EXPLORING WEATHER TRENDS

1.Extract data from database - Extracted the data from database using the following SQL queries

Select * from city_data where city='Boston';   *//to get city wide temperature data//*

Select * from global_data;   *//to get global temperature data//*

2.Open CSV – Opened CSV file in R using the following code

data = read.csv("C:/Users/avala/Downloads/results.csv")   *//loads global data//*

data1 = read.csv("C:/Users/avala/Downloads/results (1).csv")   *//loads boston data//*

Calculated moving averages in Excel and then loaded the files into R. Calculated 7 day moving averages using AVERAGE() function in Excel for both the temp columns.

Global moving avg calculation snapshot

C8			×	✓	fx	=AVERAGE(B2:B8)		

	A	B	C	D	E	F	G
1	year	avg_temp	moving_avg				
2	1750	8.72					
3	1751	7.98					
4	1752	5.78					
5	1753	8.39					
6	1754	8.47					
7	1755	8.36					
8	1756	⚠ 35	8.078571429				
9	1757	9.02	8.121428571				
10	1758	6.74	7.944285714				
11	1759	7.99	8.26				
12	1760	7.19	8.088571429				
13	1761	8.77	8.131428571				
14	1762	8.61	8.167142857				
15	1763	7.5	7.974285714				
16	1764	8.4	7.885714286				
17	1765	8.25	8.101428571				
18	1766	8.41	8.161428571				
19	1767	8.22	8.308571429				
20	1768	6.78	8.024285714				
21	1769	7.69	7.892857143				
22	1770	7.69	7.92				
23	1771	7.85	7.841428571				
24	1772	8.19	7.832857143				

Boston moving avg calculation snapshot

E8			×	✓	f_x	=AVERAGE(D2:D8)		

	A	B	C	D	E	F	G
1	year	city	country	avg_temp	moving_avg		
2	1743	Boston	United Sta	1.19			
3	1744	Boston	United Sta	9.63			
4	1745	Boston	United Sta	-1.37			
5	1746	Boston	United States				
6	1747	Boston	United States				
7	1748	Boston	United States				
8	1749	Boston	United States	⚠	3.15		
9	1750	Boston	United Sta	7.88	5.38		
10	1751	Boston	United Sta	8.6	5.036666667		
11	1752	Boston	United Sta	0.36	5.613333333		
12	1753	Boston	United Sta	7.35	6.0475		
13	1754	Boston	United Sta	7.75	6.388		
14	1755	Boston	United Sta	4.28	6.036666667		
15	1756	Boston	United Sta	7.76	6.282857143		
16	1757	Boston	United Sta	6.65	6.107142857		
17	1758	Boston	United Sta	6.09	5.748571429		
18	1759	Boston	United Sta	6.8	6.668571429		
19	1760	Boston	United Sta	5.53	6.408571429		
20	1761	Boston	United Sta	8.05	6.451428571		
21	1762	Boston	United Sta	7.42	6.9		
22	1763	Boston	United Sta	4.99	6.504285714		
23	1764	Boston	United Sta	7.36	6.605714286		
24	1765	Boston	United Sta	6.73	6.697142857		
25	1766	Boston	United Sta	7.96	6.862857143		

3.Create a line chart – Created a line chart in R using the following code

In order to compare the temperatures, we need data that pertains to the same year in both the sets.

To achieve that, created an inner join, joining by year, to get rows in both the tables when the year matches.

```
library(ggplot2)
library(tidyverse)
final_data <- data %>% inner_join(data1, by=c("year" = "year"))
```

```
> head(final_data)
  year avg_temp.x moving_avg.x  city          country avg_temp.y moving_avg.y
1 1750       8.72           NA Boston United States       7.88     5.380000
2 1751       7.98           NA Boston United States       8.60     5.036667
3 1752       5.78           NA Boston United States       0.36     5.613333
4 1753       8.39           NA Boston United States       7.35     6.047500
5 1754       8.47           NA Boston United States       7.75     6.388000
6 1755       8.36           NA Boston United States       4.28     6.036667
> |
```

final_data <- final_data %>% rename(Global_Avg=moving_avg.x, Boston_Avg=moving_avg.y)

#Approach 1

ggplot(final_data, aes(x=year)) + geom_line(aes(y = Global_Avg), color = "darkred") +

geom_line(aes(y = Boston_Avg), color="steelblue") + labs(y="Moving avg", title = "Global avg temp vs Boston avg temp") + theme_minimal()

final_data <- final_data %>% select(year, Global_Avg, Boston_Avg) %>% gather(key = "variable", value = "value", -year)

#With legends

ggplot(final_data, aes(x = year, y = value)) + geom_line(aes(color = variable)) + scale_color_manual(values = c("darkred", "steelblue")) + labs(y="Moving avg", title = "Global avg temp vs Boston avg temp") +

theme_minimal()
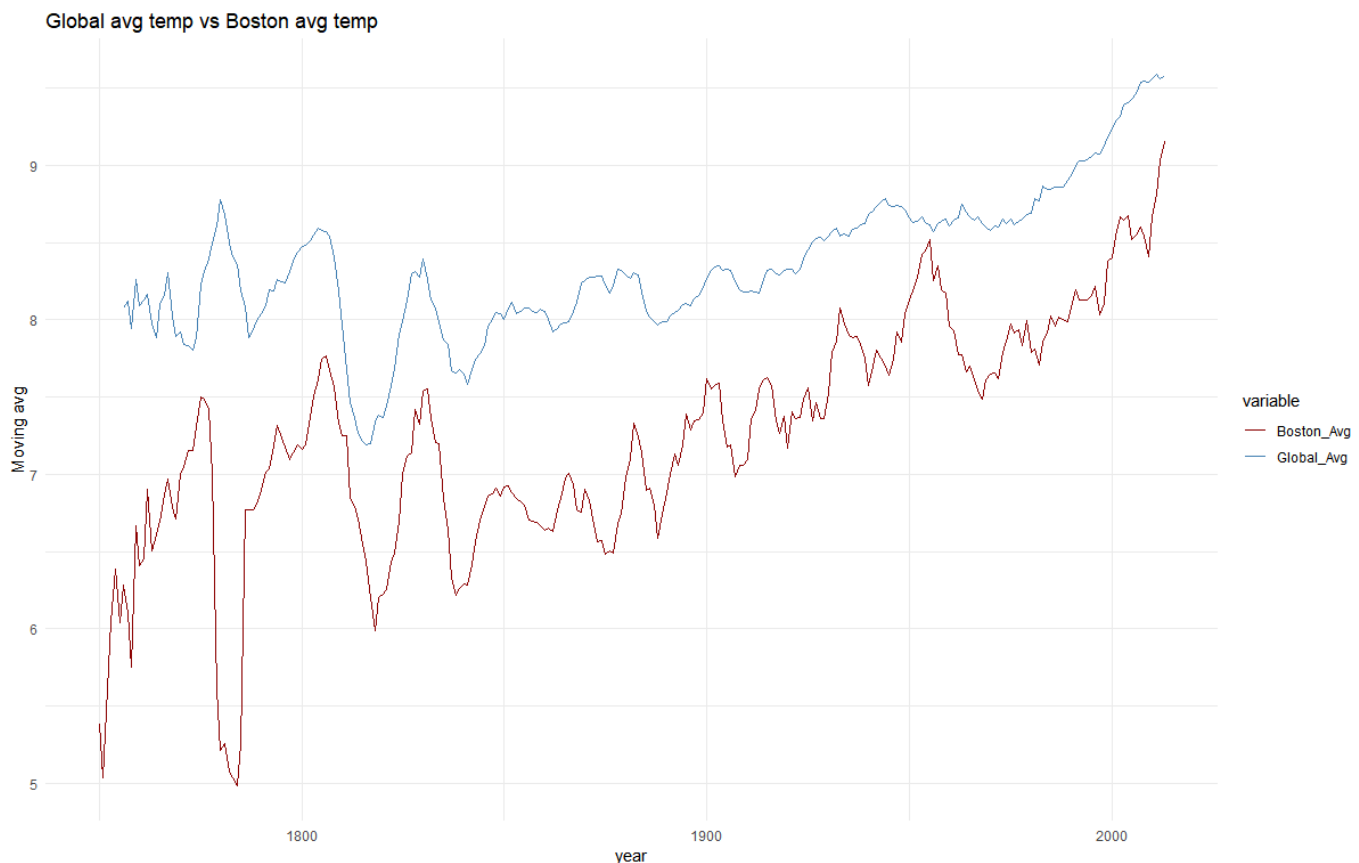
4.Make Observations

1. There's a drastic difference in the period before 1800 where global average temperature has been in the range of 8-9 while Boston avg is less ranging from 5-7. This could be because Boston is generally very cold all year long.

2. There has been only 2 major downfalls in Boston avg one before the year 1800 and one between the years 1800 and 1900.

3. One similarity is that both the averages kept increasing right immediately after the downfall in early 1800's.

4. Global avg has been a smooth curve while Boston avg showed some drastic changes(peaks and lows).