Title Placeholder

Load in Specific Packages

```r
library(readr)
library(tidyr)
library(dplyr)
```

```
## 
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
## 
##     filter, lag

## The following objects are masked from 'package:base':
## 
##     intersect, setdiff, setequal, union
```

```r
library(here)
```

```
## here() starts at C:/Users/romin/ToyRepo
```

```r
library(lemon)
library(kableExtra)
```

```
## 
## Attaching package: 'kableExtra'

## The following object is masked from 'package:dplyr':
## 
##     group_rows
```

```r
library(ggplot2)
library(reshape())
```

```
## 
## Attaching package: 'reshape'

## The following object is masked from 'package:dplyr':
## 
##     rename

## The following objects are masked from 'package:tidyr':
## 
##     expand, smiths
```

```r
library(hexbin)
library(data.table)
```

```
## 
## Attaching package: 'data.table'

## The following object is masked from 'package:reshape':
```

```
##
##     melt

## The following objects are masked from 'package:dplyr':
##
##     between, first, last
```
```r
library(GGally)
```
```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```
```r
library(formattable)
library(viridis)
```
```
## Loading required package: viridisLite
```
```r
library(TTR)
library(zoo)
```
```
##
## Attaching package: 'zoo'

## The following objects are masked from 'package:data.table':
##
##     yearmon, yearqtr

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```
```r
library(ggrepel)
```

Load in the data

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

Remove Uncessary Data for Analysis

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

Find Average of All Measures by Year

```r
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

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

Find Last Data Points for Data

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

Display Summary of All Data

```r
inputFile <-"reportReg.pdf"
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")
```

```
## Warning in geom_pointline(data = avgYearly, aes(y = AvgYear)): `geom_pointpath`
## and `geom_pointline` have been soft-deprecated. A replacement can be found in
## ggh4x::geom_pointpath.
```

```
## Warning in geom_pointline(data = movingAvg, aes(y = MA, col = area_name)):
## `geom_pointpath` and `geom_pointline` have been soft-deprecated. A replacement
## can be found in ggh4x::geom_pointpath.
```

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

## [1] "C:\\Users\\romin\\ToyRepo\\Models\\reportReg_output.pdf"

# ggplot(cancerReg,aes(x=year, y= measure, col=area_name)) + geom_line() + geom_point()
# ggplot(cancerReg, aes(x=area_name, y=measure, fill=area_name)) + geom_violin()
# ggparcoord(data= cancerReg, columns = c(), groupColumn = "area_name")
# ggplot(cancerReg, aes(fill=year, y=measure, x=area_name)) + geom_bar(position = "stack", stat = "iden
# ggplot(cancerReg, aes(x = year, y = measure, fill = area_name)) + geom_area(alpha=0.6 , size=.5, col
# scale_fill_viridis(discrete = T)
# ggplot(cancerReg, aes(x = year, y = measure, col = area_name)) +
# geom_point(size = 3, position = position_jitter(h = 0.50, w = 0.50), alpha = 0.5) +
# geom_smooth(se=FALSE, size = 0.7 )
#
# ----------
# cancerReg <- cancerReg %>%
# arrange(year) %>%
# group_by(area_name) %>%
# mutate(MA = rollmean(measure, k = 14, fill = NA, align = "right"))
# finalValues <- cancerReg %>%
# group_by(area_name) %>%
# summarise(
# lastMA = dplyr::last(measure),
# lastYear=dplyr::last(year)
# )
# finalValues
#
# options(repr.plot.width =20, repr.plot.height =20)
```

```r
# subData <- cancerReg[seq(0, nrow(cancerReg), 14), ]
# subData <- subData %>% select(-area_name)
# print(subData)
# ggplot(cancerReg, aes(x = year)) +
#     geom_point(data = subData, aes(x=year, y = MA), color = "black", size = 2) +
#     geom_line(data = subData, aes(y = MA), colour = "black", linetype = "dashed", size = 0.7) +
#     geom_pointline(aes(y = measure, col = area_name), linewidth = 0.5, alpha = 0.5) +
#     # geom_text_repel(data = cancerReg, aes(y = measure, label = area_name), nudge_x = 1, na.rm = TRU

#     geom_text_repel(data = finalValues, aes(
#         x = lastYear,
#         y = lastMA,
#         label = area_name,
#         color=area_name),
#         size = 2.5,
#         fontface = "bold",
#         nudge_y = 7.6,
#         direction = "y",
#         hjust= -0.8,
#         segment.linetype=2,
#         segment.size = 0.5,
#         segment.curvature=0
#         ) +
#     theme(legend.position = "none")
#-------------------------------------------------------------------
# overallPlot <- overallPlot + geom_text_repel(data = finalValues, aes(x = lastYear, y = lastMA, label
# overallPlot
```

```r
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

```r
healthBoardData <- cancerReg %>%
    group_by(year) %>%
    mutate(Percent = measure / sum(measure) * 100)
healthBoardData <- cancerReg %>% pivot_wider(names_from = year, values_from = measure)
healthBoardData <- healthBoardData %>%
    mutate(
        Average = rowMeans(select(., `2003`:`2020`), na.rm = TRUE)
    )
kable(healthBoardData, format = "latex", booktabs = TRUE)
```

area_code	area_name	2003	2004	2005	2006	2007	2008	2009	2010	2011	201
S08000015	NHS Ayrshire & Arran	649.7	642.8	631.6	626.1	636.4	648.2	652.2	643.8	643.2	654.
S08000016	NHS Borders	614.6	616.8	632.7	641.5	643.5	627.5	633.8	614.8	614.5	608.
S08000017	NHS Dumfries & Galloway	620.5	645.2	639.4	642.1	613.4	619.8	609.1	628.9	617.6	602.
S08000019	NHS Forth Valley	651.8	648.2	642.8	618.3	633.2	638.9	649.8	634.1	620.3	617.
S08000020	NHS Grampian	617.8	601.5	589.7	586.6	595.9	605.0	605.4	608.5	606.2	606.
S08000022	NHS Highland	602.9	594.5	598.9	613.3	627.4	627.6	625.8	625.4	616.8	614.
S08000024	NHS Lothian	655.1	660.8	659.4	666.1	668.7	683.4	681.6	691.8	684.3	691.
S08000025	NHS Orkney	589.0	582.4	587.1	598.1	587.4	589.3	543.0	569.2	531.3	549.
S08000026	NHS Shetland	616.0	636.7	571.8	574.3	562.6	594.9	596.0	663.4	685.6	731.
S08000028	NHS Western Isles	652.0	628.1	610.3	664.4	707.2	726.0	715.0	674.4	646.0	651.
S08000029	NHS Fife	611.4	614.0	630.7	641.7	648.9	643.9	640.4	650.0	651.3	659.
S08000030	NHS Tayside	611.2	607.3	600.6	596.0	617.6	621.1	613.7	603.1	606.9	616.
S08000031	NHS Greater Glasgow & Clyde	697.3	693.9	695.0	687.6	698.1	710.9	715.4	714.4	712.3	713.
S08000032	NHS Lanarkshire	633.7	632.7	643.5	637.3	640.5	649.8	665.6	675.0	677.5	672.

`healthBoardData`

```
## # A tibble: 14 x 21
##    area_code area_name   `2003` `2004` `2005` `2006` `2007` `2008` `2009` `2010`
##    <chr>     <chr>        <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>
##  1 S08000015 NHS Ayrshi~   650.   643.   632.   626.   636.   648.   652.   644.
##  2 S08000016 NHS Borders   615.   617.   633.   642.   644.   628.   634.   615.
##  3 S08000017 NHS Dumfri~   620.   645.   639.   642.   613.   620.   609.   629.
##  4 S08000019 NHS Forth ~   652.   648.   643.   618.   633.   639.   650.   634.
##  5 S08000020 NHS Grampi~   618.   602.   590.   587.   596.   605    605.   608.
##  6 S08000022 NHS Highla~   603.   594.   599.   613.   627.   628.   626.   625.
##  7 S08000024 NHS Lothian   655.   661.   659.   666.   669.   683.   682.   692.
##  8 S08000025 NHS Orkney    589    582.   587.   598.   587.   589.   543    569.
##  9 S08000026 NHS Shetla~   616    637.   572.   574.   563.   595.   596    663.
## 10 S08000028 NHS Wester~   652    628.   610.   664.   707.   726    715    674.
## 11 S08000029 NHS Fife      611.   614    631.   642.   649.   644.   640.   650
## 12 S08000030 NHS Tayside   611.   607.   601.   596    618.   621.   614.   603.
## 13 S08000031 NHS Greate~   697.   694.   695    688.   698.   711.   715.   714.
## 14 S08000032 NHS Lanark~   634.   633.   644.   637.   640.   650.   666.   675
## # i 11 more variables: `2011` <dbl>, `2012` <dbl>, `2013` <dbl>, `2014` <dbl>,
## #   `2015` <dbl>, `2016` <dbl>, `2017` <dbl>, `2018` <dbl>, `2019` <dbl>,
## #   `2020` <dbl>, Average <dbl>
```

#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