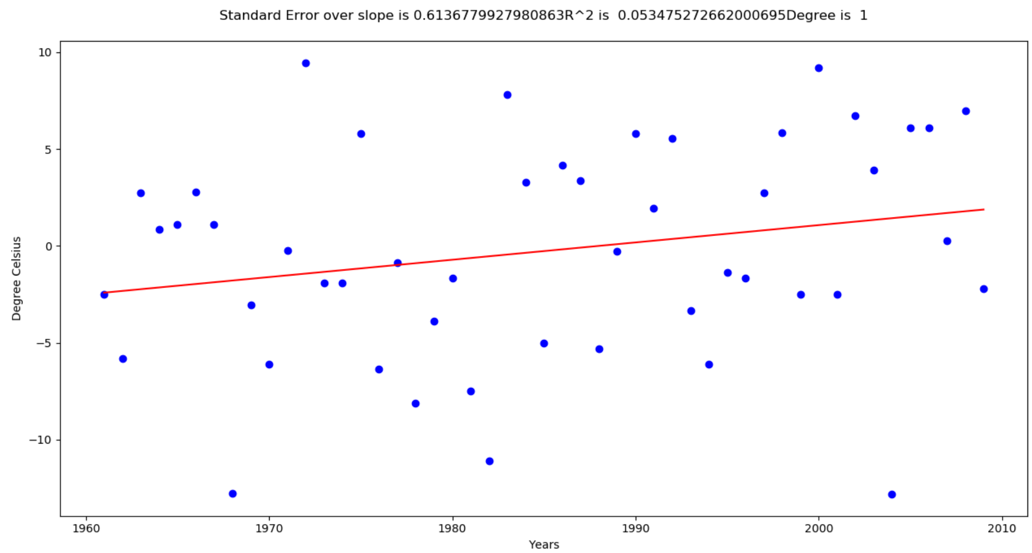
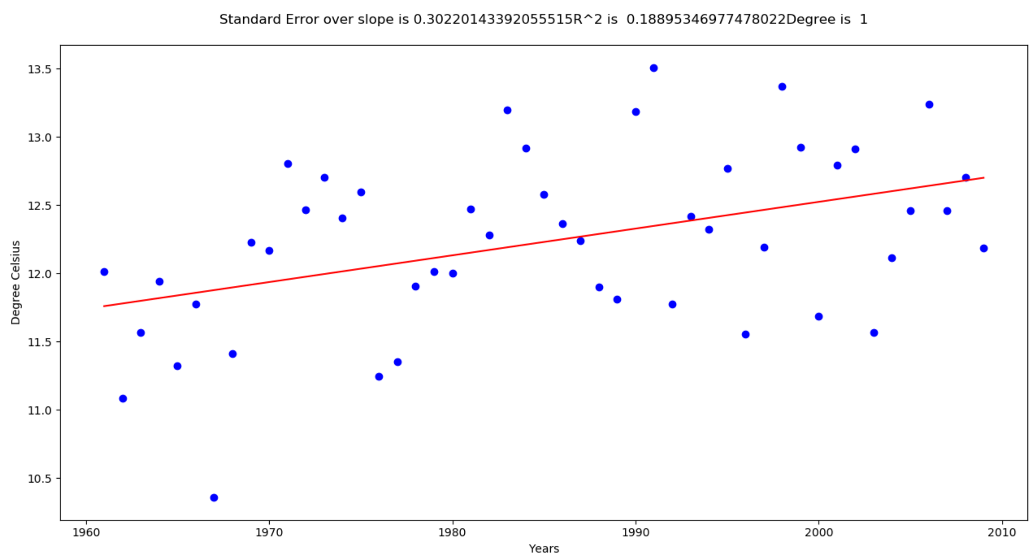
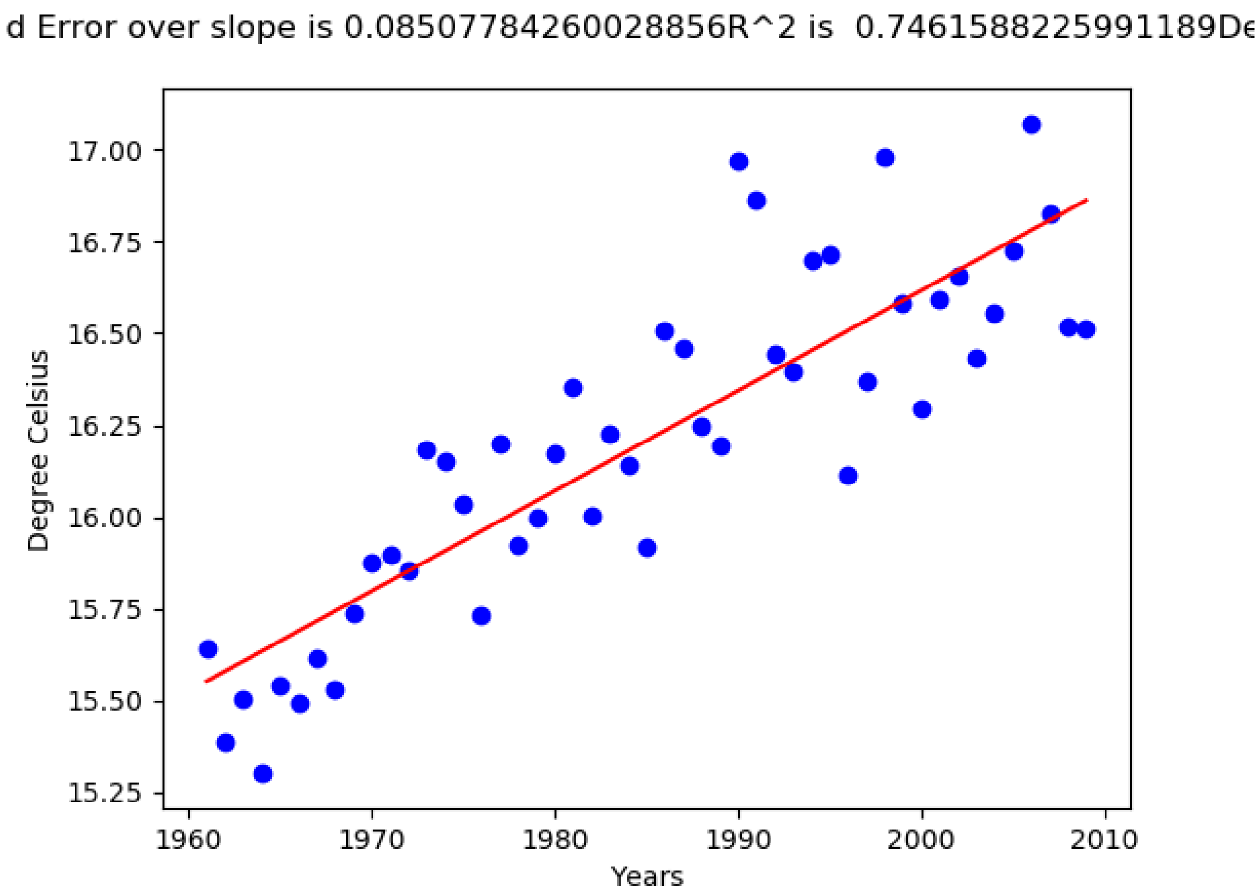
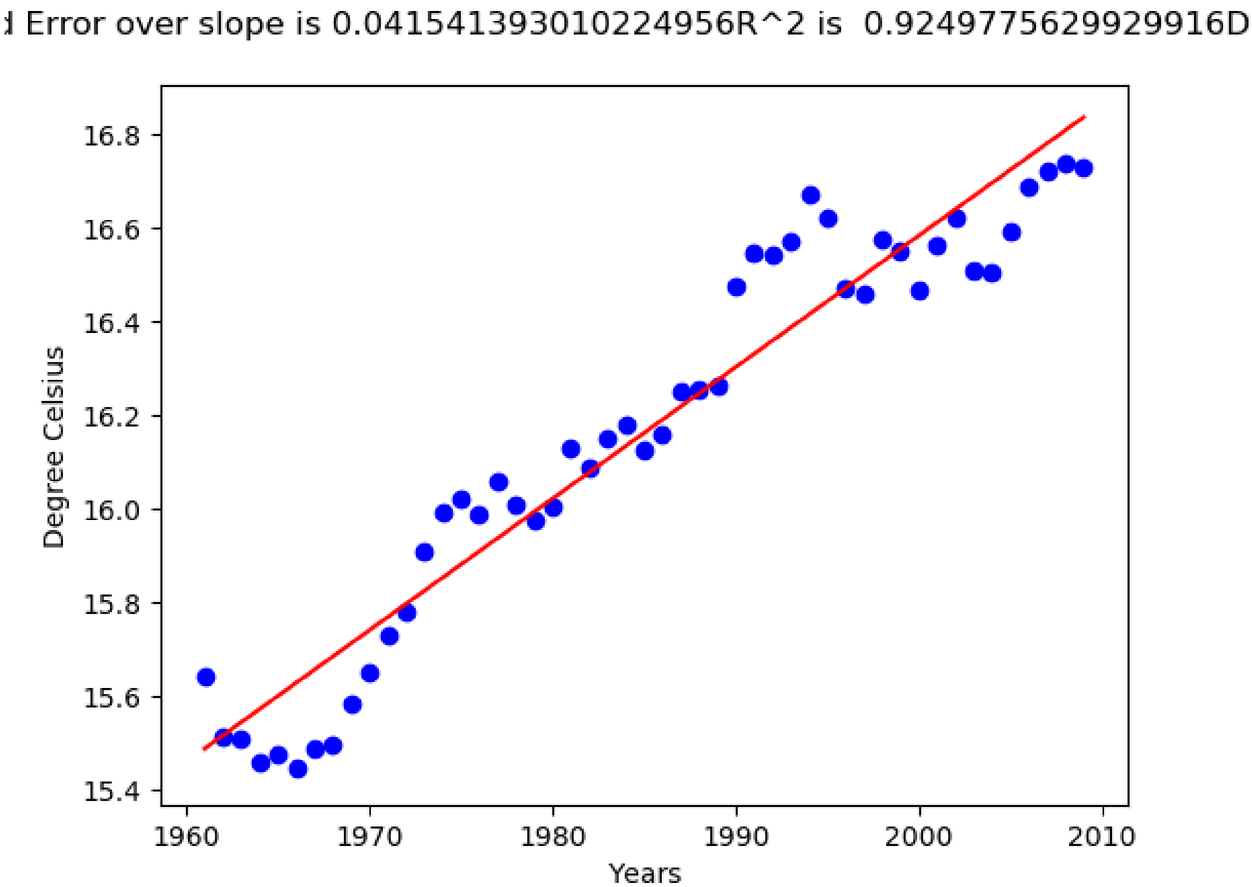


The R^2 value for Jan 10th across the years is significantly less than the R^2 value for the average across the years, we can see this reflect on the best fit line itself. This is because there can be more variations across the temperature on a specific day every year over let's say the average temperature for that year. This supports the changing temperature theory, because on a year we have a better fit however across the years for the same day (if temperature wasn't changing) we would have the same or similar R^2 values, but we don't, showing climate change. We can see that both the models are linearly increasing, for the second one(graph with the average yearly temperature) the standard error over slope is less than 0.5 which means that it is a trend.However, one can refute the point by saying that the first graph shows a value above 0.5 so it might not be a trend, but even if we take average of those values we do get a number below 0.5, which is showing that the temperature of New York is increasing and it is a trend. Therefore, we can infer that it is a trend that global warming is happening across the years. The R^2 value for the first one is 0.05 and the second one is 0.19. The yearly average smooths out the day to day noise. Therefore, the first graph is more noisy

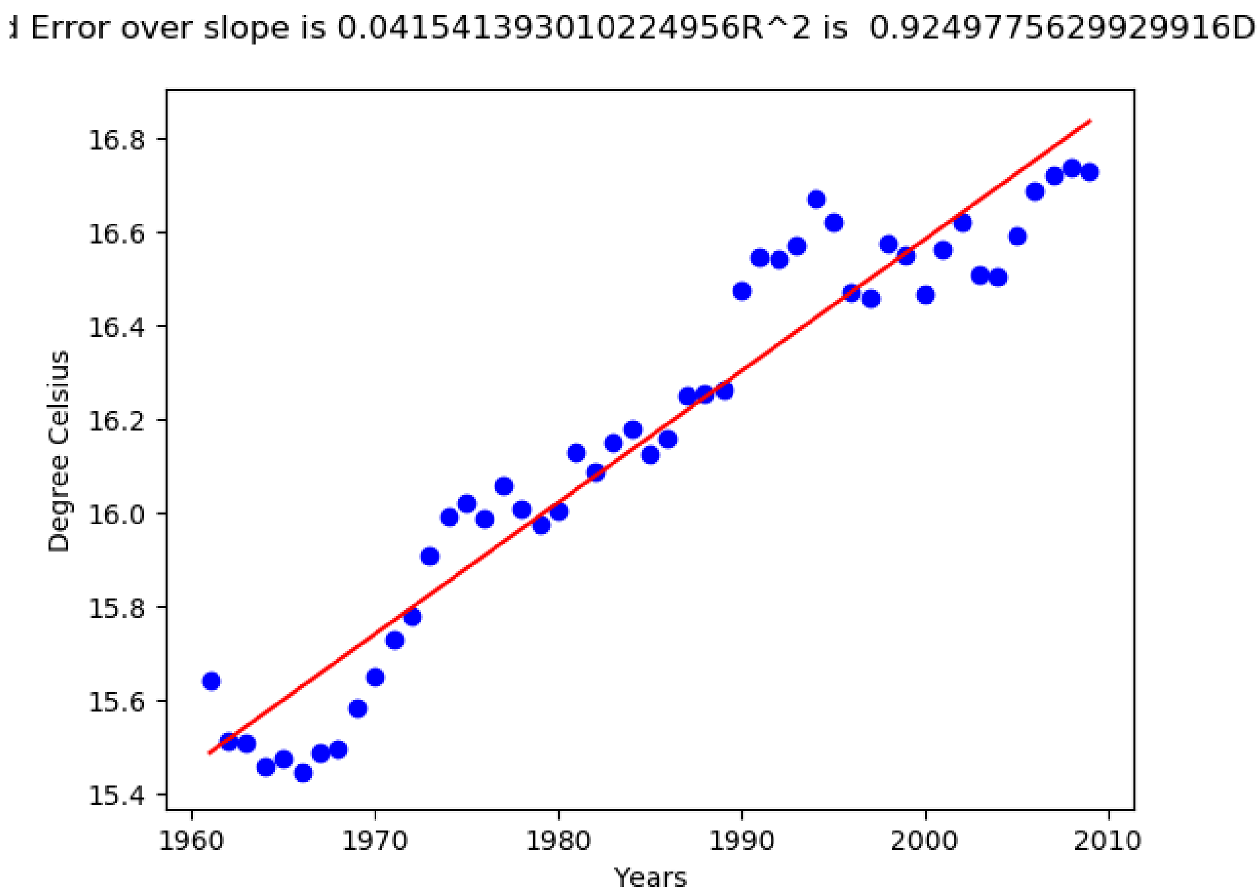

Standard Error over slope is 0.30220143392055515R^2 is 0.18895346977478022Degree is 1
13.5
13.0
12.5
12.0
11.5
11.0
10.5
Degree Celsius
1960
1970
1980
1990
2000
2010
Years

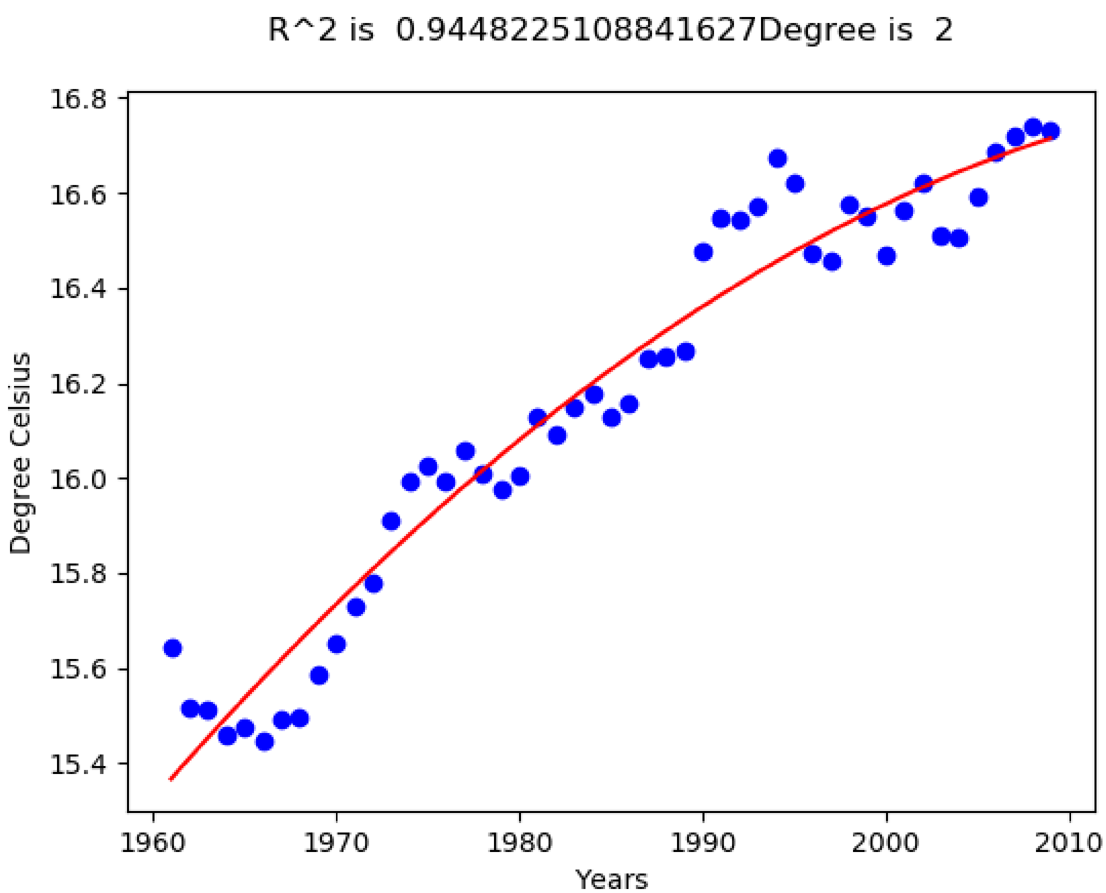


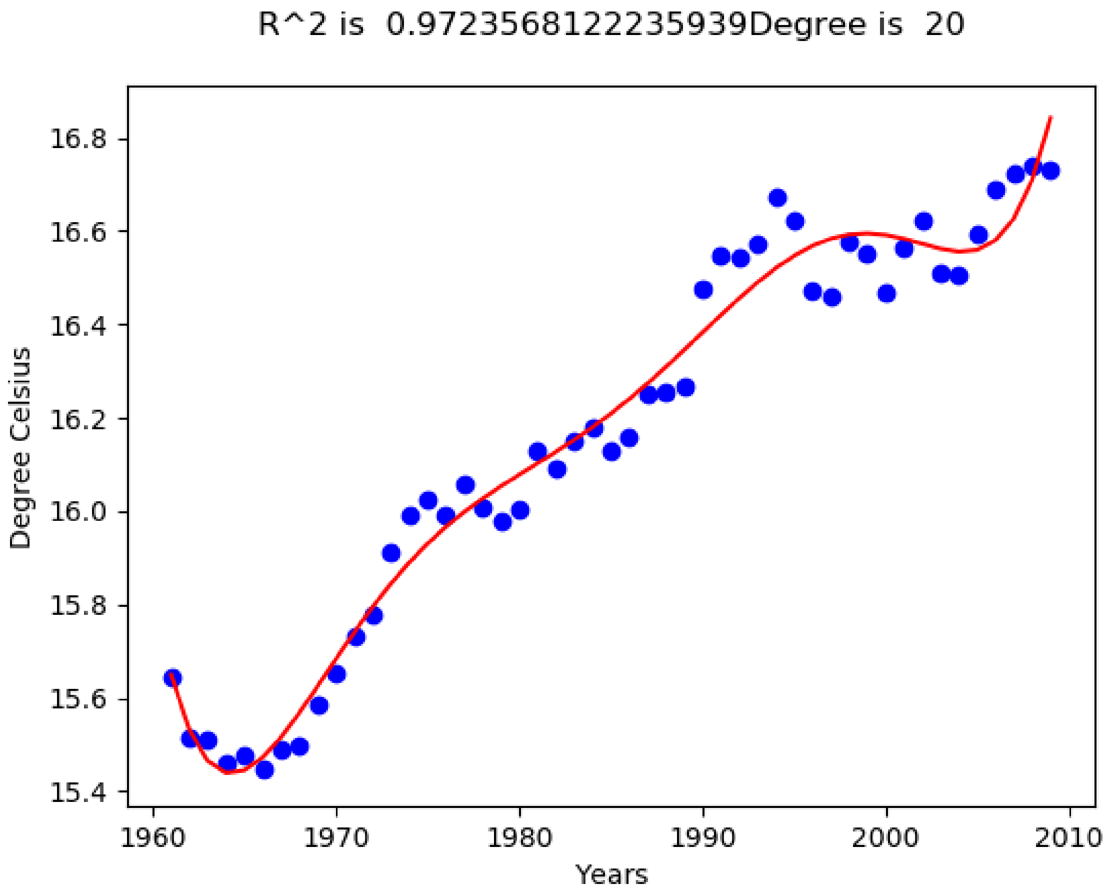
So, this is the average over multiple cities and this supports our claim of global warming even more. The Graph has a good fit of 0.746 and it is significantly less noisier compared to when we did the graph for only New York. Because of the high R^2 value we can see the resulting fit to be better. The reason why we have less noise is just because we have more data points therefore we can make the model train with more data points.More data points also smoothens out any local extremities. The R^2 value would vary if we had more cities let's say like a 100 cities, however that also varies from where we are getting the cities from. Let's say if we get 50 cities from Europe and then 50 from Africa, the R^2 value would decrease because there would be more noise, it's definitely going to be harder to find a trend across cities from different geographic location. Therefore, the vice versa is true, if we had cities from the same location and more of them then the R^2 would actually increase as we have more data points to train our model with similar trends. One final thing to note is the SE/Slope is 0.08 and the graph is linearly increasing, therefore strongly suggesting that global warming is a trend



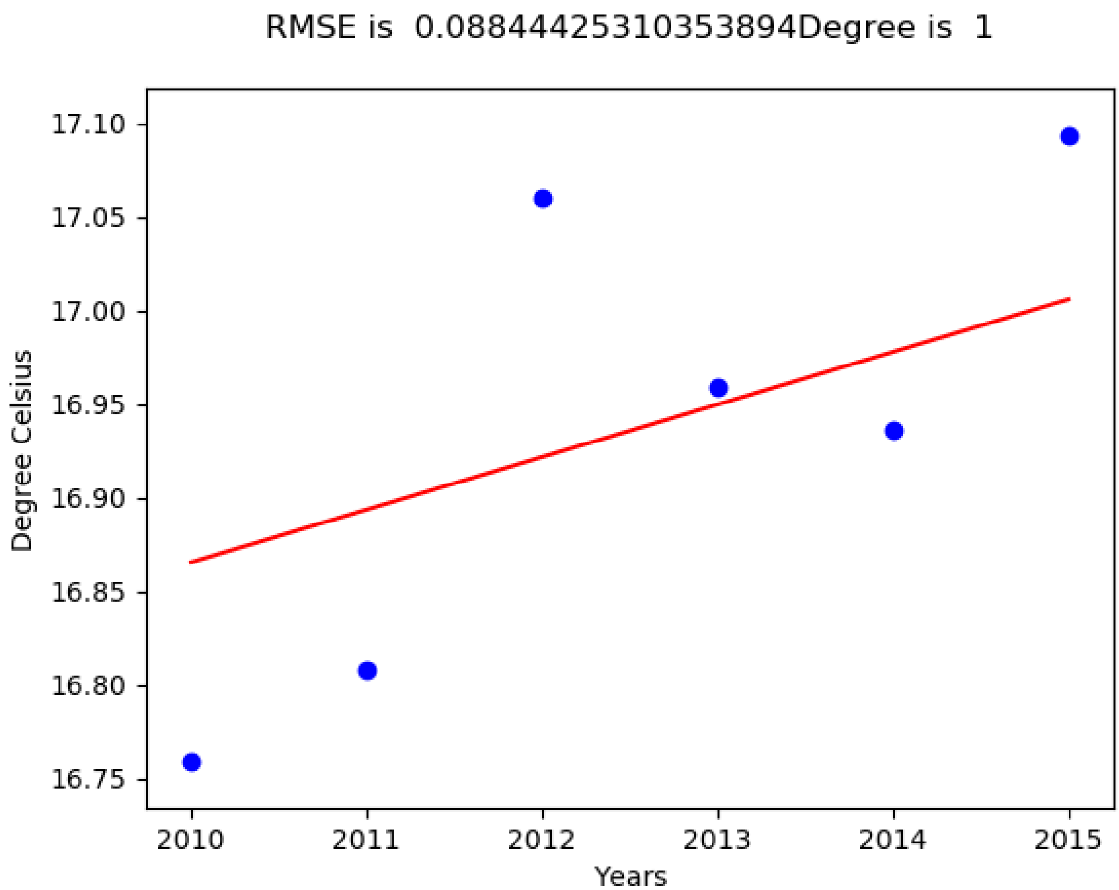
The R^2 value for this graph is even better than the previous in particular it is 0.92. This is because, we are doing an average of a 5 year window therefore reducing any local noise even more which is the reason why we have a better fit. This also supports global warming because the SE over slope is less than 0.5 ie 0.04 and it is a linearly increasing model

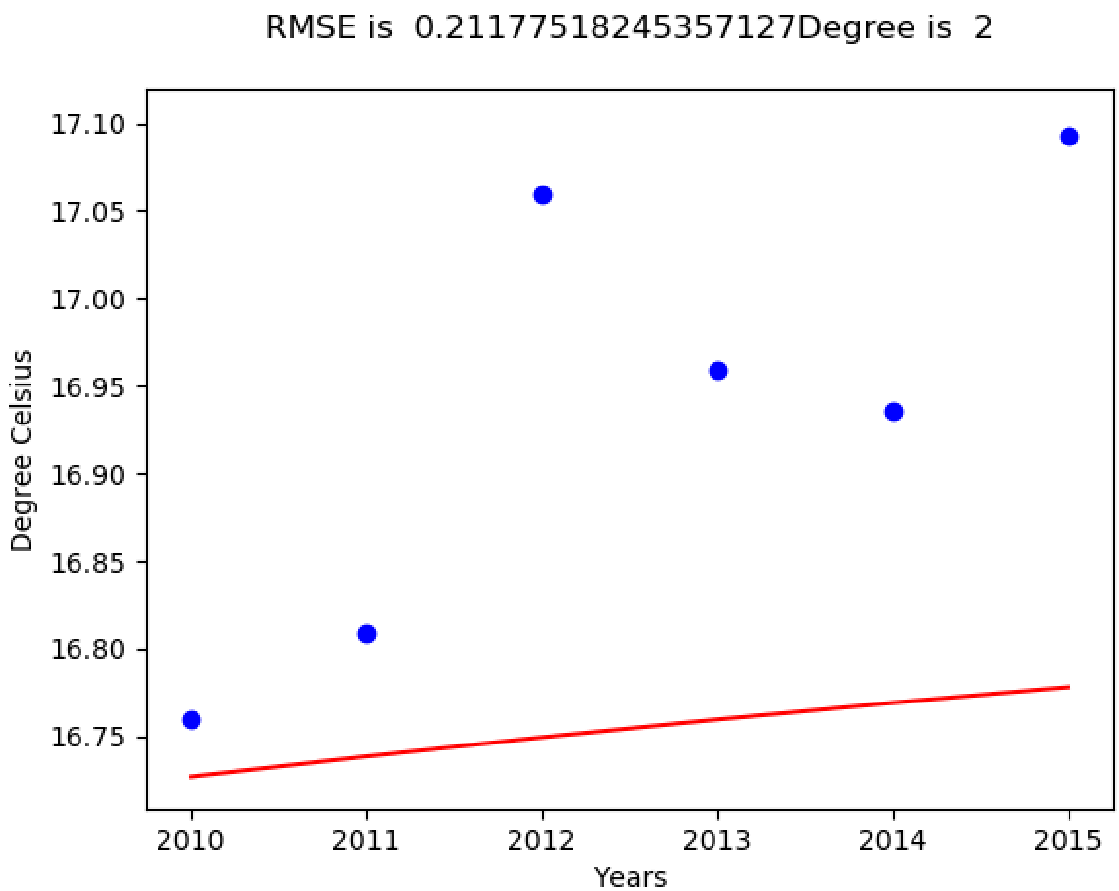

d Error over slope is 0.041541393010224956R^2 is 0.9249775629929916D
Degree Celsius
Years
16.8
16.6
16.4
16.2
16.0
15.8
15.6
15.4
1960
1970
1980
1990
2000
2010

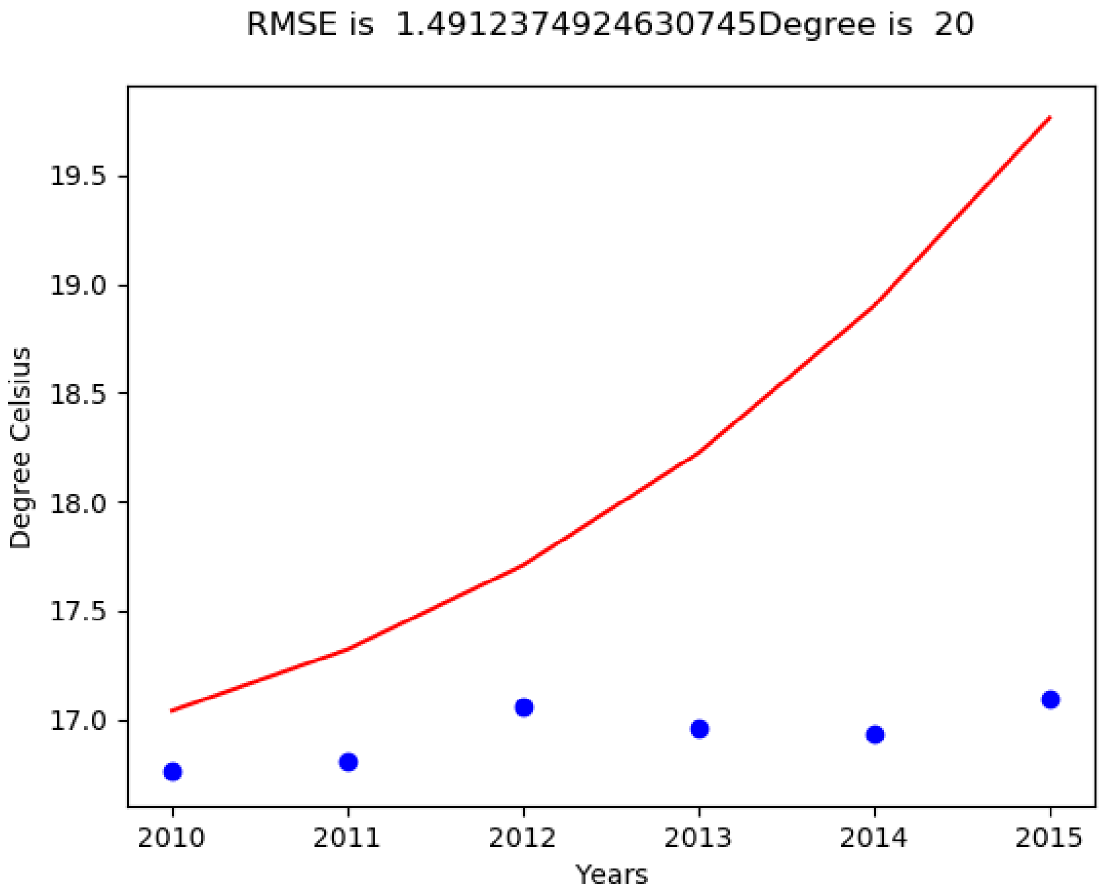

R^2 is 0.9448225108841627Degree is 2
16.8
16.6
16.4
16.2
16.0
15.8
15.6
15.4
Degree Celsius
1960
1970
1980
1990
2000
2010
Years



So, the higher the degree the better the value for R^2 and we can see that trend. So, for degree one we have R^2 equals 0.92, for degree 2 it is 0.944 and for degree 20 it is 0.97. This makes sense because in higher degrees, we have more variables, so the linear regressions adjusts it to fit all the data points, however is that a good predictor? We cannot tell unless we try it on a different dataset, because for this dataset the higher the degree the better the graph fit will be. So, in the case of degree 20 we can see a case of overfitting because it's also adjusting to the noise itself. We can test this hypothesis if we try the model on a different dataset

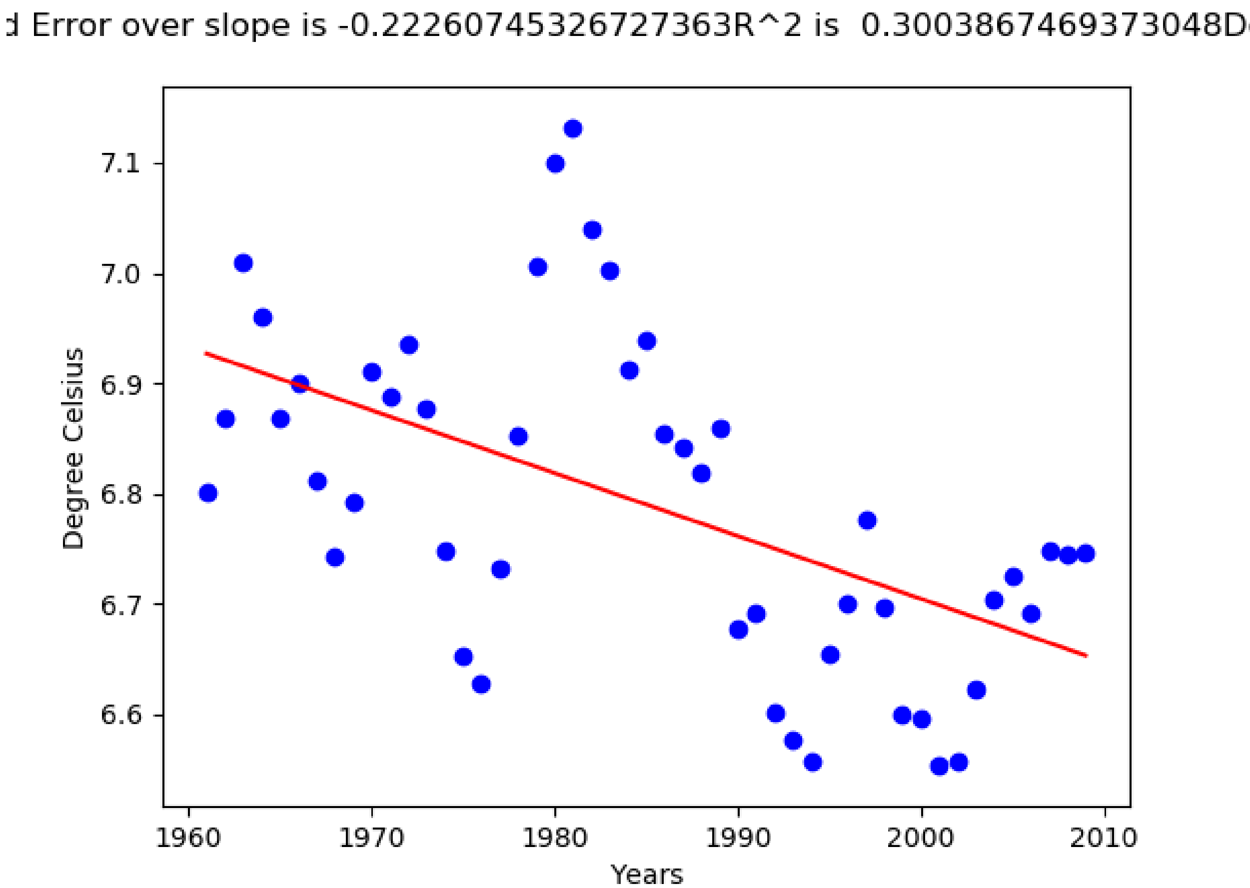

RMSE is 0.08844425310353894Degree is 1
17.10
17.05
17.00
16.95
16.90
16.85
16.80
16.75
Degree Celsius
2010
2011
2012
2013
2014
2015
Years


RMSE is 0.21177518245357127Degree is 2
Degree Celsius
Years
17.10
17.05
17.00
16.95
16.90
16.85
16.80
16.75
2010
2011
2012
2013
2014
2015



Degree 1 performed the best here, the RMSE for degree 1 is 0.08 for degree 2 it is 0.21 and for degree 20 it's 1.49. This is a big contrast to whatever we saw before in which degree 20 performed the best. This is because degree 20 was fitting to the noise in the previous data set. However, when it got introduced to a new dataset, the noise is different, therefore it is not a good predictor. Degree 1 on the other hand captured the general trend better

If we used the New York model, the result would be worse, this is because that model is only trained on data from New York. It will not be a good representative of how the 5 year window average would be for the 21 cities where the temperatures can have a vast difference from New York. Therefore, the RMSE will increase



The graph doesn’t match with our claim, in fact it contradicts it and shows that the temperature variation is not infact getting larger with time as it is a linearly decreasing model. I feel like a better way to take standard deviation is to take it city wise instead of all 21 cities or group them geographically as it would produce more consistent training data.