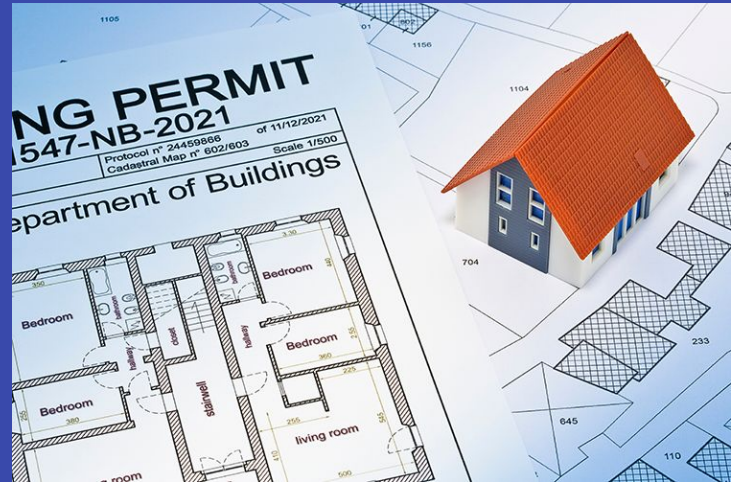


Unraveling Vancouver's Housing Challenge through Building Permits Analysis

Tatyana Mozgacheva
Data Science, 2024, Brainstation



The problem

Overview

Vancouver faces a major housing issue with high property prices and a shortage of affordable homes.

Problem statement

This project aims to reveal how the permit process impacts housing shortages.

Users

- City planners
- Officials
- Applicants
- Developers

The Big Ideas

Predicting Permit Processing Times

- Analyze past building permits and estimate processing times



Spatial Analysis

- Map building permits to see where construction is happening in the city.
- Identify busy development areas and spots that need more focus.

Potential Impact

This project helps Vancouver residents by analyzing data to address a critical housing issue. The insights gained can inform policymakers and urban planners, potentially leading to more accessible housing solutions.



EDA

Main Question of interest:

Does the time it takes to obtain a permit relate to factors like property values, type of work, property use, and location, etc?

Dataset Description

Building permits are required for new buildings, additions or alterations to existing buildings, and for demolitions or salvage and abatement work.

Starting Date: 2016

Ending Date: 2023

Geospatial Coverage:

Vancouver, Canada

Number of records: 40k

Source:

<https://opendata.vancouver.ca/explore/dataset/issued-building-permits/information/>

Records example

Permit Number	PermitNumberCreatedDate	IssueDate	PermitElapsedDays	ProjectValue	TypeOfWork	Address	ProjectDescription	PermitCategory	Applicant	ApplicantAddress	PropertyUse	SpecificUseCategory	BuildingContractor	BuildingContractorAddress	IssueYear	GeographicalArea	Geometry	YearMonth	geopoint_2d
DB-2023-00041	2023-01-05	2023-04-28	113	198847.5	New Building	7475 DUMFRIES STREET, Vancouver, BC	Low Density Housing - New Building - To constr...	New Build - Standalone Laneway	Shalindro Dosanjh	1185 49th Ave West\r\nVancouver,, BC V6M 2P9	Dwelling Uses	Laneway House	Amcata Homes Developments Ltd	1185 W 49TH AV\r\nVancouver, BC V6M 2P9	2023	Victoria-Fraserview	{"coordinates": [-123.0755831, 49.2165995], "typ...	2023-04	49.2165995, -123.0755831
BP-2023-00049	2023-01-06	2023-03-29	82	0.0	Salvage and Abatement	1137 E 29TH AVE, Vancouver, BC	Low Density Housing - Salvage and Abatement - ...	NaN	John Keen DBA : Farpoint Architectural	1662 West 75th Ave.\r\nVancouver, BC V6P 6G2	Dwelling Uses	Single Detached House	Vancouver Excavating and Contracting Ltd	25 Howard Ave\r\nBurnaby, BC V5B 3P3	2023	Kensington-Cedar Cottage	{"coordinates": [-123.081827, 49.244843], "typ...	2023-03	49.244843, -123.081827

Data preparation and cleaning

Formatting & Validity

PermitNumberCreatedDate,
IssueDate to the **Datetime**
type

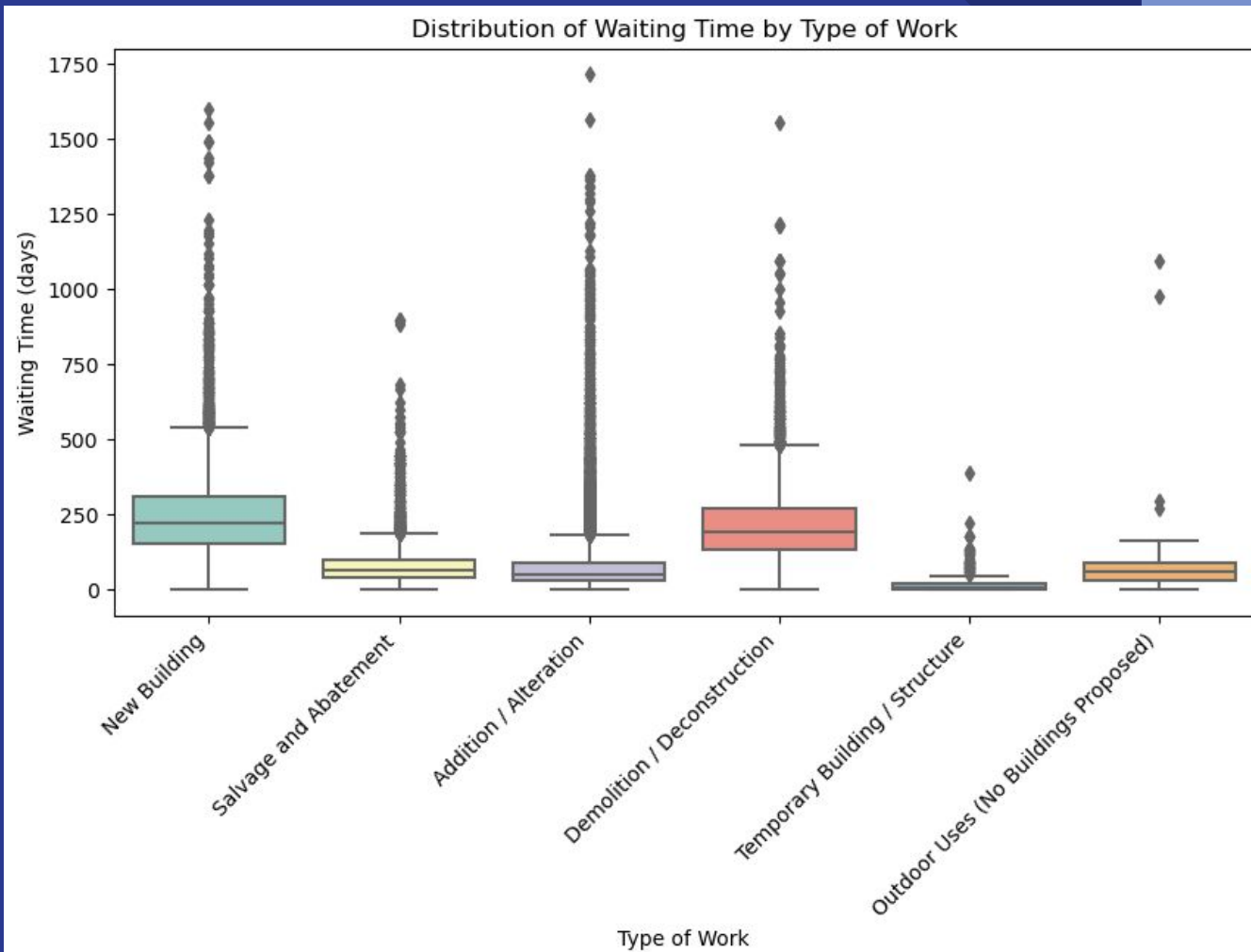
Checking for duplicates

No duplicates

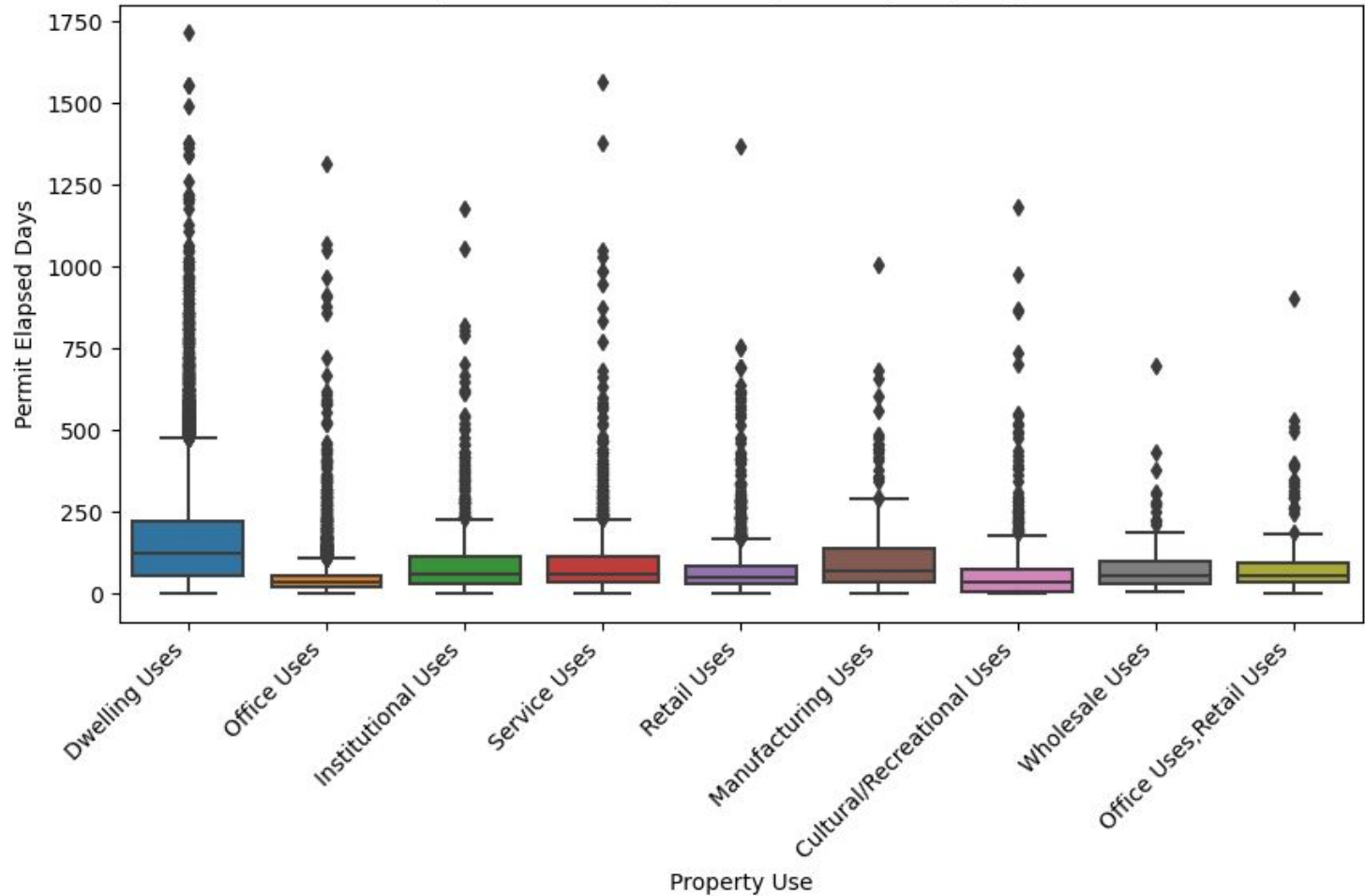
Missing Data

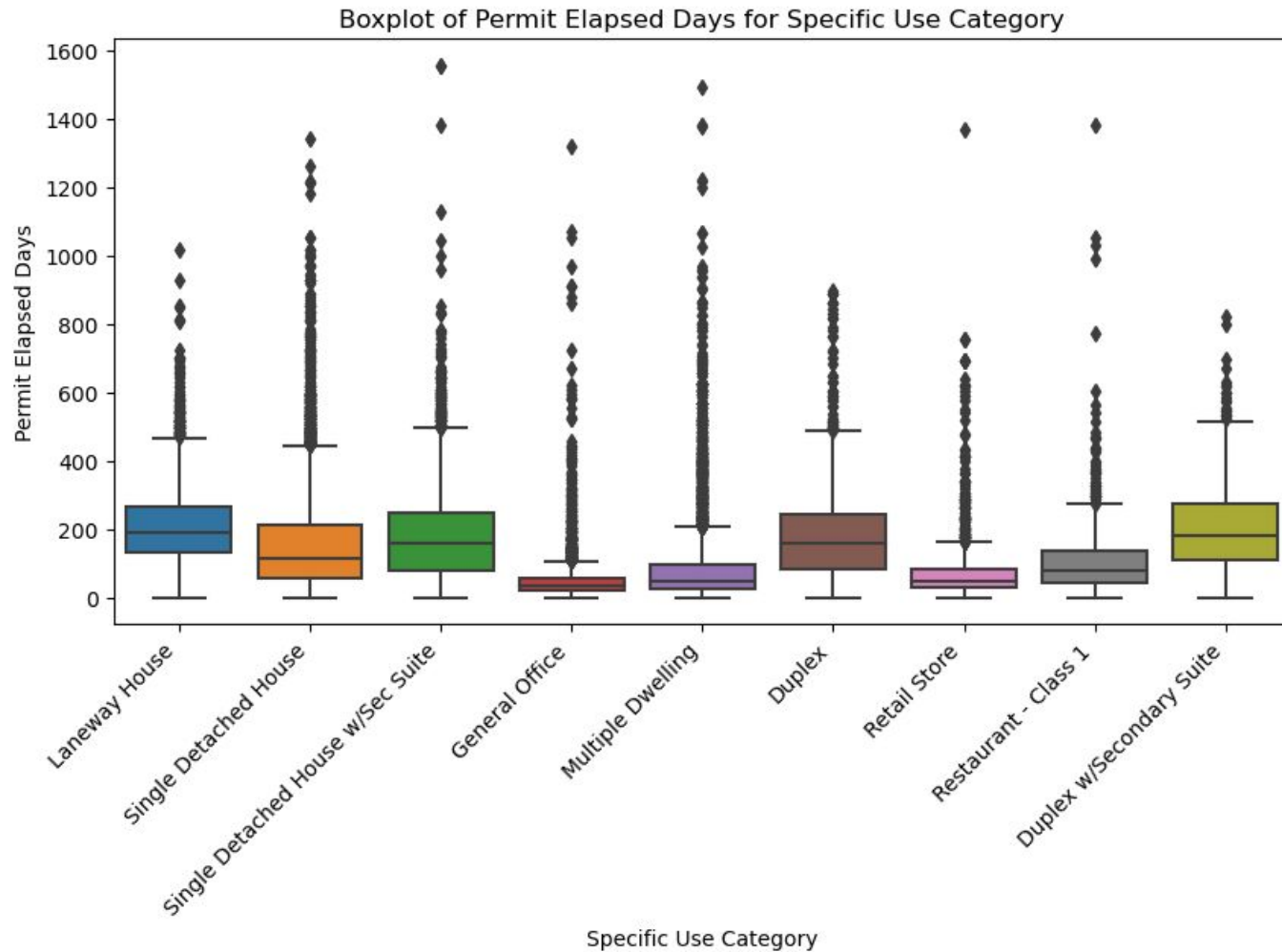
Address	0.401719
PermitCategory	43.007011
ApplicantAddress	0.266107
PropertyUse	0.012794
SpecificUseCategory	0.017911
BuildingContractor	37.810245
BuildingContractorAddress	56.885523
GeoLocalArea	1.194923
Geom	1.174454
geo_point_2d	1.174454

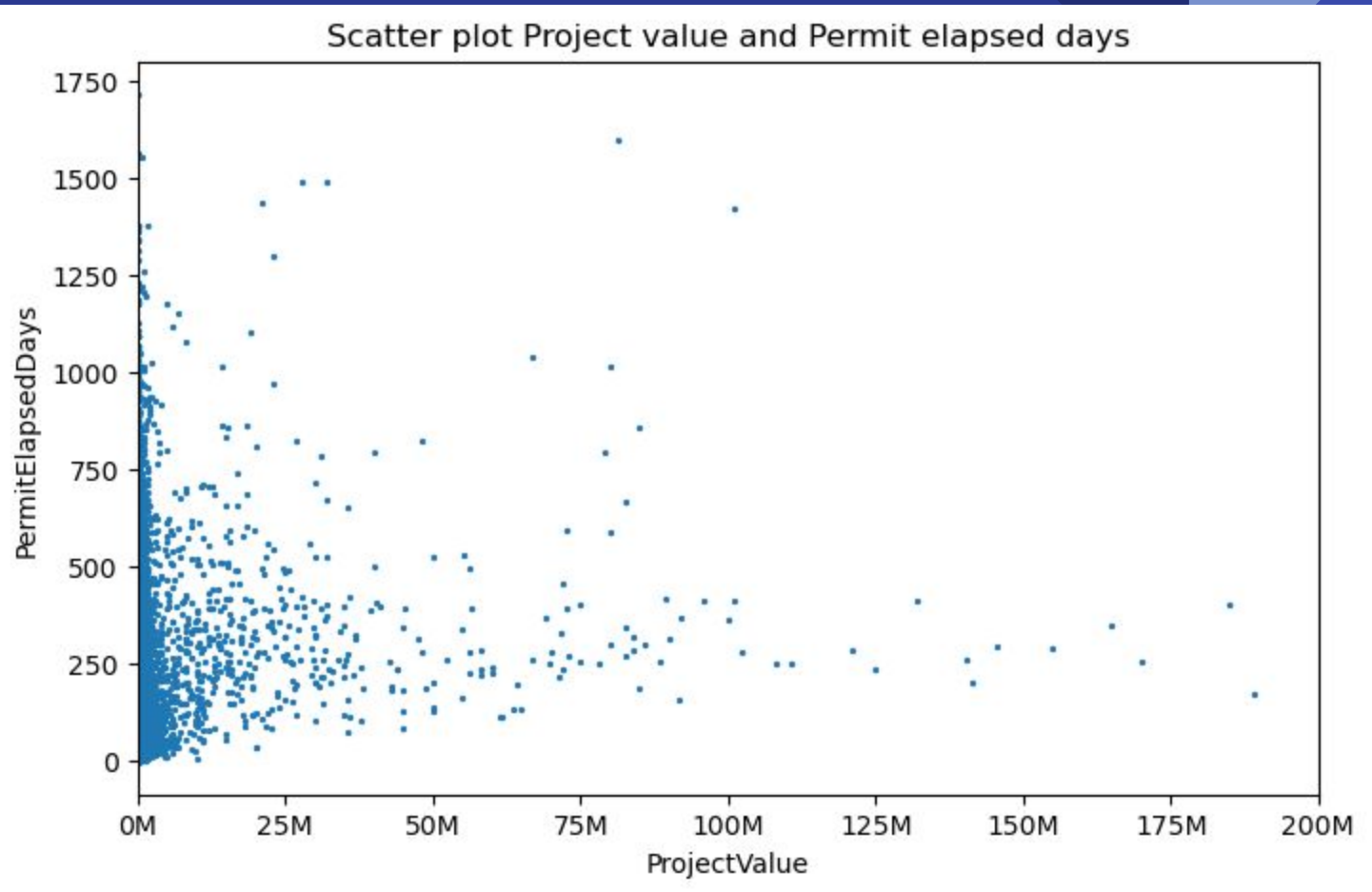
Remove rows <3% and **PermitCategory,**
BuildingContractor, BuildingContractorAddress



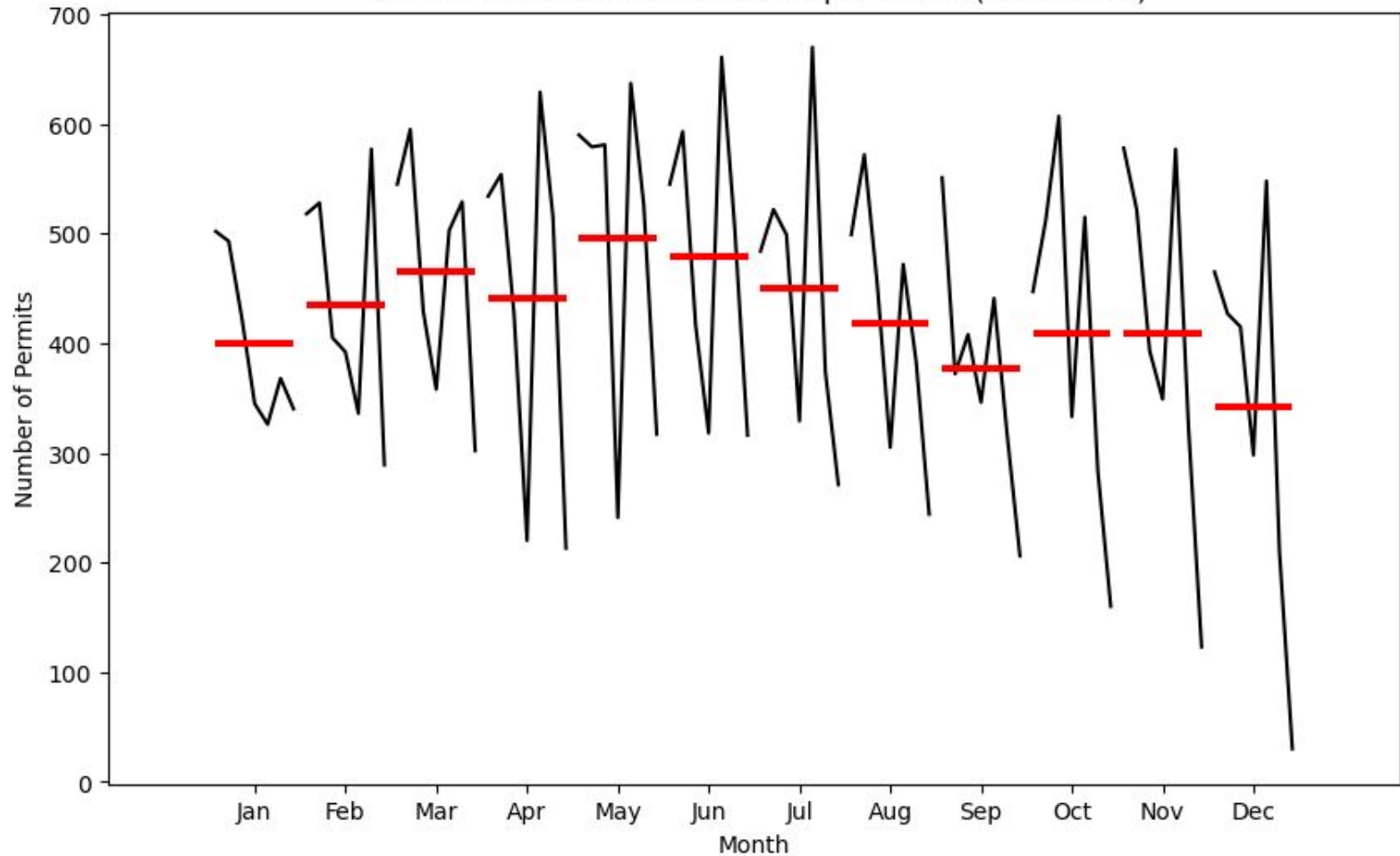
Boxplot of Permit Elapsed Days for Top Property Uses



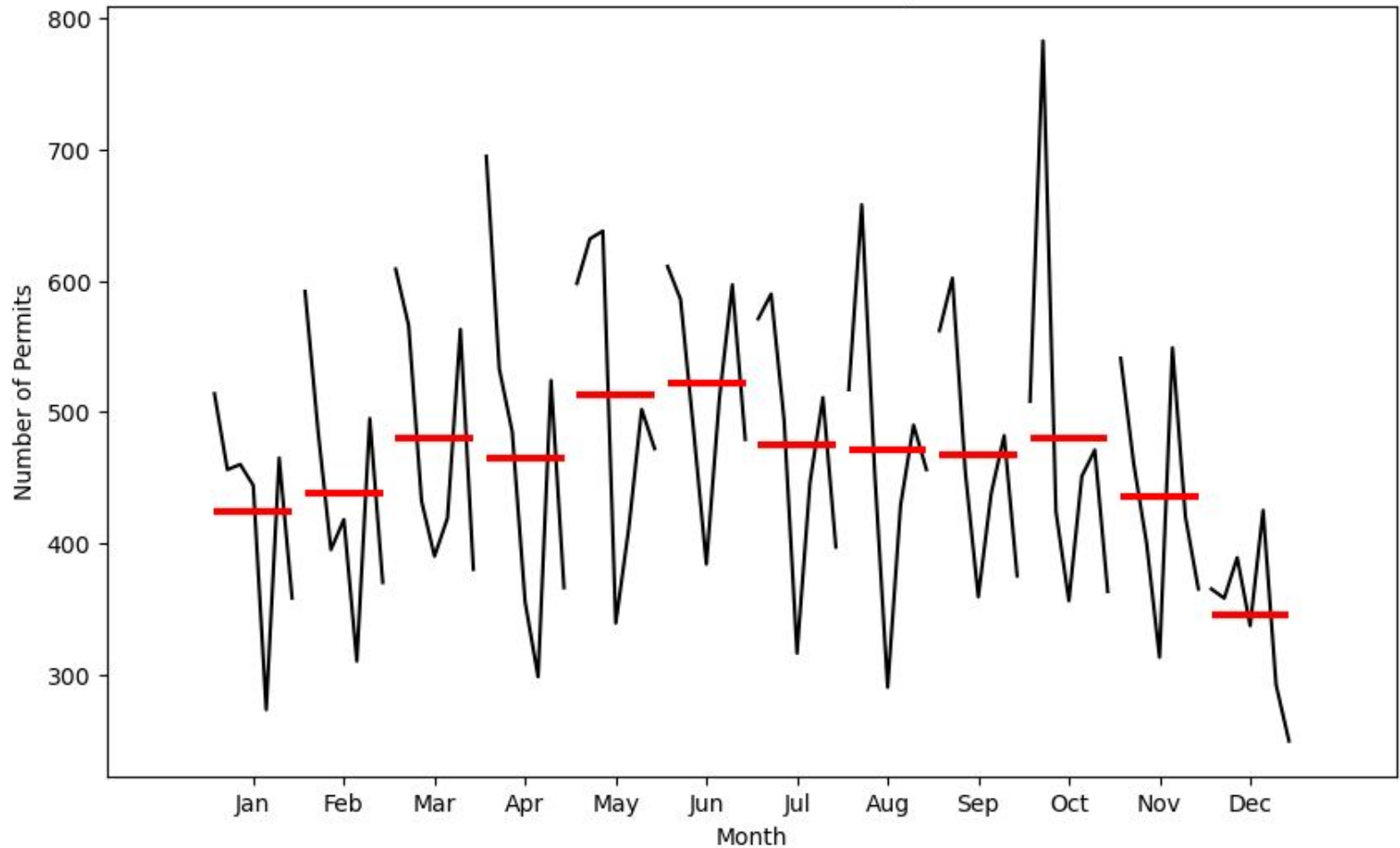




Seasonal Plot of Created Permits per Month (2017-2023)



Seasonal Plot of Issued Permits per Month



Feature Engineering

One Hot Encoding

TypeOfWork

PropertyUse (top 12 frequent)

SpecificUseCategory (top 22)

GeoLocalArea

Drop columns

IssueYear

IssueDate

PermitNumber

YearMonth

Geom

Bag-of-Words (Count)

ProjectDescription

Applicant

Baseline models

Linear Regression

Training Metrics:

Mean Absolute Error: 64

Mean Squared Error: 11026

R-squared: 0.43

Testing Metrics:

Mean Absolute Error: 64.5

Mean Squared Error: 11252

R-squared: 0.39

Ridge

Training Metrics:

Mean Absolute Error: 64

Mean Squared Error: 11032

R-squared: 0.427

Testing Metrics:

Mean Absolute Error: 64.4

Mean Squared Error: 11226

R-squared: 0.39

Decision Tree Regressor

Training Metrics:

Mean Absolute Error: 66.6

Mean Squared Error: 11797

R-squared: 0.388

Testing Metrics:

Mean Absolute Error: 66.1

Mean Squared Error: 11948

R-squared: 0.355

Next steps

- **Hyperparameter tuning**
- **Cross-validation**
- **Linear Regression with Time Features**
- **Seasonal ARIMA**
- **GBoost or LightGBM with Time Features**
- **LSTM Neural Networks**