Claim Prediction for Flood Insurance

Abhishek Verma

Table of Content

Presentation Duration : 20 min

Overview

- There is a need to process flood insurance claims quickly and efficiently as volume of flood claims rises drastically in the event of natural disaster requiring fast and prudent processing.

  https://www.nytimes.com/2017/11/04/business/a-broke-and-broken-flood-insurance-program.html

- Based on the patterns of the past claims amount paid, the geographic location of property, property condition, insurance coverage carried etc a reasonable prediction of the claim amount can be made.

Table of Content

Presentation Duration : 20 min

Data Source: Nation Flood Insurance Policy Dataset

FEMA dataset that includes more than two million flood insurance claims records dating back to 1978.

[FEMA Publishes NFIP Claims and Policy Data]

FIMA NFIP Redacted Claims Data Set

Federal Insurance & Mitigation Administration National Flood Insurance Program (FIMA NFIP) Redacted Claims Dataset

Congress passed the National Flood Insurance Act (NFIA), 42 U.S.C. 4001 in 1968, creating the National Flood Insurance Program (NFIP) in order to reduce future flood losses through flood hazard identification, manage floodplain, and provide insurance protection. The Department of Housing and Urban Development (HUD) originally administered the NFIP, and Congress subsequently transferred the NFIP to FEMA upon its creation in 1979. FEMA and insurance companies participating in FEMA's Write Your Own (WYO) program offer NFIP insurance coverage for building structures as well as for contents and personal property with the building structures to eligible and insurable properties. The WYO program began in 1983 with NFIP operating under Part B of the NFIA and allows FEMA to authorize private insurance companies to issue the Standard Flood Insurance Policy (SFIP) as FEMA's fiduciary and fiscal agent. FEMA administers NFIP by ensuring insurance applications are processed properly; determining correct premiums; renewing, reforming, and cancelling insurance policies; transferring policies from the seller of the property to the purchaser of the property in certain circumstances; and processing insurance claims.

FEMA

Data Source: Key Information on Flood Claims

- Claim Amount Paid on Building Damage
- Claim Amount Paid on Content Damage

- Coverage Amount on Building
- Coverage Amount on Content

- Time of Flood
- Property Age
- Start Date of Policy

- Zip code
- County
- State

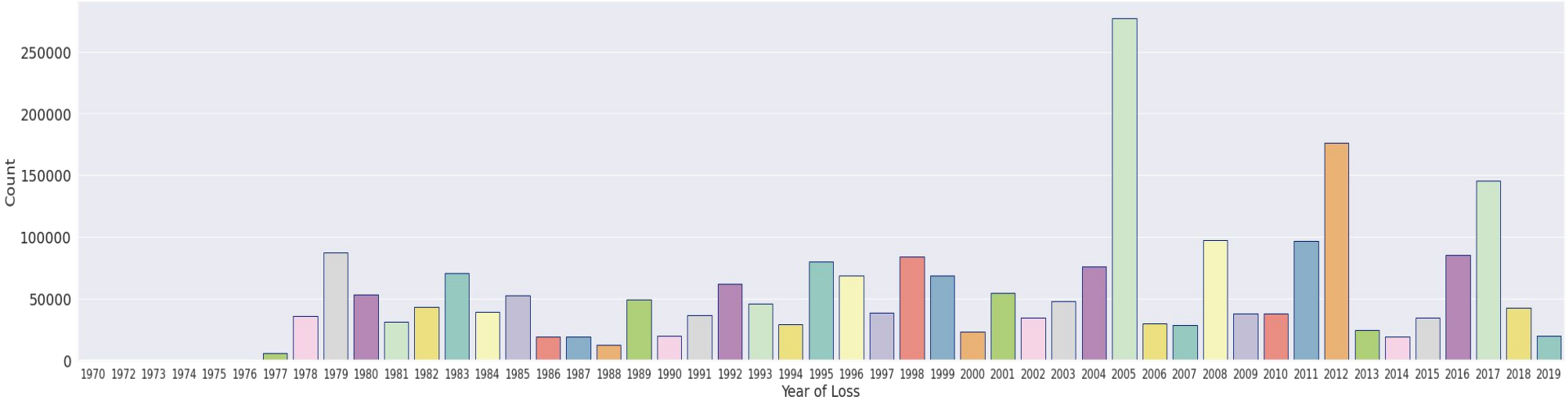
- Flood Zone Code

- Condominium (Y/N)
- Elevated Building (Y/N)
- Primary Residence (Y/N)
- House of Worship (Y/N)
- Agriculture Structure (Y/N)
- Non Profit Structure (Y/N)
- Small Business (Y/N)

Table of Content

Presentation Duration : 20 min

Key Insights: Number of Property Loss with Year



- Recent three instances of heavy property losses occurred in the years of major hurricane events; Hurricane Katrina (2005), Hurricane Sandy (2012) and Hurricane Harvey (2017)

Key Insights: Number of Property Loss with Year

2005 Atlantic hurricane season

Season summary map

Seasonal boundaries	
First system formed	June 8, 2005
Last system dissipated	January 6, 2006 (record latest, tied with 1954)
Strongest storm	
Name	Wilma (Most intense hurricane in the Atlantic basin)
• Maximum winds	185 mph (295 km/h) (*1-minute sustained*)
• Lowest pressure	882 mbar (hPa; 26.05 inHg)
Seasonal statistics	
Total depressions	31 (record high)
Total storms	28 (record high)
Hurricanes	15 (record high)
Major hurricanes (Cat. 3+)	7 (record high)
Total fatalities	3,912 total
Total damage	$171.755 billion (2005 USD) (Second-costliest tropical cyclone season on record)

2012 Atlantic hurricane season

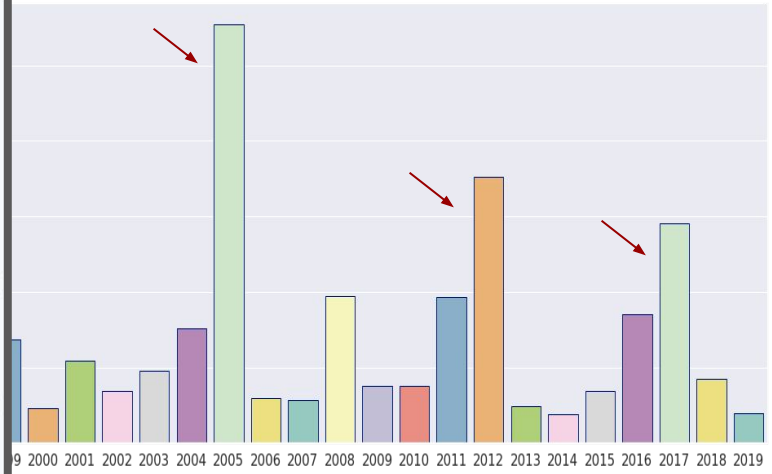
Season summary map

Seasonal boundaries	
First system formed	May 19, 2012
Last system dissipated	October 29, 2012
Strongest storm	
Name	Sandy
• Maximum winds	115 mph (185 km/h) (*1-minute sustained*)
• Lowest pressure	940 mbar (hPa; 27.76 inHg)
Seasonal statistics	
Total depressions	19
Total storms	19
Hurricanes	10
Major hurricanes (Cat. 3+)	2
Total fatalities	355 total
Total damage	≥ $72.32 billion (2012 USD) (Third-costliest tropical cyclone season on record)

2017 Atlantic hurricane season

Season summary map

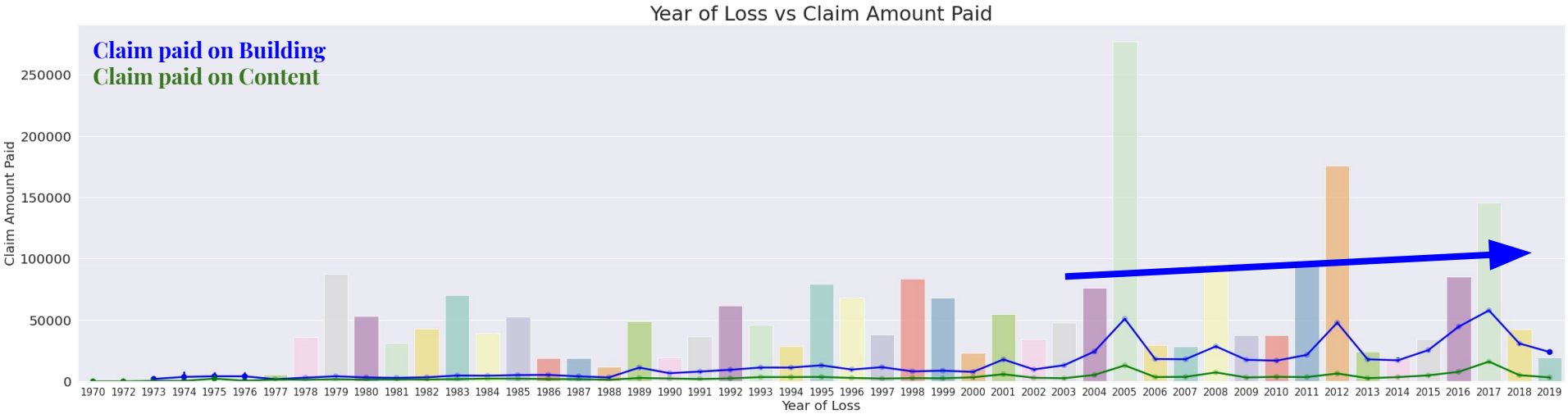
Seasonal boundaries	
First system formed	April 19, 2017
Last system dissipated	November 9, 2017
Strongest storm	
Name	Maria[nb 1]
• Maximum winds	175 mph (280 km/h) (*1-minute sustained*)
• Lowest pressure	908 mbar (hPa; 26.81 inHg)
Seasonal statistics	
Total depressions	18
Total storms	17
Hurricanes	10
Major hurricanes (Cat. 3+)	6
Total fatalities	3,364 total
Total damage	≥ $294.92 billion (2017 USD) (Costliest tropical cyclone season on record)
Related articles	
• Timeline of the 2017 Atlantic hurricane season	
• 2017 Pacific hurricane season	



Top **three** costliest hurricane season

Source: Wikipedia

Key Insights: Trend of Claim Amount

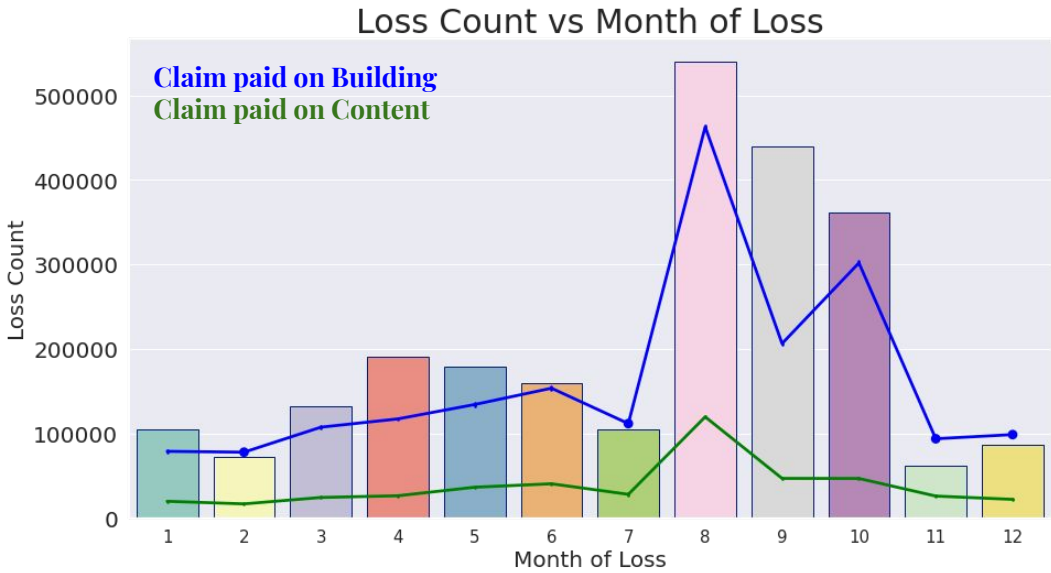


- Upward trend in claim amount paid through flood insurance on damaged properties
- Increase in paid claim amounts correlates with major hurricane seasons

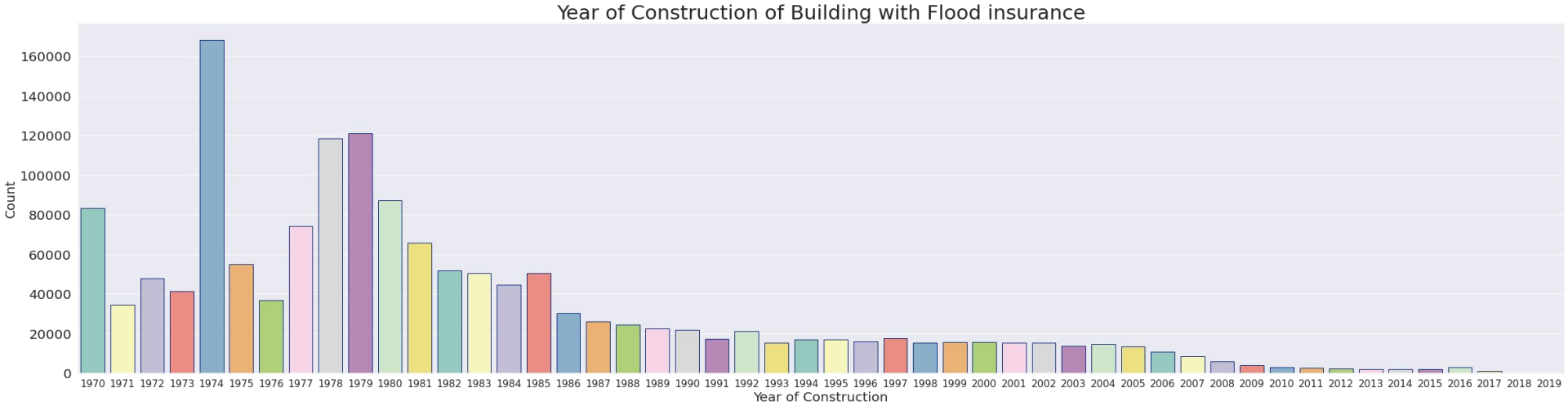
Key Insights: Trend of Claim Amount (Month)

- Majority of claims come from the month of August, September and October

- Highly correlated with the timing of major hurricane events.

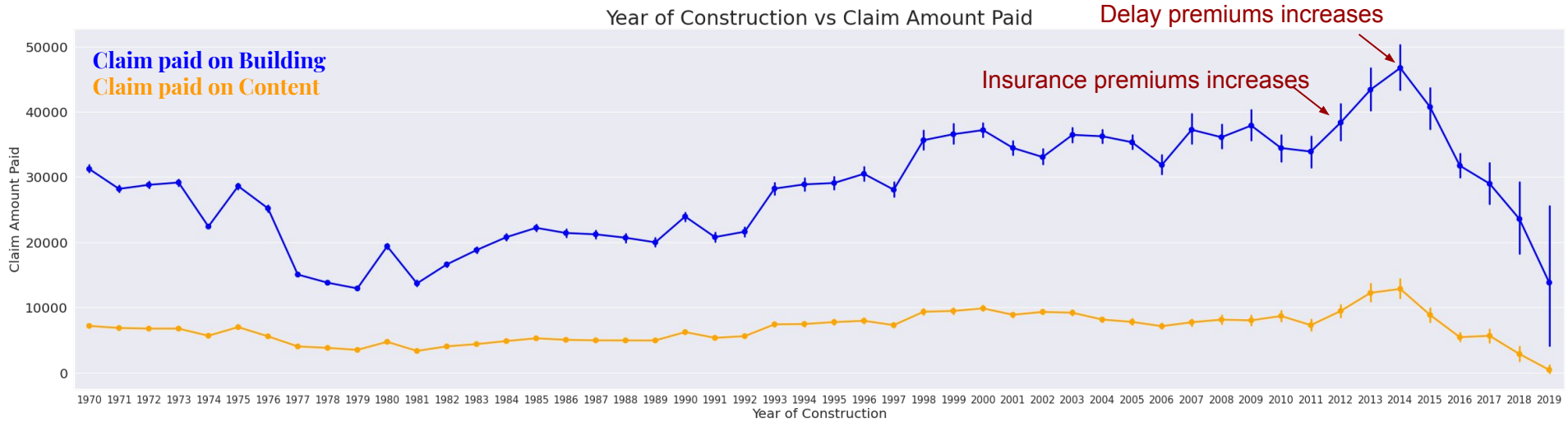


Key Insights: Year of Construction



- Major flood insurance claims made on older property (1975 - 1990)
- Downward trend on claims made on newer property hinting at better community planning to mitigate flood risks

Key Insights: Year of Construction

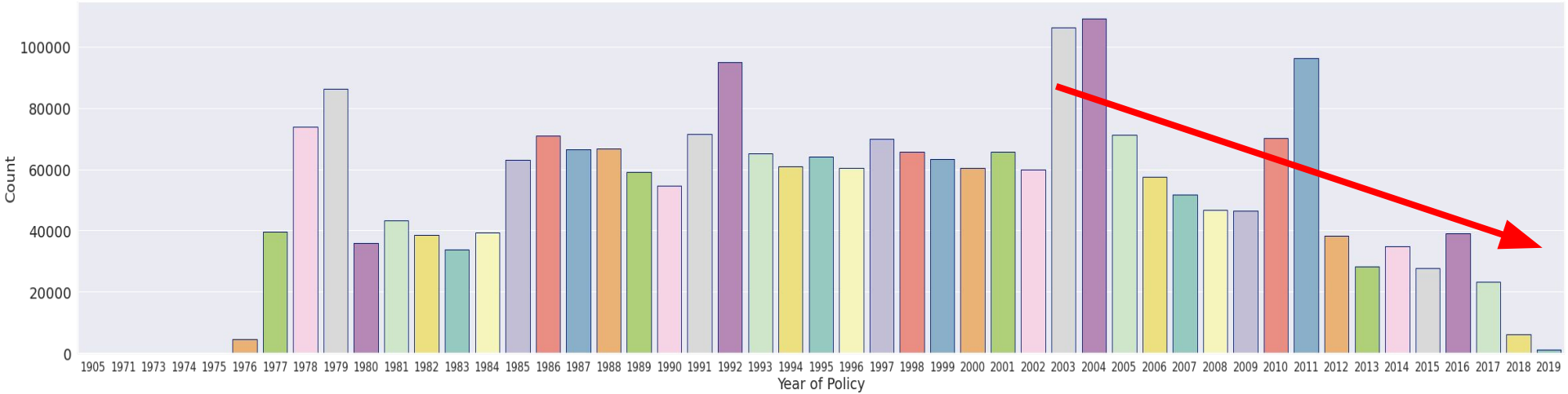


Biggert–Waters Flood Insurance Reform Act of 2012 [edit]

The Biggert–Waters Flood Insurance Reform Act of 2012 was "designed to allow premiums to rise to reflect the true risk of living in high-flood areas."[1] The bill was supposed to deal with the "insolvency" of the National Flood Insurance Program by requiring the premiums to reflect real flood risks.[3] The result was a 10 fold increase in premiums.[1] At present, $527 billion worth of property is in the coastal floodplain.[3] The federal government heavily underwrites the flood insurance rates for these areas. The law "ordered FEMA to stop subsidizing flood insurance for second homes and businesses, and for properties that had been swamped multiple times."[6] These changes were to occur gradually over the course of five years. FEMA was also instructed to do a study on the affordability of this process, a study which it has failed to complete.[6]

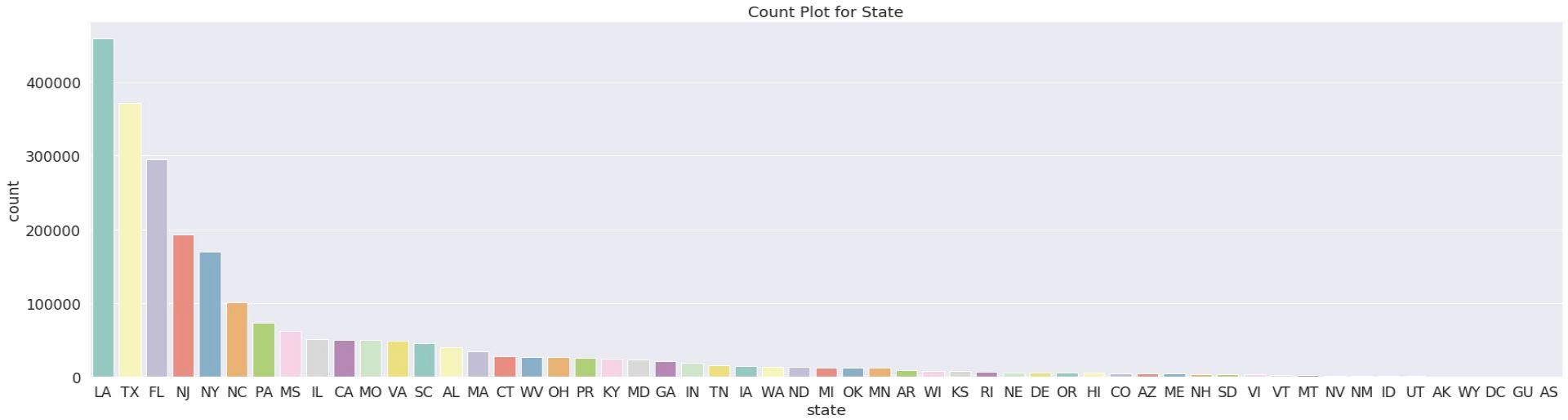
The **Homeowner Flood Insurance Affordability Act of 2014** (S. 1926) was a United States Congress bill that would have delayed the increases in flood insurance premiums that were part of the Biggert–Waters Flood Insurance Reform Act of 2012.[1][2] The reforms from that law were meant to require flood insurance premiums to actually reflect the real risk of flooding, which led to an increase in premiums.[3] At the time of the bill, the National Flood Insurance Program was $24 billion in debt.[3]

Key Insights: Insurance Policy Years



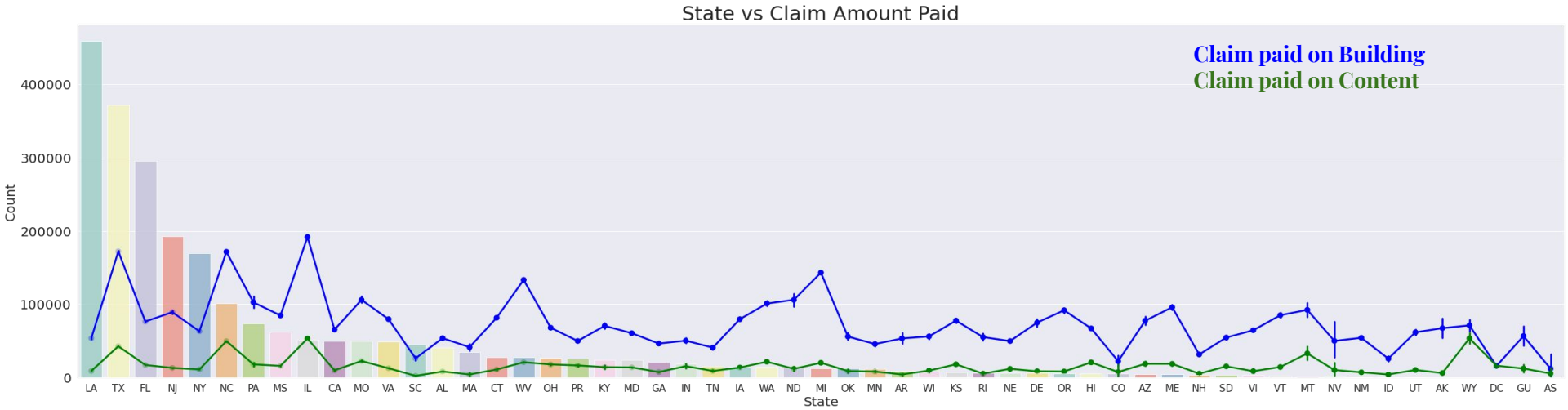
- Downward trend in number of flood insurance policies carried through 50 years

Key Insights: State-wise Number of Claims



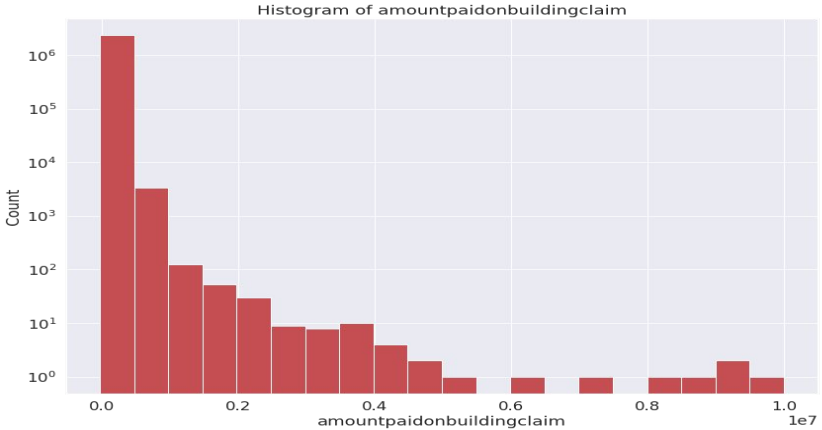
- More number of flood claims coming from coastal states. Top six states include Louisiana, Texas, Florida, New Jersey, New York and North Carolina.

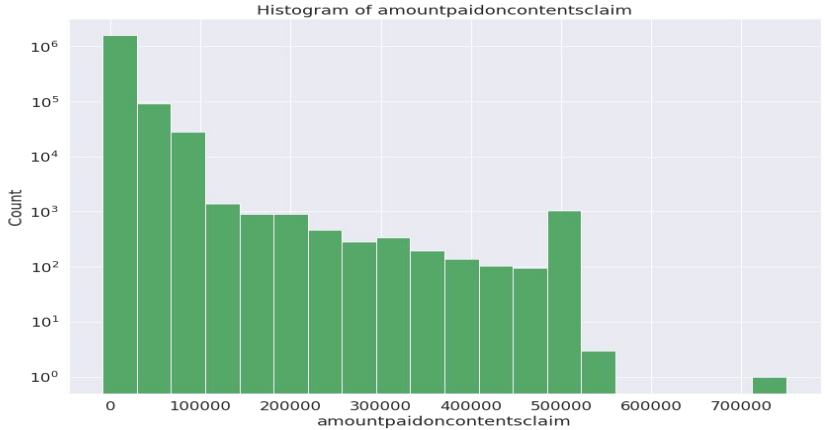
Key Insights: State-wise Number of Claims

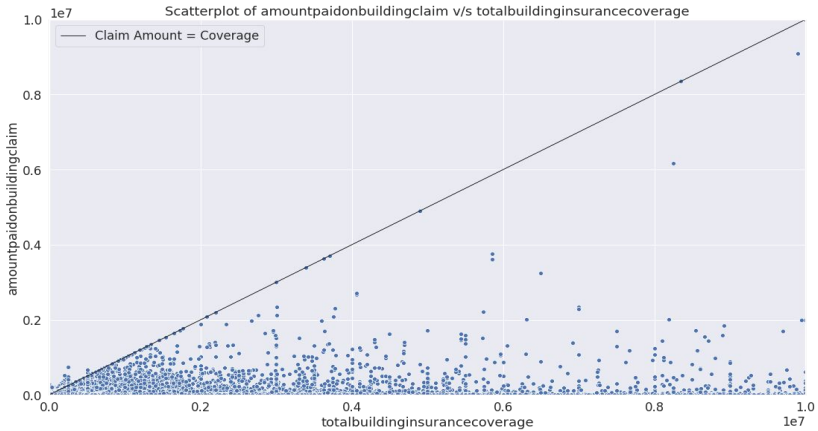


- Variation among states in paid insurance amount.

Coverage versus Claims paid







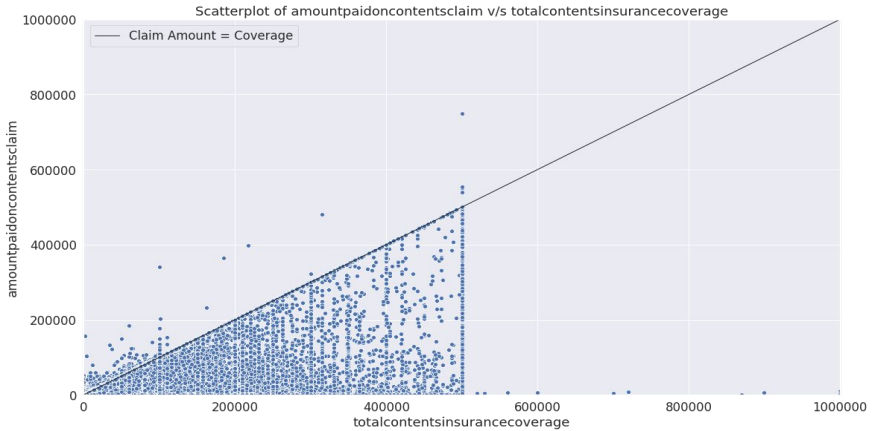
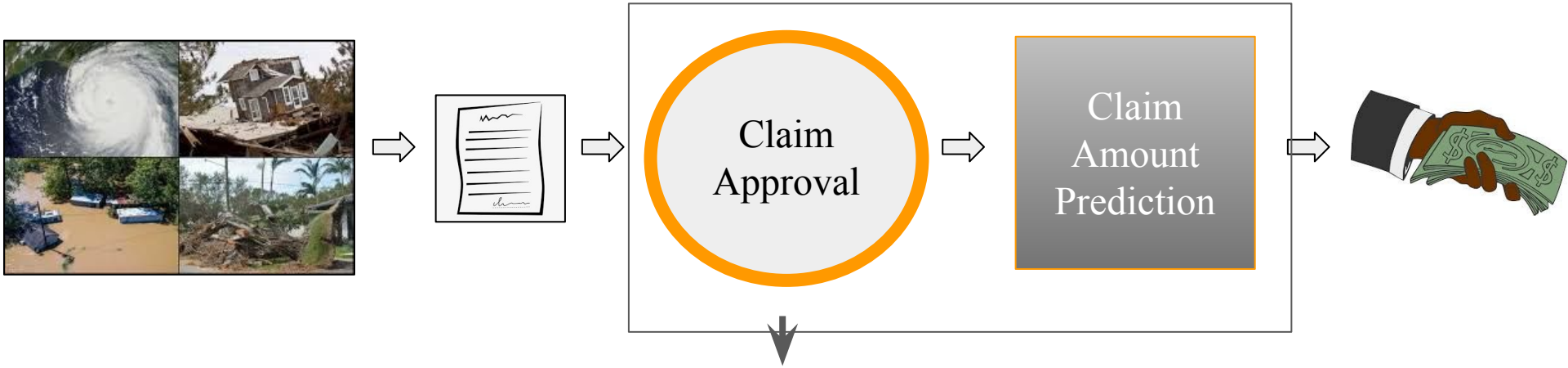


Table of Content

Presentation Duration : 20 min

Supervised ML Modelling: 2 Step

- Build supervised modelling to predict claim amount using information from FEMA dataset such as year of loss event, property location, property condition etc

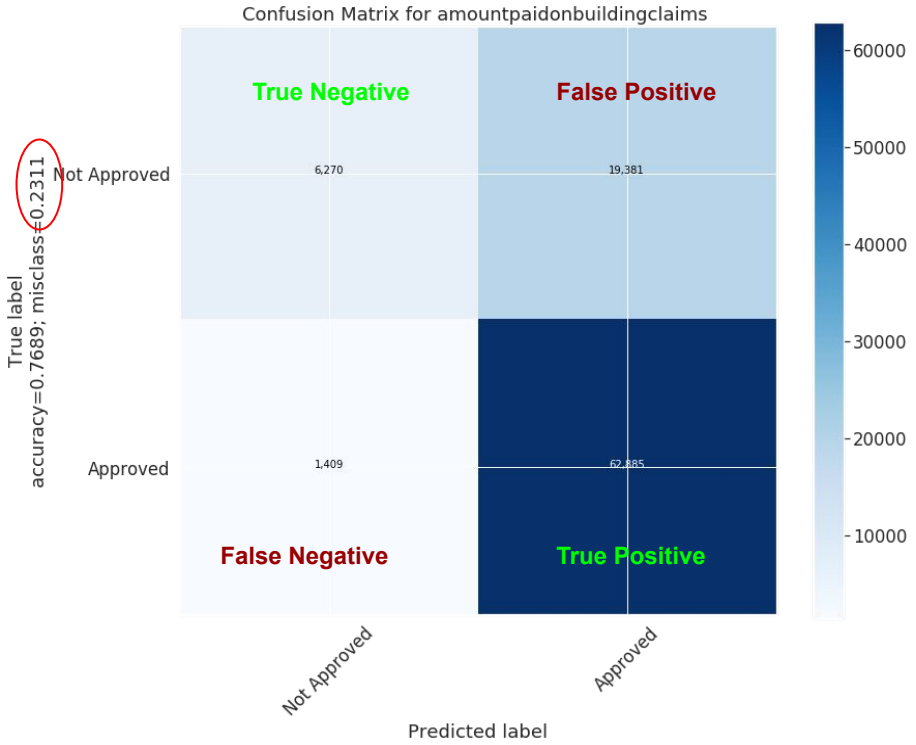


Is a claim acceptable for payout ??

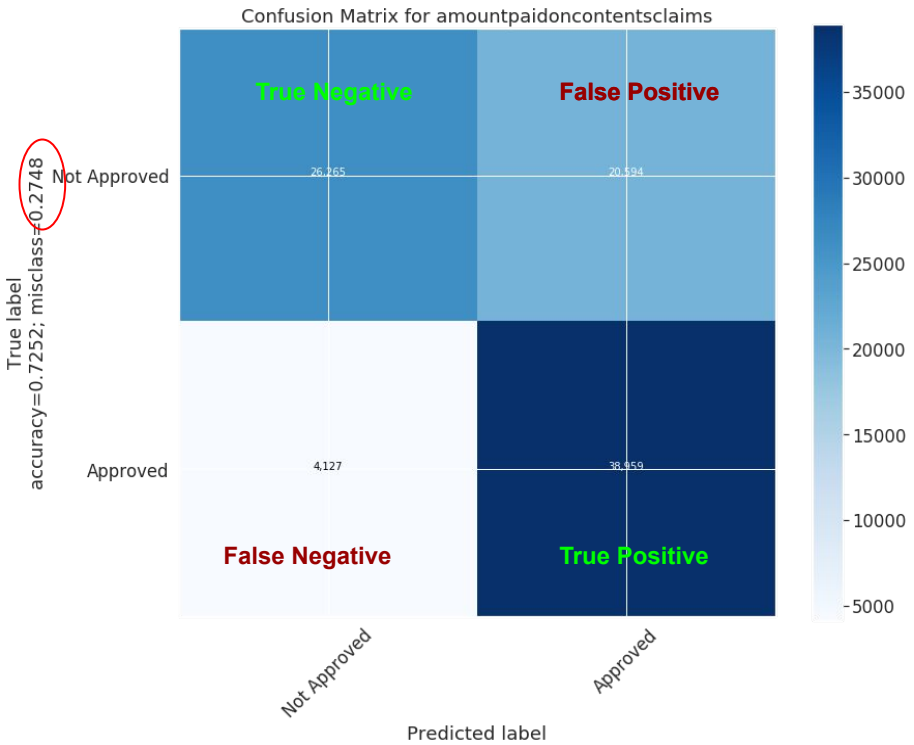
Performed on Individual States

Random Forest Classification

Building Damage



Content Damage



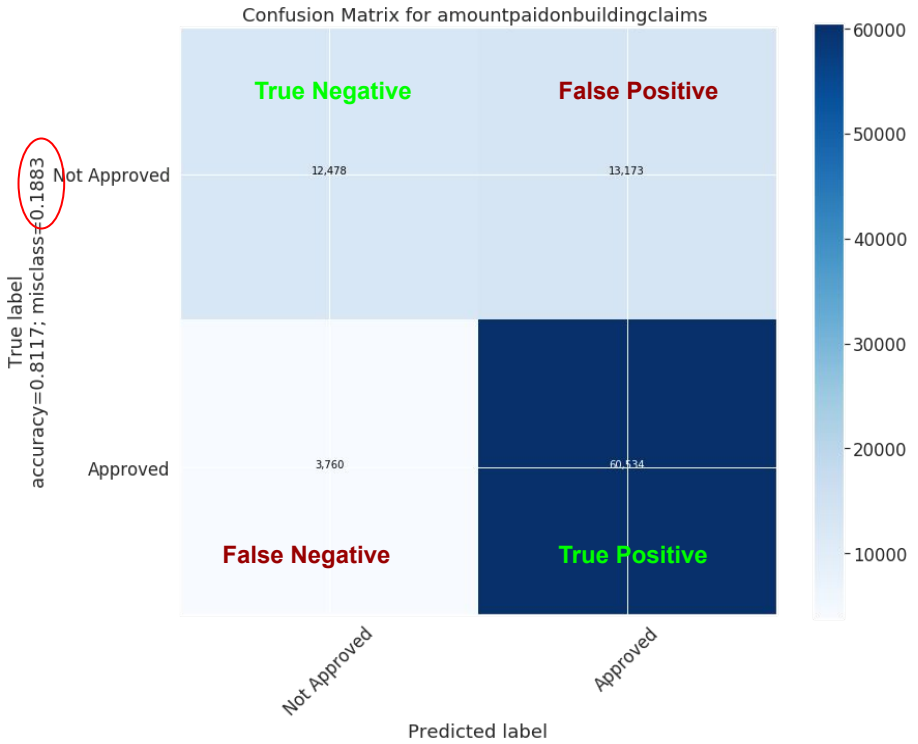
Performed on Individual States

Gradient Boost Classification

Building Damage

Reduced False Positive

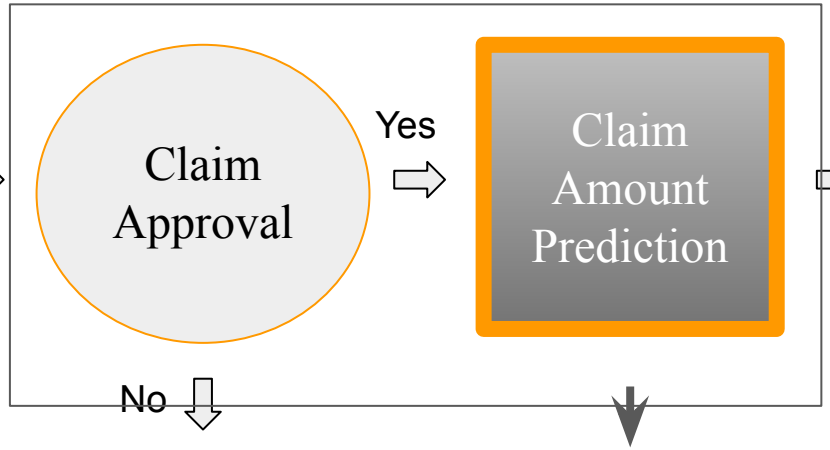
Content Damage



Performed on Individual States

Supervised ML Modelling: 2 Step

- Build supervised modelling to predict claim amount using information from FEMA dataset such as year of loss event, property location, property condition etc

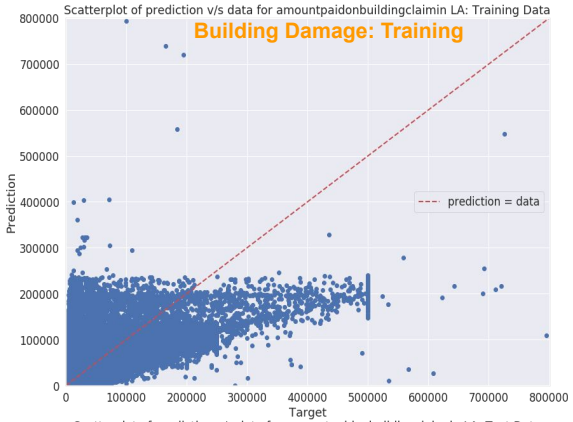


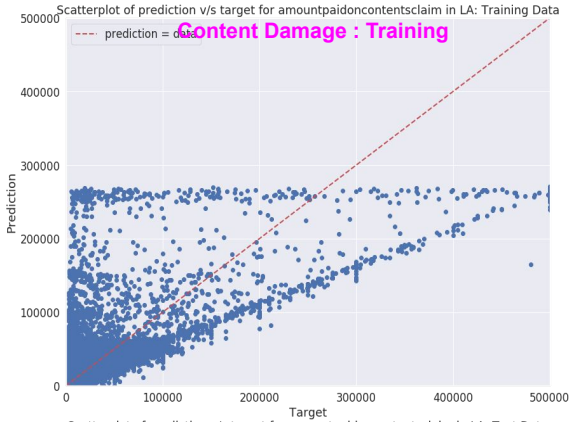
Performed on Individual States

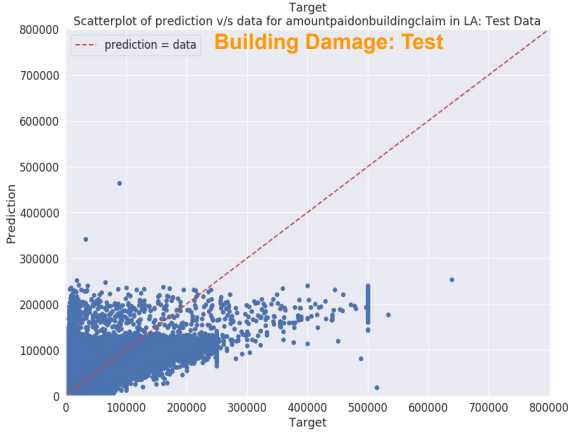
Linear Regression Performance

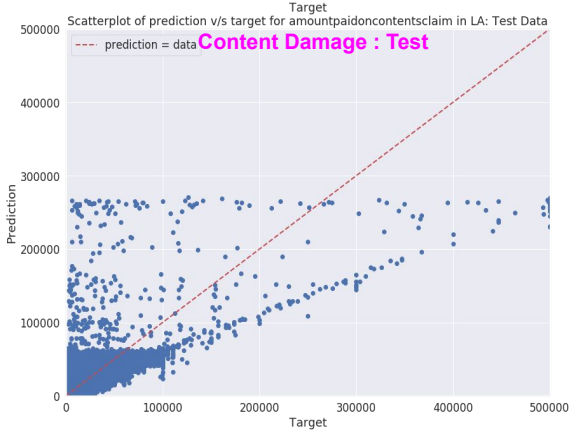
	Training	Test
R Square	0.677	0.631
RMS Error	40789.23	42323.43

	Training	Test
R Square	0.672	0.612
RMS Error	17819.11	19919.19





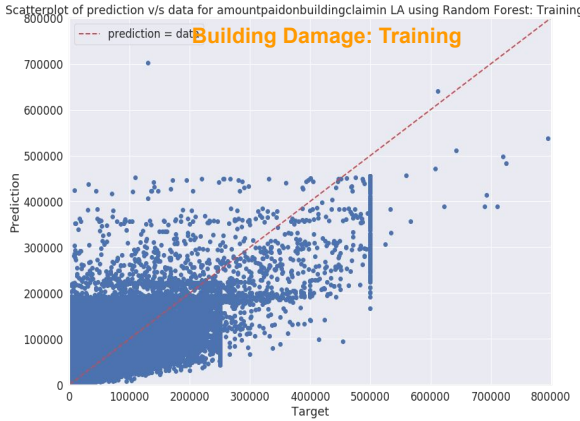


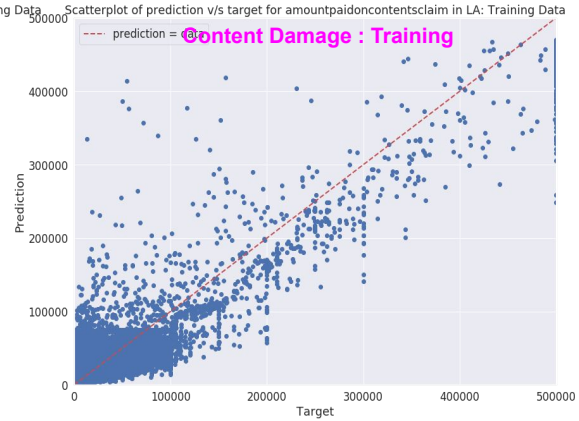


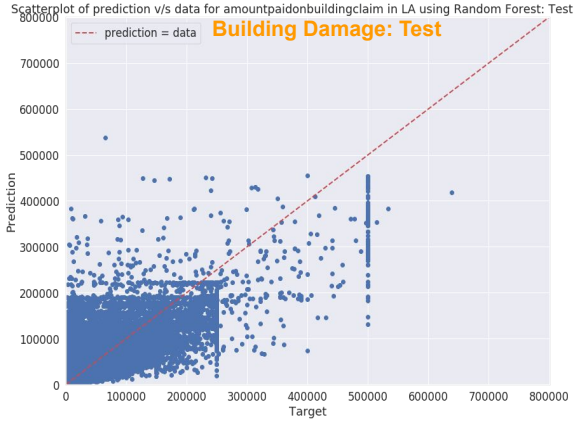
Random Forest Performance

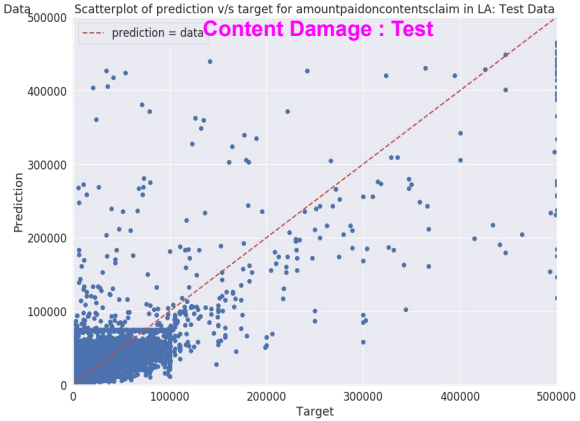
	Training	Test
R Square	0.743	0.713
RMS Error	30032.47	32173.71

	Training	Test
R Square	0.772	0.658
RMS Error	12922.82	16739.56





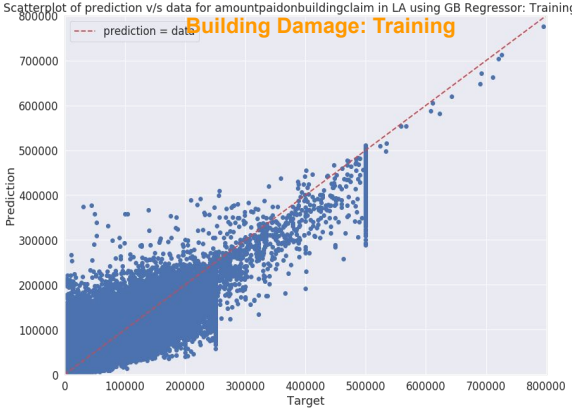


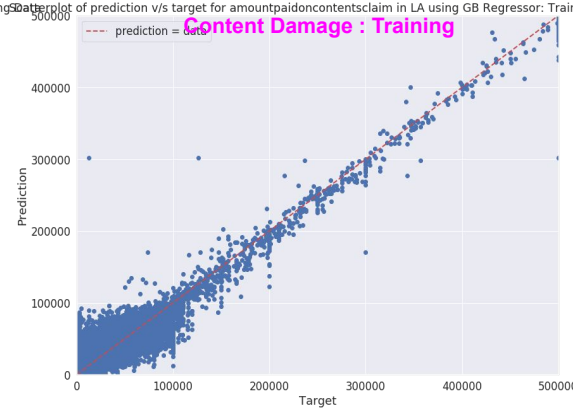


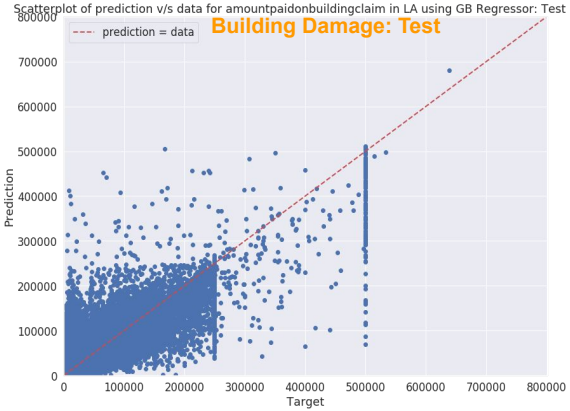
Gradient Boosting Performance

	Training	Test
R Square	0.859	0.772
RMS Error	21899.32	28708.95

	Training	Test
R Square	0.898	0.688
RMS Error	8834.17	15988.67







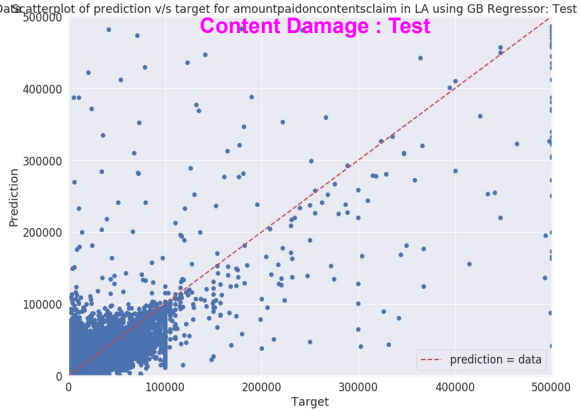
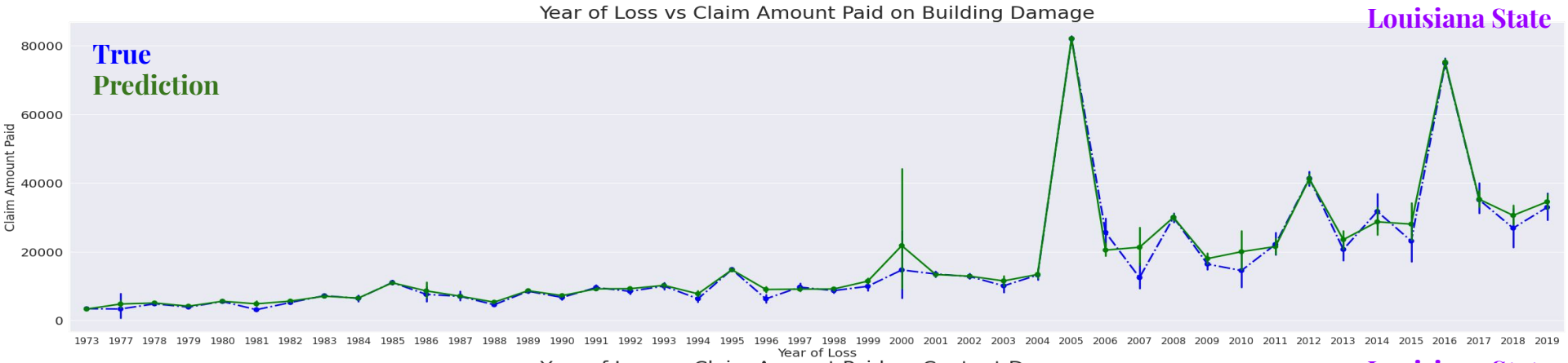
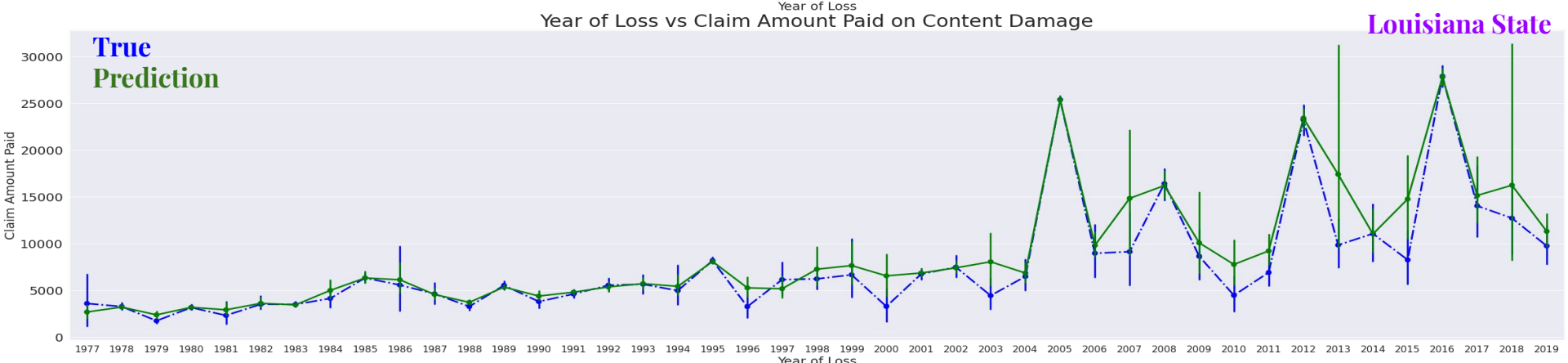


Table of Content

Presentation Duration : 20 min

Model Evaluation (Test Dataset)





Modelling: Limitation

- Model evaluation suggest overfitting issues with gradient boosting method despite better accuracy.

- Prediction accuracy in content claim is relatively low.

Modelling: Future Improvements

- Strong correlation between natural disaster events (hurricane) and mean claim amount which is needed to be addressed in input.

- Include premium information of each flood insurance policy.

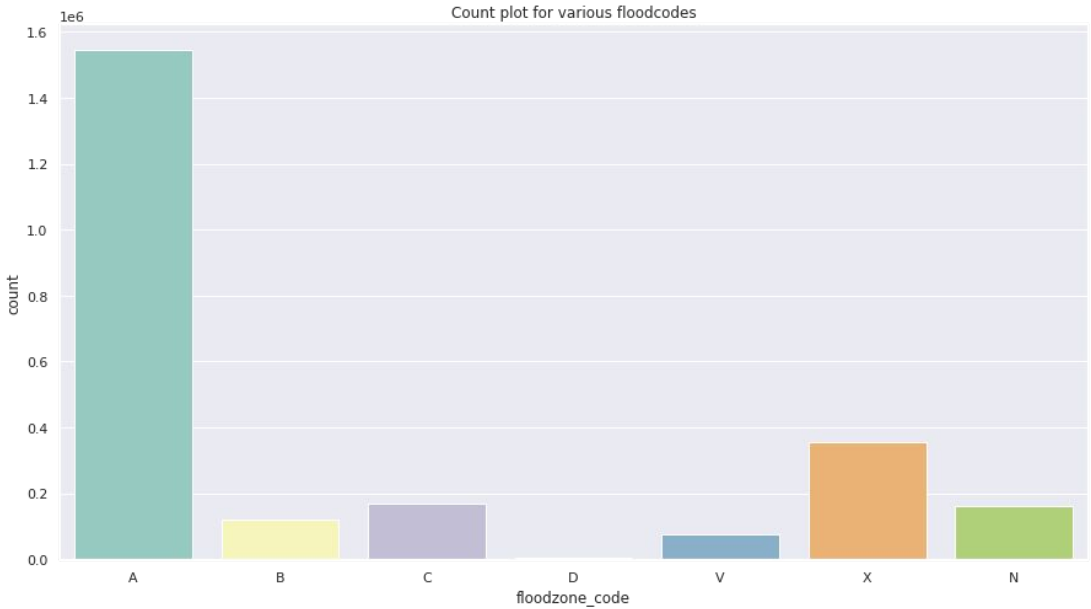
Conclusion

- Building a prediction model is useful in the process of issuing claim amounts in the event of natural disaster (Hurricane, rainfall etcs) which require many working hours and resource in underwriting and claim evaluation.

Questions ?

Appendix

Categorical Variable: Flood-Zone Code



The flood-zone can be broadly classified:

1. Special Flood: A
2. Moderate Flood: B
3. Minimum Flood: C
4. Possible Flood: D
5. Velocity Flood: V
6. Moderate/Minimum Flood: X
7. Not specified: N

Categorical Variable: Property Attributes