Appendix III - Analyse the Data

Jamie Cash

2023-07-06

Load the processed data

Weather station data provided by the MET office is provided on their website in txt files. This data was prepared and processed to create two summary datasets: seasonal temperatures by year; and annual temperatures by station. Load these datasets.

```
seasonal <- read.csv('seasonal_summary.csv')
head(seasonal)
```

```
##   year season high  low average_max average_min
## 1 1865 Autumn 22.6  3.1   15.577778    7.500000
## 2 1865 Summer 22.3 10.8   20.988889   11.655556
## 3 1866 Autumn 17.0  3.7   13.800000    7.300000
## 4 1866 Spring 15.7  1.7   12.255556    4.133333
## 5 1866 Summer 21.9 10.5   20.222222   11.411111
## 6 1866 Winter  9.9  0.9    8.822222    2.866667
```

```
station <- read.csv('annual_station_summary.csv')
head(station)
```

```
##   year        station_name station_long station_lat high  low total_rainfall
## 1 1978           Aberporth     -4.56999    52.13914 16.3  1.5          734.6
## 2 1978              Armagh     -6.64866    54.35234 18.1  0.4          773.7
## 3 1978 Ballypatrick Forest     -6.15336    55.18062 15.6  0.5            0.0
## 4 1978            Bradford     -1.77234    53.81341 17.7 -0.9          869.8
## 5 1978             Braemar     -3.39635    57.00612 16.6 -8.4          923.9
## 6 1978            Camborne     -5.32656    50.21782 17.5  5.0          380.1
```

Add calculations

Add mean, rolling mean and delta to station data. Once calculated, we only need last year for plotting on maps, so filter.

```
# Calculate means, rolling means and deltas.
station_means_2022 <- station %>%
  group_by(station_name) %>%
  arrange(year) %>%
  mutate(
```

```
      high_mean=mean(high, na.rm=TRUE),
      low_mean=mean(low, na.rm=TRUE),
      rainfall_mean=mean(total_rainfall, na.rm=TRUE),

      high_5yr_roll=rollapplyr(high, 5, partial=TRUE, fill=NA, FUN=function(x) mean(x, na.rm=TRUE)),
      low_5yr_roll=rollapplyr(low, 5, partial=TRUE, fill=NA, FUN=function(x) mean(x, na.rm=TRUE)),
      rainfall_5yr_roll=rollapplyr(total_rainfall, 5, partial=TRUE, fill=NA, FUN=function(x) mean(x, na

      high_delta=high_5yr_roll-high_mean,
      low_delta=low_5yr_roll-low_mean,
      rainfall_delta=rainfall_5yr_roll-rainfall_mean
    ) %>%
  filter(year==2022)


head(station_means_2022)
```

```
## # A tibble: 6 x 16
## # Groups:   station_name [6]
##    year station_name        station_long station_lat  high   low total_rainfall
##   <int> <chr>                      <dbl>       <dbl> <dbl> <dbl>          <dbl>
## 1  2022 Aberporth                  -4.57        52.1  20.6   2.3           724.
## 2  2022 Armagh                     -6.65        54.4  21.8   0.6           859.
## 3  2022 Ballypatrick Forest        -6.15        55.2  18.6   2            1217
## 4  2022 Bradford                   -1.77        53.8  22.7   0.4           845.
## 5  2022 Braemar                    -3.40        57.0  19.4  -3.1           964.
## 6  2022 Camborne                   -5.33        50.2  22.4   3.8           998.
## # i 9 more variables: high_mean <dbl>, low_mean <dbl>, rainfall_mean <dbl>,
## #   high_5yr_roll <dbl>, low_5yr_roll <dbl>, rainfall_5yr_roll <dbl>,
## #   high_delta <dbl>, low_delta <dbl>, rainfall_delta <dbl>
```

Create map to plot data against

```
world <- ne_countries(scale = "medium", returnclass = "sf")

uk_map <- ggplot(data=world,) +
    geom_sf(fill='#59A608') +
    xlab("") +
    ylab("") +
    theme(axis.text.x=element_blank() , axis.ticks.x=element_blank(),
        axis.text.y=element_blank() , axis.ticks.y=element_blank(),
        panel.background = element_rect(fill = 'lightblue', colour = 'lightblue'),
        panel.grid.major=element_blank(), panel.grid.minor=element_blank()  ) +
    coord_sf(xlim = c(2, -11), ylim = c(49, 59), expand = FALSE)


uk_map
```

Answer the questions

Have the seasons shifted?

- Is the average highest temperature for December, January and February across all UK locations for the last 5 years higher than the historical average?

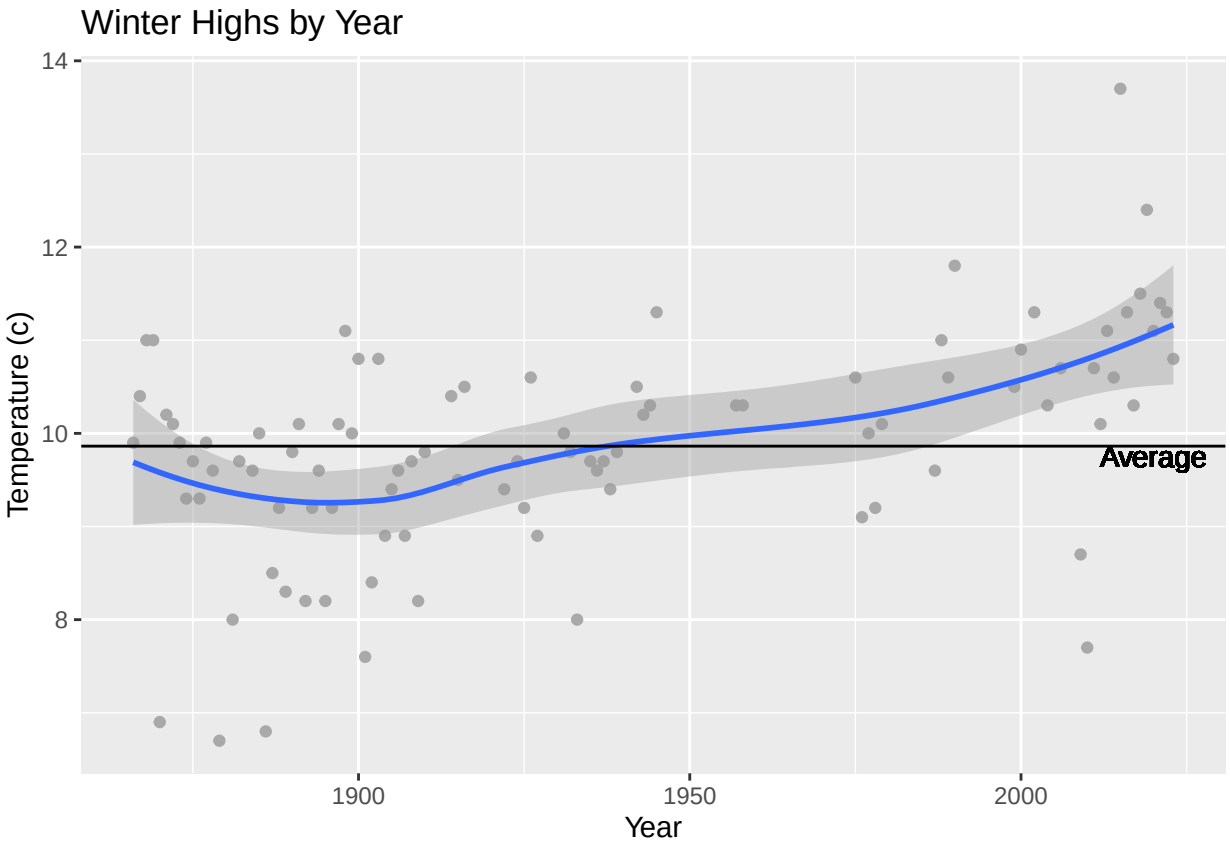
Chart

```
winter <- seasonal %>% filter(season=='Winter')
avg = mean(winter$high, na.rm=TRUE)

plot <- winter %>%
  ggplot(aes(x=year, y=high)) +
  geom_point(color='darkgray') +
  geom_smooth() +
  labs(title="Winter Highs by Year", x="Year", y="Temperature (c)") +
  geom_hline(yintercept=avg, show.legend="Average") +
  geom_text(aes(2020, avg,label="Average", vjust=1))

ggsave("winter_highs.png", plot)

plot
```



- Is the average lowest temperature for March, April and May across all UK locations for the last 5 years lower than the historical average?

Historical mean of spring lows

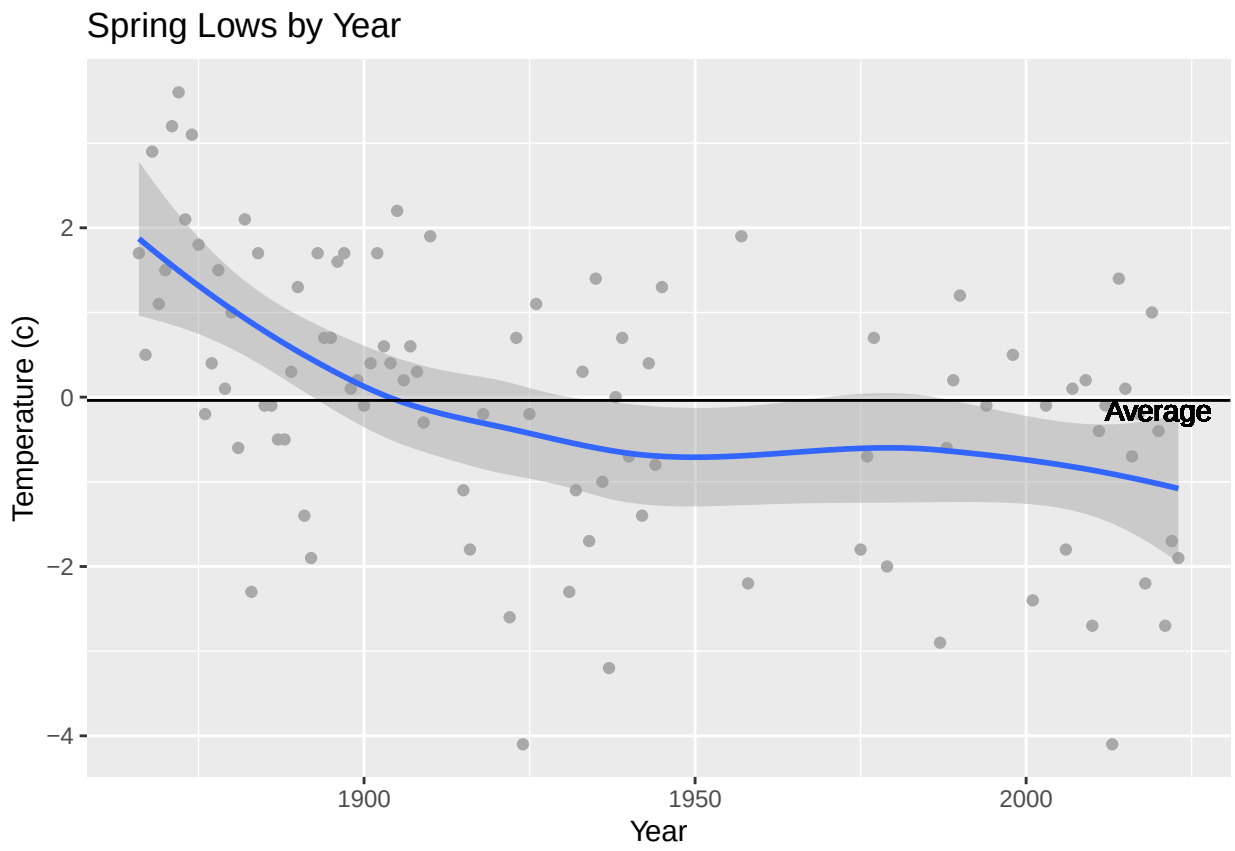
Chart

```
spring <- seasonal %>% filter(season=='Spring')
avg = mean(spring$low, na.rm=TRUE)

plot <- spring %>%
  ggplot(aes(x=year, y=low)) +
  geom_point(color='darkgray') +
  geom_smooth() +
  labs(title="Spring Lows by Year", x="Year", y="Temperature (c)") +
  geom_hline(yintercept=avg, show.legend="Average") +
  geom_text(aes(2020, avg,label="Average", vjust=1))

ggsave("spring_lows.png", plot)

plot
```



Are we gettig higher temperature peaks?

- For each location, is the average highest annual temperature over the last 5 years higher than the historical average?

Get the min and max values for the scale.

```
min(station_means_2022$high_delta)
```

```
## [1] 0.3856522
```

```
max(station_means_2022$high_delta)
```

```
## [1] 1.903478
```

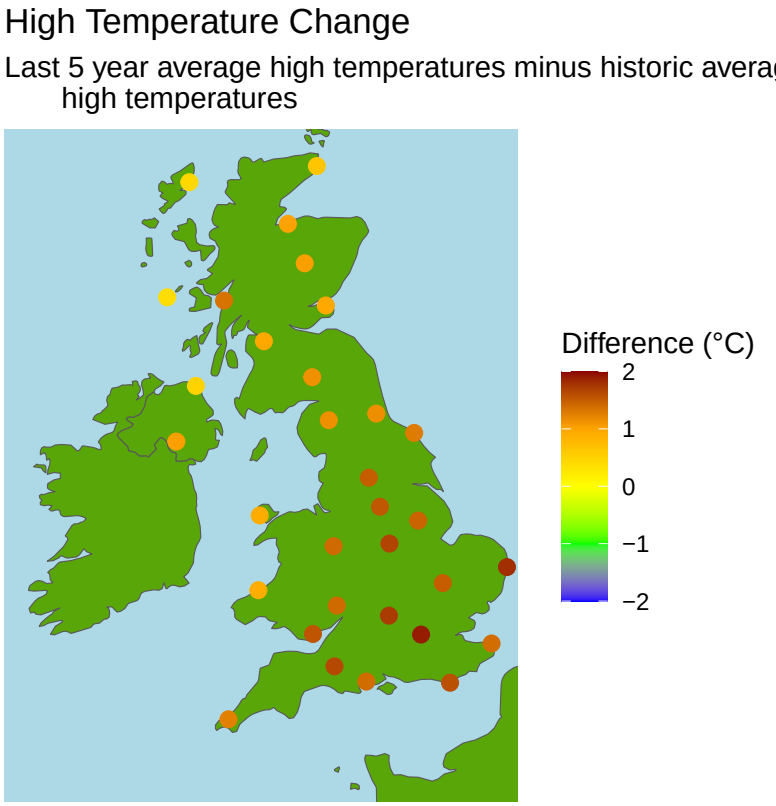
Plot the deltas on the chart. Scale should be even. We will use -2 to 2 as 1.9 is the max.

```
high_delta_map <- uk_map +
  geom_point(data = station_means_2022,
             mapping = aes(x=station_long, y=station_lat, colour=high_delta),
             size=2.5) +
  scale_colour_gradientn(colors=c("blue", "green", "yellow", "orange", "darkred"),
                         limits = c(-2,2)) +
```

```
  labs(title="High Temperature Change",
       color="Difference (°C)",
       subtitle="Last 5 year average high temperatures minus historic average
       high temperatures")


ggsave("station_high_deltas.png", high_delta_map)

high_delta_map
```



Are we getting lower temperature dips?

- For each location, is the average lowest annual temperature over the last 5 years lower than the historical average?

Plot the deltas on the chart. Scale should the same as the previous chart for consistency.

Get the min and max values for the report.

```
min(station_means_2022$low_delta)
```

```
## [1] -0.09782609
```

```
max(station_means_2022$low_delta)
```

```
## [1] 0.9243478
```

Some stations have a lower low. Which ones? We will include these in the report.

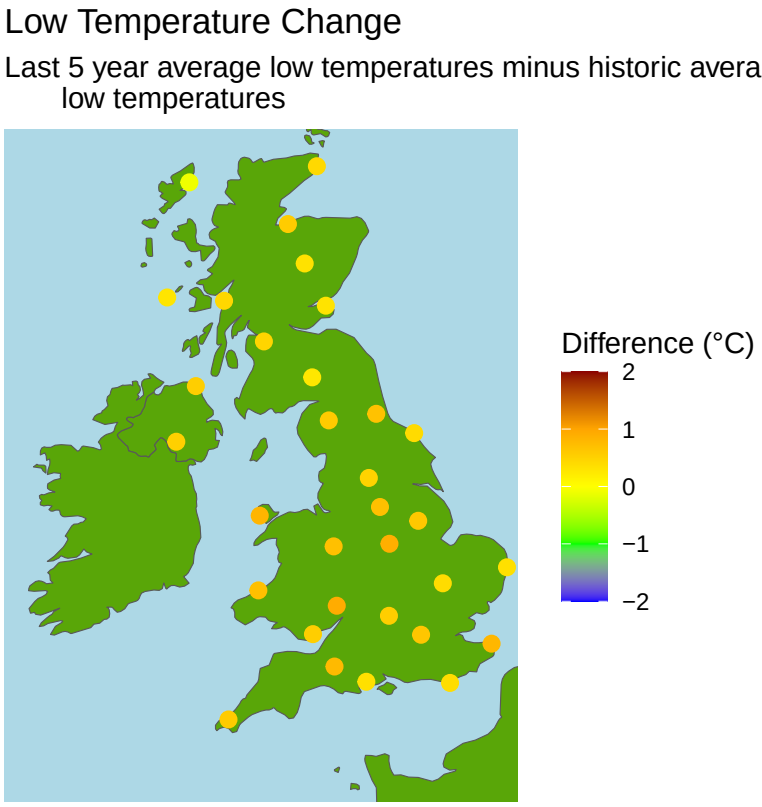
```
station_means_2022 %>% filter(low_delta <=0)
```

```
## # A tibble: 1 x 16
## # Groups:   station_name [1]
##    year station_name      station_long station_lat  high   low total_rainfall
##   <int> <chr>                    <dbl>       <dbl> <dbl> <dbl>          <dbl>
## 1  2022 Stornoway Airport        -6.32        58.2  16.7   1.8          1316.
## # i 9 more variables: high_mean <dbl>, low_mean <dbl>, rainfall_mean <dbl>,
## #   high_5yr_roll <dbl>, low_5yr_roll <dbl>, rainfall_5yr_roll <dbl>,
## #   high_delta <dbl>, low_delta <dbl>, rainfall_delta <dbl>
```

```
low_delta_map <- uk_map +
  geom_point(data = station_means_2022,
             mapping = aes(x=station_long, y=station_lat, colour=low_delta),
             size=2.5) +
  scale_colour_gradientn(colors=c("blue", "green", "yellow", "orange", "darkred"),
                         limits = c(-2,2)) +
  labs(title="Low Temperature Change",
       color="Difference (°C)",
       subtitle="Last 5 year average low temperatures minus historic average
       low temperatures")


ggsave("station_low_deltas.png", low_delta_map)

low_delta_map
```



Are we getting more or less rainfall?

- For each location, is the average yearly rainfall for the last 5 years lower or higher than the historical average?

Get the min and max values for the scale.

```
min(station_means_2022$rainfall_delta)
```

```
## [1] -93.75087
```

```
max(station_means_2022$rainfall_delta)
```

```
## [1] 292.1735
```

Plot the deltas on the chart. Scale should be even. We will use -300 to 300 as highest delta is 292.

```
rainfall_delta_map <- uk_map +
  geom_point(data = station_means_2022,
             mapping = aes(x=station_long, y=station_lat, colour=rainfall_delta),
             size=2.5) +
  scale_colour_gradientn(colors=c("yellow", "brown", "green"),
                         limits = c(-300,300)) +
```

```
  labs(title="Rainfall Change",
       color="Difference (mm)",
       subtitle="Last 5 year average rainfall minus historic average rainfall")


ggsave("station_rainfall_deltas.png", rainfall_delta_map)

rainfall_delta_map
```