Heat Islands: The Association Between Distance from City Center and Land Surface Temperature in July 2020

Mili Sloan

Research Statistics 2

Mrs. Rachtchinine

May 30, 2025

Rationale

In a period of urbanization, with over 50% of the population living in cities and a changing climate, understanding the impact of urbanization on the local climate is imperative. The construction of urban areas introduces materials to a region that are different from the native plant life or ground covering, altering the interaction between the atmosphere and the land and changing the area's climate. The phenomenon of observing increased temperatures in urban areas produces an area known as an urban heat island (UHI). The creation of UHIs has implications far beyond an increase in temperature; UHIs contribute to a warming global climate, the severity of heat waves, increased energy demand, and unpredictable climate patterns, including but not limited to severe thunderstorms. The potential impacts of UHIs and the recent and continued growth of cities result in understanding the severity and variation of UHIs being essential to mitigating their impacts and reducing the overall temperature of urban areas and their contribution to global climate change.

Recent studies of UHIs focus on seasonal and day/night variations and the percentage of the ground covered in vegetation and impervious surfaces or surfaces that rainwater cannot penetrate. Less investigation has been done to address the severity of temperature changes across the city center, the city outskirts, the suburbs, and the rural areas surrounding an urban center. The two most populous U.S. cities are New York City (8,478,072 people) and Los Angeles (3,878,704 people). This paper aims to investigate the association between distance from the center of an urban area as the explanatory variable and temperature in the two most populous U.S. cities as the response variable. Based on previous studies of UHIs, a negative association between distance from city center and temperature is expected, but an urban cooling phenomenon has been observed and could occur.

Literature Review

Beginning with a case study of a UHI, Farhan Ul Moazzam et al. investigate a UHI on Jeju Island, Republic of Korea, in the years 2002, 2011, and 2021, focusing on the influence of different land types on the temperature in various regions of the island. Land surface temperature (LST) measurements, which are defined as the temperature of the air directly above the surface of the Earth, generated from measurements by the Landsat satellite series, serve as a proxy for the temperature across Jeju Island. LST is used since data from on-ground weather stations is unavailable. Images taken by the Landsat satellites also provide data on the land types across the island, including vegetated, impervious, and barren areas. The researchers investigated temperatures during March for the years 2002, 2011, and 2021 and found a clear increase in average temperature across the island from 2002 to 2021, which is attributed to climate change and the urbanization of the island increasing the percentage of the land occupied by impervious surfaces. To confirm an association between increased LST and urbanization, Farhan Ul Moazzam et al. map the average temperature in March of 2002, 2011, and 2021 to the locations where temperatures are observed and map the primary land type to each area on Jeju Island. By mapping the temperatures to their measurement locations, the association between increasing urbanization and increasing LST becomes clear, as land that switched from vegetated or barren to urbanized experienced a larger increase in LST than land that did not change its classification. Between 2002 and 2021, the percentage of land area covered by impervious surfaces increased due to the urbanization of Jeju Island, while the percentage of barren land decreased and the percentage of vegetation decreased slightly but remained nearly constant. The researchers expect that the increase in the UHI phenomenon on Jeju Island between 2002 and 2021 is less pronounced than elsewhere due to the conversion of barren land to impervious land rather than vegetated land to impervious land. Impervious and barren land absorb similar amounts of solar energy as heat, while impervious and vegetated land absorbs significantly different amounts of solar energy as heat, explaining the previous finding.

A review of the literature on UHIs from Deilami et al. provides greater insight into the causes of UHIs. The research attributes three main causes to the greater temperature of impervious land or urban areas compared to vegetated areas, including the lower albedo of the urban construction materials, the decreased sky exposure caused by high-rise buildings, and the decreased latent heat of urban building materials. Albedo describes the proportion of light a material or surface reflects, with a higher albedo corresponding to a greater amount of light reflected. Compared to vegetation, urban construction materials often have a low albedo, meaning less light is reflected and more is absorbed as heat energy, contributing to a temperature increase in urban areas. The lack of exposure of all areas of an urban region to the sky decreases the amount of absorbed heat dissipated at night, increasing the temperature of the urban area overall. Latent heat refers to the amount of energy that two materials can transfer to each other without changing their temperatures. The decreased latent heat of urban building materials compared to soil or vegetation results in less heat absorption before the material changes temperature. The inability of urban construction materials to store water also decreases the latent heat of an urban area due to the high latent heat of water compared to the latent heat of urban building materials. The dependence of UHIs on the amount of solar radiation absorbed as heat results in a significant change in the behavior of UHIs between daytime and nighttime. As albedo, sky exposure, and latent heat are each related to the amount of urban building materials present in an area, the percent of an area that is urbanized or constructed with urban materials is often a predictor for the severity of UHIs.

According to Deilami et al., current research on UHIs focuses on satellite technologies, UHI modeling, and UHI mitigation policies rather than on the variation of the temperature in an area with spatial and temporal variables. Out of the literature that focuses on the formation of UHIs, the most common areas of focus include vegetation cover of an area, impervious or built-up cover of an area, seasonal and diurnal variations, and population density. Out of each of the previous potential causes, the study identifies the percentage of an area that is covered in vegetation and the percentage of the area that is covered in impervious surfaces as the most common variables used to explain the temperature variations that contribute to UHIs, which suggests that there is a lack of literature investigating the scope of UHIs and their variations between city center, city outskirts, suburbs, and rural areas in a quantitative manner. As a result of the absence of abundant literature discussing the variation of UHIs with distance to the city center, this paper aims to identify if such an association exists and the strength of that association.

Methods

As this study focuses on the UHI phenomenon in cities in the U.S., it uses data from the Atmospheric Infrared Sounder (AIRS) instrument on the EOS-Aqua satellite, which images the continental U.S. twice daily. Due to the potential for the intensity of a UHI to vary across multiple years, as seen in Farhan Ul Moazzam et al.'s work, this paper selects data from the most recent available year, 2020. Data from July is used to further reduce the potential for the seasonal variability of UHIs discussed in Deilami et al.'s work to influence the results of the linear regression analysis. July is chosen over other months due to the increased intensity of UHIs in the summer season, discussed by Deilami et al., and the greater strain that UHIs place on an area in the summer when UHIs can result in a severe increase in energy demand. To reduce natural daily variations in temperature, data from a single day is used to investigate the UHI phenomenon. Due to the potential for cloud cover to reduce the accuracy of LST measurements, as seen in Farhan Ul Moazzam et al.'s research, historical weather data for New York City and Los Angeles identifies the day with the least precipitation in July 2020, July 11, 2020, and July 10 2020 respectively. Finally, the twice daily measurements of LST allow for the investigation of the UHI phenomenon in the daytime and nighttime; therefore, the data is further separated to create a data set for daytime and nighttime LST in each location.

The variables this paper will use include LST generated from images taken by the EOS-Aqua satellite during its pass over a certain area. Since the EOS-Aqua satellite does not constantly take images of the continental U.S., the images used to generate LST measurements are considered a sample, and the LST measurements are considered a sample by extent. The LST measurements are generated from measurements of the light emitted from the area being imaged, referred to as the radiance, and serve as a proxy for measuring UHIs since the intensity of a UHI is directly proportional to the temperature of an area. Latitude and longitude measurements will be used to determine distance from the city center. Each LST measurement can be pinpointed to a location 25 kilometers away from the location of the previous measurement, which is expected to be a high enough resolution to understand the variation of LST with distance to the city center. The city center is identified using the reported geographical center of the city, which is Bushwick, Brooklyn for New York City, and Beverly Hills, Los Angeles. The linear regression will be completed with the distance from the city center as the explanatory variable and the LST as the response variable. The measurements of LST used in the study will be limited to within 250 kilometers of the geographical center of New York City to include an adequate amount of urban, suburban, and rural land. The measurements are limited to within 100 kilometers of the Los Angeles city center due to the decreased land area of the Los Angeles area. The units of measure for distance are kilometers, and the units for LST are Kelvin. The sample data for each location are displayed below.

Table 1, Distance from the Los Angeles City Center and Daytime and Nighttime LST Measurements

Distance from Los Angeles City Center (km)	Daytime LST (K)	Nighttime LST (K)
95.53238469	307.5521851	293.1611938
44.77694639	307.1604919	297.4156494
59.65390149	307.6119995	292.6880188
78.58949912	308.1652222	292.7667542
99.28808404	308.7728271	292.8330383
11.15759468	307.374939	292.9727478
24.7498042	307.8042297	293.0758667
46.49314648	308.2753906	292.2791748
69.07721954	308.8140259	293.5315552
91.88409	309.3984985	293.5968628
71.87754715	308.1441956	293.8727417
49.7936651	308.4146423	293.9907837
29.15978391	308.7397461	292.4375305
16.70458948	309.1218872	292.5513
27.67062437	308.5632935	292.6530151
48.06402505	308.8904114	292.7420044
70.08919449	309.4113464	293.1065063
92.58877827	309.9682312	293.17099
82.77644484	308.6720581	293.4822388
64.59842694	308.9640198	293.5848389
50.48450028	309.2991943	292.0159912

Distance from Los Angeles City Center (km)	Daytime LST (K)	Nighttime LST (K)
95.53238469	307.5521851	293.1611938
44.77694639	307.1604919	297.4156494
44.48745502	309.6822815	292.1164856
49.64178547	309.1235962	295.5891724
63.27875999	309.59198	295.677124
81.23304838	310.0914612	298.09552
86.09306254	309.4154968	294.7813416
76.10918623	309.7629395	294.8674
72.28252175	309.1557617	295.0537
75.55450378	309.5757446	295.1394348
85.11031404	310.0301819	295.2247314
99.14938806	310.5062561	297.6425171

Table 2, Distance from the New York City Center and Daytime and Nighttime LST Measurements (Shortened to 30 Values)

Distance from New York City center (km)	Daytime LST (K)	Nighttime LST (K)
242.5046	302.6635	295.8461
241.9877	301.8575	295.2968
229.2723	302.0492	295.4867
235.8337	301.3173	294.9672
220.77	301.4896	295.1501
206.8056	301.6744	295.3398
174.0697	302.3016	295.944
234.6839	300.8329	294.6597
201.1616	302.3222	295.0186
171.6726	302.6902	297.6713
159.1551	302.8925	297.8719
148.6384	301.9351	295.809
140.5723	302.1605	296.0188
238.5896	300.4048	296.0518

219.7619	300.5374	296.2207
201.4106	300.683	296.3965
183.6786	302.0127	297.1691
150.9412	302.3668	297.5518
136.5913	302.562	297.7511
124.2255	301.5976	295.6867
114.4886	301.816	295.8946
108.0935	302.0455	296.106
247.2905	300.0333	295.786
227.2823	300.1458	295.9476
207.5069	300.2715	296.493
188.0381	300.4101	296.6685
168.982	301.7327	294.7954
150.4956	301.8965	296.7393
132.8173	302.0726	296.9322
116.3163	301.0894	297.1298

The daytime and nighttime satellite measurements for each city occur in the same locations and share the same values. The summary statistics of the distance from the Los Angeles and New York City centers are seen below. The distributions of distances from the Los Angeles and New York City centers are left-skewed, indicating that most measurements taken by the satellite are far from the city center. Due to the larger size of the New York City area considered, the summary statistics are larger for New York City than for Los Angeles. The interquartile range (IQR) for New York City is 95.818 km, which is larger than the 36.283 km IQR for Los Angeles. There are no outliers in the distance from the city center for New York City and Los Angeles by the 1.5 x IQR test.

Figure 1, Distribution of Distance from Los Angeles City Center (Top), Distribution of Distance from New York City Center (Bottom)

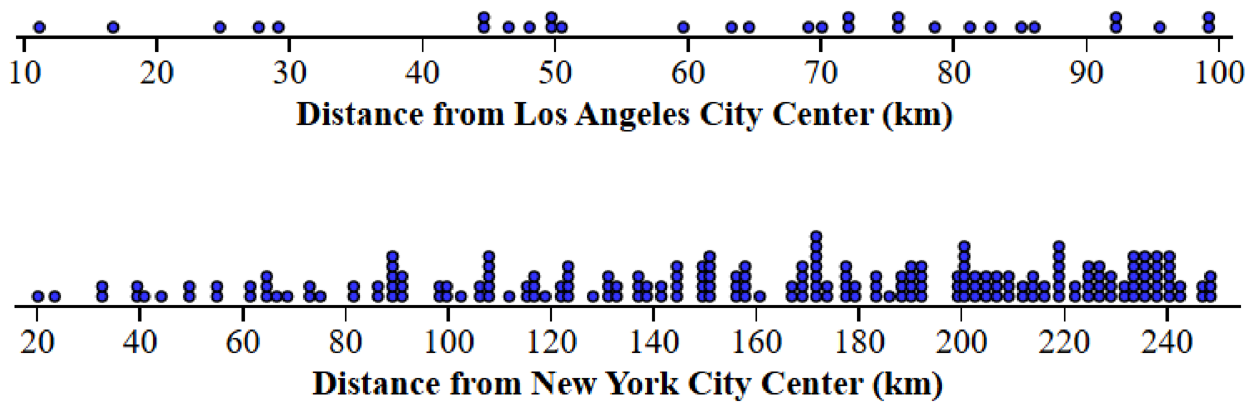


Table 2, Summary Statistics of Distance from Center of Los Angeles

mean	standard deviation	minimum	Q_1	median	Q_3	maximum
63.16	24.727	11.158	46.493	69.077	82.776	99.288

Table 3, Summary Statistics of Distance from Center of New York City

mean	standard deviation	minimum	Q_1	median	Q_3	maximum
163.184	59.785	20.164	118.416	172.597	214.234	249.754

The LST does vary between daytime and nighttime in each location. The summary statistics for the daytime and nighttime LST in Los Angeles and New York City are seen below. The distribution of daytime LST in Los Angeles is symmetric and has an IQR of 1.301 K. The nighttime LST in Los Angeles is more variable than daytime LST, with a roughly symmetric distribution and a larger IQR of 2.312 K. The daytime and nighttime LST in New York City follow a symmetric distribution, with the nighttime temperatures being more variable, with a larger IQR of 1.434 K, compared to the daytime IQR of 1.352. By the 1.5 x IQR test, there are no outliers in the distribution of daytime or nighttime LST in Los Angeles or New York City. In each location, the center of distribution of daytime LST is larger than the distribution of nighttime LST. Daytime Los Angeles is centered on a larger LST value than daytime New York City, and nighttime New York City is centered on a larger LST value than nighttime Los Angeles.

Figure 2, Distribution of LST for Daytime Los Angeles (Top Left), Distribution of LST for Nighttime Los Angeles (Bottom Left), Distribution of LST for Daytime New York City (Top Right), Distribution of LST for Nighttime New York City (Bottom Right)

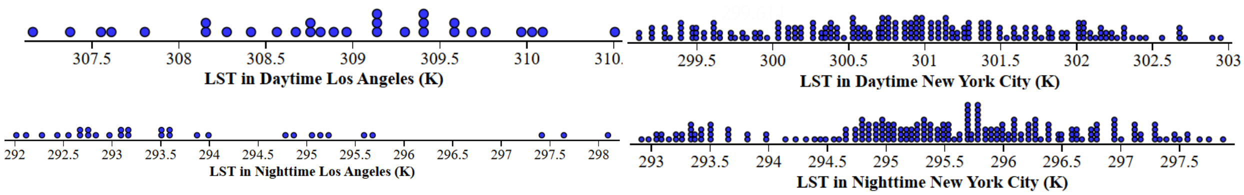


Table 5, Summary Statistics for Daytime LST in Los Angeles

mean	standard deviation	minimum	Q_1	median	Q_3	maximum
308.905	0.852	307.16	308.275	308.964	309.576	310.506

Table 6, Summary Statistics for Nighttime LST in Los Angeles

mean	standard deviation	minimum	Q_1	median	Q_3	maximum
293.939	1.634	292.016	292.742	293.482	295.054	298.096

Table 7, Summary Statistics for Daytime LST in New York City

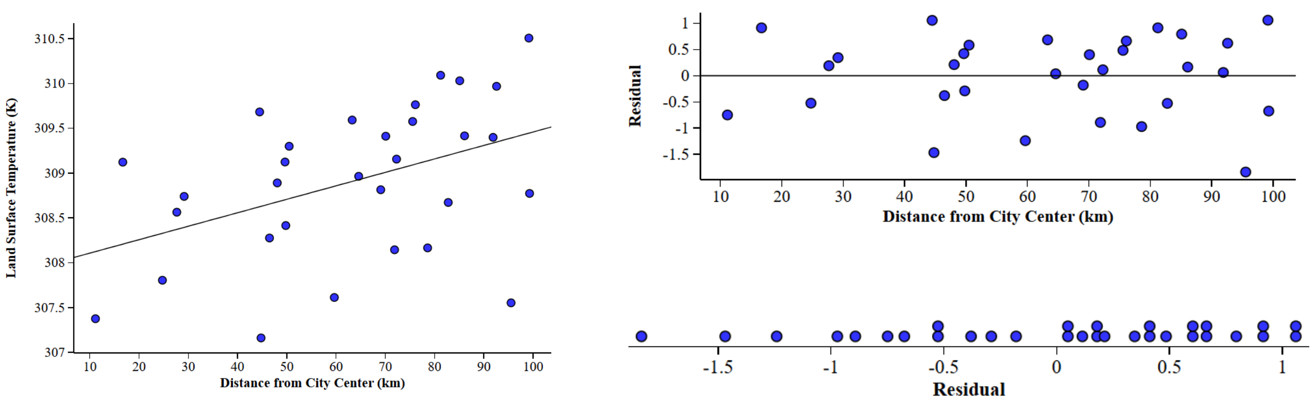
mean	standard deviation	minimum	Q_1	median	Q_3	maximum
300.893	0.918	299.12	300.233	300.909	301.585	302.947

Table 8, Summary Statistics for Nighttime LST in New York City

mean	standard deviation	minimum	Q_1	median	Q_3	maximum
295.48	1.186	292.916	294.865	295.541	296.299	297.872

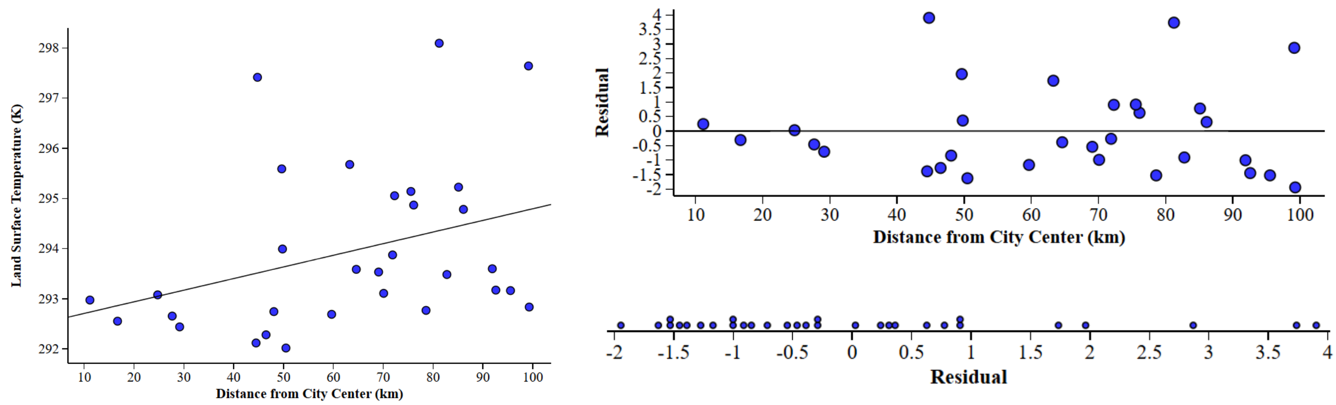
The independence condition is met for all data sets since each satellite measurement is distinct from the others, and one measured distance from the city center has no influence on the other measured distances from the city center. Due to the repeated appearance of the satellite in the same location at the same time twice daily to measure LST and the launch of the satellite at a time that can be considered random, the sample of LSTs used for the study can be considered a subset of a systematic random sample, satisfying the randomness condition.

Figure 3, LST with Respect to Distance from City Center in Daytime Los Angeles (Left), Residual Plot for LSRL of LST with Respect to Distance from City Center in Daytime Los Angeles (Top Right), Dot Plot of Residuals for LSRL of LST with Respect to Distance in City Center in Daytime Los Angeles (Bottom Right)



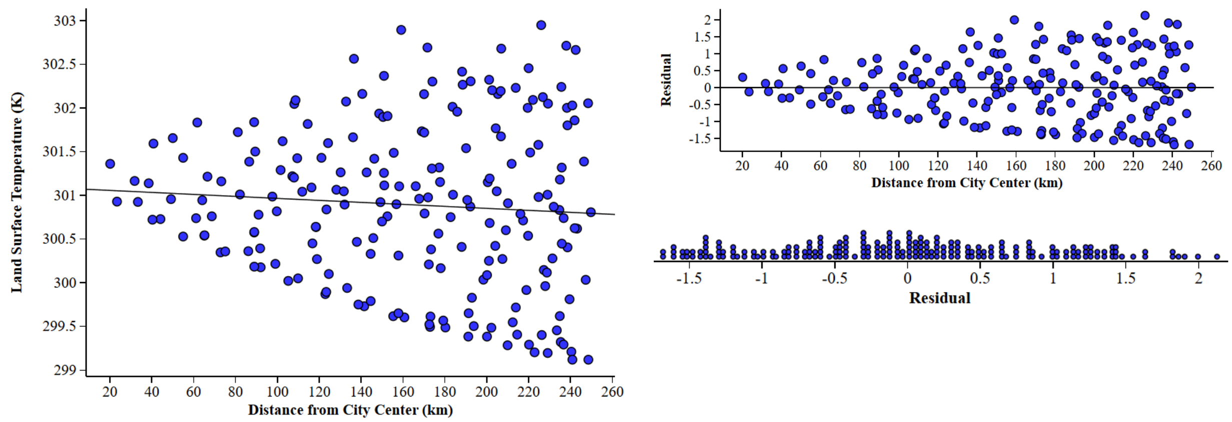
The plot of LST in daytime Los Angeles with respect to distance from the Los Angeles city center follows a linear pattern, and there is no clear pattern in the residuals produced by the least squares regression line (LSRL), indicating that a linear model is a good fit for the data. The dot plot of residuals is roughly symmetric, does not have a strong skew, has an IQR of 1.149, and, based on the 1.5 x IQR test, does not have any outliers, indicating that the distribution of residuals is approximately Normal. The plot of the residuals resulting from the LSRL for LST and distance from city center does not show a clear pattern and is a random scatter around zero, indicating an equal variance of the residuals with all explanatory variable values. The r value is 0.436, and the r^2 value is 0.19, indicating a positive association between daytime LST and distance from the Los Angeles city center and an LSRL that accounts for 19% of the variation of LST from its mean value.

Figure 4, LST with Respect to Distance from City Center in Nighttime Los Angeles (Left), Residual Plot for LSRL of LST with Respect to Distance from City Center in Nighttime Los Angeles (Top Right), Dot Plot of Residuals for LSRL of LST with Respect to Distance in City Center in Nighttime Los Angeles (Bottom Right)



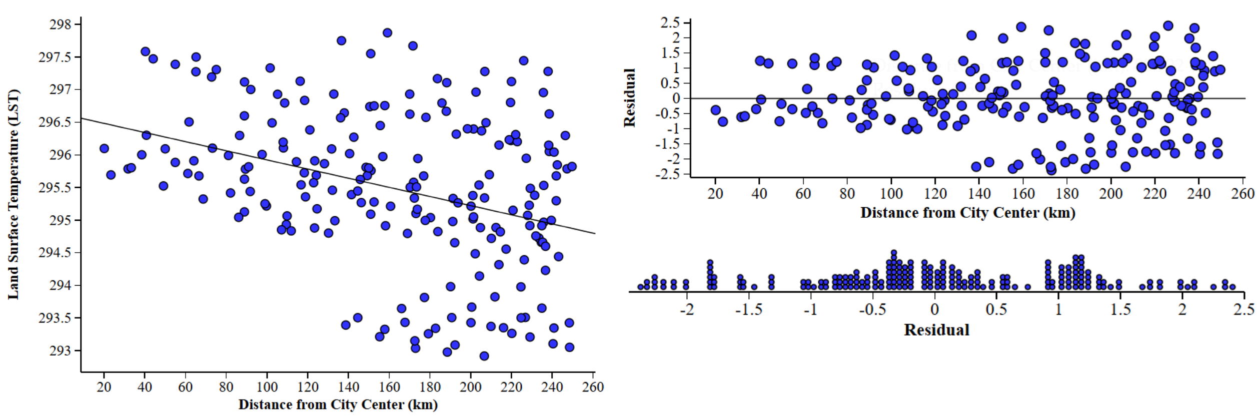
The dot plot of residuals for LSRL for LST with respect to distance from the Los Angeles city center during the nighttime is roughly symmetric and, based on the 1.5 x IQR test, has two outliers on the upper end. Due to the small sample size, the Normality conditions will still be considered satisfied. The plot of LST in nighttime Los Angeles with respect to distance from the Los Angeles city center follows a linear pattern. There is a slight pattern in the residuals produced by the LSRL for LST with respect to distance from the city center, with the residuals slightly increasing as distance from the center increases; it is slight enough to proceed with caution, and the linearity and equal variance conditions are met. The r value is 0.351 and the r^2 value is 0.123, indicating the linear regression between LST and the distance from the Los Angeles city center accounts for 12.3% of the variation in LST from its mean value.

Figure 5, LST with Respect to Distance from City Center in Daytime New York City (Left), Residual Plot for LSRL of LST with Respect to Distance from City Center in Daytime New York City (Top Right), Dot Plot of Residuals for LSRL of LST with Respect to Distance in City Center in Daytime New York City (Bottom Right)



The plot of LST in daytime New York City with respect to distance from the New York City center does not follow a linear pattern, regardless of the transformations applied. The residuals plot is not a random scatter around zero since the value of the residuals increases as distance from the city center increases, suggesting that a linear regression is an inappropriate model to account for the variation of LST with distance from city center and a more complex model is needed.

Figure 6, LST with Respect to Distance from City Center in Nighttime New York City (Left), Residual Plot for LSRL of LST with Respect to Distance from City Center in Nighttime New York City (Top Right), Dot Plot of Residuals for LSRL of LST with Respect to Distance in City Center in Nighttime New York City (Bottom Right)



The plot of LST in nighttime New York City with respect to distance from the New York City center follows a linear pattern and there is a slight pattern in the residuals produced by the LSRL for LST, with the residuals slightly increasing as distance from the center increases. The pattern is slight enough that the linear regression analysis will continue with caution. The dot plot of residuals is roughly symmetric and has no outliers, satisfying the Normality condition. The plot of the residuals resulting from the LSRL for LST with respect to distance from city center, slightly increases as distance from the city center increases, but the pattern is not distinct, and the plot is a random scatter around zero, indicating that there is equal variance of the residuals with all values of the explanatory variable. The r value from the linear regression is -0.353 and the r^2 value is 0.125, indicating that the LSRL accounts for 12.5% of the variation in the LST values from the mean LST in nighttime New York City on July 11 2020.

Results

A t-test for population slope will be carried out for daytime and nighttime Los Angeles and nighttime New York City. The population of interest is the true slope of the LSRL between LST and distance from the city center for daytime and nighttime Los Angeles on July 10, 2020 and nighttime New York City on July 11, 2020. The null hypothesis is that each population slope is zero; the alternative hypothesis is that each population slope is not equal to zero. The chosen significance level is $\alpha = 0.05$.

The LSRL between LST and distance from the Los Angeles city center in the daytime has a slope of 0.01502 K/km, indicating that for each increase in distance by 1 km, the LST is predicted to increase by 0.01502 K. The p-value for the sample slope is 0.014. At a confidence level of 95%, the plausible value of the population slope is from 0.0028 K/km to 0.02734 K/km. Since the p-value = 0.014 < 0.05 = α, the null hypothesis is rejected; there is significant evidence that there is an association between LST and distance.

The LSRL between LST and distance from the Los Angeles city center at nighttime has a slope of 0.0234 K/km, indicating that for each increase in distance by 1 km, the LST is predicted to increase by 0.0232 K. The p-value for the sample slope is 0.053. At a confidence level of 95%, the plausible value for the population slope is from 0.0008 K/km to 0.0456 K/km. Since the p-value = 0.053 > 0.05 = α, the null hypothesis cannot be rejected; there is not significant evidence that there is an association between LST and distance.

The LSRL between LST and distance from the New York City center in the nighttime has a slope of -0.007 K/km, meaning that for each increase in distance from the New York City center by 1 km, the LST is predicted to decrease by 0.007 K. The p-value for the sample slope is less than 0.001. At a confidence level of 95%, the plausible value for the population slope is from -0.0090 K/km to -0.0050 K/km. Since p-value < 0.001 < 0.05 = α, the null hypothesis is rejected; there is convincing evidence that there is an association between LST and distance.

Although the plausible slopes are not large for any data set, they result in a predicted temperature increase one-hundred kilometers outside of Los Angeles and two-hundred-fifty kilometers outside of New York City that is the same magnitude as the temperature differences found by Farhan Ul Moazzam et al. and Deilami et al., indicating that there is a practical significance to the results of the linear regression.

Conclusion

A surprising result of the study is the positive value of the slopes calculated for the samples taken from Los Angeles. The UHI phenomenon suggests that densely populated, urban areas, with a decreased percentage of land covered in vegetation and a greater percentage of land covered in impervious surfaces, will lead to an increase in the LST surrounding the center of the urban area. The opposite association is seen in the slopes calculated for Los Angeles, with a positive slope indicating that temperatures increase as distance to the center of Los Angeles increases on July 10, 2020. The increase in temperature as urbanization decreases is briefly mentioned in the paper by Deilami et al. as the urban cool island (UCI) phenomenon. UCIs often occur in arid areas due to the replacement of barren land with impervious surfaces rather than the replacement of vegetation with impervious surfaces. As discussed in Deilami et al.'s research, vegetation has a higher albedo than barren land and impervious surfaces, leading to the decrease in temperature with distance from a city center typical for an UHI. The replacement of barren land with impervious surfaces is seen to decrease the extremity of UHIs, which is seen in Farhan Ul Moazzam et al.'s work with Jeju Island, where barren land, rather than vegetation, is replaced by impervious surfaces characteristic of urban building materials, decreasing the severity of the increase of the UHI phenomenon. Los Angeles' location near the Pacific Ocean could contribute to the UCI phenomenon observed in the regression analysis, since, according to research by Rajeswari et al., the presence of surfaces that absorb a high degree of solar radiation as heat (either barren land or urban building materials) increases the intensity of a sea breeze. A sea breeze refers to cool wind moving from over water towards the land induced by a temperature difference between the water and land and a sea breeze might be the cause of the positive association between LST and distance from the Los Angeles city center.

Differences between the LSRL slope for daytime and nighttime Los Angeles indicate differences in the UHI phenomenon between daytime and nighttime. From reviewing Deilami et al.'s research, the UHI phenomenon is expected to be more intense at night. Although the data suggests Los Angeles is a UCI rather than a UHI, the slope in the daytime is b = 0.01502 K/km and the slope in the nighttime is b = 0.0234 K/km, which suggests that the true slope of the LSRL for LST with respect to distance from the Los Angeles city center on July 10 2020 is larger in the nighttime. While the greater slope is explainable for a UHI, it is less common for a UCI, since the intensity of a sea breeze decreases at night as the water and land become similar temperatures, which should decrease the slope of the LSRL for LST with respect to distance from the Los Angeles city center. The greater slope in nighttime Los Angeles is a surprising result of this study and does not have a current explanation in literature of UCIs.

Despite the larger sample slope for nighttime Los Angeles, the slope is not significant at a $\alpha = 0.05$ level, which suggests that the comparison between the nighttime and daytime slopes could be irrelevant. Further samples and potentially a more complex model should be used to determine the true difference between the intensity of the Los Angeles UCI between daytime and nighttime.

The negative association between distance from the New York City center and LST can be explained by similar reasoning as the positive association between distance from the Los Angeles city center and LST. Unlike Los Angeles, New York City and the surrounding area is not an arid environment and can be considered part of the temperate forest biome. Due to the climate differences between Los Angeles and New York City, the ground outside of the urbanized areas is vegetated in New York City rather than barren land like in Los Angeles. Vegetation and impervious surfaces used as urban building materials have different albedos, like vegetation reflecting more sunlight and having a higher albedo than impervious surfaces, explaining the decrease in temperature as distance from New York City increases since the percent of the ground covered in impervious surfaces decreases and the percent covered in vegetation increases, overall increasing the albedo of the ground and decreasing the percent of solar radiation that it absorbs, which in turn decreases LST. The difference between the direction of the association in Los Angeles and New York City demonstrates the influence that local climate can have on the UHI phenomenon.

Comparing the p-value for the significance tests performed for daytime New York City and Los Angeles, New York City has a p-value less than 0.001 and Los Angeles has a p-value of 0.014. The p-value indicates the probability of getting a sample size of $n = 206$ and $n = 31$ respectively with the slope calculated from the regression assuming that the true population slope is zero. The lesser p-value assigned to New York City suggests that the slope for the LSRL between LST and distance from the city center is more significant there than in Los Angeles, which is reflective of the unique climates of each of the locations. As discussed earlier, the UHI phenomenon depends on albedo and there is greater difference in the albedo of vegetation and urban building materials than barren land and urban building materials. The New York City area has a greater amount of ground covered in vegetation than Los Angeles, which accounts for the more significant slope value for New York City, suggesting that it is a more pronounced UHI than the UCI that occurs in Los Angeles.

The major takeaway from the linear regressions and slope significance tests for LST as explained by distance from the New York City and Los Angeles centers is the dependence of the micro-climates urban areas create on the pre-existing climate of an area. Los Angeles is an arid environment located close to the Pacific Ocean, which results in a UCI rather than a UHI and New York City is a temperate forest environment with vegetation, resulting in a greater significance of the slope for the LST as explained by distance to the New York City center.

Reflection

Through completing a study on UHIs using linear regression, I learned to better work with and interpret variables with satellite data, better understood the conditions needed for linear regression, and learned what information must be included in the study and what information can be omitted. Before completing the data analysis portion of the project, I never considered the independence of satellite measurements, which I work with often, and the type of sample that satellite measurements would be, but due to this project, I am more confident in using statistics analyses when working with satellite data. Through checking the conditions for each data set I used in the project, I also realized that each condition will not always be met perfectly, like the presence of outliers in the residuals dot plot for the nighttime Los Angeles data, but in certain situations, an analysis can still proceed. I learned to be more critical when checking conditions, focusing not on just checking a box but on considering the size of a data set and when I must consider a failed condition further before the study can proceed. I also realized what information is valuable in a final paper and that it is unnecessary to include all steps of the research process since some steps do not contribute meaningfully to the final result. For example, I learned to exclude the results of failed linear transformations. Overall, the project allowed me to consider the process of completing a linear regression in more depth and better understand how to structure a paper, which will become useful later as I use statistical analyses as a researcher.

Bibliography

AIRS Project Instrument Suite. (n.d.). Jet Propulsion Laboratory California Institute of Technology. Retrieved May 21, 2025, from https://airs.jpl.nasa.gov/mission/airs-project-instrument-suite/eos-aqua-satellite/

Deilami, K., Kamruzzaman, M., & Liu, Y. (2018). Urban heat island effect: A systematic review of spatio-temporal factors, data, methods, and mitigation measures. *International Journal of Applied Earth Observation and Geoinformation*, *67*, 30-42. https://doi.org/10.1016/j.jag.2017.12.009

Moazzam, M. F. U., Doh, Y. H., & Lee, B. G. (2022). Impact of urbanization on land surface temperature and surface urban heat island using optical remote sensing data: A case study of Jeju Island, Republic of Korea. *Building and Environment*, *222*, 109368. https://doi.org/10.1016/j.buildenv.2022.109368

Population Growth Reported Across Cities and Towns in All U.S. Regions. (2025, May 15). United States Census Bureau. Retrieved May 21, 2025, from https://www.census.gov/newsroom/press-releases/2025/vintage-2024-popest.html

Williams, E. (n.d.). *Latitude/Longitude Distance Calculator*. National Hurricane Center and Central Pacific Hurricane Center. Retrieved May 21, 2025, from https://www.nhc.noaa.gov/gccalc.shtml