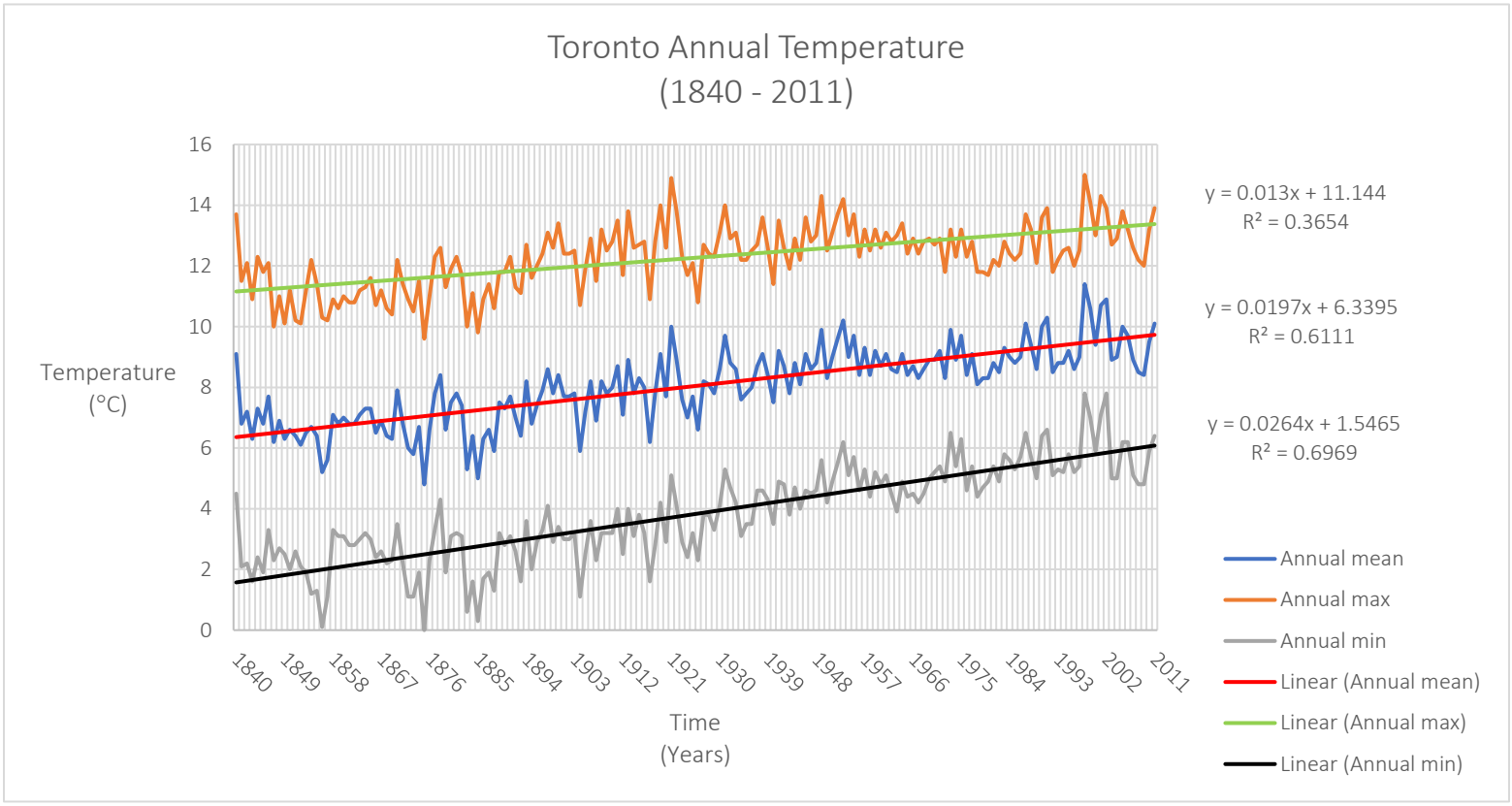
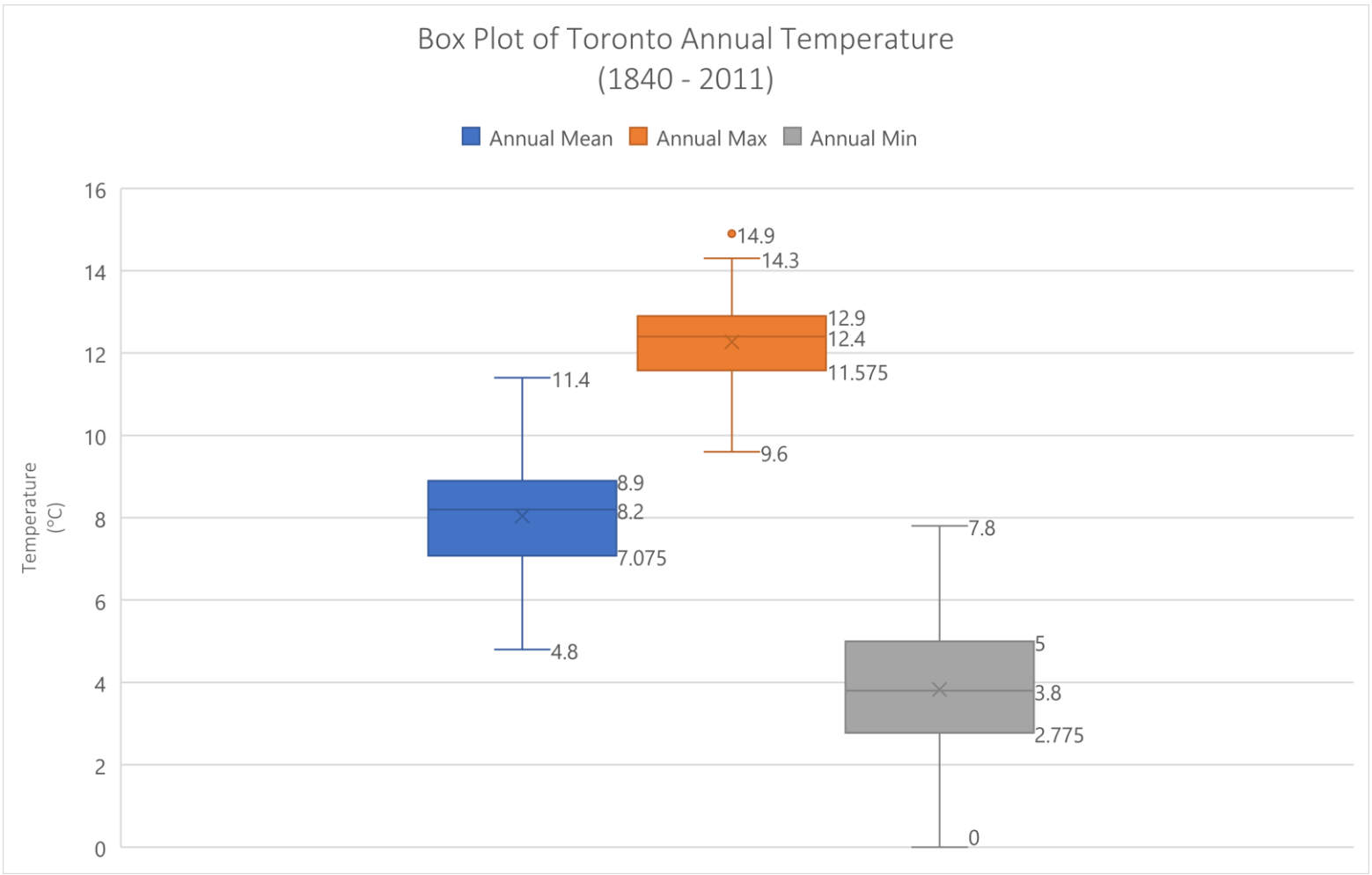
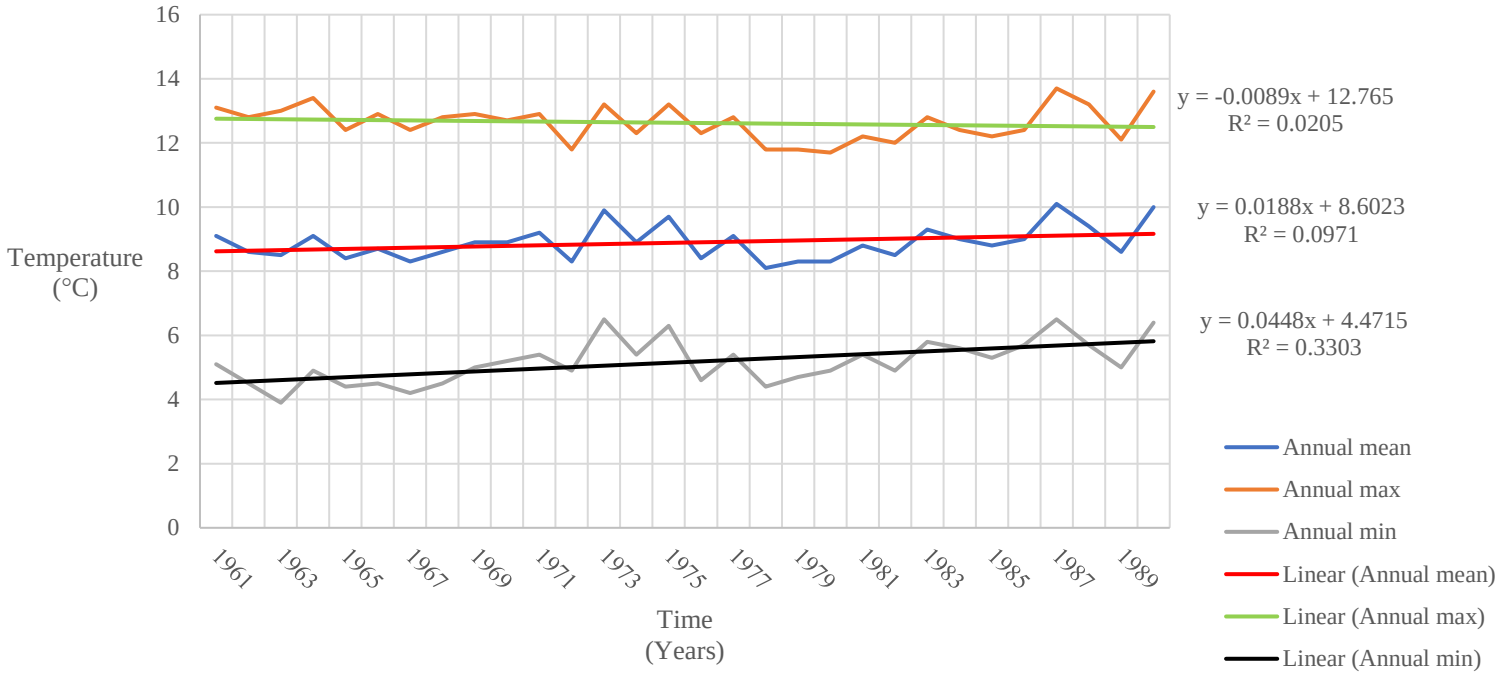
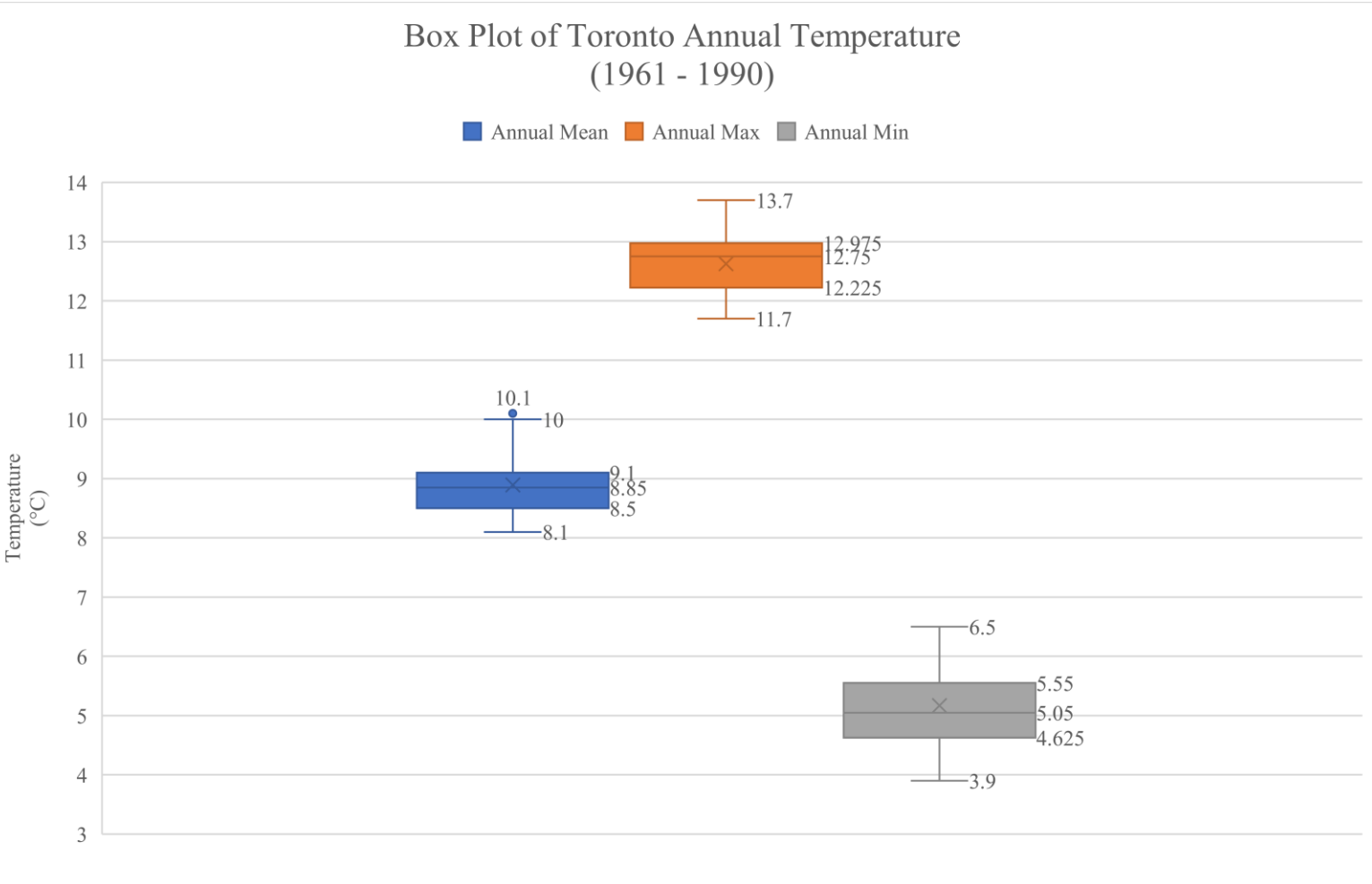
Table of Contents Pg #:

Matthew Kwok
1003284279

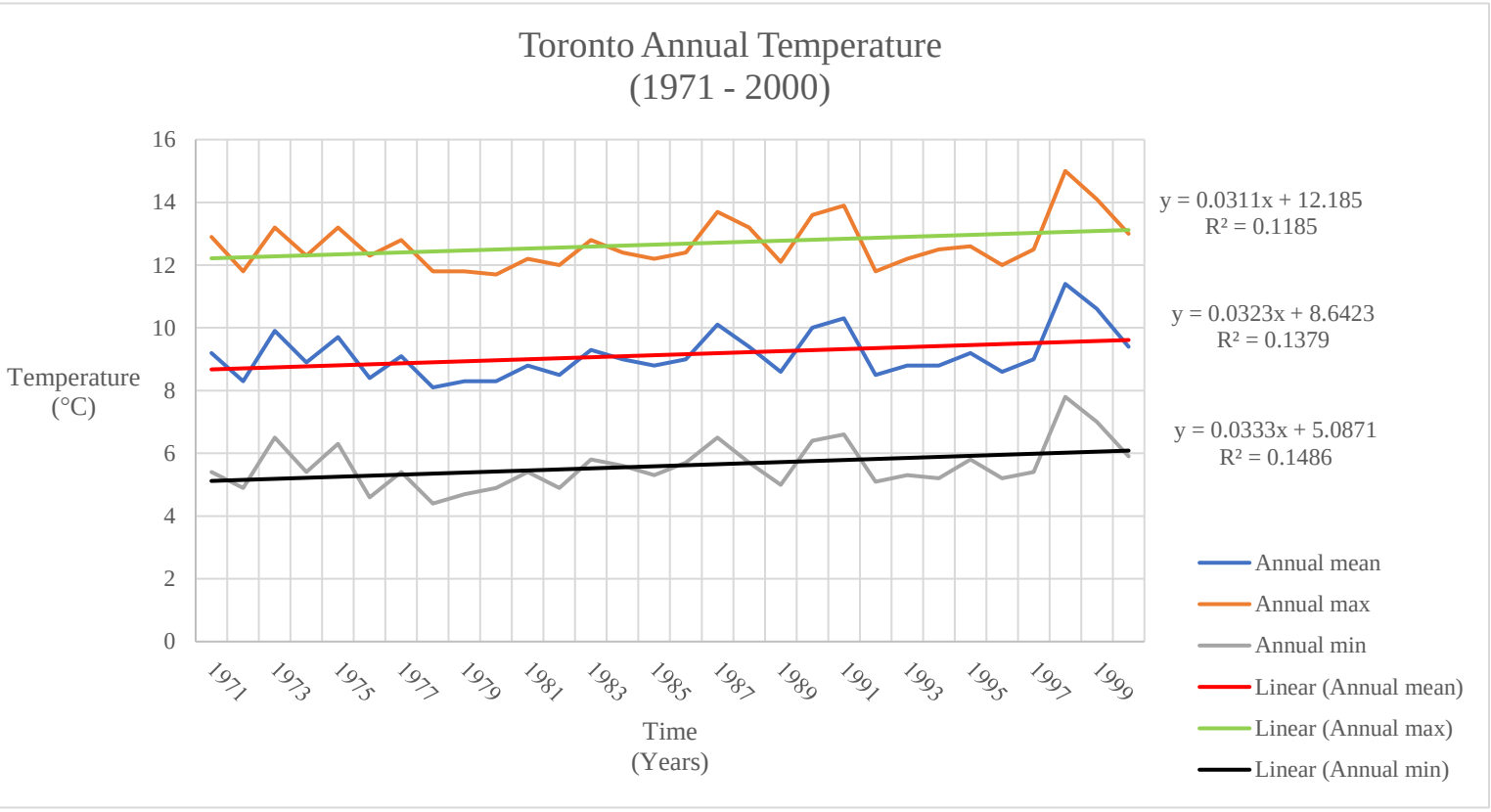


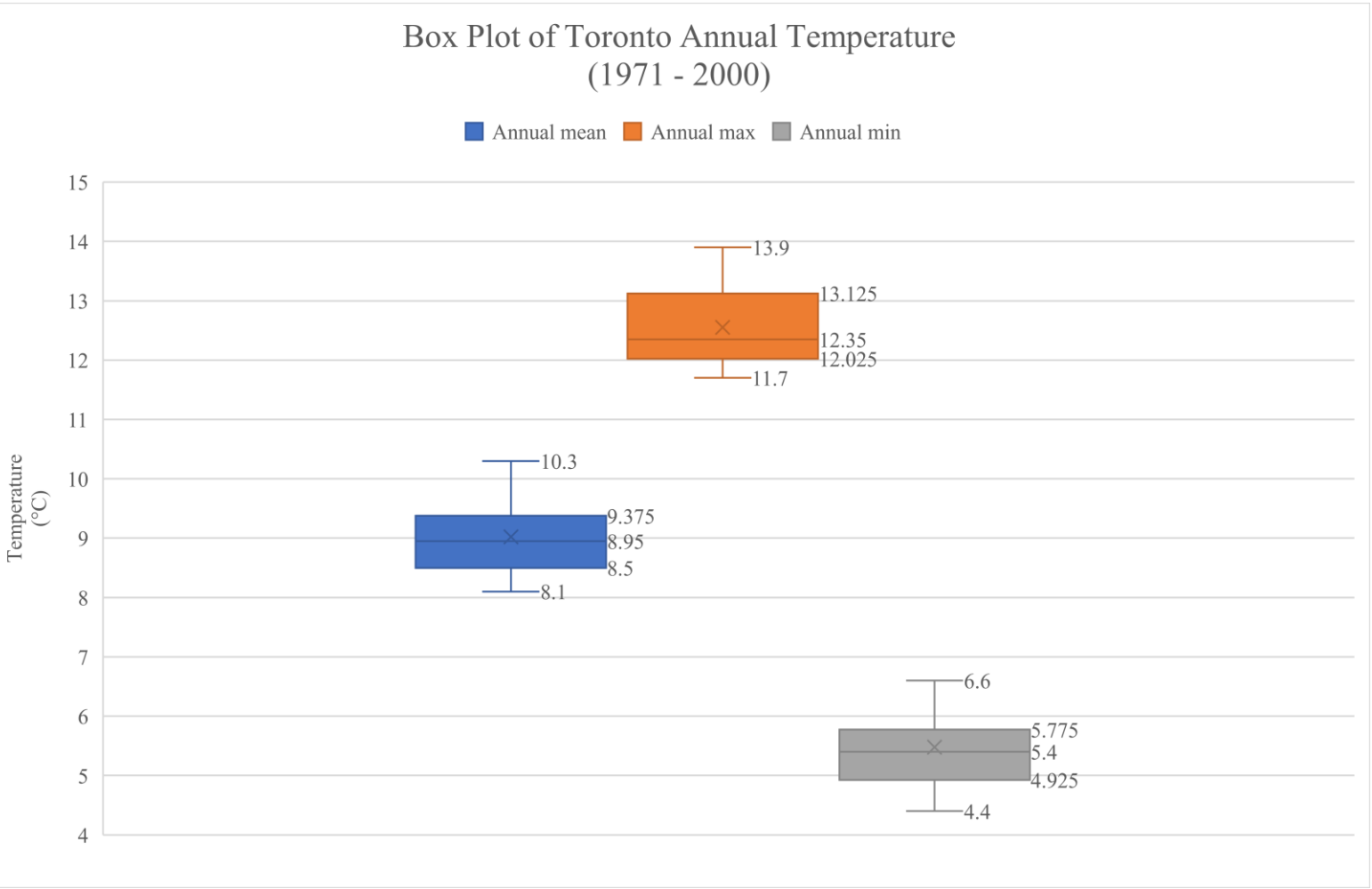


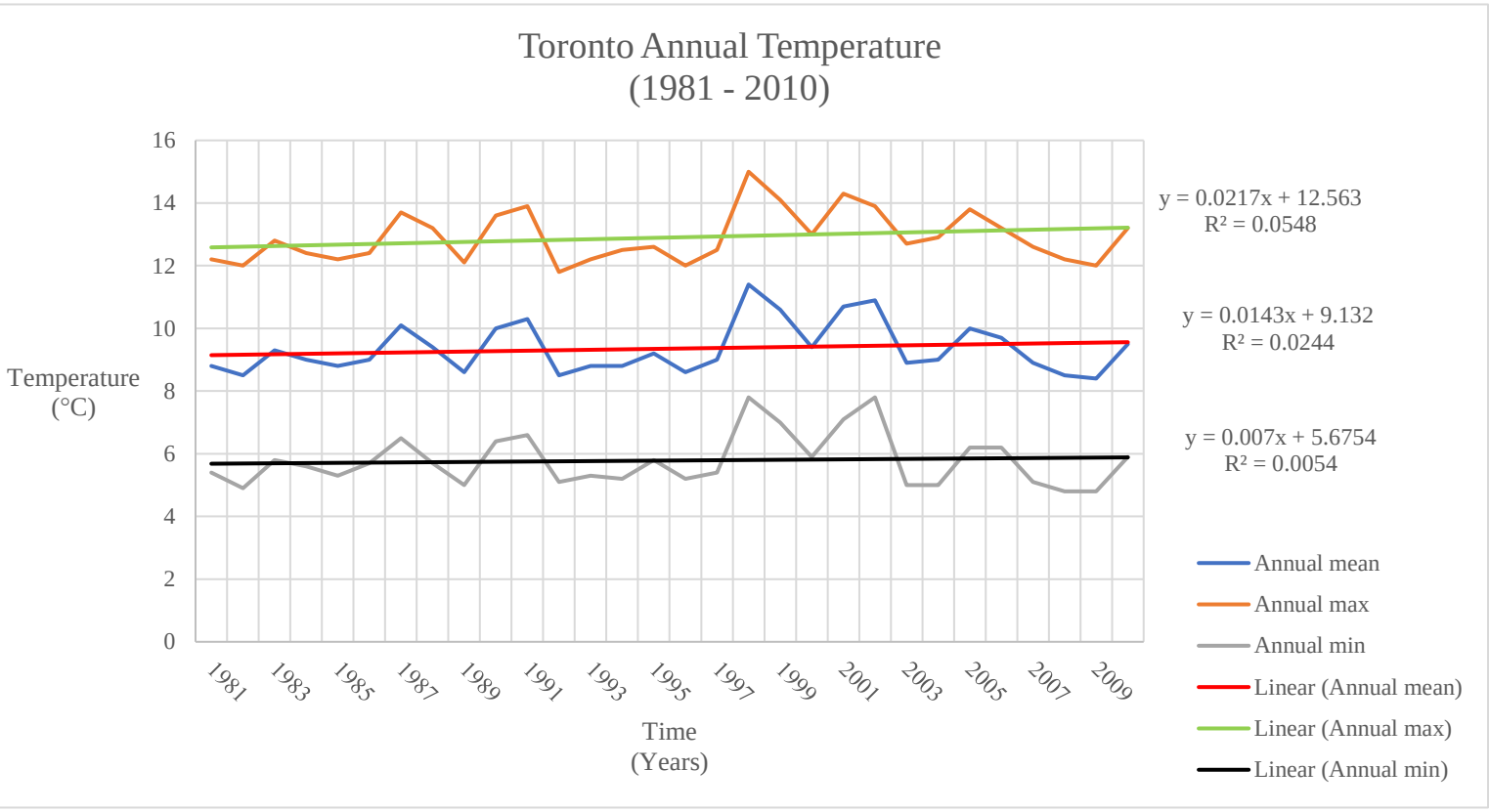

Toronto Annual Temperature
(1961 - 1990)
Temperature (°C)
Time (Years)
y = -0.0089x + 12.765
R² = 0.0205
y = 0.0188x + 8.6023
R² = 0.0971
y = 0.0448x + 4.4715
R² = 0.3303
Annual mean
Annual max
Annual min
Linear (Annual mean)
Linear (Annual max)
Linear (Annual min)

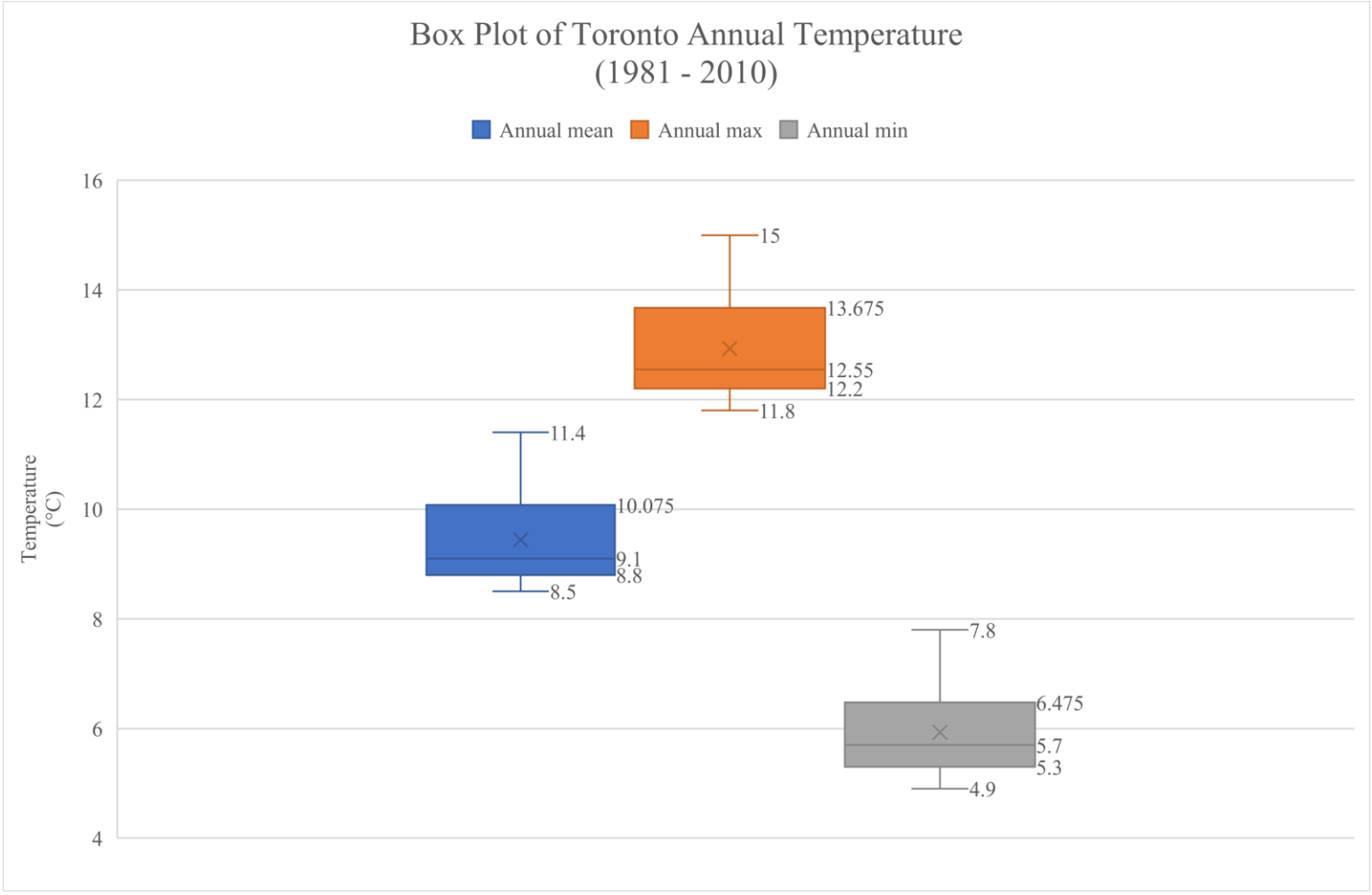

Box Plot of Toronto Annual Temperature
(1961 - 1990)
Annual Mean
Annual Max
Annual Min
Temperature (°C)
13.7
12.975
12.75
12.225
11.7
10.1
10
9.1
8.85
8.5
8.1
6.5
5.55
5.05
4.625
3.9

Matthew Kwok
1003284279






Toronto Annual Temperature
(1981 - 2010)
Temperature (°C)
Time (Years)
y = 0.0217x + 12.563
R² = 0.0548
y = 0.0143x + 9.132
R² = 0.0244
y = 0.007x + 5.6754
R² = 0.0054
Annual mean
Annual max
Annual min
Linear (Annual mean)
Linear (Annual max)
Linear (Annual min)


Box Plot of Toronto Annual Temperature
(1981 - 2010)
Annual mean
Annual max
Annual min
Temperature (°C)
15
13.675
12.55
12.2
11.8
11.4
10.075
9.1
8.8
8.5
7.8
6.475
5.7
5.3
4.9

Matthew Kwok
1003284279

For 1840 - 2011

Regression Analysis of Annual Mean 1840 – 2011:

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regressio	1	164.5747	164.5747	267.1457	1.06255E-36
Residual	170	104.7282	0.616048		
Total	171	269.303			

	P-value
Intercept	6.50209E-27
X Variable 1	1.06255E-36

Regression Analysis of Annual Max 1840 – 2011

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regressio	1	71.41961474	71.41961474	97.89702184	1.62661E-18
Residual	170	124.0214899	0.729538176		
Total	171	195.4411047			

	P-value
Intercept	1.20598E-06
X Variable 1	1.62661E-18

Regression Analysis of Annual Min 1840 – 2011

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regressio	1	294.6561325	294.6561325	390.91004	6.24639E-46
Residual	170	128.1408442	0.753769672		
Total	171	422.7969767			

	P-value
Intercept	5.72947E-42
X Variable 1	6.24639E-46

For 1961 - 1990

Regression Analysis of Annual Mean 1961 – 1990

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regressio	1	0.792365	0.792365	3.011853	0.093652971
Residual	28	7.366302	0.263082		
Total	29	8.158667			

	P-value
Intercept	0.197735724
X Variable 1	0.093652971

Regression Analysis of Annual Max 1961 – 1990

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.179759733	0.179759733	0.586703244	0.45010617
Residual	28	8.578906934	0.306389533		
Total	29	8.758666667			

	P-value
Intercept	0.199716982
X Variable 1	0.45010617

Regression Analysis of Annual Min 1961 – 1990

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regressio	1	4.520863181	4.520863181	13.81048255	0.000894356
Residual	28	9.165803485	0.327350124		
Total	29	13.68666667			

	P-value
Intercept	0.00157851
X Variable 1	0.000894356

For 1971 - 2000

Regression Analysis of Annual Mean 1971 – 2000

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	2.348397108	2.348397108	4.477624245	0.043364691
Residual	28	14.68526956	0.524473913		
Total	29	17.03366667			

	P-value
Intercept	0.080318785
X Variable 1	0.043364691

Regression Analysis of Annual Max 1971 – 2000

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	2.16776	2.167759733	3.765594824	0.062447286
Residual	28	16.11891	0.575675248		
Total	29	18.28667			

	P-value
Intercept	0.134328894
X Variable 1	0.062447286

Regression Analysis of Annual Min 1971 – 2000

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	2.492779755	2.49278	4.888869232	0.03536681
Residual	28	14.27688691	0.509889		
Total	29	16.76966667			

	P-value
Intercept	0.052639744
X Variable 1	0.03536681

For 1981 - 2010

Regression Analysis of Annual Mean 1981 – 2010

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.45846941	0.45846941	0.70086292	0.409582793
Residual	28	18.31619726	0.654149902		
Total	29	18.77466667			

	P-value
Intercept	0.578300139
X Variable 1	0.409582793

Regression Analysis of Annual Max 1981 – 2010

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	1.059595106	1.059595106	1.624753839	0.212907115
Residual	28	18.26040489	0.652157318		
Total	29	19.32			

	P-value
Intercept	0.378337919
X Variable 1	0.212907115

Regression Analysis of Annual Min 1981 – 2010

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.108975528	0.108975528	0.151409793	0.700135316
Residual	28	20.15269114	0.719738969		
Total	29	20.26166667			

	P-value
Intercept	0.82195051
X Variable 1	0.700135316

Question 1:

For each period 1840-2011, 1961-1990, 1971-2000, and 1981-2010 their associated annual temperature distributions, box plots, and regression analyses are displayed. In addition, Pearson R-values, mean values, and standard deviations are provided at the bottom of the page.

According to the box plots, majority of the temperature distributions for annual Tmean, Tmax and Tmin have no significant outliers. However, one outlier exists in the box plot for Tmax during 1840-2011 where there was an annual max temperature of 14.9 °C; this outlier represents an abnormally hot day. Another outlier exists in the 1961-1990 data set where there was a Tmean of 10.1 °C.

In terms of the Pearson R-value, it indicates the strength of a linear relationship. For periods 1961-1990, 1971-2000, and 1981-2010 the R-value is close to zero which indicates there is a very poor linear relationship between annual temperatures and time. The Pearson R-values for 1840-2011 could be argued to be close to R = 1 which would indicate a good positive linear relationship between annual temperature and time. However, noticing that the three R^2 - values for period 1840 – 2011 are <70%, it does not further support a linear relationship conclusion.

In terms of the statistical significance of the data, we look at Significance F values and P-values. If we are assuming a significance level of $\alpha = 0.05$, then a Significance F-value > 0.05 would imply that the data is statistically insignificant and if F-value < 0.05 then the data is statistically significant. Just like the F-values, the statistically significant data are p-values that are less than the significance value ($\alpha = 0.05$). At first glance, the data for 1840 – 2011 have the smallest F-values thus making it the most statistically significant. The least significant data set would be for 1981 – 2010, with respect to the F-values. For p-values, the 1840 – 2011 time period still has the smallest. Thus, 1840 – 2010 data is the most statistically significant data in terms of both the significance f-value and p-value both being less than 0.05. In contrast, the least significant data set in terms of the p-value and the significance f-value would be the time period from 1981-2010.

Overall, there is a general increasing trend in average yearly temperature, maximum yearly temperature and minimum yearly temperature, as seen in 1840-2011 graph of yearly temperature. Since it was concluded that the data set for 1840-2011 was statistically significant, this increase shows some concern for the climate in Toronto.

Question 2

Data for 1840 – 2011:

	Pearson R Value	Mean	Std. Deviation
Annual Mean	0.781737599	8.043605	1.2549387
Annual Max	0.604506253	12.26686	1.069079247
Annual Min	0.834818002	3.826744	1.572417546

Data for 1961 - 1990

	Pearson R-value	mean	Std. deviation
Annual Mean	0.311639864	8.893333333	0.530408647
Annual Max	-0.143260746	12.62666667	0.549566182
Annual Min	0.574727313	5.166666667	0.6869891

Data for 1971 - 2000

	Pearson R-value	Mean	Std. Deviation
Annual Mean	0.371305768	9.143333333	0.766399254
Annual Max	0.344301011	12.66666667	0.794087346
Annual Min	0.385549141	5.603333333	0.760436958

Data for 1981 - 2010

	Pearson R-value	mean	Std. Deviation
Annual Mean	0.156267644	9.353333333	0.804613136
Annual Max	0.234188956	12.9	0.816214982
Annual Min	0.073337636	5.783333333	0.835869703

The mean values found on above are also the 30-year average temperatures for their respective time periods. For time periods 1971 – 2000, 1961 – 1990, and 1981 – 2010 all have similar 30-year average

mean values. The outlier are the man values for 1840 – 2011. One explanation for why 1840 – 2011 mean values are different requires understanding of how mean values are calculated. Considering that 1840 – 2011 is the largest time scale it includes a wide variety of data points including very large and very small values. These large and small values are used in calculated the mean and thus would skew the mean values. In other words, mean/averages are sensitive to large and small values such that it would show a true mean. A more representative average would be finding the median values which ignores large and small values and focuses on the most middle data point.

If baseline options were limited to 1961 – 2000 and 1981 – 2010, then based on all the statistical information I would pick 1961 – 2000 baseline values. Firstly, the temperature graphs for 1961 – 2000 are less steep compared to 1981 – 2010; showing that there is less temperature change in the 1961 – 2000 data set which follows the definition of a good baseline. Also, the mean temperature for 1961 – 2000 is the lowest which also fits under the definition of a good baseline which is to have a reference point where the climate did not change significantly. In addition, the significance F-value and P-values for 1961 – 2000 are much closer to being less than the significance level of 0.05 which makes the data set more statistically significant.

Question 3

Merged – Toronto City 1981 – 2010 Average Temperature: 9.5 °C

CANGRD 1981 – 2010 Average Temperature: 8.7 °C

1981 – 2010 Tmean = 9.3533

Briefly, all three values not the same. The Merged-Toronto Tmean is about the same as the 1981-2010 mean. However, the CANGRD Tmean is much more different than the other two values. One of the explanations for this is that CANGRD data is provided by the National Resources Canada which covers the entire Canadian landmass. Since CANGRD data considers other areas, its mean value would not be representative of Toronto's mean temperature. One of the reasons why the 1981-2010 Tmean is smaller than the Merged-Toronto data is the increasing urbanization in the GTA (Mohsin & Gough, 2009). The increasing urbanization causes the winter annual minimum temperature to increase (Mohsin & Gough, 2009). It is also further elaborated in the paper, written by Mohsin and Gough, that there are substantial temperature changes during 1989-2000 (2009). It should be noted that this increasing urbanization is occurring in the suburbs of Toronto since southern Toronto was not reported to have increase in Tmean due to the cooling affects of Lake Ontario (Mohsin & Gough, 2009). To be specific, the continued urbanization of the suburbs has contributed a minimum of 30% to the total increase in Tmean in Toronto (Mohsin and Gough, 2009).

Mohsin, T. Gough, A, W. (2009, October 2). Trend Analysis of long-term temperature time series in the Greater Toronto Area (GTA). *Theoretical and Applied Climatology, 101*(3-4), 311-327. doi: 10.1007/s00704-009-0214-x