Title Placeholder

Load in Specific Packages

```
suppressWarnings({
    library(readr)
    library(tidyr)
    library(dplyr)
    library(here)
    library(lemon)
    library(kableExtra)
    library(ggplot2)
    library(reshape())
    library(hexbin)
    library(data.table)
    library(GGally)
    library(formattable)
    library(viridis)
    library(TTR)
    library(zoo)
    library(ggrepel)
    library(grid)
})
```

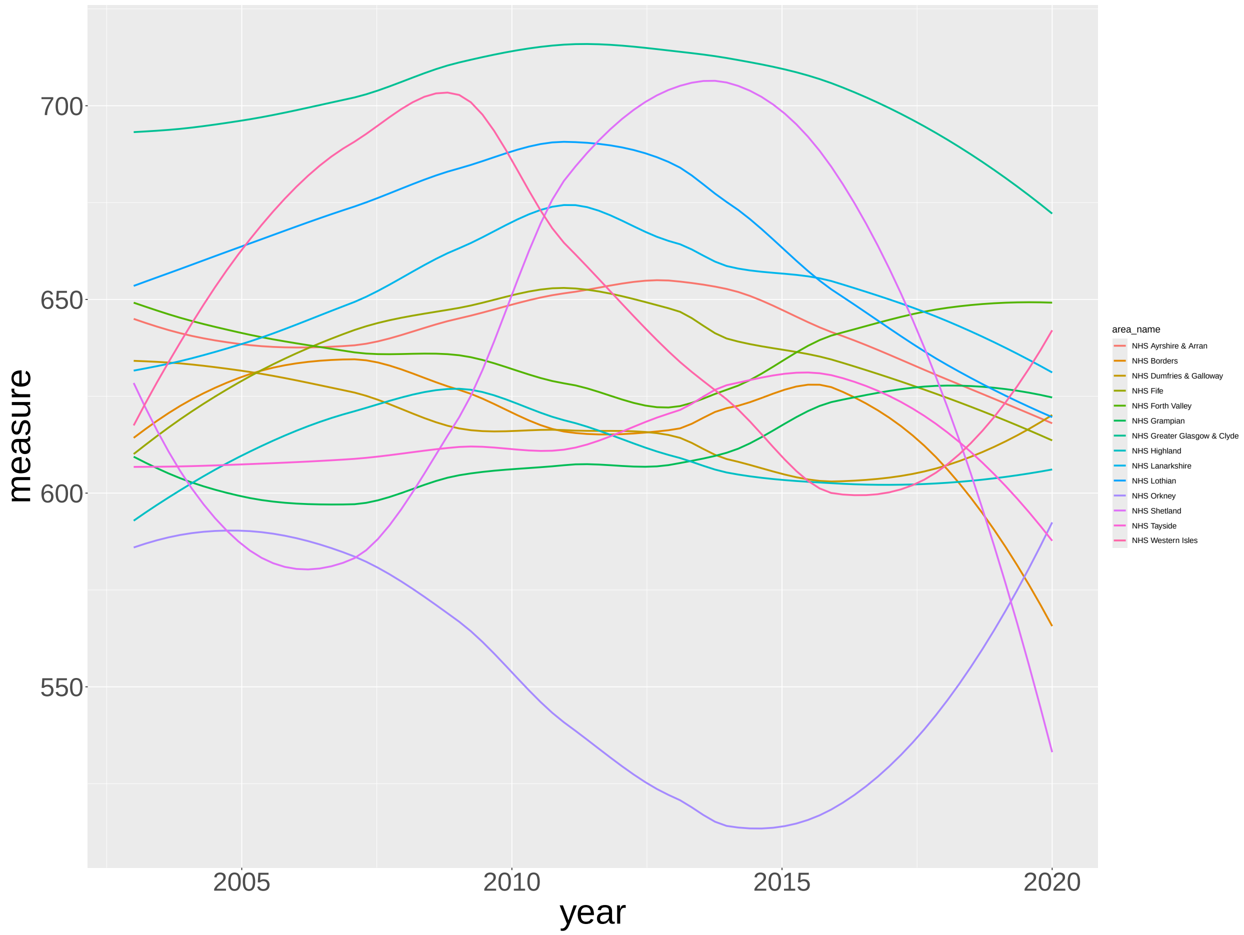
Load in the data

```
cancerReg <- read.csv("C:\\Users\\romin\\ToyRepo\\Models\\cancerReg.csv")
```

Remove Uncessary Data for Analysis

```
cancerReg <- cancerReg %>% select(-period,-area_type,-type_definition,-indicator, -upper_confidence_inte
```

Display All Data Points



Find Average of All Measures by Year

```
avgYearly <- cancerReg %>%
    group_by(year) %>%
    mutate(AvgYear = mean(measure, na.rm = TRUE)) %>%
    select(-area_name, -measure, -area_code)
```

Calculate Moving Average for Each Health Board

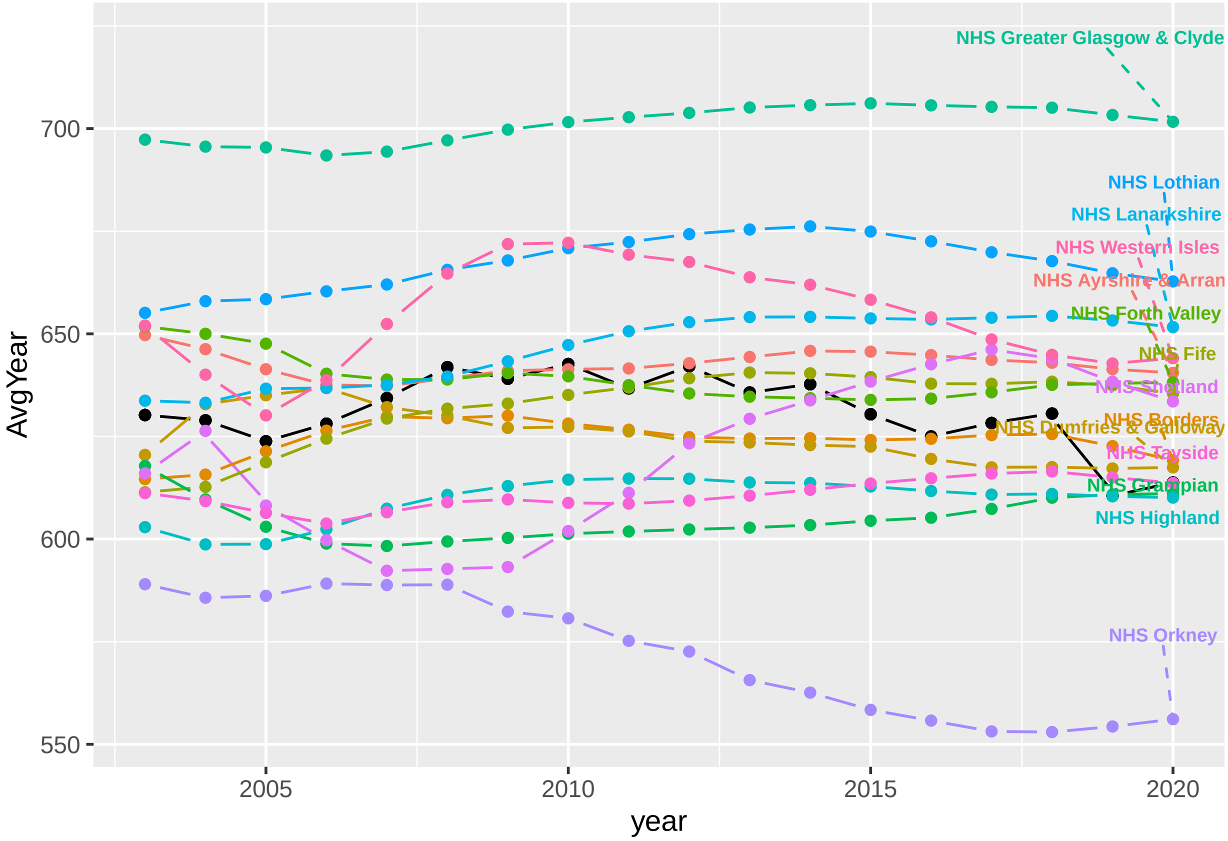
```
movingAvg <- cancerReg %>%
    group_by(area_name) %>%
    arrange(year) %>%
    mutate(MA = cumsum(measure) / row_number())
```

Find Last Data Points for Data

```
finalValues <- movingAvg %>%
    group_by(area_name) %>%
    summarise(
        lastMA = dplyr::last(MA),
        lastYear=dplyr::last(year)
    )
```

Display Summary of All Data

```
    ggplot(data = cancerReg, aes(x = year)) +
        geom_pointline(data = avgYearly, aes(y = AvgYear)) +
        geom_pointline(data = movingAvg, aes(y = MA, col = area_name)) +
        geom_text_repel(
            data = finalValues, aes(
                x = lastYear,
                y = lastMA,
                label = area_name,
                color = area_name
            ),
            size = 2.5,
            fontface = "bold",
            nudge_y = 20.6,
            direction = "y",
            hjust = -0.7,
            segment.linetype = 2,
            segment.size = 0.5,
            segment.curvature = 0
        ) +
        theme(legend.position = "none")
```



Calculate Differences Function

```
    sigPercent <- data.frame(
        area_name = character(),
        year = integer(),
        percentNum = numeric(),
        stringsAsFactors = FALSE
    )
    boardAvg <- function(currBoard, currVal, currYear) {
        currAvgYear <- filter(movingAvg, area_name == currBoard & year == currYear) %>% select(MA)
        numCurrAvgYear <- gsub("[^0-9.]", "", currAvgYear$MA)
        numCurrAvgYear <- as.numeric(numCurrAvgYear)
        diffVal <- currVal - numCurrAvgYear
        percentVal <- ((diffVal / numCurrAvgYear) * 100)

        if (percentVal >= 3 || percentVal <= -3) {
            sigPercent <- sigPercent %>% add_row(area_name = currBoard, year = currYear, percentNum = r
        }
        return(sigPercent)

        # WHile the moving average does not provide a direct estimation of the predictied values it sti
    }
```

Calculate Differences

```
    healthBoards <- unique(cancerReg$area_name)
    totalYears <- unique(cancerReg$year)
    for (currBoard in healthBoards) {
        for (currYear in totalYears) {
            currVal <- subset(cancerReg, year == currYear & area_name == currBoard)
            currVal <- select(currVal, -area_code, -area_name, -year)
            currVal <- as.numeric(currVal)
            sigPercent <- boardAvg(currBoard, currVal, currYear)
        }
    }
    sigPercent <- sigPercent %>% arrange(desc(year))
```

```
colourCells <- function(values, average){
    diffVal <- values - average
    if (diffVal>=0 & diffVal<=20){
        return(paste0("\\cellcolor{green!," ,round(diffVal/100), "}"))
    }
}
```

Summary Table of Data Within Graph

```
# inputFile <- "reportReg.pdf"
# sigPercentWide_colored <- sigPercentWide %>%
#     mutate(across(everything(), ~ cell_spec(.,
#                                              color = ifelse(. < 0, "red", "black"),
#                                              background = ifelse(. < 0, "lightpink", "white"))))
```

```r
sigPercentWide <- sigPercent %>% pivot_wider(
    names_from = year,
    values_from = percentNum
)
# sigPercentWide <- sigPercentWide %>% mutate(
#   across(-1,

#       ~ cell_spec(., color = ifelse(is.na(.), "black", ifelse(. < 0, "green", "red")))
#   )
#   )
# sigPercentWide <- sigPercentWide %>% mutate(across(everything(), ~ replace_na(., 0)))
kable(sigPercentWide, format = "latex", booktabs = TRUE, longtable=TRUE) %>%
    kable_styling(latex_options = c("striped", "hold_postion")) %>%
    row_spec(0, bold = TRUE) %>%
     kableExtra::landscape()
```

area_name	**2020**	**2019**	**2018**	**2017**	**2016**	**2015**	**2014**	**2013**	**2012**	**2011**	**2010**	**2009**	**2008**	**2007**	**2006**	**2005**
NHS Borders	-9.54	-7.56	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NHS Lothian	-5.11	-7.11	-4.85	-5.55	-4.63	NA	NA	NA	NA	NA	3.12	NA	NA	NA	NA	NA
NHS Orkney	5.48	3.89	NA	-6.66	-6.13	-9.02	-5.95	-12.29	-4.10	-7.63	NA	-6.75	NA	NA	NA	NA
NHS Shetland	-12.88	-14.09	-5.15	7.74	8.53	8.62	7.78	9.53	17.39	12.16	10.21	NA	NA	-5.01	-4.24	-5.98
NHS Western Isles	3.38	-5.25	-8.71	-11.61	-8.55	-6.64	-3.04	-5.60	NA	-3.48	NA	6.42	9.23	8.40	4.02	-3.15
NHS Fife	-5.27	NA	NA	NA	-3.25	NA	NA	NA	3.17	NA	NA	NA	NA	3.11	NA	NA
NHS Tayside	-3.77	-3.74	NA	NA	NA	3.02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NHS Greater Glasgow & Clyde	-4.01	-4.05	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NHS Lanarkshire	-4.22	NA	NA	NA	NA	NA	NA	NA	3.04	4.13	4.29	3.47	NA	NA	NA	NA
NHS Ayrshire & Arran	NA	-3.96	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NHS Forth Valley	NA	NA	4.26	3.40	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	-3.43	NA
NHS Grampian	NA	NA	6.72	4.95	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NHS Dumfries & Galloway	NA	NA	NA	-4.66	-6.24	NA	NA	NA	-3.39	NA	NA	NA	NA	NA	NA	NA
NHS Highland	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3.29	NA	NA

```
# qpdf::pdf_rotate_pages(inputFile, pages = 4, angle = 90)
```

#Note for next time: what I want to do at this point is to show the changing colours as a difference change if its only within a small amount of chaning values then ignore the calues and do not #colour the cell, otherwise red fir a rise and green for a fall