PIERCING THE SKY RESPONSIBLY:

Predicting NYC Upcoming Skyscrapers' Compliance with 2030 Carbon Goals

PIERCING THE SKY RESPONSIBLY:

Predicting NYC Upcoming Skyscrapers' Compliance with 2030 Carbon Goals

**New York City
Law Department**

Local Law 97

**NYC local law 97, passed in April 2019, aims to reduce carbon emissions by 40% in 2030**

- Group A (Residential): 0.00675 kgCO2e/ft²
- Group B (Business): 0.00846 kgCO2e/ft²
- Group E (Educational): 0.00574 kgCO2e/ft²
- Group I-1 and I-2 (Institutional): 0.00987 kgCO2e/ft²
- Group M (Mercantile): 0.01181 kgCO2e/ft²
- Group R-1 (Residential Hotels): 0.00987 kgCO2e/ft²
- Group S (Storage): 0.02003 kgCO2e/ft²

Over 70 percent of New York's greenhouse gas emissions come from buildings
- NYC Mayor's Office of Climate and Environmental Justice

We know that most buildings in Manhattan are high-rise…

- Do they meet this goal?
- What types of buildings produce most greenhouse gas emissions?
- How could future NYC skyscrapers work towards this goal?
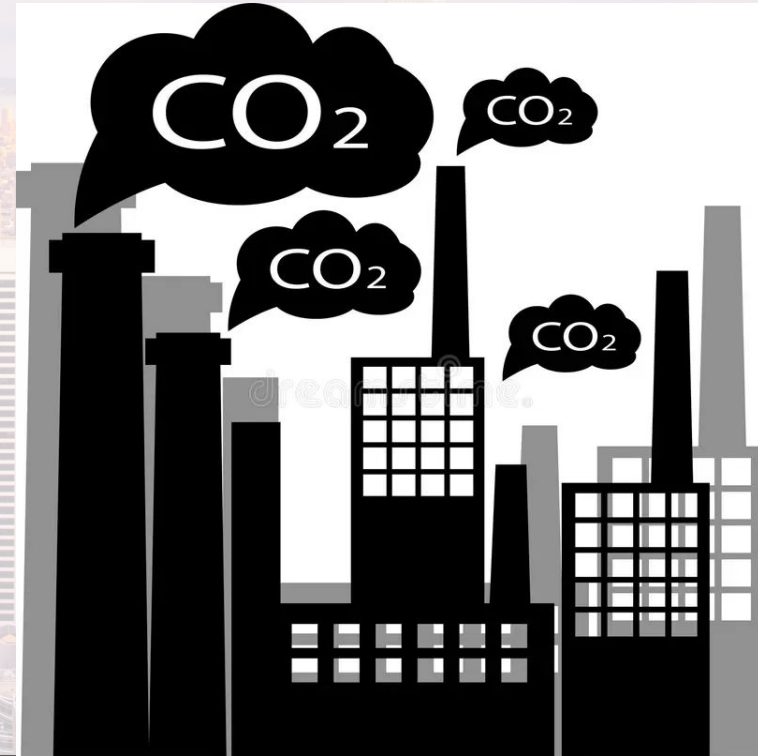
Data Preparation

Data Exploration
- Loading the dataset & understanding basic information about it

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 29842 entries, 0 to 29841
Data columns (total 59 columns):
```

Data Cleaning
- Replacing "Not Available" / "Not Applicable" columns with Nan
- Selecting the most relevant columns through manual scrapping
- Removing rows that had >= 10 missing data columns

Imputation to fill in missing values
- KNN imputation to fill missing values given our assumption that GHG emissions of a certain property would be similar to other properties that are nearby

Exploratory Data Analysis: Total GHG Emissions by Borough

Distribution of Total GHG Emissions (Metric Tons CO2e) by Borough
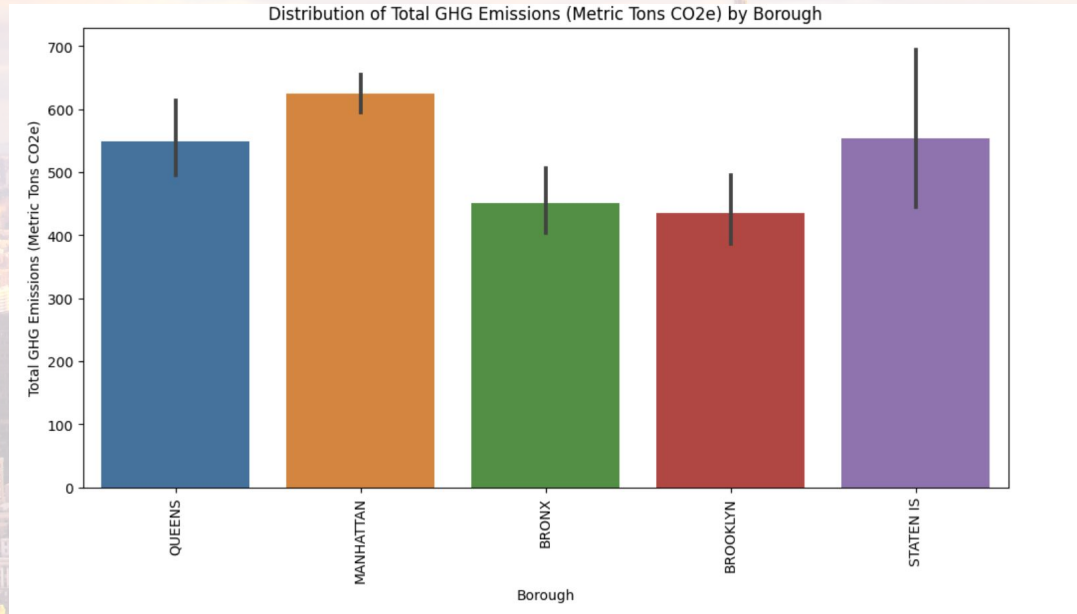
Manhattan (59 km^2), followed by Staten Island (152 km^2) and Queens (280 km^2)

Since Manhattan produces the most GHG Emissions Intensity, we will focus our analysis on Manhattan.

Exploratory Data Analysis: ENERGY STAR Score and GHG Emissions

Correlation between STAR score and total GHG emissions intensity
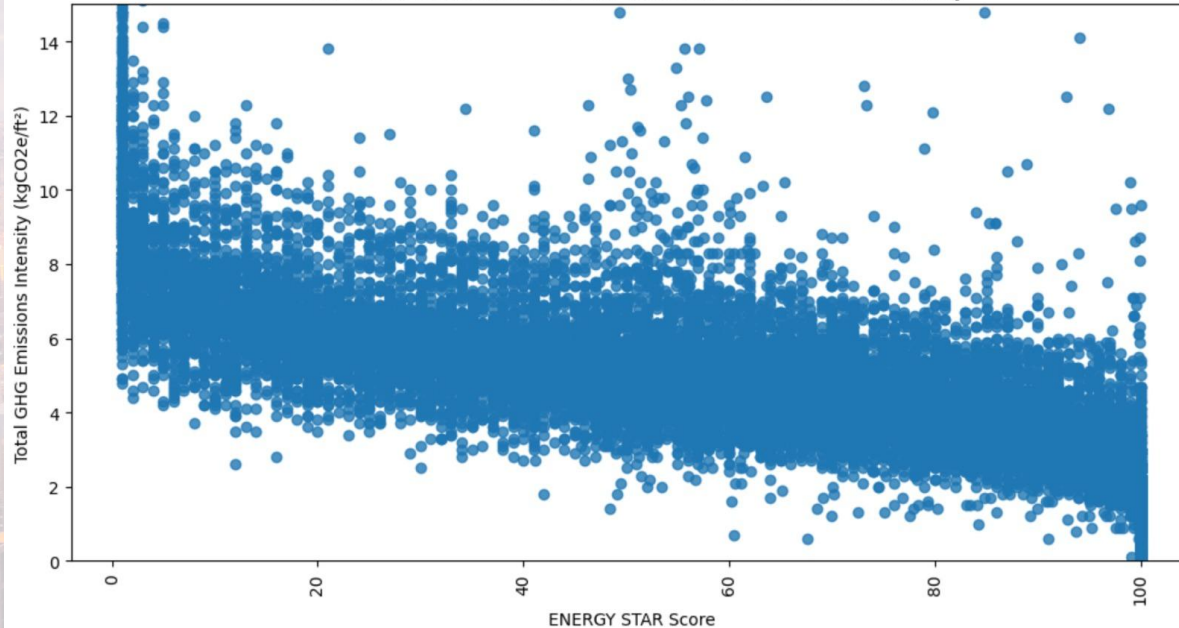
As we can see, Total GHG Emissions intensity and ENERGY STAR Score are negatively correlated

Exploratory Data Analysis: ENERGY STAR Score across building types

Distribution of ENERGY STAR Scores by Property Type
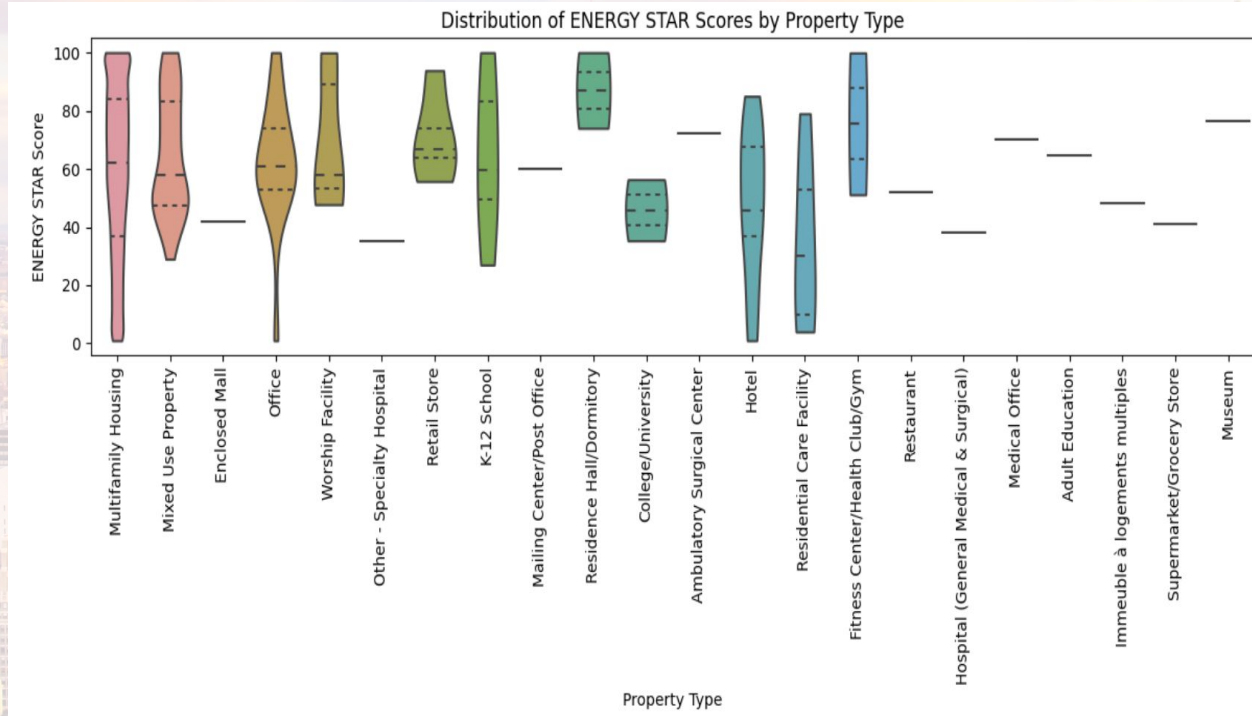
Multifamily housing has lowest star scores alongside hotels and residential care facilities.

Exploratory Data Analysis: ENERGY STAR Score and EUI

- **site EUI:** energy use intensity used on-sit at a building

- **source UI:** total amount of energy used to operate building, both on-site and off

- The difference between site and source EUI gives an indication of the amount of energy lost during transmission

Distribution of Absolute_DIFF by ENERGY STAR Score
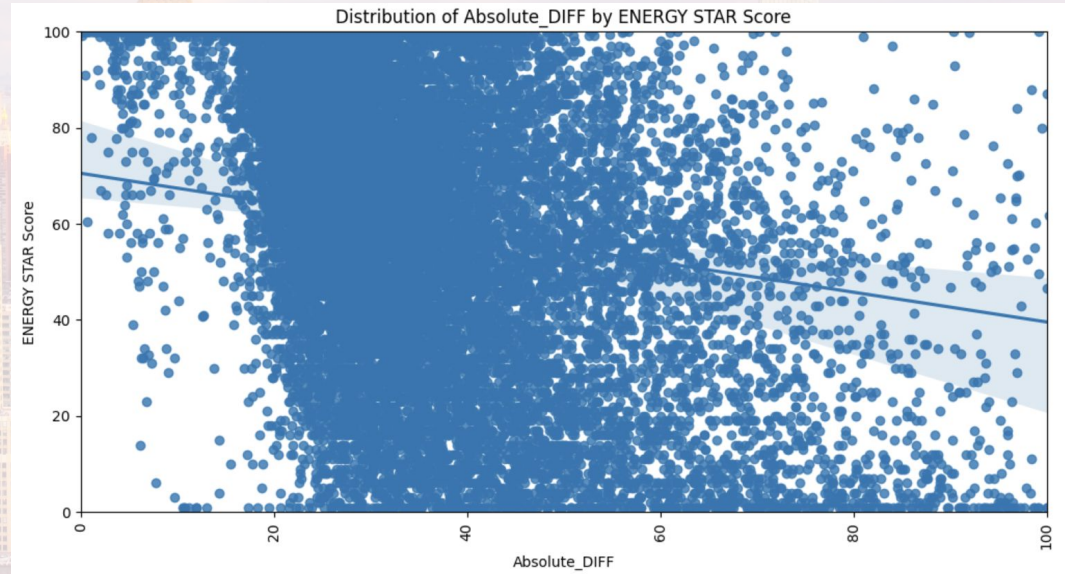
As we can see, multifamily housing produces most GHG emissions

Distribution of Total GHG Emissions Intensity (kgCO2e/ft²) by Primary Property Type - Self Selected
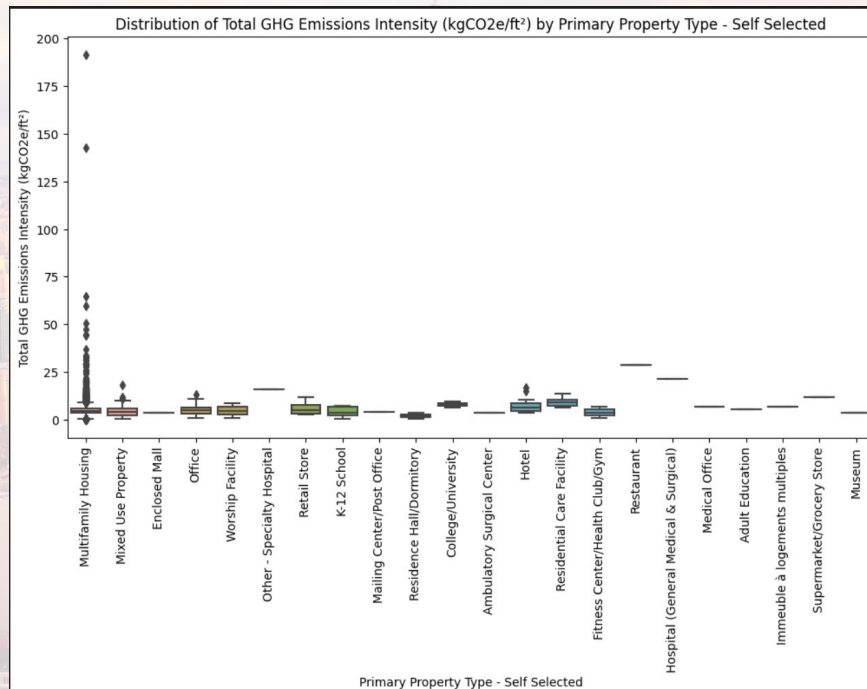
Understanding whether buildings are meeting the GHG emission limits based on property type.

```
fail
True        17806
False           2
Name: count, dtype: int64
                            Property Name Primary Property Type – Self Selected
6779    2022 Copy of Clearview 6 Bldg A–1                   Multifamily Housing
16854           708–716 OCEAN COMMUNITY CORP               Multifamily Housing
```

```
        Property Id                       Property Name  \
6779       21322144  2022 Copy of Clearview 6 Bldg A–1
16854       6669870         708–716 OCEAN COMMUNITY CORP

            NYC Building Identification Number (BIN)       City  \
6779    4443389;4443384;4443385;4443388;4458409;444338...   Queens
16854                                       3327291  BROOKLYN

        Postal Code Primary Property Type – Self Selected  \
6779         11357                 Multifamily Housing
16854        11226                 Multifamily Housing
```

Only 2 of the buildings fall within the Local Law 97 regulations

A Look Into These Residential Buildings

NYC Department of Buildings
Property Profile Overview

158-18 CROSS ISLAND PARKWAY SR SOUTH   **QUEENS 11357**   **BIN# 4443389**

CROSS ISLAND PARKWAY   158-18 - 158-18

Health Area : 222	**Tax Block** : 4742
Census Tract : 1029	**Tax Lot** : 22
Community Board : 407	**Condo** : NO
Buildings on Lot : 7	**Vacant** : NO

View DCP Addresses...   **Browse Block**

Low-Rise Residential Building in Queens

A Look Into These Residential Buildings

Multi-story Low-Rise Building with a Smaller Floor Area

NYC Department of Buildings

Property Profile Overview

706 OCEAN AVENUE		BROOKLYN 11226		BIN# 3327291
OCEAN AVENUE	706 - 712	Health Area	: 7210	Tax Block	: 5123
		Census Tract	: 510.01	Tax Lot	: 31
		Community Board	: 314	Condo	: NO
		Buildings on Lot	: 2	Vacant	: NO

View DCP Addresses...   **Browse Block**

706 ocean avenue

708 Ocean
New York

Google

Nov 2022

NYC Department of Buildings

DOB Violation Display for 010112BENCH02002

Premises: 706 OCEAN AVENUE BROOKLYN		BIN: 3327291	Block: 5123	Lot: 31

Issue Date:	01/01/2012	Violation Category:	V - DOB VIOLATION - ACTIVE
Violation Type:	BENCH - FAILURE TO BENCHMARK		
Violation Number:	02002	Device No.:	
OATH/ECB No.:			
Infraction Codes:			
Description:	FAILURE TO FILE BENCHMARKING REPORT OF ENERGY USE AS PER AD. CODE SEC. 28-309.4		

Disposition:

Code:		Date:	
Inspector:			
Comments:			
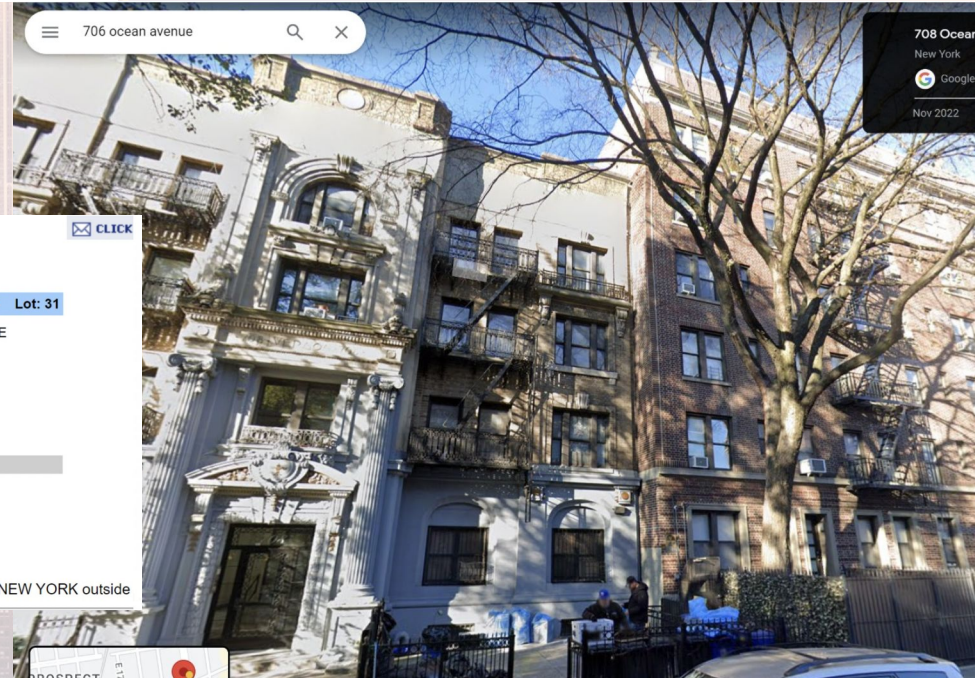
ns please review these Frequently Asked Questions, the Glossary, or call the 311 Citizen Service Center by dialing 311 or (212) NEW YORK outside

To predict what factors lead to higher GHG emissions amongst multifamily housing, we made use of multiple regression models.

```
Cross-Validation Scores (Without Scaling):
KNN: 0.654218562676791
OLS: 0.6773510175383961
Ridge: 0.6773524705140357
Lasso: 0.6775105091230527

Cross-Validation Scores (With Scaling):
KNN: 0.65483416417694
OLS: 0.6773510175383741
Ridge: 0.6773723225776165
Lasso: 0.677570670035135
Best Model Overall: Lasso
Best Model Without Scaling: Lasso
Best Model With Scaling: Lasso
```

```
Best Tuned KNN Mean CV Score: 0.6686673650029825
Best Tuned Ridge Mean CV Score: 0.6775631141633838
Best Tuned Lasso Mean CV Score: 0.6794305322559795
Best Overall Tuned Model: Lasso with Mean CV Score of 0.6794305322559795
```

We made use of 4 regression models - KNN, OLS, Ridge and Lasso - and tested each model's accuracy with scaling and tuning. Here, X consists of variables like gross floor area, bedroom density, building height while Y is the gross GHG emissions. Our best model achieved **~68% accuracy**.

To predict what factors lead to higher GHG emissions amongst multifamily housing, we made use of logistic regression model.

```
Accuracy: 0.6209994385176867
Confusion Matrix:
[[2212    0]
 [1350    0]]
Classification Report:
              precision    recall  f1-score   support

           0       0.62      1.00      0.77      2212
           1       0.00      0.00      0.00      1350

    accuracy                           0.62      3562
   macro avg       0.31      0.50      0.38      3562
weighted avg       0.39      0.62      0.48      3562
```

We also made use of a logistic regression, where X consists of variables like gross floor area, bedroom density, building height while Y is a binary variable - whether gross GHG is above or below the national median. We achieved an accuracy score of 72%

262 FIFTH AVENUE

- 305 meter residential skyscraper (by 2024)

Based on current energy systems and consumption, future skyscrapers will not meet these goals by 2030, unless a radical change is made

Perhaps, the best way to ensure that is to increase the penalty, currently add 268$ per ton of CO2 over the limit