Unraveling Vancouver's Housing Challenge through Building Permits Analysis

Tatyana Mozgacheva
Data Science, 2024, Brainstation

The problem

Overview

Vancouver is currently grappling with a significant housing challenge marked by soaring property prices and a scarcity of affordable housing.

Problem statement

This project aim to uncover insights into how the permitting process contributes to or mitigates the housing shortage.

Users

City planners and officials can enhance workflow efficiency, allocate resources effectively, and optimize the permitting process using predictions.

Applicants and developers benefit from knowing estimated permit issuance times, allowing for more accurate construction timeline planning.

The Big Ideas

Predicting Permit Processing Times

Utilize machine learning algorithms to analyze historical building permits data and predict processing times.This can assist in proactive planning and resource allocation.

Spatial Analysis

Use geospatial data to visualize the distribution of building permits across the city. This can help in identifying areas with concentrated development and areas that may need more attention.

Potential Impact

The societal value of this project lies in addressing a critical issue that directly impacts the quality of life for Vancouver residents. By quantifying the scale of the problem through data analysis, I can provide valuable information to policymakers and urban planners and any residents. This could lead to more informed decisions on housing development, potentially alleviating the shortage and making housing more accessible.

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

PermitNumber	PermitNumberCreatedDate	IssueDate	PermitElapsedDays	ProjectValue	TypeOfWork	Address	Project Description	PermitCategory	Applicant	ApplicantAddress	PropertyUse	SpecificUseCategory	BuildingContractor	BuildingContractorAddress	IssueYear	GeoLocalArea	Geom	YearMonth	geo_point_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": [-123.07558831, 49.2165995], "t...	2023-04	49.2165995, -123.0755831
BP-2023-00049	2023-01-06	2023-03-29	82	0.0	Salvage and Abatement	1137 E 29TH AVENUE, 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

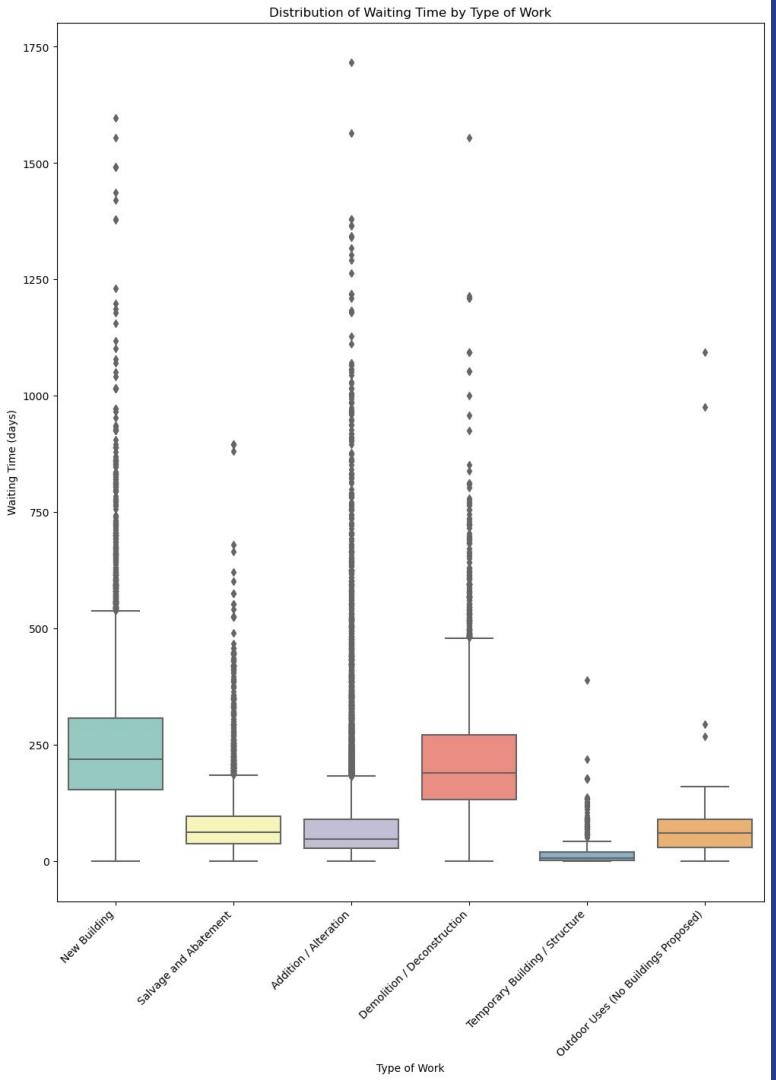
Checking for duplicates

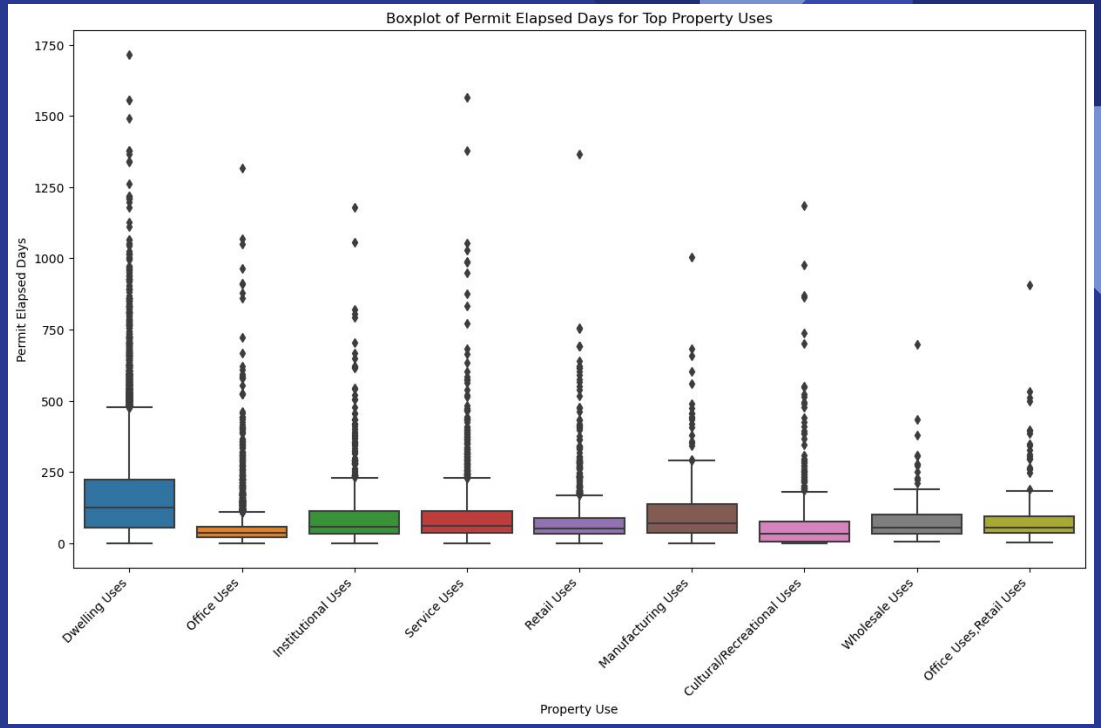
No duplicates

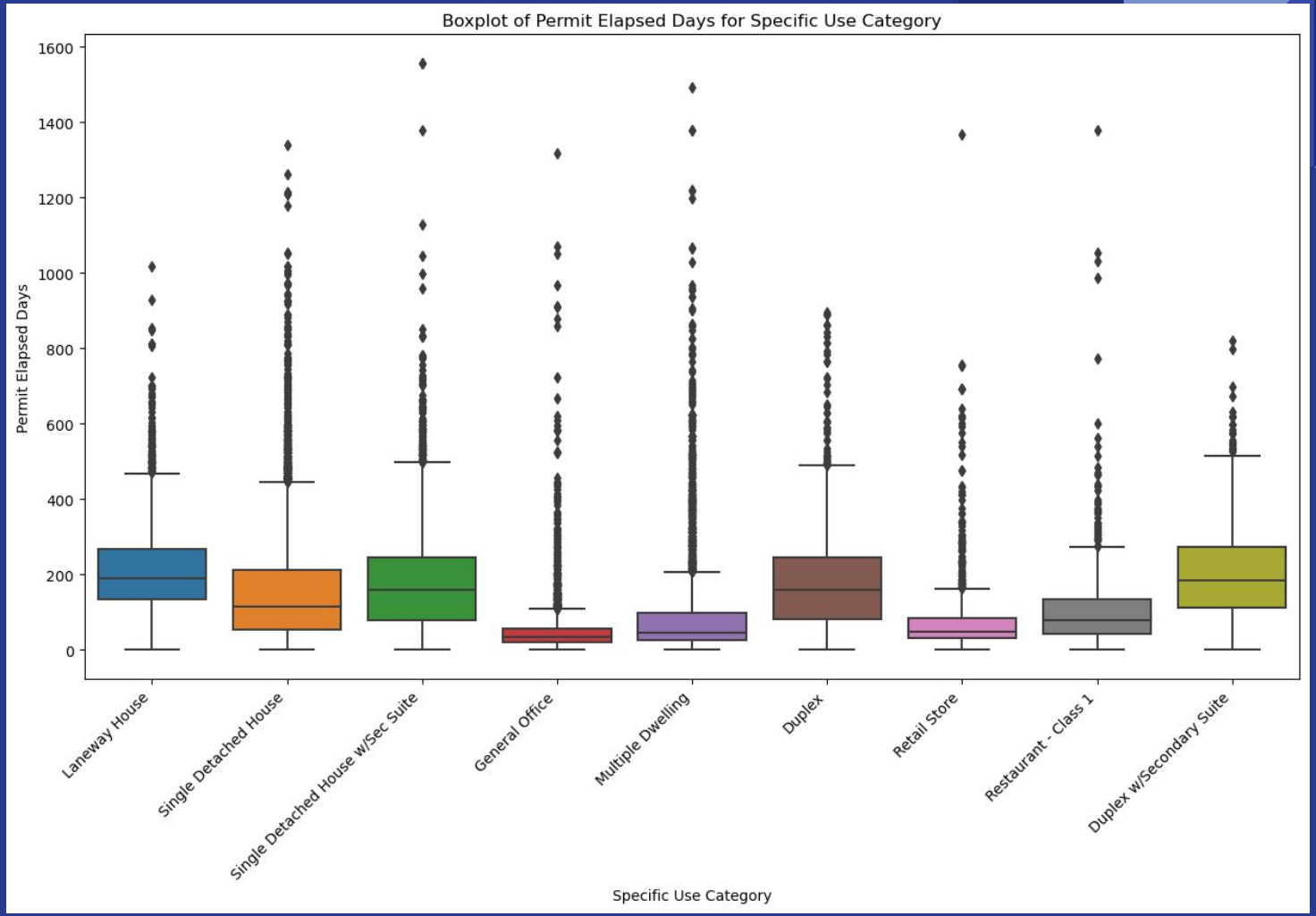
Missing Data

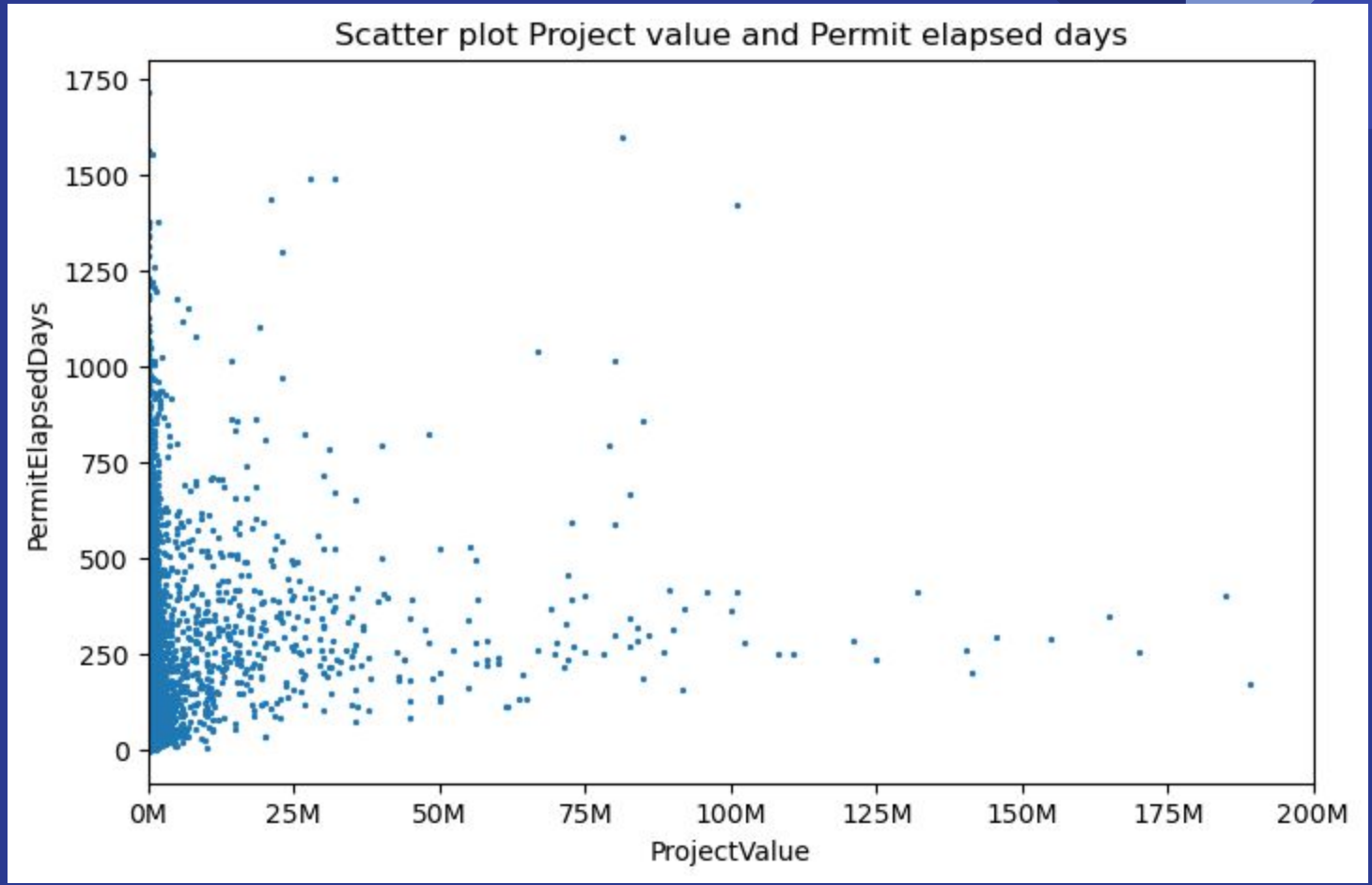
Address	0.401719
PermitCategory	43.007011
ApplicantAddress	0.266107
PropertyUse	0.012794
SpecificUseCategory	0.017911
BuildingContractor	37.810245
BuildingContractorAddress	56.885523
GeoLocalArea	1.194923
Geom	1.174454
geo_point_2d	1.174454

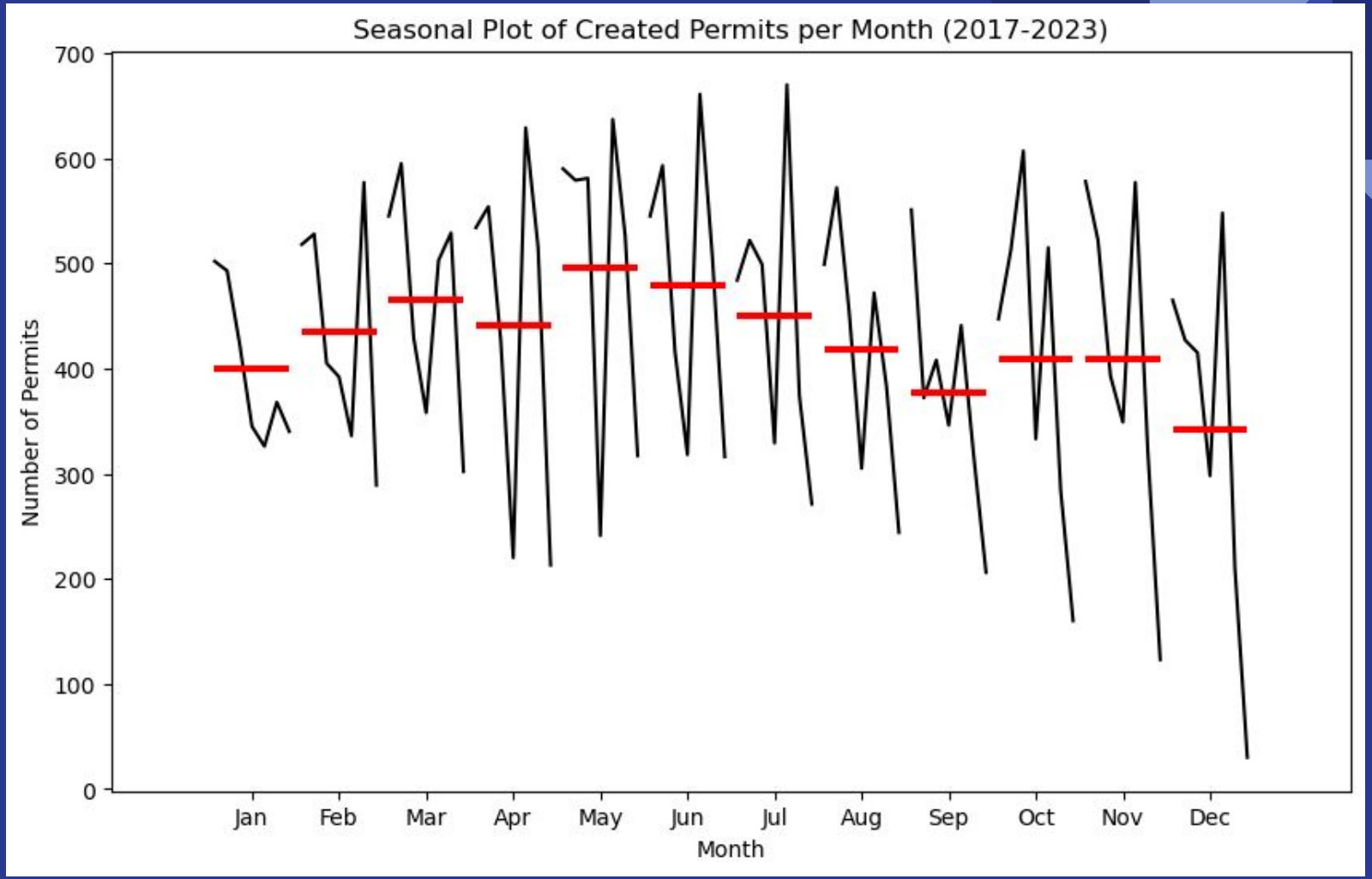
Remove rows <3% and **PermitCategory, BuildingContractor, BuildingContractorAddress**

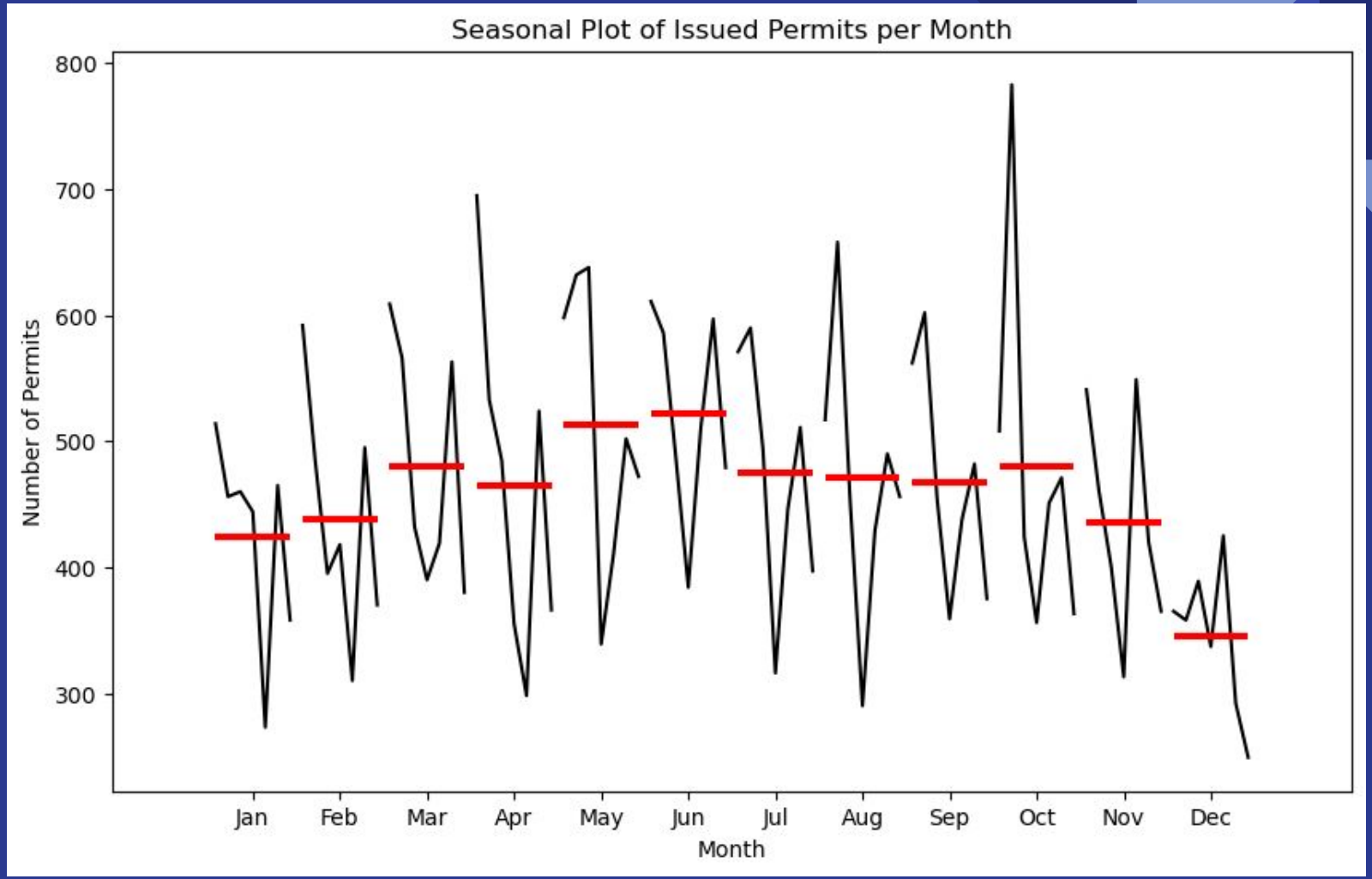

Distribution of Waiting Time by Type of Work
Waiting Time (days)
0
250
500
750
1000
1250
1500
1750
New Building
Salvage and Abatement
Addition / Alteration
Demolition / Deconstruction
Temporary Building / Structure
Outdoor Uses (No Buildings Proposed)
Type of Work


Boxplot of Permit Elapsed Days for Top Property Uses
Permit Elapsed Days
0
250
500
750
1000
1250
1500
1750
Dwelling Uses
Office Uses
Institutional Uses
Service Uses
Retail Uses
Manufacturing Uses
Cultural/Recreational Uses
Wholesale Uses
Office Uses,Retail Uses
Property Use


Boxplot of Permit Elapsed Days for Specific Use Category
Permit Elapsed Days
0
200
400
600
800
1000
1200
1400
1600
Laneway House
Single Detached House
Single Detached House w/Sec Suite
General Office
Multiple Dwelling
Duplex
Retail Store
Restaurant - Class 1
Duplex w/Secondary Suite
Specific Use Category


Scatter plot Project value and Permit elapsed days
PermitElapsedDays
1750
1500
1250
1000
750
500
250
0
0M
25M
50M
75M
100M
125M
150M
175M
200M
ProjectValue


Seasonal Plot of Created Permits per Month (2017-2023)
Number of Permits
700
600
500
400
300
200
100
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month


Seasonal Plot of Issued Permits per Month
Number of Permits
800
700
600
500
400
300
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month

Next steps

Feature Selection:
Select relevant features for predicting '**PermitElapsedDays**.' Focus on features that have a strong correlation with the target variable.

Data Splitting:
Split the dataset into training and testing sets.

Baseline Regression Model:
Pick a simple regression model like Linear Regression as a baseline.

Evaluate Baseline Model:
Evaluate the model's performance on the testing set using metrics such as Mean Absolute Error (MAE) or Root Mean Squared Error (RMSE).