



TOO HOT TO HANDLE?

AN ASSESSMENT OF THE RELATIONSHIP BETWEEN HEAT WAVES AND CRIMINAL ACTIVITY

C'ASIA GRIFFIN

STEPHANIE JANSSON

SARAH LOVING

MENTOR: DR. MARY KATHRYN COWLES

HEAT WAVES AND CRIMINAL ACTIVITY

- Heat waves: Two or more consecutive days during which the maximum temperature of each day was above the 95th percentile in a given region.
- Purpose: This project seeks to verify whether there is an association between heat waves and crime rates.
- Significance: The more knowledge that people have concerning the influence of heat waves on crime, mental health, etc., the better equipped they will be to combat what negative effects heat waves may entail.



METHODOLOGY



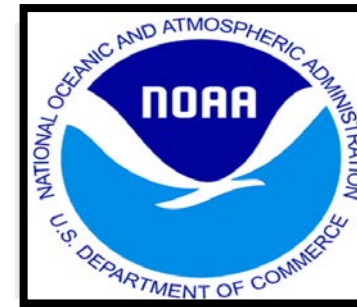
DATA SOURCES



Crime Data: National Incident-Based Reporting System (NIBRS)

As reported by local police departments from 2001-2012

- Iowa City (IA) Police Dept
- McKinney (TX) Police Dept
- Charleston (SC) County Police



Weather Data: National Oceanic and Atmospheric Administration (NOAA)

Weather data sourced from local weather stations from 2001-2012

- Iowa City (IA) Municipal Airport
- McKinney (TX) Municipal Airport
- Charleston (SC) International Airport

VARIABLES OF INTEREST

Independent Variable: daily maximum temperature used in two ways

- Quantitative variable
- Binary Heatwave variable
 - "0" = No Heat Wave Present
 - "1" = Heat Wave Present

Response Variable: the reported crime count for each day

- Discrete counts
- Three categories of crimes were examined separately:
 - Assault
 - Homicide
 - Theft/Larceny

CITIES OF INTEREST

Iowa City, Iowa

- Population: 75,798 citizens

McKinney, Texas

- Population: 181,330 citizens

Charleston, South
Carolina

- Population: 130,113 citizens

Goal: To assess the relationship between the number of crimes and temperature

Problem: Simple linear regression model requires that errors are normally distributed

- CANNOT assume this for our data

Problem: Response variable measured in counts

- Discrete rather than continuous
- Cannot take negative or non-integer values
- Variance increases as the mean increases

Solution: A special case of the Generalized Linear Model (GLM) -- Poisson regression

GLM POISSON REGRESSION



STATISTICAL CHALLENGES

TEMPORAL AUTOCORRELATION

DAYS CLOSER TOGETHER EXPERIENCE SIMILAR AMOUNTS OF CRIMES REGARDLESS OF WHAT HEATWAVES ARE DOING

VIOLATES INDEPENDENCE ASSUMPTION REQUIRED BY A LINEAR MODEL



HIERARCHICAL GENERALIZED LINEAR MODEL (HGLM)

Reasons for use

- Relaxes the assumption that error terms are independent
- Enables modeling of temporal autocorrelation

Unfortunate side effects

- HGLM generally results in a larger p-value than it ought to

'hglm' function from R package 'hglm' was used to conduct the hierarchical generalized linear regression analysis
R package 'spdep' was used to prepare our data for the HGLM analysis

THE MODEL

- **Hierarchical GLM Poisson Regressions**

$$\text{Log}(E(Y_i)) = \beta_0 + \beta_1 x_i + \phi_i$$

- **Interpretation:**

- $X = \text{TMAX}$: For each one-degree increase in maximum daily temperature (TMAX), we expect the crime count to be multiplied by e^{β_1}
- $X = \text{Heatwave}$: If heatwave is present, we expect the crime count to be multiplied by e^{β_1}
- The ϕ_i captures the random effects for each time point in order to account for spatial autocorrelation

ASSAULTS: METHODS AND ANALYSIS



IOWA CITY, IOWA

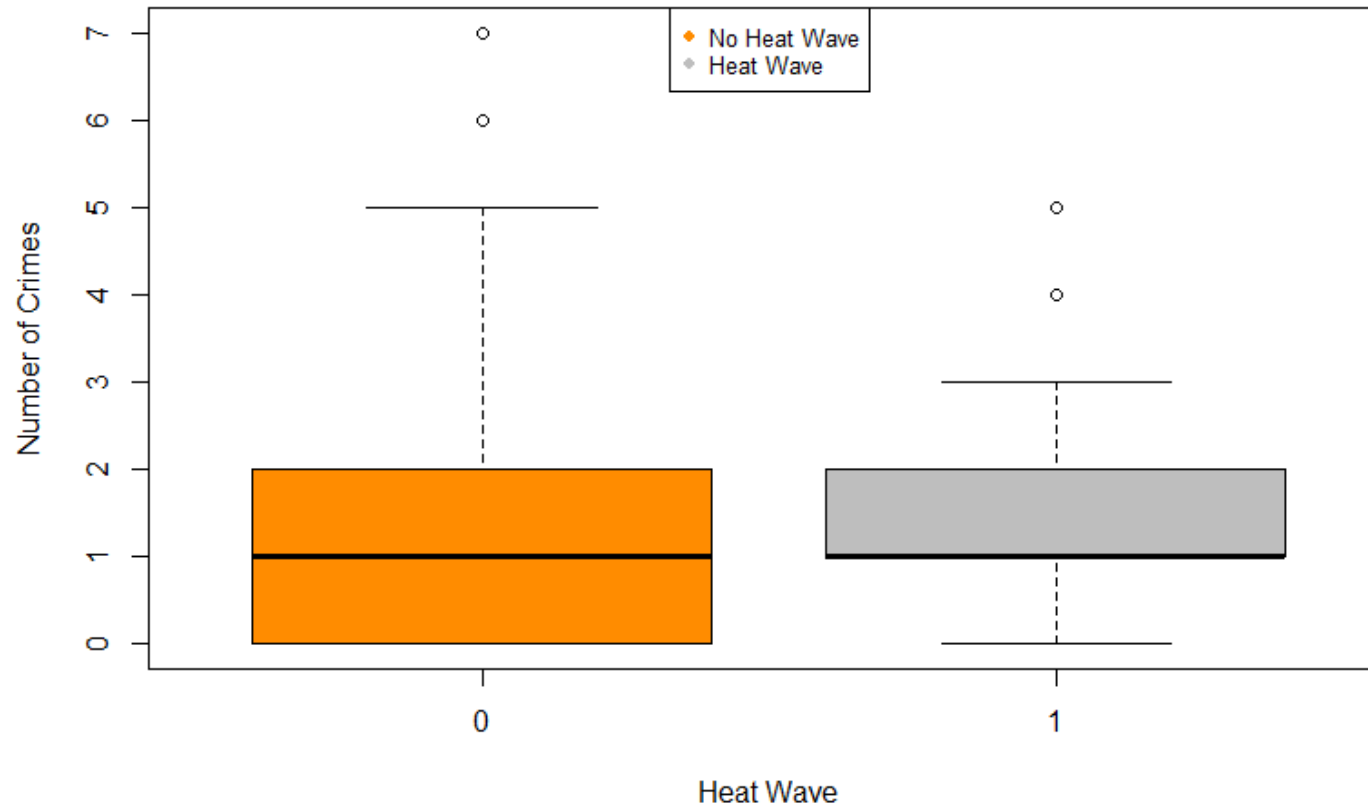
Durbin-Watson* test: serial autocorrelation

Use of HGLM to account for it

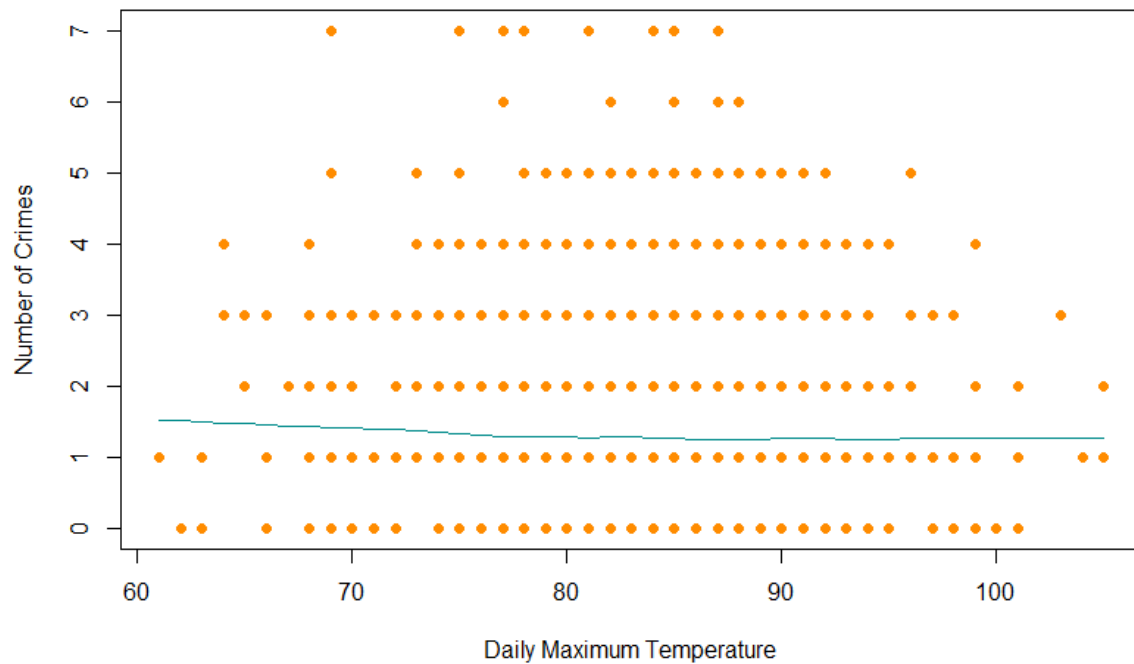
No evidence to support an association between presence of heatwaves and the number of assaults in Iowa City

* 'dwtest' function from R package 'lmtest' was used to conduct the Durbin-Watson test

IOWA CITY: ASSAULTS DURING HEAT WAVES



Iowa City, Iowa Assaults Results Table and Scatterplot for Continuous Temperature



	Estimate	Standard Error	P-Value
Heat Wave (Binary)	-0.03170	0.09374	0.735
TMAX (Temperature)	-0.003836	.004744	0.419

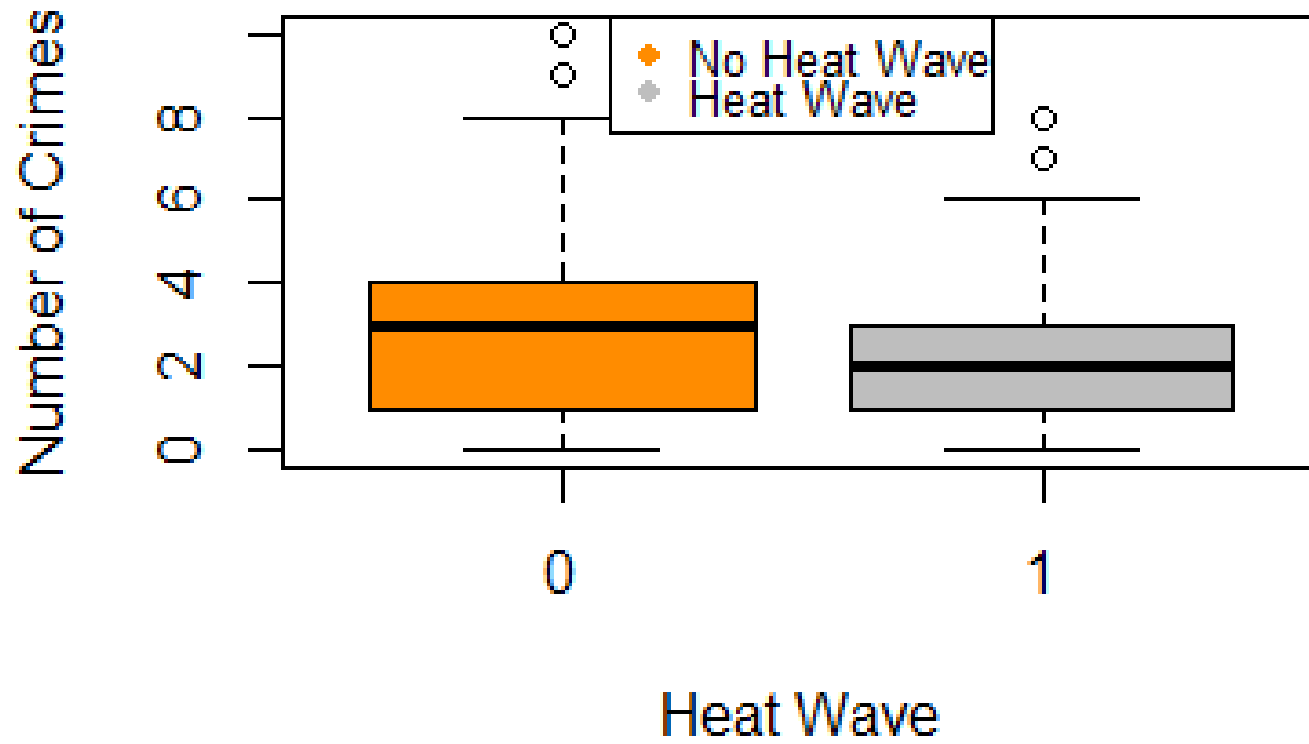
MCKINNEY, TEXAS

The Durbin-Watson* test showed that serial autocorrelation was high in the Assaults category. Thus, hierarchical general linear regression was used

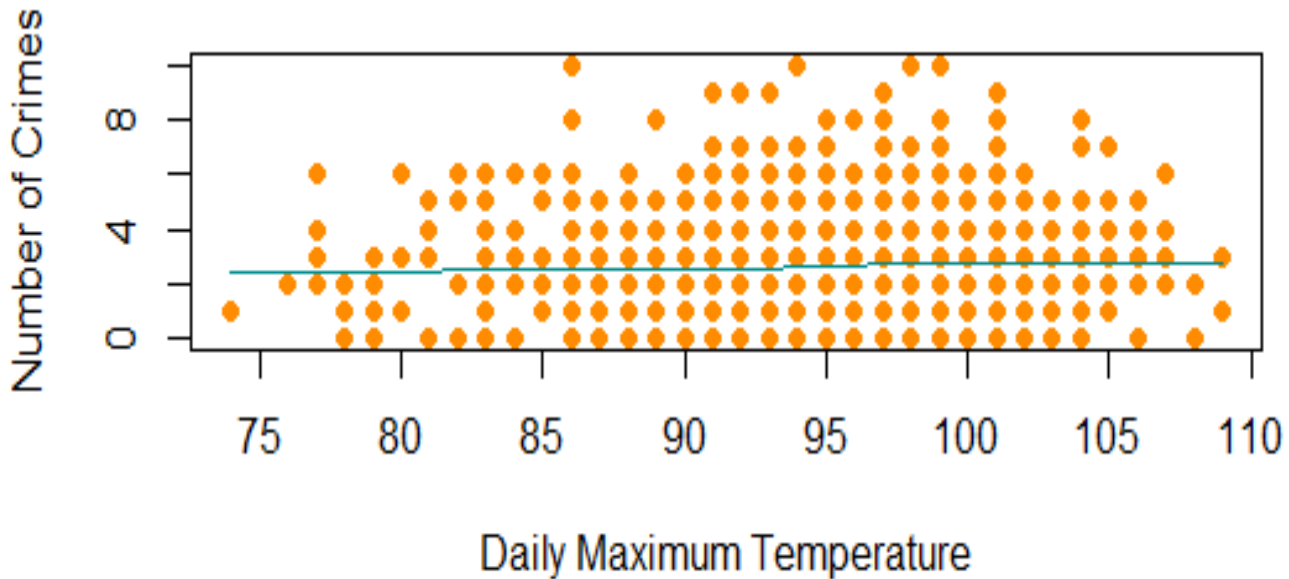
Analysis showed no significant association between heat waves and assaults.

BUT: a marginally significant positive relationship between temperature and number of crimes

* 'dwtest' function from R package 'lmtest' was used to conduct the Durbin-Watson test



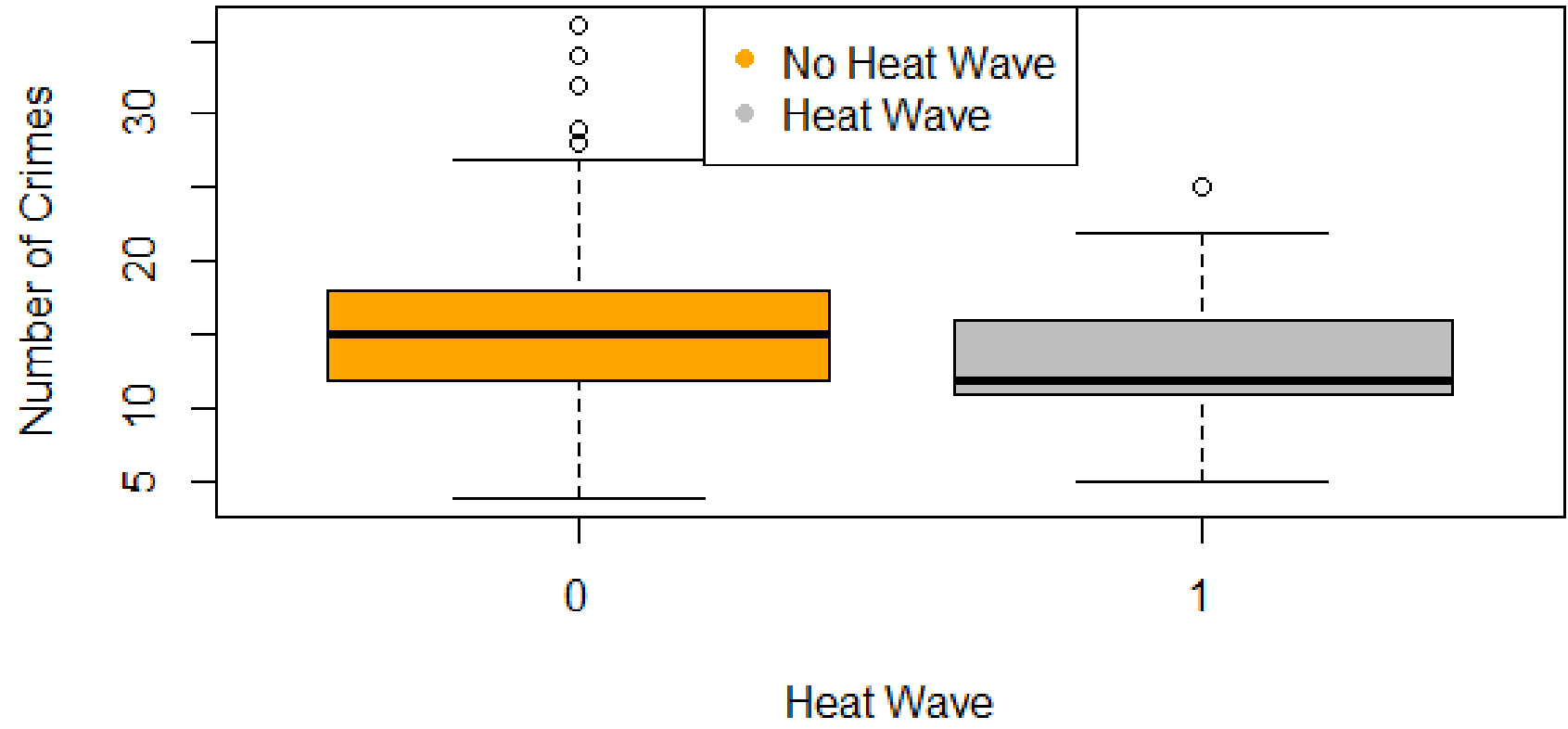
MCKINNEY, TX Assaults Results Table and Scatterplot for Continuous Temperature



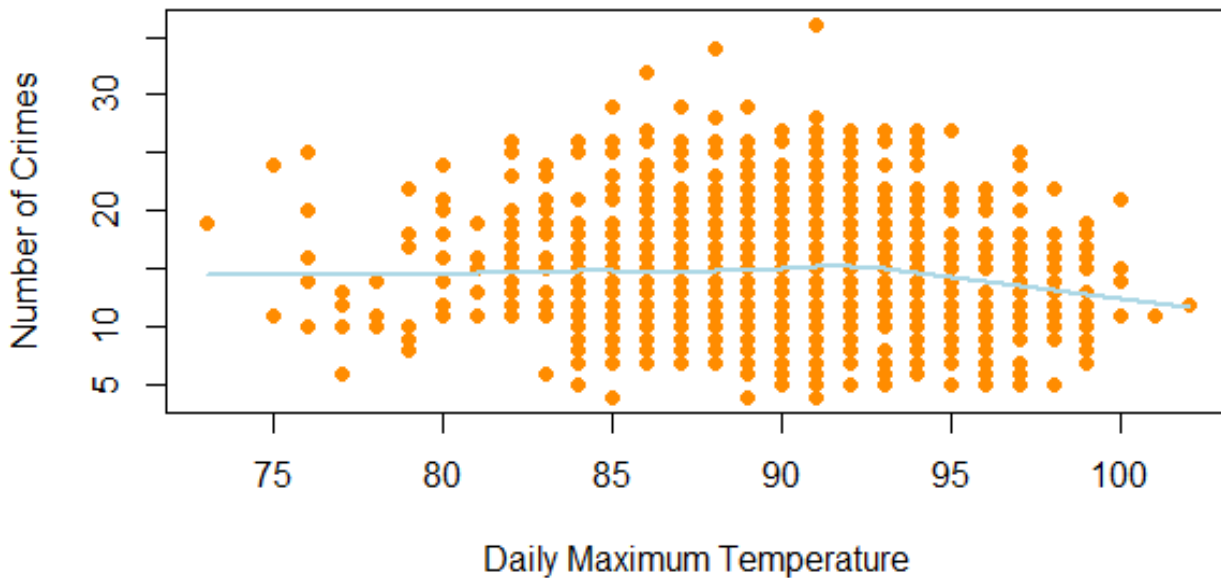
	Estimate	Standard Error	P-value
Heat Wave (Binary)	-0.1321	0.1255	0.293
TMAX (Temperature)	0.007063	0.004025	0.0797

- Using Poisson regression with no adjustment for temporal autocorrelation, we found a significant negative relationship between maximum temperature and daily number of assaults

CHARLESTON,
SOUTH
CAROLINA




Charleston, SC Assaults Results Table and Scatterplot for Continuous Temperature



	Estimate	Standard Error	P-Value
Heat Wave (Binary)	-0.13584	0.03584	0.00015** *
TMAX (Temperature)	-0.00358	0.00163	0.02790

Analyses were run for both robberies and homicides in McKinney and Iowa City



However, there were too few incidents of these types of crimes to justify any results

CONCLUSION

In McKinney and Iowa City, we did not find a significant association between heatwaves and criminal activity

In Charleston, we found a significant negative association between heatwaves and criminal activity

LIMITATIONS OF THE STUDY

We only investigated three categories of crimes for three small cities

- Not enough data for most categories because we chose relatively peaceful, small cities not known for criminal activity
- Larger cities may yield significant results

Only examined heatwaves during summer breaks

- Example: Iowa City is a college town; much of the population is gone during these months

Limitations may have led to unexpected signs on coefficients and their lack of significance

FUTURE STUDIES

- Run similar analysis on larger U.S. cities to see whether there are comparable results
- Find alternatives to the HGLM software for accommodating serial autocorrelation in regression models
- Analyze the potential association between mental health and heat waves

ACKNOWLEDGMENTS

- We would like to thank our awesome research mentor, Dr. Kate Cowles
- Our amazing Professor and Director, Dr. Gideon Zamba
- The extraordinary T.A's, Melissa and Biyue
- The Iowa Summer Institute in Biostatistics (ISIB) program
- And the National Heart Lung and Blood Institute (NHLBI), grant # HL147231



National Heart, Lung,
and Blood Institute

REFERENCES

- National Oceanic and Atmospheric Administration. <https://www.ncdc.noaa.gov/cdo-web/search?datasetid=GHCND>
- National Incident-Based Reporting System. <https://crime-data-explorer.fr.cloud.gov/downloads-and-docs>
- Bowles, Erik. *Classifying Heat Waves in the United States*. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.880.4387&rep=rep1&type=pdf>
- Lars Ronnegard, Xia Shen and Moudud Alam (2010) hglm: A Package for Fitting Hierarchical Generalized Linear Models. *The R Journal*, 2(2): 20-28.
- Bivand, Roger S. and Wong, David W. S. (2018) Comparing implementations of global and local indicators of spatial association *TEST*, 27(3), 716-748.
- R package 'lmtest' used: Achim Zeileis, Torsten Hothorn (2002). Diagnostic Checking in Regression Relationships. *R News* 2(3), 7-10.



THANK YOU SO VERY MUCH!!!
QUESTIONS?



DURBIN-WATSON TEST (FOR LINEAR REGRESSION)

- The Durbin-Watson test is used to detect whether there is serial autocorrelation
 - If the test yields a low p-value, there's autocorrelation
 - Thus, we had to use the HGLM

WILCOXON RANK-SUM TEST

Goal: to compare frequency of crimes on Heatwave days vs. Non-heatwave days

Problem: This test can only be used if there is not a significant amount of serial autocorrelation