Financial Impact of Hurricanes on US Publicly-Traded REITs from 1995-2020

Team A - December 16, 2021

S&P Global

Team A

Project Management	Data Team	Modeling Team
Maddie	**Aansh** - Property & company data expert	**Zixin**
	Uli - Hurricane data expert	**Zilin**
Advisors		
Vincent Agboto - Professor	**Jingshu** - Data joining & visualization	**Yicheng**
Allen Chang - Associate	**Raven** - Data joining & visualization	**Zehua**

Background

Hurricane **rainfall** levels are expected to **increase** on average by the end of the century[1]

Since 1980, hurricanes in the US have caused at least **$945.9 billion in damage**[2]

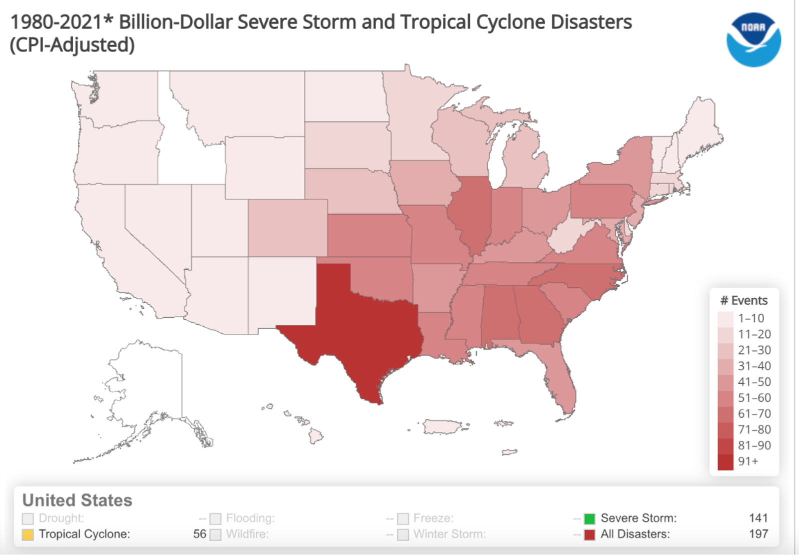
Hurricane **intensities** are expected to **increase** on average by the end of the century[1]

Businesses want to know how this will impact them - and S&P wants to be able to answer this question for their clients

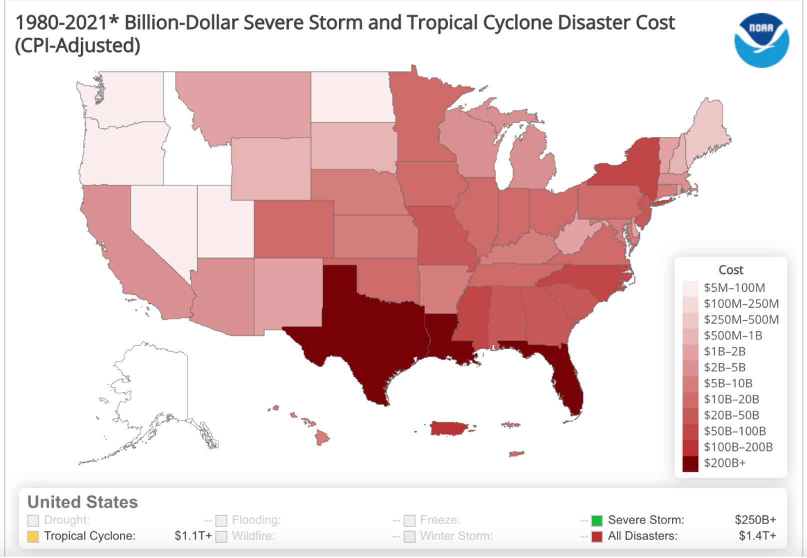
Source:
1. Geophysical Fluid Dynamics Laboratory (https://www.gfdl.noaa.gov/global-warming-and-hurricanes/)
2. National Oceanic and Atmospheric Administration (https://coast.noaa.gov/states/fast-facts/hurricane-costs.html)

Background

Number of Hurricanes per State



Cost of Hurricanes per State



NOAA National Centers for Environmental Information (NCEI) U.S. Billion-Dollar Weather and Climate Disasters (2021). https://www.ncdc.noaa.gov/billions/, DOI: 10.25921/stkw-7w73

Research Question

What is the relationship between hurricane intensity and the financial loss of impacted companies?

- Hurricane intensity measured by maximum wind speed per year and number of hurricane encounters
- Impacted companies defined to be any companies within a 50 mile radius of hurricane landfall
- Financial loss defined to be asset write downs

Methodology

1

Understand the structure of the data:
- What financial data is available over which years
- How hurricane data is presented
- Features to join the financial data with hurricane data
- Specific attention to geolocation (longitude and latitude of hurricane or property)

2

Define scope of research:
- Years: 1995-2020 based on available data
- Publicly traded companies because those are the ones with more data available
- Hurricanes which pass through the US because that's the country we will look at economic control factors for

3

Define analytical methods to use for each research question:
- Goal is to understand real estate sector losses due to hurricanes
- Pull macroeconomic data (GDP, inflation rate) to use as control variables
- Build regression models on key variables

4

Iteratively evaluate and improve models based on:
- Predictive ability of models
- Interpretability of models

Data - Hurricanes

IBTrACS data of North Atlantic Area[1]

- Hurricane data from 1842-2021
- Intensity (wind speed) data less reliable because does not address differences in measuring techniques across times and places
- **Includes indicators of which storms hit the US**
- Most useful data attribute will be the landfall information

Indicator of Landfall Information

ADT-HURSAT Data[2]

- Hurricane data from 1978-2017
- Addresses inconsistencies in measuring techniques present in other datasets
- Created by applying the advanced Dvorak Technique to a globally homogenized record of geostationary satellite imagery
- **Lacks indication of whether or not storms made landfall at all**
- Most useful data attribute will be wind speed

More Reliable Wind Speed Data

Source:
1. National Oceanic and Atmospheric Administration (https://www.ncdc.noaa.gov/ibtracs/index.php?name=ib-v4-access)
2. Proceedings of the National Academy of Sciences of the United States (https://www.pnas.org/content/suppl/2020/05/12/1920849117.DCSupplemental)

Initial Data Cleaning - Hurricanes

Data Cleaning Methods

- Recall, the only attribute we want to use from ADT-HURSAT[1] data in our final model is wind speed
- First, we **filter** our IBTrACS dataset to include only hurricanes that made landfall
- Next, we **merge** two datasets by storm_id to have storm surface wind speed from ADT-HURSAT data for hurricanes that have landfalls
- Lastly, our dataset contains the wind speed data from ADT-HURSAT dataset before 2017, and from IBTrACS dataset for 2017-2020

After cleaning, here's the shape of our hurricane dataset

```
df_hurricane.shape
```

```
(7457, 12)
```

Here's an example of what it looks like

```
df_hurricane.head()
```

	stormid	oceanbasin	surfacewindspeed	latitude	longitude	year	month	day	hour	name	datetime	uniqueID
0	1995154N17276	NaN	25.0	17.4	-84.3	1995	6	2	0.0	ALLISON	1995-06-02 00:00:00	0

Source:
1. Proceedings of the National Academy of Sciences of the United States (https://www.pnas.org/content/suppl/2020/05/12/1920849117.DCSupplemental)
2. National Oceanic and Atmospheric Administration (https://www.ncdc.noaa.gov/ibtracs/index.php?name=ib-v4-access)

Initial Data Analysis - Hurricanes

Hurricane Data Source – IBTrACS v.04

- Preemptively, we specify the assortment of values that should be treated as null values.

```
table_na_values=['-999.','-999','-999.000', '-1', '-1.0','0','0.0']
```

- Select only columns needed
- Filter data for hurricane that cause landfall with with category greater than or equal to zero

```
# for LANDFALL* Minimum distance to land over next 3 hours ( = 0 means landfall)
# USA_STATUS  (HU,HR - hurricane)
# USA_SSHS >=0  (Saffir-Simpson Hurricane Scale information based on the wind speed provided by the US agency wind
```

```
df_clean02 = pd.read_csv('IBTrACS v04/IBTrACS_v04_clean02.csv')
df_clean02.shape
```

(45043, 15)

SID	SEASON	NUMBER	BASIN	SUBBASIN	NAME	ISO_TIME	NATURE	LAT	LON	LANDFALL	USA_STATUS	USA_WIND	USA_SSHS
1851175N26270	1851	5	NaN	GM	NOT_NAMED	1851-06-25 00:00:00	TS	27.5333	-94.2667	150	HU	80	1
1851175N26270	1851	5	NaN	GM	NOT_NAMED	1851-06-25 03:00:00	TS	27.7013	-94.6988	125	HU	80	1
1851175N26270	1851	5	NaN	GM	NOT_NAMED	1851-06-25 06:00:00	TS	27.8000	-95.0800	97	HU	80	1
1851175N26270	1851	5	NaN	GM	NOT_NAMED	1851-06-25 09:00:00	TS	27.8616	-95.4384	82	HU	80	1
1851175N26270	1851	5	NaN	GM	NOT_NAMED	1851-06-25 12:00:00	TS	27.9000	-95.7333	59	HU	80	1

Raw Data Size:
163 columns
701,349 rows

Clean Data Size:
15 columns
45,045 rows

Hurricane Data Source – ADT-HURSAT

Kossin et al (2020) Data

```
!ls 'Kossin'/*.csv | head
```

```
Kossin/DataKossinLongFormat.csv
Kossin/DataKossinLongFormatNonNA.csv
Kossin/pnas.1920849117.sd01.csv
Kossin/pnas.1920849117.sd02.csv
Kossin/pnas.1920849117.sd03.csv
Kossin/pnas.1920849117.sd04.csv
Kossin/pnas.1920849117.sd05.csv
Kossin/pnas.1920849117.sd06.csv
Kossin/pnas.1920849117.sd07.csv
Kossin/pnas.1920849117.sd08.csv
```

Raw Data Size:
2102 columns
4180 rows

Clean Data Size:
10 columns
1,048,575 rows

Data Cleaning

- Similar with IBTRaACS : specify null values.
- Pivoting Data
- Add Category based on Wind Speed
- Notes : we can not identify which one causing landfall, since there is no variable indicated landfall

	StormID	Basin	Latitude1	Latitude2	Latitude3	Latitude4	Latitude5	Latitude6	Latitude7	Latitude8	...	Hour291	Hour292	Hour293	Hour294
0	1978151N15260	EP	15.30	15.30	15.30	15.60	16.20	16.50	17.76	18.00	...	NaN	NaN	NaN	NaN
1	1978168N11242	EP	11.20	11.51	12.40	12.62	13.54	14.13	15.00	15.05	...	NaN	NaN	NaN	NaN
2	1978168N14254	EP	13.71	13.71	13.71	13.69	13.41	13.30	13.17	12.52	...	NaN	NaN	NaN	NaN
3	1978173N25274	NaN	25.20	25.20	25.20	25.60	26.00	26.11	26.50	26.60	...	NaN	NaN	NaN	NaN
4	1978178N14260	EP	13.77	13.77	13.77	13.77	14.00	14.33	15.48	15.60	...	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
4175	2017333N06082	NI	6.00	6.00	6.00	6.00	5.90	6.10	6.20	6.50	...	NaN	NaN	NaN	NaN
4176	2017340N09089	NI	8.50	8.80	9.80	10.00	11.10	12.20	12.80	13.40	...	NaN	NaN	NaN	NaN
4177	2017347N11129	WP	10.90	10.90	10.90	10.90	10.90	11.25	11.39	11.24	...	NaN	NaN	NaN	NaN
4178	2017354N08134	WP	8.20	8.20	8.20	8.20	8.20	8.50	8.69	8.61	...	NaN	NaN	NaN	NaN
4179	2017360S14124	SI	-14.36	-14.36	-14.36	-14.36	-14.36	-15.00	-15.80	-16.40	...	NaN	NaN	NaN	NaN

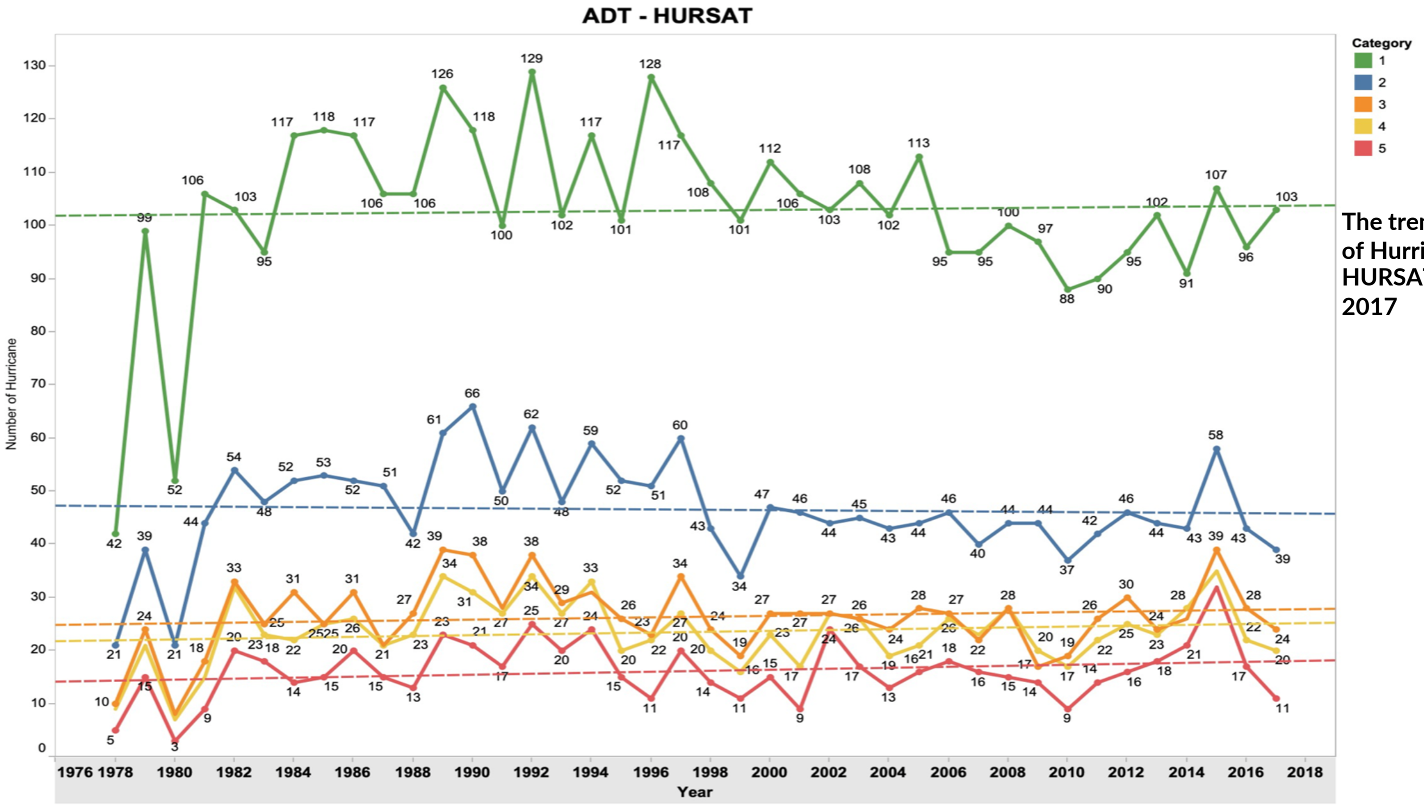
4180 rows x 2102 columns

	SID	BASIN	STORM_SPEED	LAT	LON	SEASON	MONTH	DAY	HOUR	USA_SSHS
0	1978151N15260	EP	35.0	15.30	-100.25	1978.0	5.0	30.0	0.0	1
1	1978168N11242	EP	25.0	11.20	-118.00	1978.0	6.0	17.0	0.0	1
2	1978168N14254	EP	25.0	13.71	-107.84	1978.0	6.0	17.0	0.0	1
3	1978173N25274	NaN	25.0	25.20	-86.50	1978.0	6.0	21.0	0.0	1
4	1978178N14260	EP	25.0	13.77	-100.57	1978.0	6.0	26.0	0.0	1
...	...	...	...	...	...	...	...	...	...	...
1048570	2011336S06098	SI	NaN	NaN	NaN	NaN	NaN	NaN	NaN	5
1048571	2011337S13069	SI	NaN	NaN	NaN	NaN	NaN	NaN	NaN	5
1048572	2011338N06114	WP	NaN	NaN	NaN	NaN	NaN	NaN	NaN	5
1048573	2011344N12117	WP	NaN	NaN	NaN	NaN	NaN	NaN	NaN	5
1048574	2011346N03156	WP	NaN	NaN	NaN	NaN	NaN	NaN	NaN	5

1048575 rows x 10 columns

https://www.ncei.noaa.gov/data/international-best-track-archive-for-climate-stewardship-ibtracs/v04r00/access/csv/ ibtracs.ALL.list.v04r00.csv

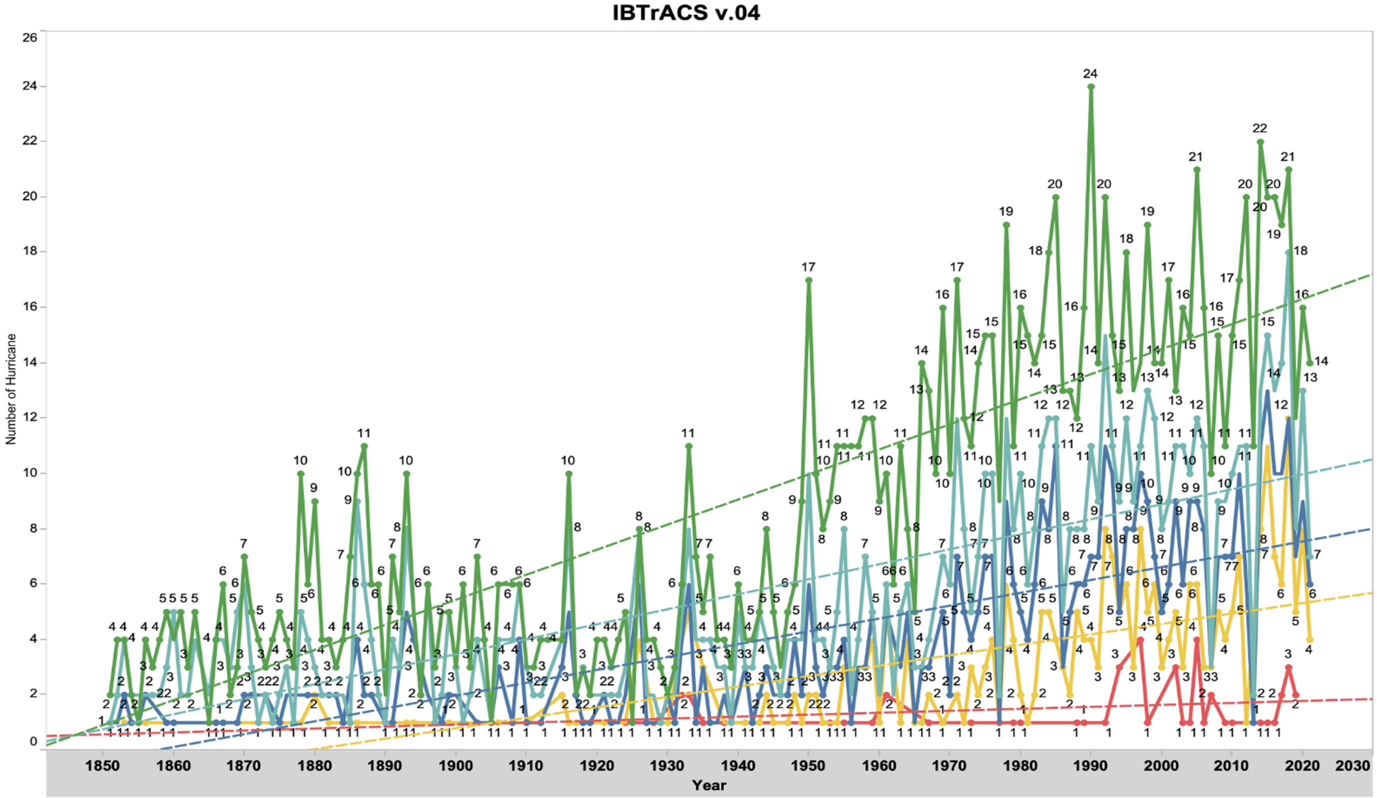
Initial Data Analysis - ADT-HURSAT



The trend of distinct count of SID (Hurricane ID) using ADT- HURSAT data for Year. Color shows details about Category. The view is filtered on Category, which keeps 1, 2, 3, 4 and 5.

The trend based on distinct count of Hurricane ID using ADT-HURSAT - global data of 1978-2017

Initial Data Analysis - IBTrACS



The trend of distinct count of SID (Hurricane ID) using (IBTrACS v04 for Season. Color shows details about Usa Sshs. The data is filtered on Landfall, which keeps 2,546 of 4,604 members. The view is filtered on Usa Sshs, which keeps 1, 2, 3, 4 and 5.

The trend based on distinct count of Hurricane ID using IBTrACS v04 - global data of 1842-2021

Data - Real Estate

Data Source

- Pulled using S&P proprietary tool, S&P Capital IQ Pro[1]
- Property-level data pulled from Asset data feature for Real Estate Properties
- Scope: US-based publicly-traded REITs - currently owned and sold properties

Key Data Attributes

- Years: 1995-2020
- Property locations (longitude and latitude)
- Total company assets
- Asset write down
- We will use property-level data to identify hurricane encounters, but will look at company-level financial data (since this is often not available on a property level)

Source:
1. S&P Capital IQ Pro (https://www.capitaliq.spglobal.com/web/client?auth=inherit#dashboard)

Initial Data Cleaning - Real Estate

Data Cleaning Methods

- First, we want to be able to track asset values yearly - so we reformat that data to be tall instead of wide and put it into its own dataset
- Then, we want to see the locations of all our assets - so we reformat that data into its own dataset as well
- In both datasets, we keep the property ID so we can later use this to link our datasets together

After cleaning, here's the shape of our assets dataset

```
df_assets_location.shape

(53230, 6)
```

Here are what the heads of those datasets look like

	pptyKey	year	writedown
0	7283	2020	0.0
1	7283	2019	0.0
2	7283	2018	0.0
3	7283	2017	0.0
4	7283	2016	0.0

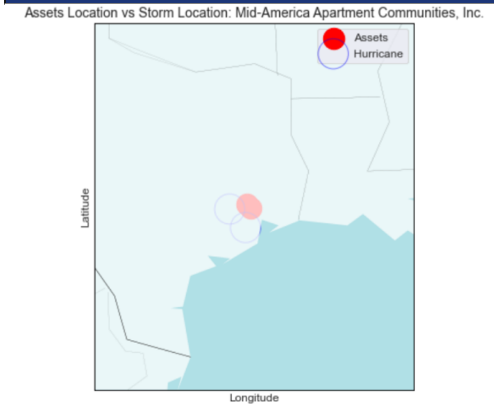
```
df_assets_location.head()
```

	InstName	pptyKey	pptyName	lat	long	reitStatus
0	Acadia Realty Trust	7283	Crescent Plaza	42.07879	-70.99062	Yes
1	Acadia Realty Trust	7292	Mark Plaza	41.25680	-75.90829	Yes
2	Acadia Realty Trust	7297	New Loudon Center	42.75452	-73.75660	Yes
3	Acadia Realty Trust	7305	Plaza 422	40.34024	-76.39804	Yes
4	Acadia Realty Trust	21685	Route 6 Mall	41.55060	-75.22972	Yes

Source: S&P Capital IQ Pro (https://www.capitaliq.spglobal.com/web/client?auth=inherit#dashboard)

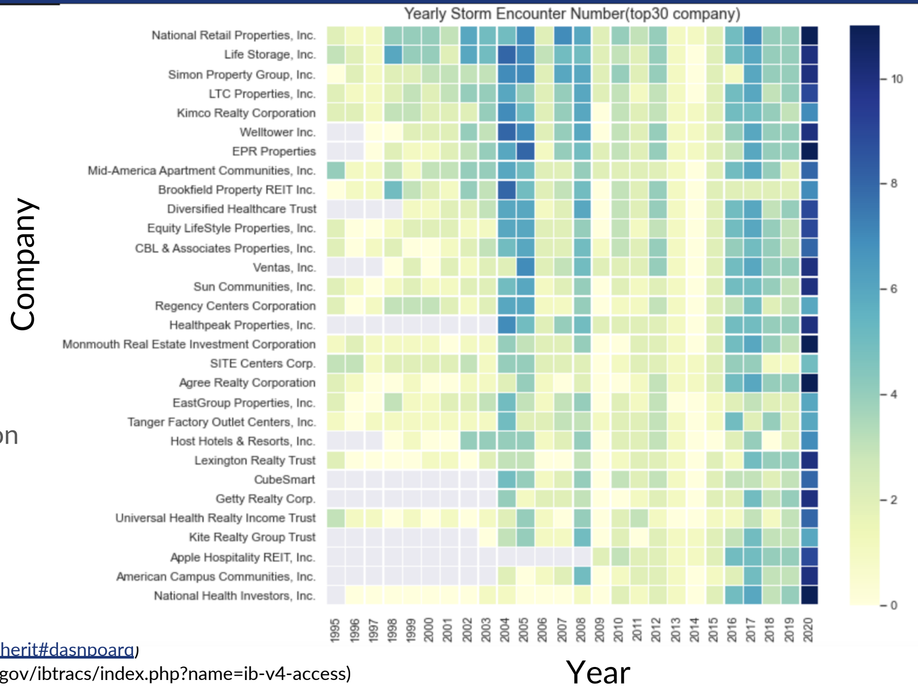
Joining Real Estate & Property Data Together

Joining Methods



- Connected via latitudes and longitudes of property location and hurricane landfall
- Looked at hurricanes with landfall within a **50 mile radius**

Top 30 Companies with the Most Storm Encounters

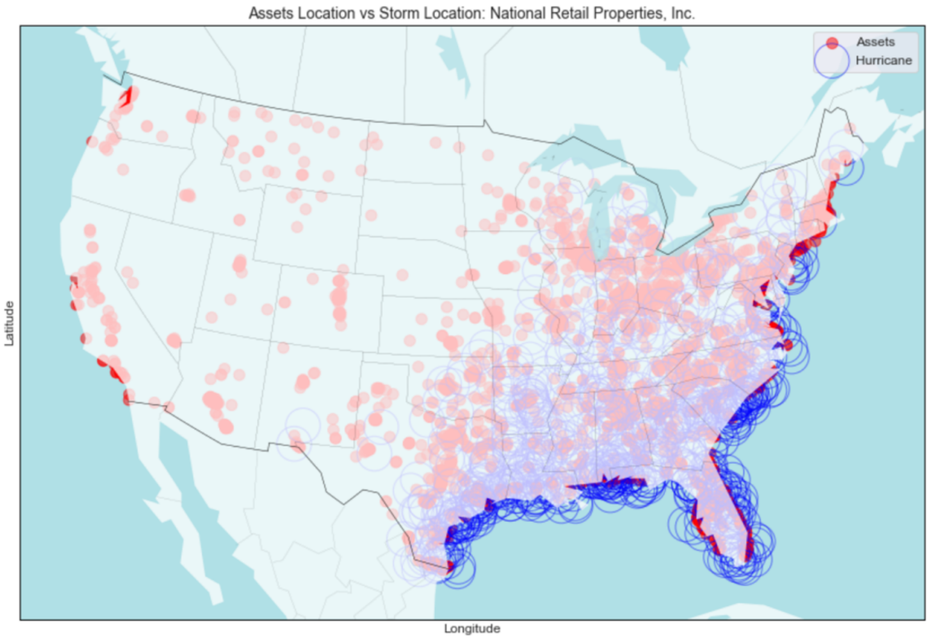


Analysis done on data from:
1. S&P Capital IQ Pro for company data (https://www.capitaliq.spglobal.com/web/client?auth=inherit#dasnpoarg)
2. National Oceanic and Atmospheric Administration for hurricane data (https://www.ncdc.noaa.gov/ibtracs/index.php?name=ib-v4-access)
3. Proceedings of the National Academy of Sciences of the United States

Inspecting the Data - An Example of One Company

Data Inspection Methods

- Using our asset & hurricane datasets connected by latitude and longitude, we plotted the assets of that company and the storms that affected them
- We did this for the top five companies hit by the most storms
- The other four companies' plots are available in the appendix



Year 1995 - 2020

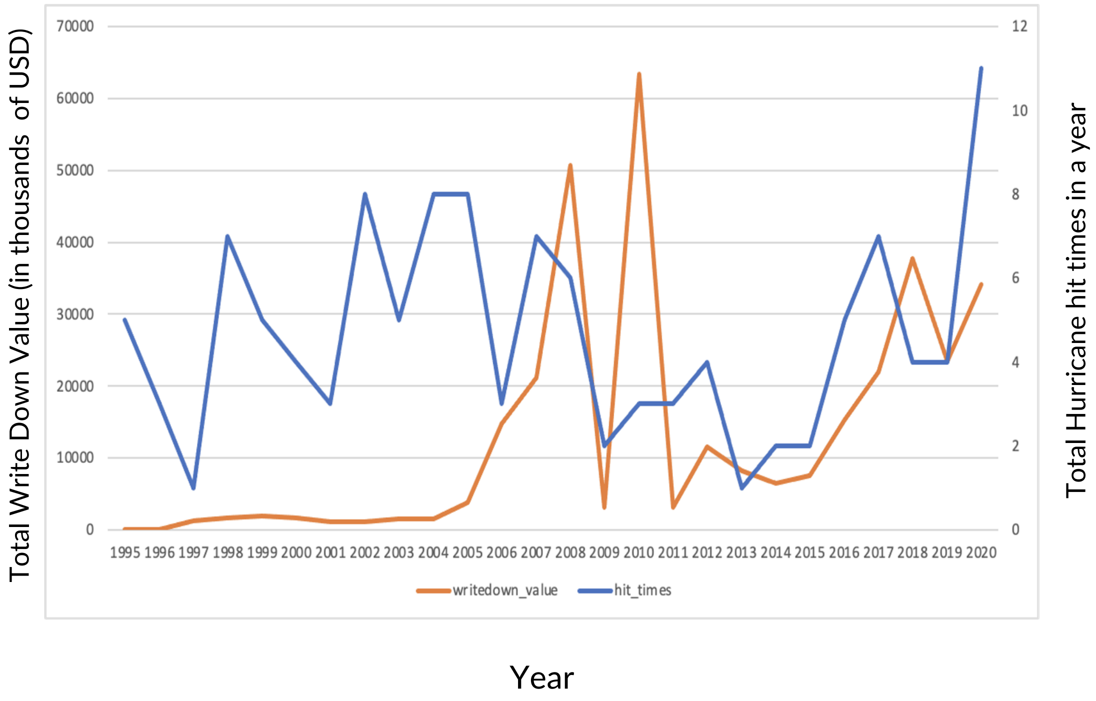
Analysis done on data from:
1. S&P Capital IQ Pro for company data (https://www.capitaliq.spglobal.com/web/client?auth=inherit#dashboard)
2. National Oceanic and Atmospheric Administration for hurricane data (https://www.ncdc.noaa.gov/ibtracs/index.php?name=ib-v4-access)
3. Proceedings of the National Academy of Sciences of the United States (https://www.pnas.org/content/suppl/2020/05/12/1920849117.DCSupplemental)

Inspecting the Data - An Example of One Company

Data Inspection Methods

- Graphing how company assets write down value changed over time for our selected companies with the most hurricane hits - only one of those graphs is shown here
- Next step will be to incorporate this data into modeling to see how it is changing in correlation with hurricane hits (and the strength of those hits)

Total Asset Write-down Value & Total Hurricane Hit times by Year: Macerich Company



Analysis done on data from:
1. S&P Capital IQ Pro for company data (https://www.capitaliq.spglobal.com/web/client?auth=inherit#dashboard)
2. National Oceanic and Atmospheric Administration for hurricane data (https://www.ncdc.noaa.gov/ibtracs/index.php?name=ib-v4-access)

Dataset after cleaning

- InstName: name of the institution
- year: year
- writedown_value: write down value of real estate that year
- inflation_rate: US inflation rate that year
- asset_value: total asset value of the company
- number_of_storm: number of storm encountered that year
- storm_experience_hour: duration of storm
- max_storm_speed: maximum value of storm wind speed
- writedown_pct: The proportion of write down value over asset value

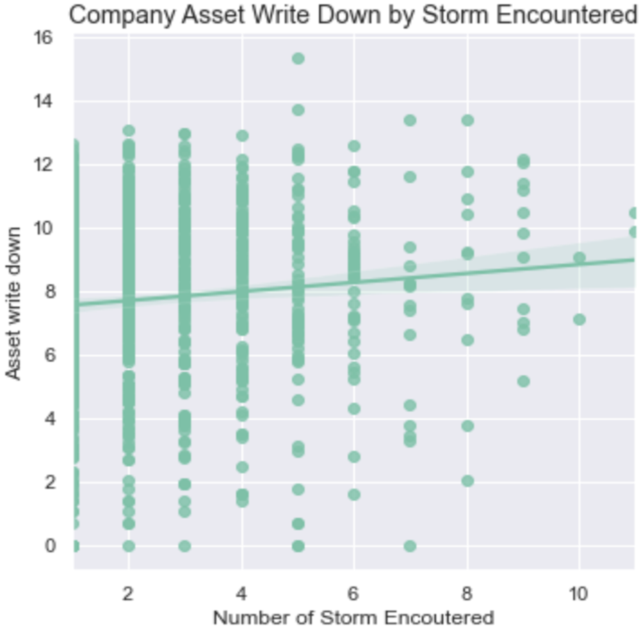
	InstName	year	writedown_value	inflation_rate	asset_value	number_of_storm	storm_experience_hour	max_storm_speed	writedown_pct
0	Acadia Realty Trust	1995	0.0	2.17	35468	0	0	0.0	0.000000
1	Agree Realty Corporation	1995	0.0	2.17	12500	2	2	0.0	0.000000
2	Alexander's, Inc.	1995	0.0	2.17	45094	0	0	0.0	0.000000
3	Apartment Income REIT Corp.	1995	0.0	2.17	18935	0	0	0.0	0.000000
4	AvalonBay Communities, Inc.	1995	810.0	2.17	167279	0	0	0.0	0.004842

Analysis done on data from:
1. S&P Capital IQ Pro for company data (https://www.capitaliq.spglobal.com/web/client?auth=inherit#dashboard)
2. National Oceanic and Atmospheric Administration for hurricane data (https://www.ncdc.noaa.gov/ibtracs/index.php?name=ib-v4-access)

All the companies which are in this dataset have been impacted by the hurricane

Inspecting Data - All Affected Companies

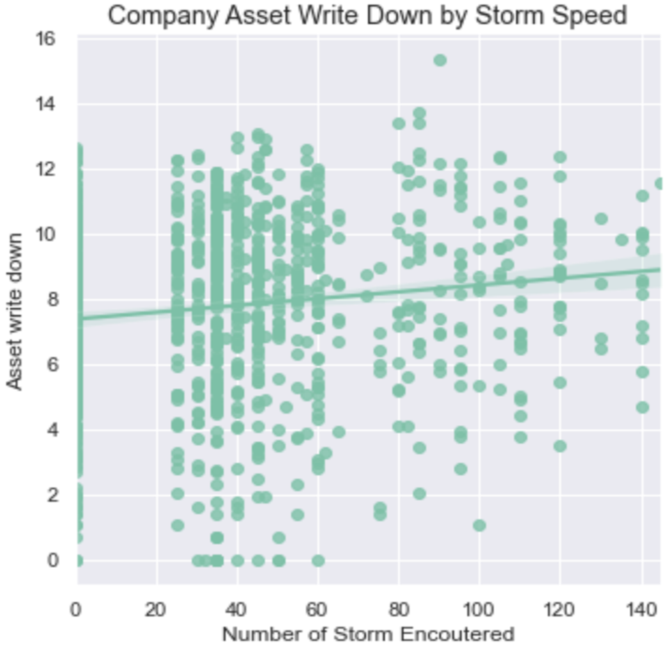
Asset Write Down VS. # of storm Encountered



210 Firms

- Affected by storms
- With writedown_value
- With max_speed

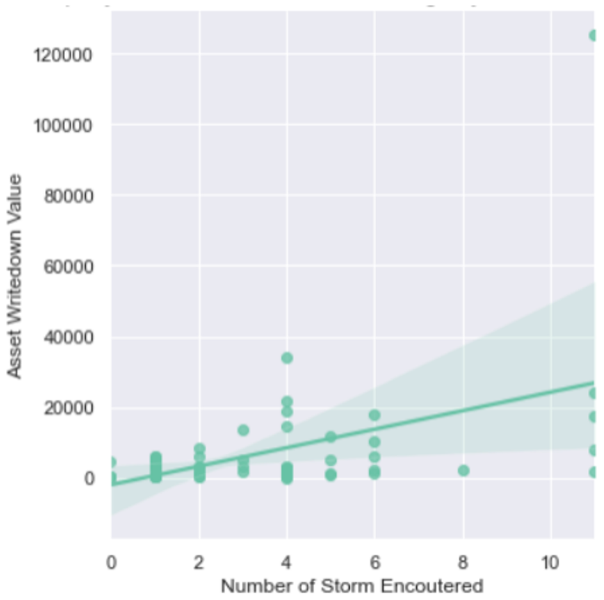
Asset Write Down VS. Storm speeds



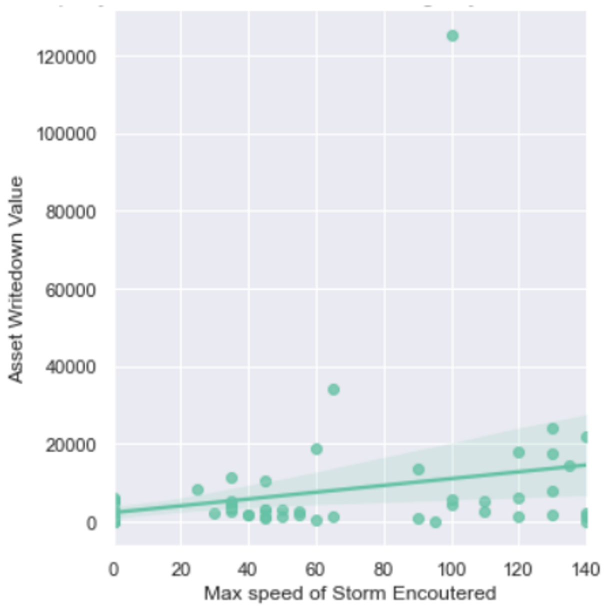
Analysis done on data from:
1. S&P Capital IQ Pro for company data (https://www.capitaliq.spglobal.com/web/client?auth=inherit#dashboard)
2. National Oceanic and Atmospheric Administration for hurricane data (https://www.ncdc.noaa.gov/ibtracs/index.php?name=ib-v4-access)

Inspecting Data - Top 5 firms

Asset Write Down VS. # of storm Encountered



Asset Write Down VS. Storm speeds



Analysis done on data from:
1. S&P Capital IQ Pro for company data (https://www.capitaliq.spglobal.com/web/client?auth=inherit#dashboard)
2. National Oceanic and Atmospheric Administration for hurricane data (https://www.ncdc.noaa.gov/ibtracs/index.php?name=ib-v4-access)

Data Preparation for Regression Modeling

- 1. Delete the records that the number of storms is 0 that year;
- 2. Delete the records that the write down value is 0 that year;
- 3. Logarithmic conversion writedown_pct and writedown_value to reduce skew distribution;
- 4. Independent variables: inflation_rate, number_of_storm, max_storm_speed, storm_experience_hour;
- 5. Dependent variables: writedown_value or writedown_pct.

	Unnamed: 0	InstName	year	writedown_value	inflation_rate	asset_value	number_of_storm	storm_experience_hour	max_storm_speed	writedown_pct
45	45	Brookfield Property REIT Inc.	1996	4.875197	2.17	628887	1	1	0.0	-8.476510
52	52	Federal Realty Investment Trust	1996	7.579679	2.17	684802	2	14	0.0	-5.857206
60	60	Macerich Company	1996	1.945910	2.17	414589	1	1	0.0	-10.989133
67	67	Pennsylvania Real Estate Investment Trust	1996	6.639876	2.17	98970	1	1	0.0	-4.862696

Linear Model

Asset Write Down VS. # of storm Encountered （210 Companies selected）

Model:	OLS	Adj. R-squared:	0.027
Dependent Variable:	writedown_value	AIC:	4828.1388
Date:	2021-12-07 10:19	BIC:	4852.6222
No. Observations:	989	Log-Likelihood:	-2409.1
Df Model:	4	F-statistic:	7.954
Df Residuals:	984	Prob (F-statistic):	2.60e-06
R-squared:	0.031	Scale:	7.6823

	Coef.	Std.Err.	t	P>|t|	[0.025	0.975]
const	7.9946	0.2517	31.7607	0.0000	7.5006	8.4885
inflation_rate	-0.3173	0.0955	-3.3228	0.0009	-0.5047	-0.1299
number_of_storm	-0.0156	0.0646	-0.2417	0.8090	-0.1423	0.1111
storm_experience_hour	0.0055	0.0030	1.8234	0.0685	-0.0004	0.0114
max_storm_speed	0.0086	0.0032	2.6953	0.0072	0.0023	0.0149

Omnibus:	56.216	Durbin-Watson:	1.887
Prob(Omnibus):	0.000	Jarque-Bera (JB):	64.847
Skew:	-0.626	Prob(JB):	0.000
Kurtosis:	3.085	Condition No.:	187

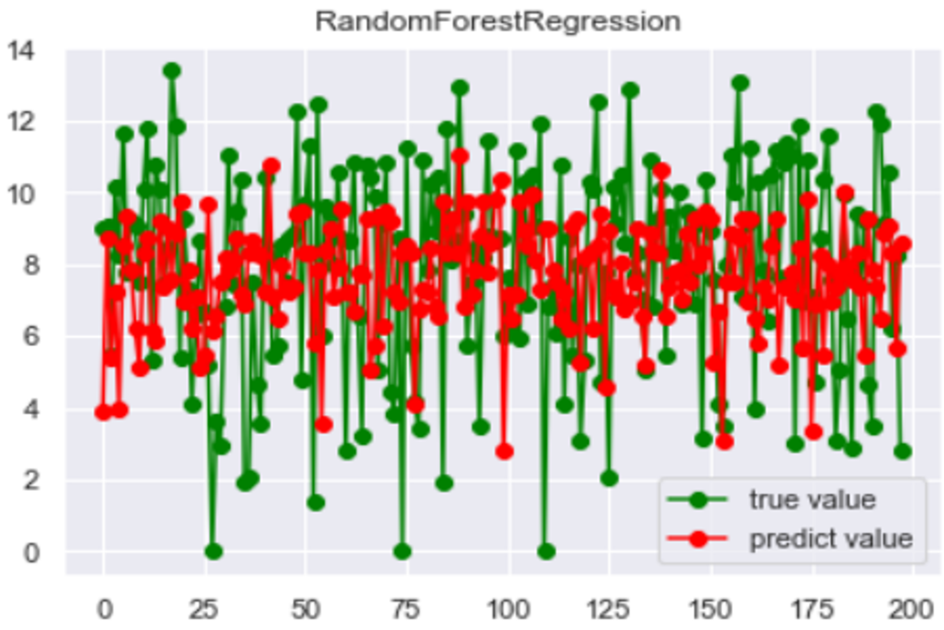
The p value of max_storm_speed is lower than 0.05, which means it is significant.

And the coefficient of it is 0.0086, which is positive. So it means that the larger the max storm speed is, the higher the write down value of the companies will be which means the companies lost more in hurricane.

However, r2 was 0.031, it means that the model was not good on predicting the results. It may not be accurate enough.

Analysis done on data from:
1. S&P Capital IQ Pro for company data (https://www.capitaliq.spglobal.com/web/client?auth=inherit#dashboard)
2. National Oceanic and Atmospheric Administration for hurricane data (https://www.ncdc.noaa.gov/ibtracs/index.php?name=ib-v4-access)

RandomForestRegressor



MAPE means Mean Absolute Percentage Error, which can somehow show the accuracy of the regression model

MAPE:41.11%

Analysis done on data from:
1. S&P Capital IQ Pro for company data (https://www.capitaliq.spglobal.com/web/client?auth=inherit#dashboard)
2. National Oceanic and Atmospheric Administration for hurricane data (https://www.ncdc.noaa.gov/ibtracs/index.php?name=ib-v4-access)

Data Preparation for Classification Modeling

- Add a new column to represent whether the company had write down value recorded or not;
- The value can be 0 if the company didn't have write down value;
- The value can be 1 if the company had write down value;
- Independent variables: inflation_rate, number_of_storm, max_storm_speed, storm_experience_hour;
- Dependent variables: whether_writedown.

InstName	year	writedown_value	inflation_rate	asset_value	number_of_storm	storm_experience_hour	max_storm_speed	writedown_pct	whether_writedown
Agree Realty Corporation	1995	0	2.17	12500	2	2	0	0.0	0
CBL & Associates Properties, Inc.	1995	0	2.17	304117	2	5	0	0.0	0
Camden Property Trust	1995	0	2.17	57839	1	1	30	0.0	0
EastGroup Properties, Inc.	1995	0	2.17	40307	2	6	0	0.0	0

Logistic Regression

Asset Write Down(0 or 1) VS. # of storm Encountered
（210 Companies selected）

```
The accuracy of logistic regression is 0.5729243786356426
(611, 1)
              precision    recall  f1-score   support

           0       0.54      0.72      0.62       275
           1       0.68      0.50      0.58       336

    accuracy                           0.60       611
   macro avg       0.61      0.61      0.60       611
weighted avg       0.62      0.60      0.59       611
```

- **Accuracy:**
- **The correctly predicted samples / all test samples**

Analysis done on data from:
1. S&P Capital IQ Pro for company data (https://www.capitaliq.spglobal.com/web/client?auth=inherit#dashboard)
2. National Oceanic and Atmospheric Administration for hurricane data (https://www.ncdc.noaa.gov/ibtracs/index.php?name=ib-v4-access)

Random Forest Classifier

Asset Write Down(0 or 1) VS. # of storm Encountered
（210 Companies selected）

The accuracy of random forest classifier is 0.5990180032733224

	precision	recall	f1-score	support
0	0.49	0.56	0.53	242
1	0.68	0.62	0.65	369
accuracy			0.60	611
macro avg	0.59	0.59	0.59	611
weighted avg	0.61	0.60	0.60	611

Analysis done on data from:
1. S&P Capital IQ Pro for company data (https://www.capitaliq.spglobal.com/web/client?auth=inherit#dashboard)
2. National Oceanic and Atmospheric Administration for hurricane data (https://www.ncdc.noaa.gov/ibtracs/index.php?name=ib-v4-access)

Areas for Further Exploration

- Given additional time we would explore the following areas
  - Extend historical data for both hurricanes and real-estate assets
  - Systematic analysis of all company statements and earnings calls for firms that experienced hurricanes
  - Conduct comparable studies in adjacent industries
    - Insurance

Lessons Learned

Assigning roles to team members kept us organized and accountable.

Refining our initial three research questions into **one question** kept us focused on the specific business need.

Through iterative presentations and feedback, we learned what visuals and presentations styles best conveyed our message:

- We should **show our complete process**, including issues, rather than only showing what worked well
- We should **start with simple examples** and then extrapolate to larger populations
- We should ensure all **sources were well-cited** and our rationale for using them explained

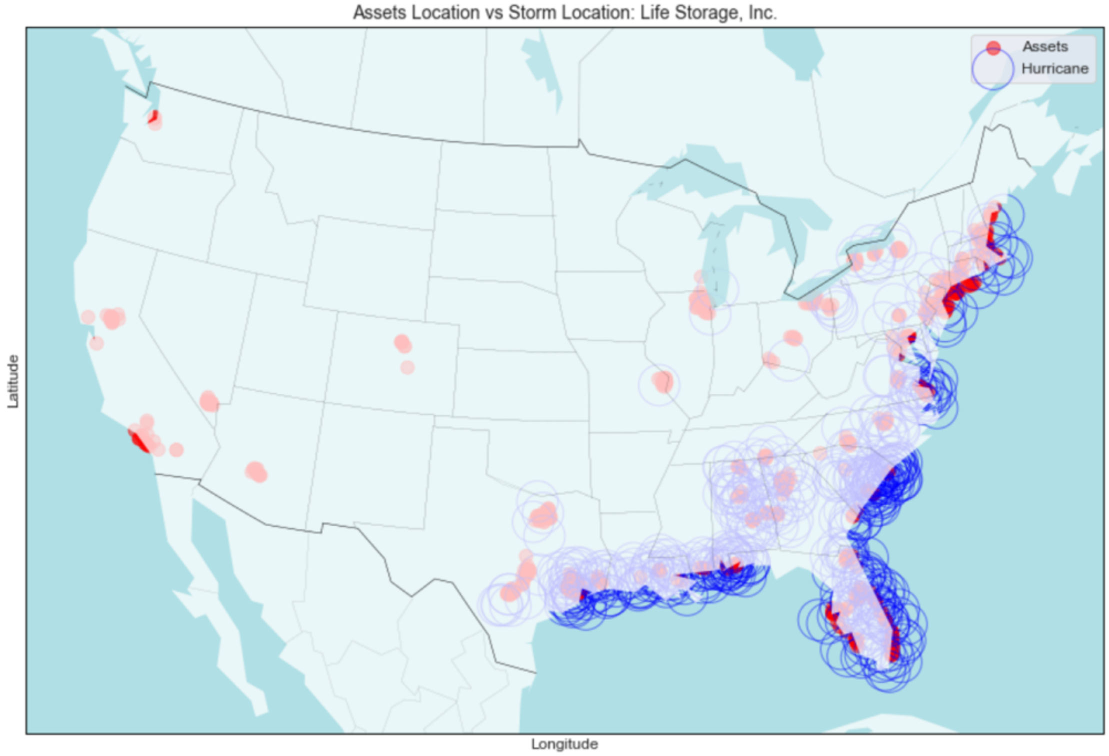
Using **business logic** was very helpful in determining which **features** to use in our model. For example, asset write down was more closely related to hurricanes than total asset value.

Some models were **computationally expensive**, and we had to consider computing power when building our models.

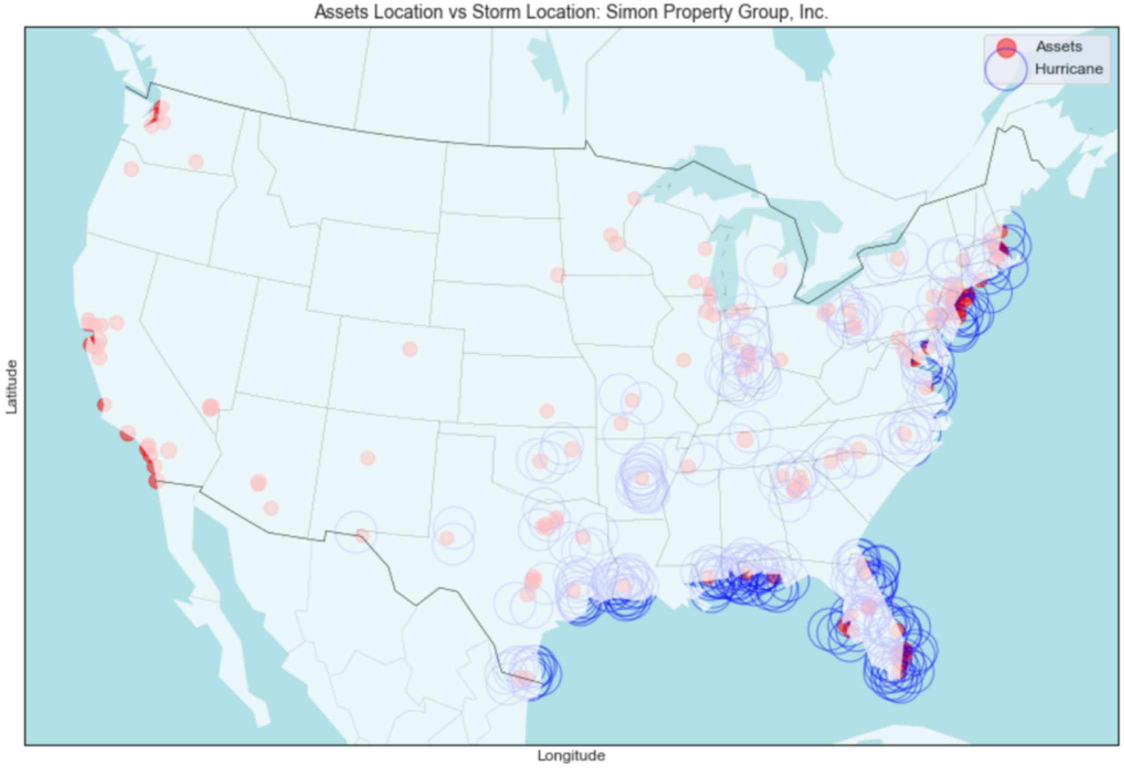
Questions?

Appendix

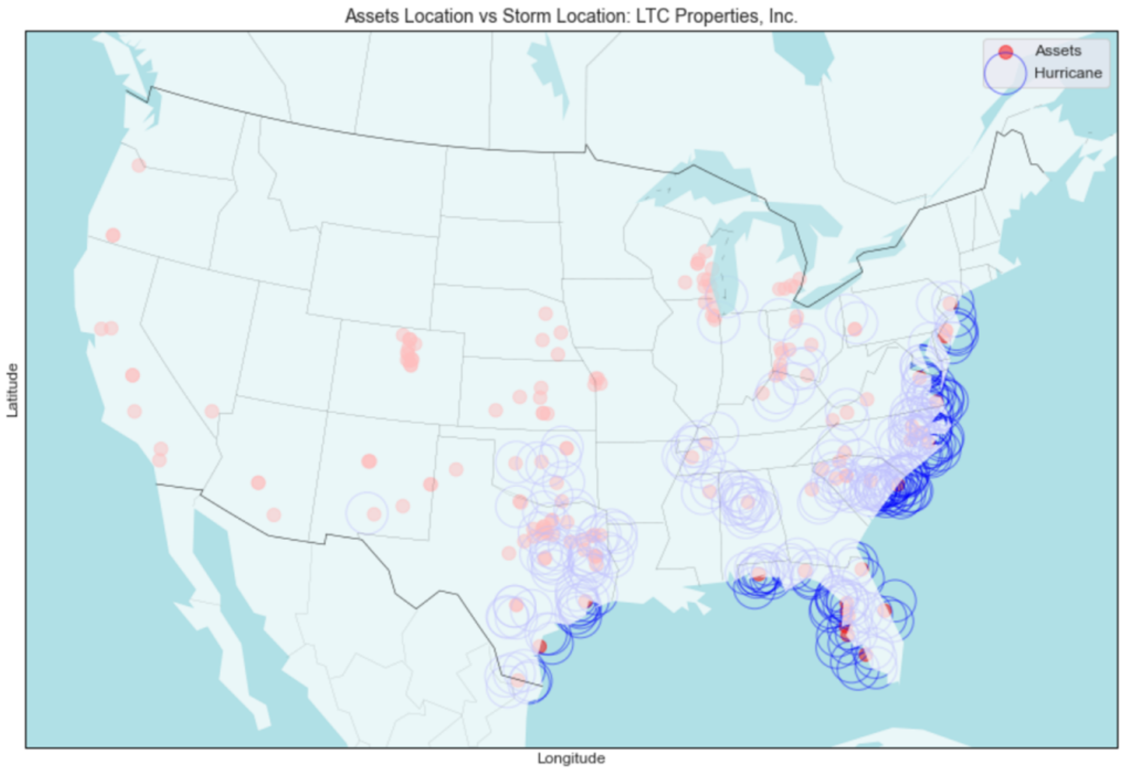
Inspecting the Data - Mapping Example #2



Inspecting the Data - Mapping Example #3



Inspecting the Data - Mapping Example #4



Inspecting the Data - Mapping Example #5