02_EDA

Tamara

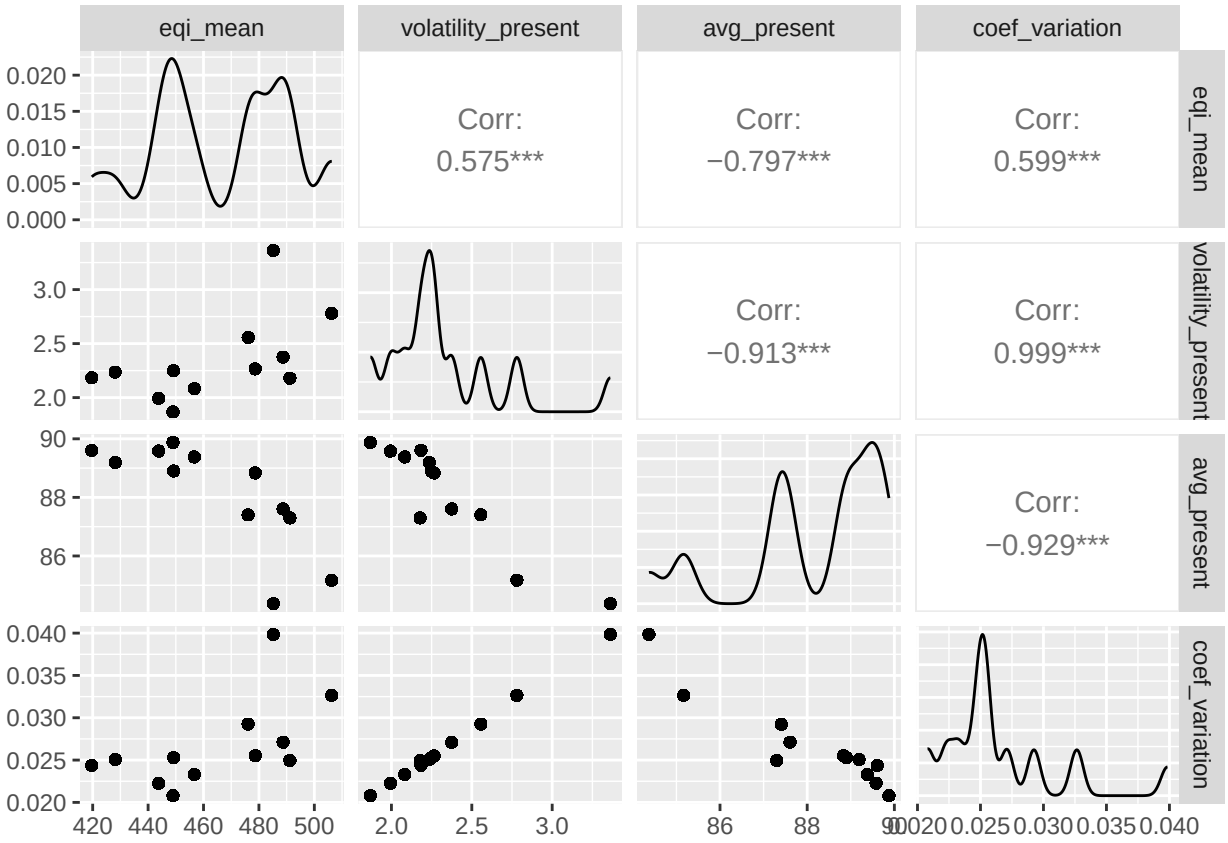
2025-06-16

Exploratory data analysis

```
## Rows: 3,144
## Columns: 13
## $ year               <dbl> 2024, 2024, 2024, 2024, 2024, 2024, 2024, 2024, 202~
## $ term               <dbl> 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, ~
## $ education_region   <chr> "Tai Tokerau", "Tai Tokerau", "Tai Tokerau", "Tai T~
## $ category           <chr> "total students", "students attending more than 90p~
## $ count              <dbl> 30315, 13815, 7280, 3954, 5266, NA, NA, NA, 101427,~
## $ percent            <dbl> NA, 45.6, 24.0, 13.0, 17.4, 83.9, 7.2, 8.9, NA, 62.~
## $ volatility_present <dbl> 2.780397, 2.780397, 2.780397, 2.780397, 2.780397, 2~
## $ avg_present        <dbl> 85.16875, 85.16875, 85.16875, 85.16875, 85.16875, 8~
## $ n_obs              <dbl> 32, 32, 32, 32, 32, 32, 32, 32, 20, 20, 20, 20, 20,~
## $ coef_variation     <dbl> 0.03264574, 0.03264574, 0.03264574, 0.03264574, 0.0~
## $ eqi_mean           <dbl> 506.2925, 506.2925, 506.2925, 506.2925, 506.2925, 5~
## $ eqi_median         <dbl> 514, 514, 514, 514, 514, 514, 514, 514, 422, 422, 4~
## $ schools_in_region  <dbl> 148, 148, 148, 148, 148, 148, 148, 148, 189, 189, 1~
```

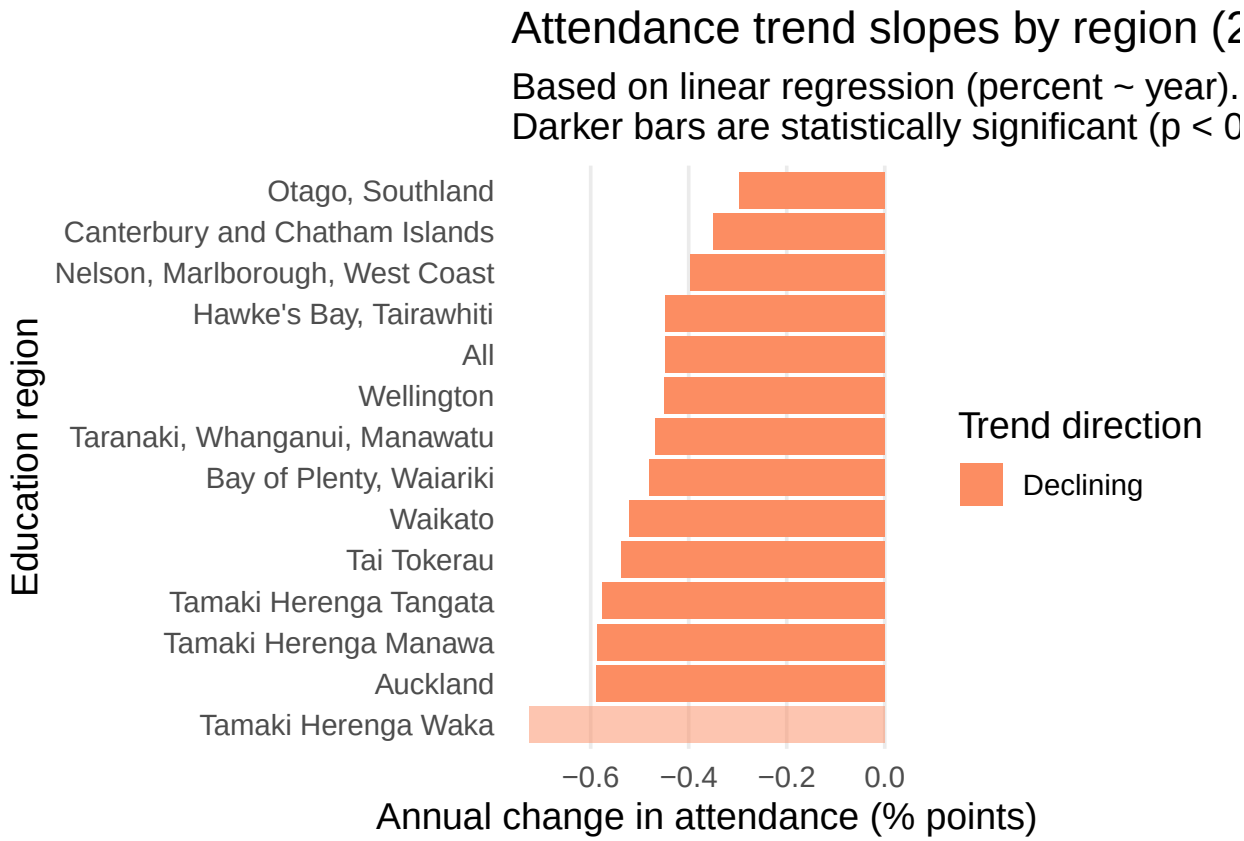
```
##               year               term   education_region           category
##                  0                  0                  0                  0
##              count            percent volatility_present        avg_present
##               1179                393                  0                  0
##              n_obs     coef_variation          eqi_mean         eqi_median
##                  0                  0                360                360
##  schools_in_region
##                360
```

```
##                      eqi_mean volatility_present avg_present coef_variation
## eqi_mean            1.0000000          0.5750981  -0.7971285      0.5986941
## volatility_present  0.5750981          1.0000000  -0.9126485      0.9989274
## avg_present        -0.7971285         -0.9126485   1.0000000     -0.9287596
## coef_variation      0.5986941          0.9989274  -0.9287596      1.0000000
```



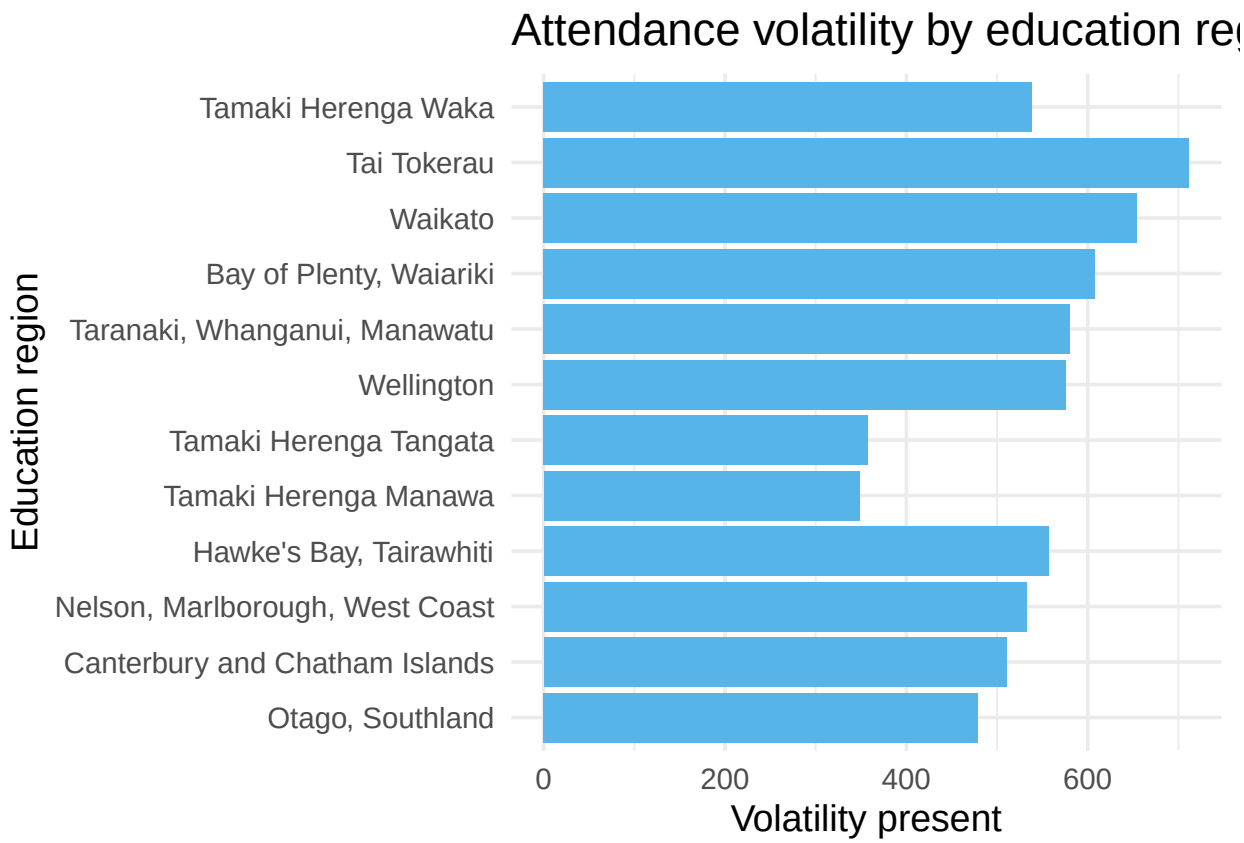

```
## Rows: 3,144
## Columns: 13
## $ year               <dbl> 2024, 2024, 2024, 2024, 2024, 2024, 2024, 2024, 202~
## $ term               <dbl> 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, ~
## $ education_region   <chr> "Tai Tokerau", "Tai Tokerau", "Tai Tokerau", "Tai T~
## $ category           <chr> "total students", "students attending more than 90p~
## $ count              <dbl> 30315, 13815, 7280, 3954, 5266, NA, NA, NA, 101427,~
## $ percent            <dbl> NA, 45.6, 24.0, 13.0, 17.4, 83.9, 7.2, 8.9, NA, 62.~
## $ eqi_mean           <dbl> 506.2925, 506.2925, 506.2925, 506.2925, 506.2925, 5~
## $ eqi_median         <dbl> 514, 514, 514, 514, 514, 514, 514, 514, 422, 422, 4~
## $ schools_in_region  <dbl> 148, 148, 148, 148, 148, 148, 148, 148, 189, 189, 1~
## $ volatility_present <dbl> 2.780397, 2.780397, 2.780397, 2.780397, 2.780397, 2~
## $ avg_present        <dbl> 85.16875, 85.16875, 85.16875, 85.16875, 85.16875, 8~
## $ n_obs              <dbl> 32, 32, 32, 32, 32, 32, 32, 32, 20, 20, 20, 20, 20,~
## $ coef_variation     <dbl> 0.03264574, 0.03264574, 0.03264574, 0.03264574, 0.0~
```

```
##               year               term   education_region           category
##                  0                  0                  0                  0
##              count            percent           eqi_mean         eqi_median
##               1179                393                360                360
##  schools_in_region volatility_present        avg_present              n_obs
##                360                  0                  0                  0
##     coef_variation
##                  0
```

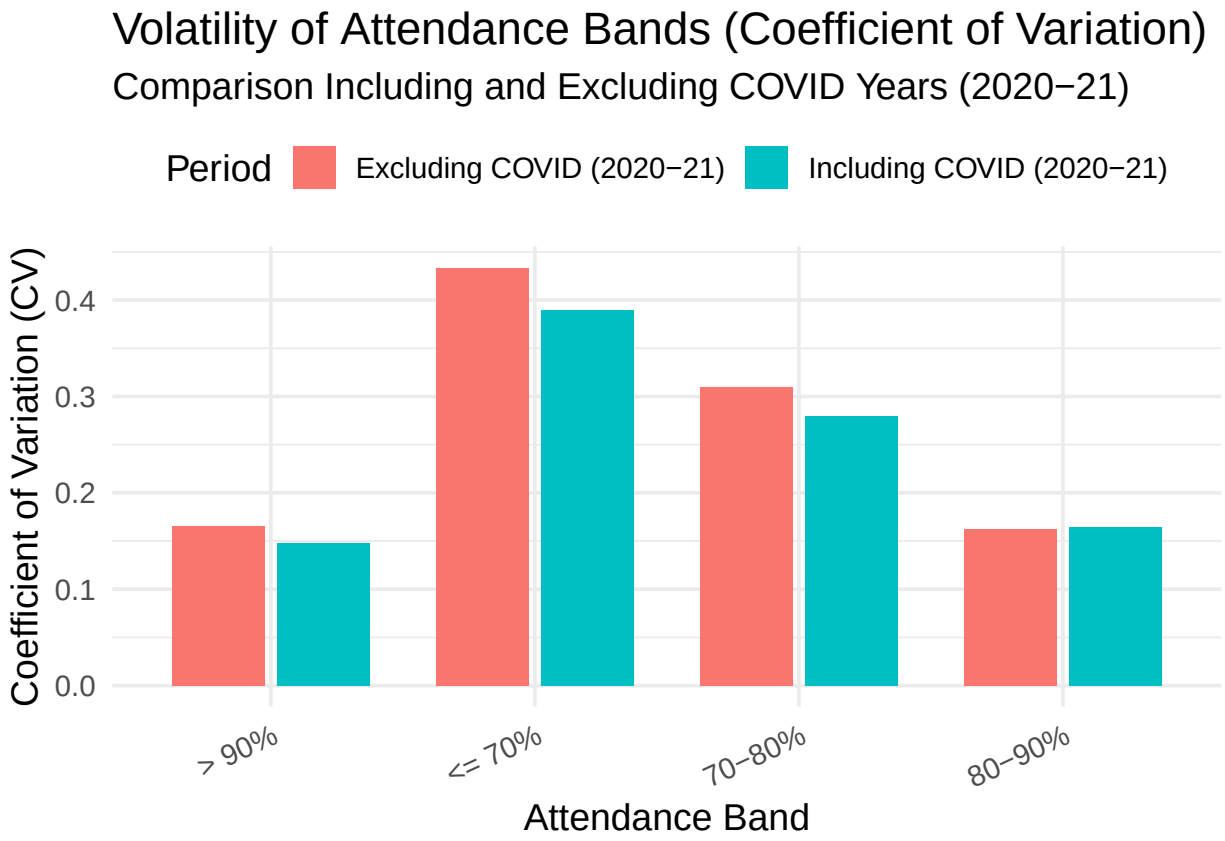



```
##
##
## --- Summary Table for Reporting ---
```

```
## # A tibble: 14 x 5
##    Region                     Annual Change (% pts~1 `P-value` `Trend Direction`
##    <chr>                                       <dbl>     <dbl> <chr>
##  1 Otago, Southland                            -0.3      0.001 Declining
##  2 Canterbury and Chatham Is~                  -0.35     0     Declining
##  3 Nelson, Marlborough, West~                  -0.4      0     Declining
##  4 Hawke's Bay, Tairawhiti                     -0.45     0     Declining
##  5 All                                         -0.45     0     Declining
##  6 Wellington                                  -0.45     0     Declining
##  7 Taranaki, Whanganui, Mana~                  -0.47     0     Declining
##  8 Bay of Plenty, Waiariki                     -0.48     0     Declining
##  9 Waikato                                     -0.52     0     Declining
## 10 Tai Tokerau                                 -0.54     0     Declining
## 11 Tamaki Herenga Tangata                      -0.58     0.041 Declining
## 12 Tamaki Herenga Manawa                       -0.59     0.033 Declining
## 13 Auckland                                    -0.59     0.002 Declining
## 14 Tamaki Herenga Waka                         -0.73     0.094 Declining
## # i abbreviated name: 1: `Annual Change (% pts)`
## # i 1 more variable: `Significant (p<0.05)` <lgl>
```


Attendance volatility by education re
Education region
Tamaki Herenga Waka
Tai Tokerau
Waikato
Bay of Plenty, Waiariki
Taranaki, Whanganui, Manawatu
Wellington
Tamaki Herenga Tangata
Tamaki Herenga Manawa
Hawke's Bay, Tairawhiti
Nelson, Marlborough, West Coast
Canterbury and Chatham Islands
Otago, Southland
0
200
400
600
Volatility present

Impact of Covid-19 disruption on data and the volatility calculation



Taranaki volatility drilldown