Client Project 12

Extracting Flood Depths from Imagery

Marina Baker, Upasana Mahanta, Lee Littlejohn
Chicago DSI 11

Project 12 - Problem Statement

Floods cause damage to infrastructure and homes. The depth of flood waters is a good indicator of the severity of damage. Floods are incredibly difficult to model, and while model outputs are useful to emergency managers, it is crucial to know the actual depth. Social media and news outlets often present pictures of floods. How can this imagery be used to estimate the depth of water in a given area?

How do we measure depth from a photograph?

- ➢ Object detection and relative visibility above water
- ➢ Manual annotation from subjective inference
- ➢ Geolocating and matching with USGS depth data

Project 12 - Methods and Challenges

The Greatest Game

Using Google, Getty, the Associated Press, Instagram, and Reddit to hunt photos with visible flooding and recognizable markers to manually tag with spatial coordinates

Project 12 - Conclusions

What we need to continue this research:

- ➢ Intact EXIF data for geolocating
- ➢ More photos by at least an order of magnitude
- ➢ Small army of underpaid students
- ➢ More hurricanes in the Social Media Age
- ➢ Time, lots of it

Utilizing Flooding Data for Analysis & Predictions

What we have:

- Mapped measurements of flood depth for each hurricane

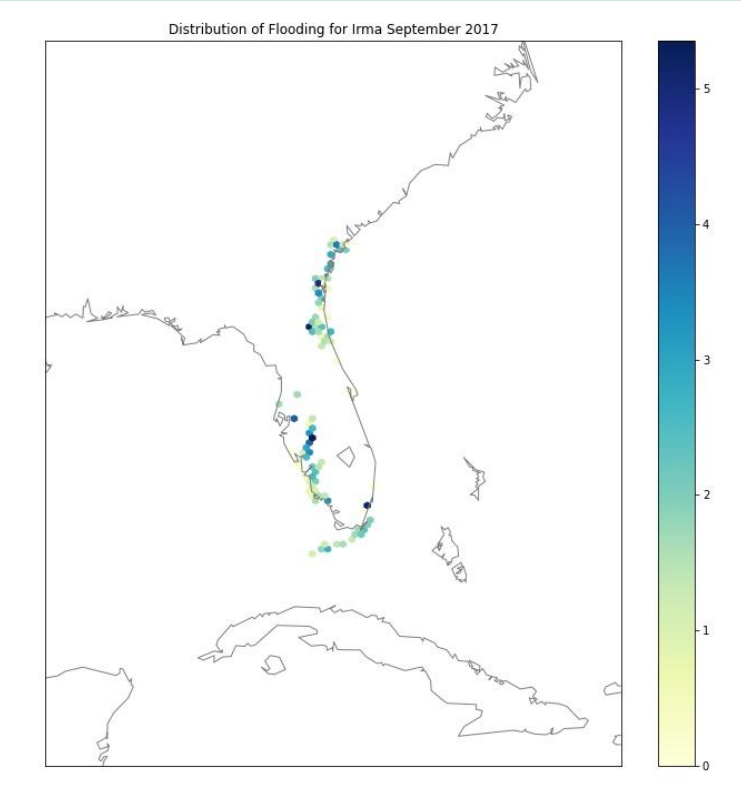
What we can find:

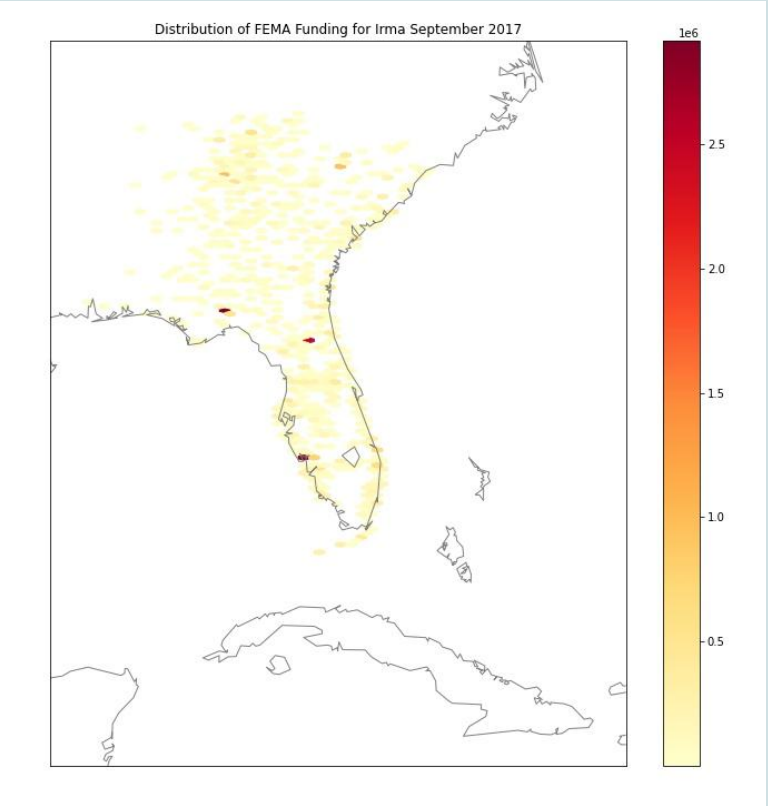
- FEMA Public Assistance Grants for each hurricane

New question:

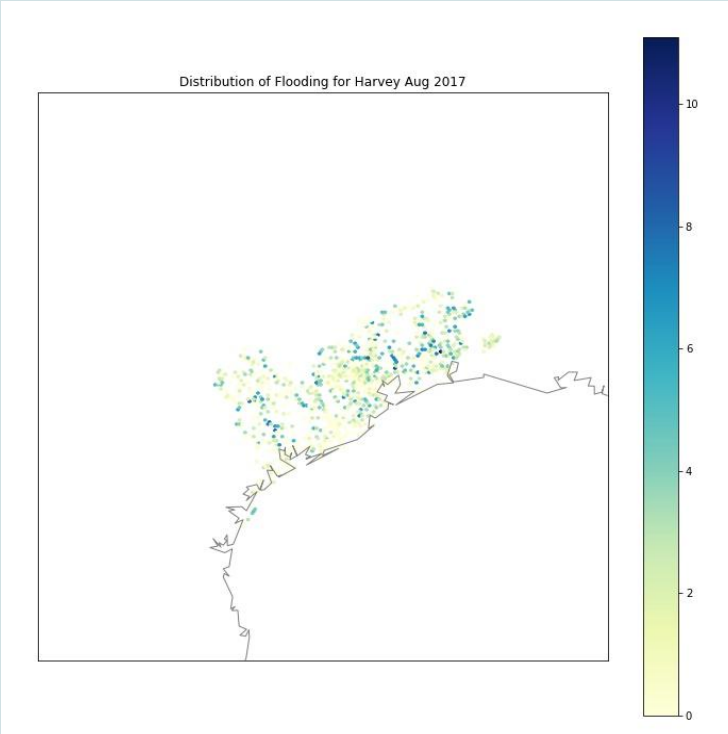
- **What is the relationship between severity of flooding of an area, and public assistance received? Is this distributed evenly or unevenly?**

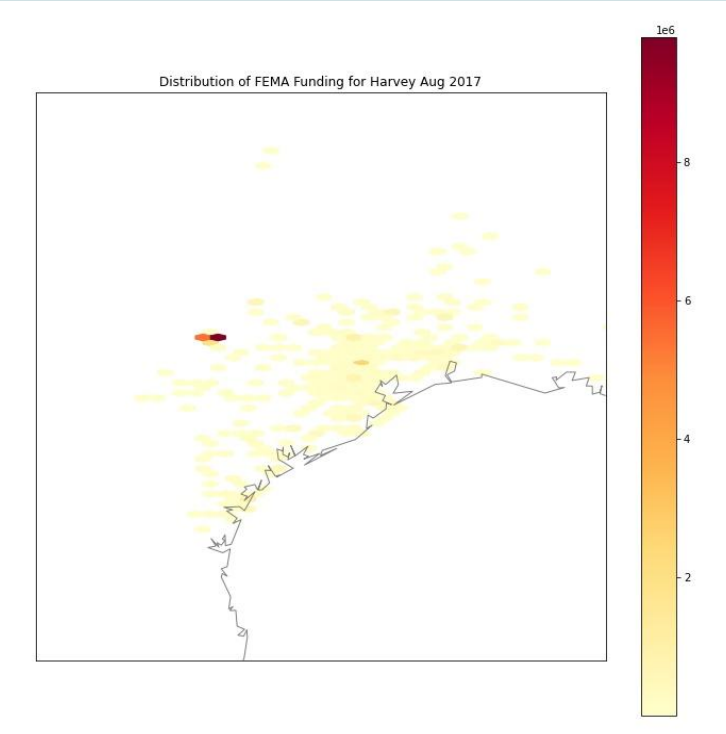
Visuals: Map of Flooding and Grants side-by-side



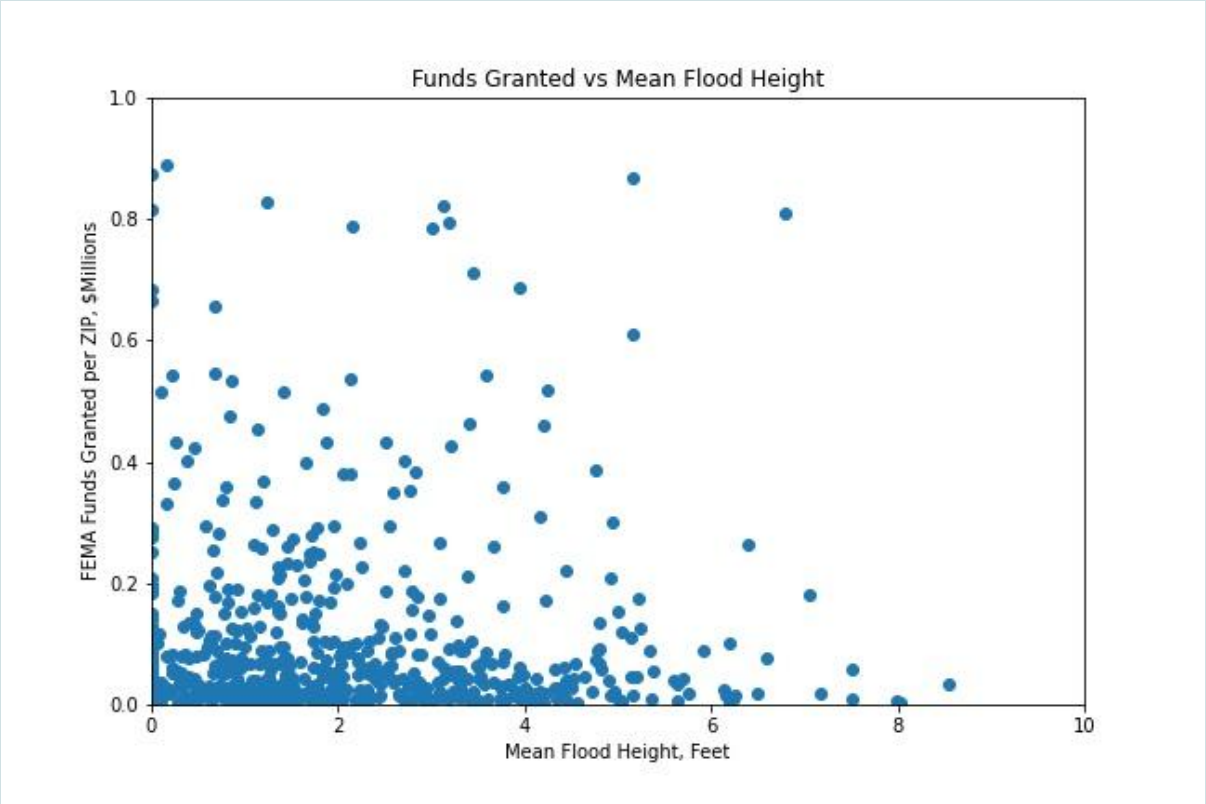


Visuals: Map of Flooding and Grants side-by-side

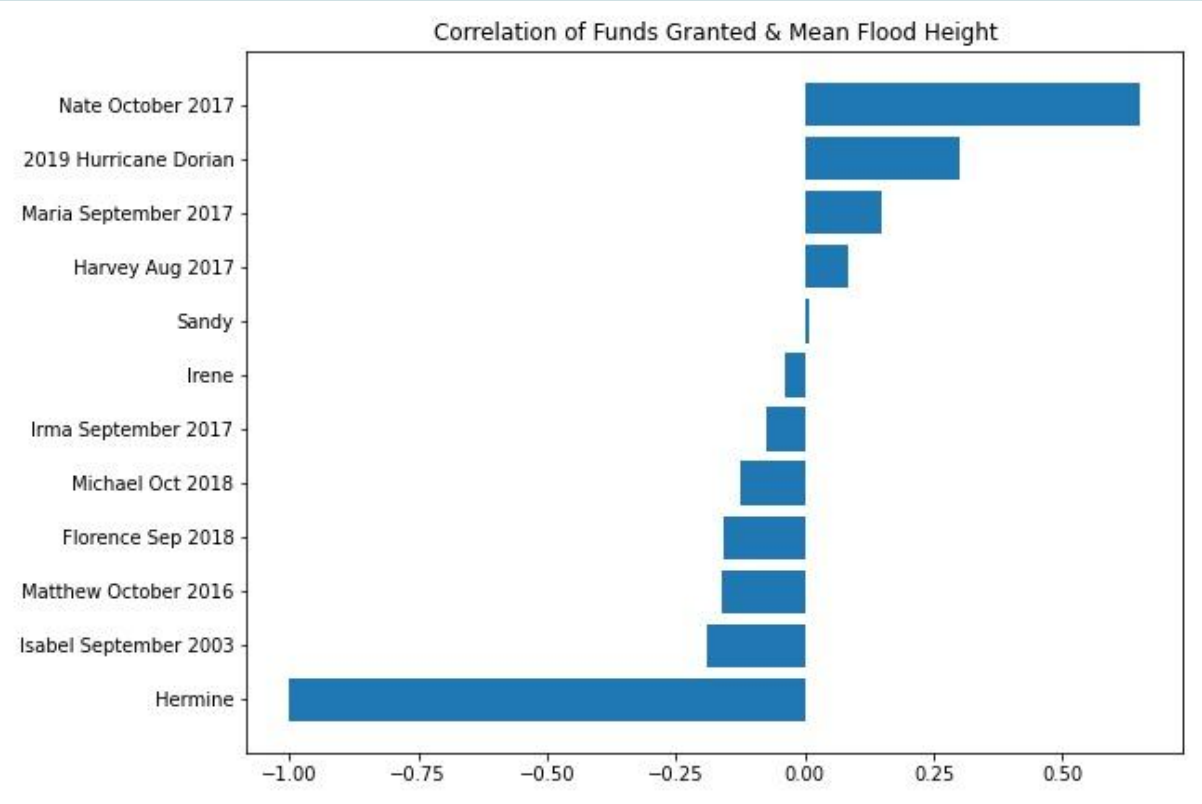




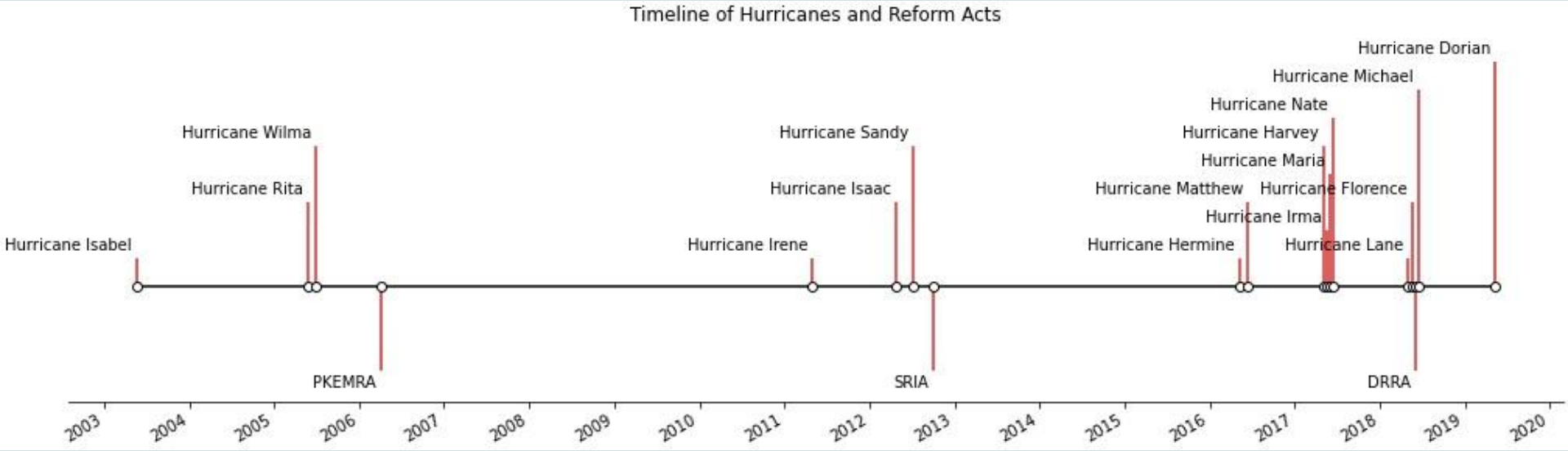
Relationship between Funding and Flood Severity



Relationship between Funding and Flood Severity



Disaster Recovery Funding Reforms



The Post Katrina Emergency Reform Act of 2006 (PKEMRA)

The Sandy Recovery Improvement Act (SRIA)

The Disaster Recovery Reform Act of 2018 (DRRA)

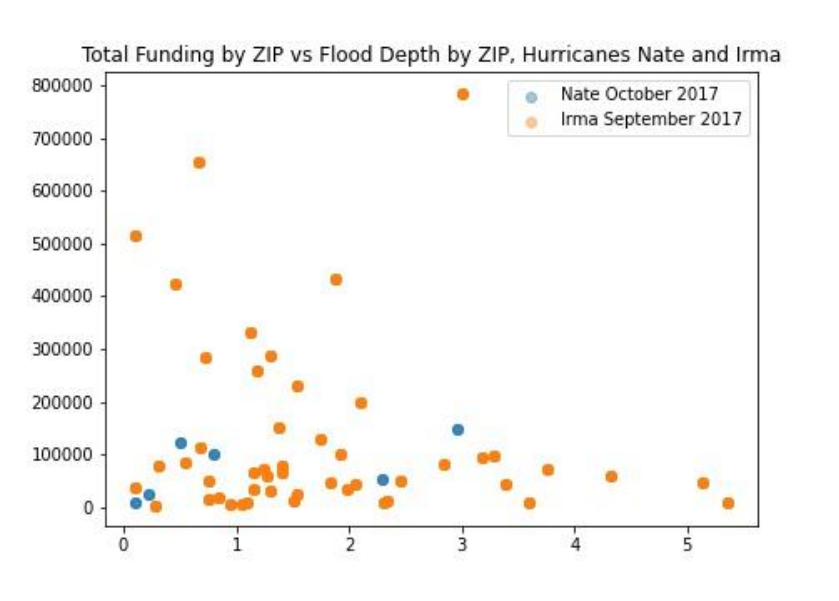
Hypothesis Test: Equality of Variance

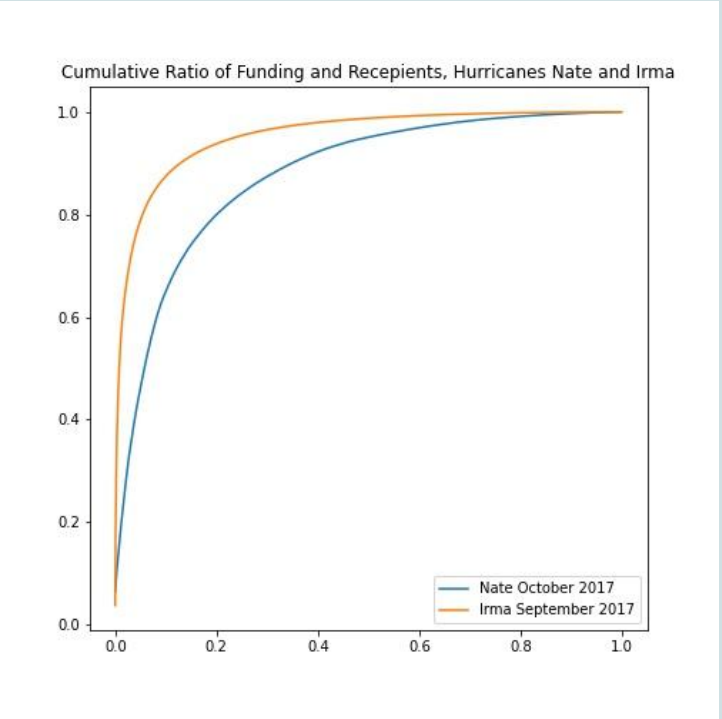
H_0: $Var(\text{funding / severity})_{\text{before reform}} = Var(\text{funding / severity})_{\text{after reform}}$

H_1: $Var(\text{funding / severity})_{\text{before reform}} \neq Var(\text{funding / severity})_{\text{after reform}}$

Reform Act	**Levene's Test *p*-value:**
The Post Katrina Emergency Reform Act of 2006 (PKEMRA)	0.231
The Sandy Recovery Improvement Act (SRIA)	0.105
The Disaster Recovery Reform Act of 2018 (DRRA)	0.741

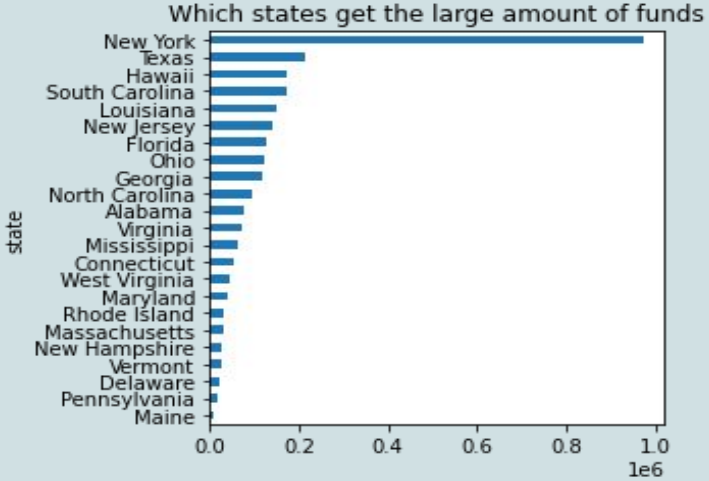
Case Studies: Hurricanes Nate and Irma





Conclusions and Recommendations

- A good dataset with annotated flood images would have *great* value
- Funding distribution varies wildly from location to location and flood to flood
- We could not prove a statistically significant change in funding distribution with the data available between the passages of major pieces of reform legislation
- We recommend more research into taking a data-based approach to appropriately distributed disaster relief funding



Next Steps

More Data!

- How damage scales with flood depth
- How economic metrics scale with flood depth (from Project 7, perhaps?)
- In-depth case studies, storm to storm, city to city

Thank You!

Appendix

Public Assistance:

The PA program offers assistance to State, local, and tribal governments after a declared major disaster or emergency for eligible disaster-related damage. In addition, certain private nonprofit organizations that operate facilities and provide services to the public typically provided by a government agency may receive public assistance. PA is a cost-shared (75 percent Federal, 25 percent non-Federal) grant program. However, States may request, and the President may grant, a higher Federal cost share in situations in which the disaster has had a significant economic impact on the State. PA provides funding to grantees for the repair and restoration of damaged public and eligible private nonprofit facilities, for emergency measures taken to protect lives and property, and in support of disaster-related debris removal. When PA is authorized, such assistance may be provided under one or more of the following seven categories:

- Category A – Debris Removal
- Category B – Emergency Protective Measures
- Category C – Repair of Roads and Bridges
- Category D – Water Control Facilities
- Category E – Public Buildings
- Category F – Public Utilities
- Category G – Other Facilities

Levene's test (Levene 1960) is used to test if *k* samples have equal variances. Equal variances across samples is called homogeneity of variance. Some statistical tests, for example the analysis of variance, assume that variances are equal across groups or samples. The Levene test can be used to verify that assumption.

Levene's test is an alternative to the Bartlett test. The Levene test is less sensitive than the Bartlett test to departures from normality. If you have strong evidence that your data do in fact come from a normal, or nearly normal, distribution, then Bartlett's test has better performance.

The Levene test is defined as:

H_0: $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$

H_a: $\sigma_i^2 \neq \sigma_j^2$ for at least one pair (i,j).

Test Statistic: Given a variable Y with sample of size N divided into k subgroups, where N_i is the sample size of the ith subgroup, the Levene test statistic is defined as:

$$W = \frac{(N-k)}{(k-1)} \frac{\sum_{i=1}^{k} N_i (\bar{Z}_{i.} - \bar{Z}_{..})^2}{\sum_{i=1}^{k} \sum_{j=1}^{N_i} (Z_{ij} - \bar{Z}_{i.})^2}$$

where Z_{ij} can have one of the following three definitions:

1. $Z_{ij} = |Y_{ij} - \bar{Y}_{i.}|$

   where $\bar{Y}_{i.}$ is the mean of the i-th subgroup.

2. $Z_{ij} = |Y_{ij} - \tilde{Y}_{i.}|$

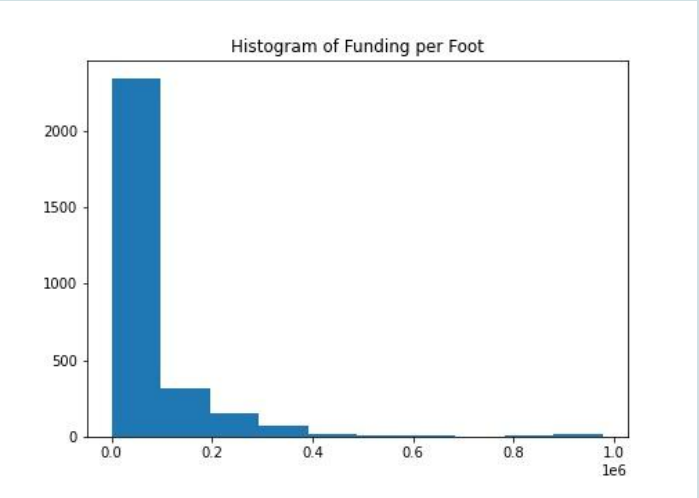
   where $\tilde{Y}_{i.}$ is the median of the i-th subgroup.

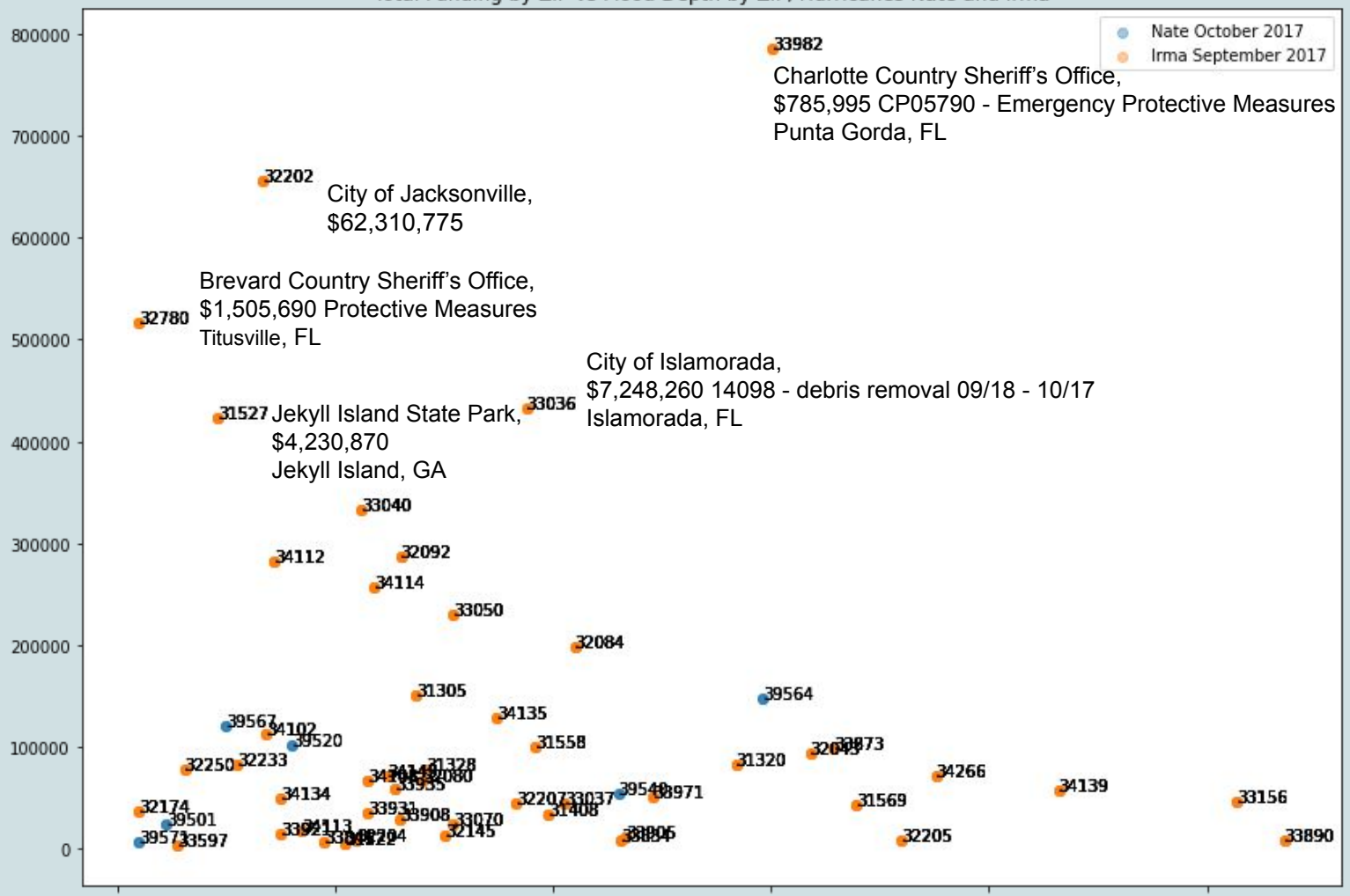
3. $Z_{ij} = |Y_{ij} - \bar{Y}'_{i.}|$

   where $\bar{Y}'_{i.}$ is the 10% trimmed mean of the i-th subgroup.

$\bar{Z}_{i.}$ are the group means of the Z_{ij} and $\bar{Z}_{..}$ is the overall mean of the Z_{ij}.

Levene's original paper only proposed using the mean. Brown and Forsythe (1974)) extended Levene's test to use either the median or the trimmed mean in addition to the mean. They performed Monte Carlo studies that indicated that using the trimmed mean performed best when the underlying data followed a Cauchy distribution (i.e., heavy-tailed) and the median performed best when the underlying data followed a χ^2_4 (i.e., skewed) distribution. Using the mean provided the best power for symmetric, moderate-tailed, distributions.




Total Funding by ZIP vs Flood Depth by ZIP, Hurricanes Nate and Irma
Nate October 2017
Irma September 2017
800000
700000
600000
500000
400000
300000
200000
100000
0
0
1
2
3
4
5
33982
Charlotte Country Sheriff's Office,
$785,995 CP05790 - Emergency Protective Measures
Punta Gorda, FL
32202
City of Jacksonville,
$62,310,775
Brevard Country Sheriff's Office,
$1,505,690 Protective Measures
Titusville, FL
32780
City of Islamorada,
$7,248,260 14098 - debris removal 09/18 - 10/17
Islamorada, FL
33036
31527
Jekyll Island State Park,
$4,230,870
Jekyll Island, GA
33040
34112
32092
34114
33050
32084
31305
39564
34135
39567
34102
39520
31558
32250
32233
31320
34266
34139
34134
32174
39501
39573
33597
33156
31569
32205
33890