\left(\frac{2\pi}{12}(x-b)\right) + c$$

to calculate what such sin wave would have been required to produce the quarterly "Boxes" we can adjust the parameters a, b, and c to minimize the difference in integral sum.

y would be a kind of rolling (sum of km)/month

(3 parameters so 3 degrees of freedom and 4 points to base off so will still be slight averaging going on)

```
a_list = seq(0,0.05,0.0001)
b_list = seq(0,12,0.01)
c_list = 1 # it is centered on 1 so does not need to change
#vector order is best a,b,c values and lowest mean squared error
best_comb = c(0,0,0,Inf)
i = 0;

sin_func = function(x, a, b, c) {
  return(a*sin(2*pi*(x-b)/12)+c)
}

for (a in a_list) {
  for (b in b_list) {
    for (c in c_list){
      i = i+1
      quart_sin = c(integrate(sin_func, 0,3, a = a, b = b, c = c)$value, integrate(sin_func, 3,6, a = a
                    integrate(sin_func, 6,9, a = a, b = b, c = c)$value, integrate(sin_func, 9,12, a = a
      error = mean((quart_sin-decomp_vkt_auck$figure*3)^2)
      #should use mean absolute or mean sqaured error
      #error = mean(abs(quart_sin-decomp_vkt_auck$figure*3))
      if (error < best_comb[4]) {
        best_comb = c(a,b,c,error)
      }
    }
  }
}
best_comb
print(i)
```

```
barplot(decomp_vkt_auck$figure, ylim = c(0.985,1.015), xpd = FALSE, names.arg = c("Q1","Q2","Q3","Q4"),
points(x = seq(0,4,0.001), sin_func(seq(0,12,0.003),best_comb[1],best_comb[2],best_comb[3]), type = "l"
```