Data Analytics (COMP47350): Homework 1

Introduction

This subject of this assignment is a randomly selected dataset of approximately 10,000 entries from the Irish Residential Property Price Register, hereafter known as the **RPPR** (link to register can be found here). The randomly selected dataset assigned to me is saved down in this directory with the title **ppr-21200122.csv**.

The purpose of this assignment is, broadly, the preparation and review of the dataset so that it is easier to work with (for example, it should be acceptable input for a machine learning algorithm). On completion, the data should be cleaned, certain features dropped / added, and correlations, or lack thereof, of certain features should be noted. The various steps are specifically set out in more detail throughout this assignment as sub-headings. The ultimate goal is to build a data analytics solution for residential property price prediction, and as such we should view all of our actions with regard to the dataset through this lens.

Please note that all references to theory are, unless otherwise stated, are obtained from Dr Georgiana Ifrim's lectures in module COMP47350.

It should also be flagged that most of the code is explained in the markdown cells in this notebook. In the event that something in the code is not particularly clear, comments will be included directly in such code cells.

Contents of this folder

This folder contains the following files:

- **COMP47350_Homework_1.ipynb**: The Jupyter Notebook version of this assignment.
- **COMP47350_Homework_1.pdf**: The PDF version of this assignment. Please note, however, that due to the size of some of the tables involved in our analysis, some of these may be cut of in the PDF version, and as such, the notebook version should be used for the purposes of the review where possible.
- **Data Quality Report.pdf**: The Data Quality Report that is referenced in this assignment. The .docx version is included also for completeness.
- **Data Quality Plan.pdf**: The Data Quality Plan that is referenced in this assignment. The .docx version is included also for completeness.
- **Graphs**: A folder containing .png versions of the various plots used for preparing the Data Quality Report. When this notebook is run, these pngs will be generated again in the root of this folder.
- **ppr-21200122.csv**: The .csv file assigned to me for the purposes of this analysis.
- **CSV_Files**: Folder containing .csv files that have been requested as output from the questions in this assignment. When this notebook is run, these csv files will be generated again in the root of this folder.
- **Property_Addresses.csv and folium.png**: Output from using the Folium library as part of this assignment - these files are explained in detail further on in this assignment.
- **Inflation.csv**: A csv file downloaded from the world bank, this is necessary for the final question - this will be explained in detail further on in this assignment.
- **requirements.txt**: A list of dependencies required to run this notebook. This was obtained as part of a "pip freeze" of a conda environment I was using for this assignment, so it contains all the dependencies required for the libraries listed within.
- **install_requirements.sh**: A small shell script that installs the dependencies listed in requirements.txt in a single click - this has been included for convenience.

Preliminary review of data source

Before beginning work on the dataset, we should first review our data source, the RPPR, and its place within the domain of the dataset, which is the Irish housing market.

The RPPR was introduced in 2011 and contains information about the properties purchased since 1 January 2010. The *Information Note* contained in the RPPR notes a couple of interesting points that we should keep in mind while working on our dataset:

- the RPPR should contain only residential properties;
- if a property sale consisted of both residential and non-residential parts, then only the residential part will be noted in the register;
- if the property is new, then the price shown in the register will be shown exclusive of VAT, which equals to 13.5%; and
- if multiple properties in one transaction, then these will all be included as a single entry in the register, although this is not a guarantee.

It is also interesting to note that the RPPR is based on the filings made with the Irish Revenue Commissioners in respect of a particular property sale, and as such, it is entirely possible that the RPPR will contain errors, making the Data Quality Report an extremely important part of this assignment (as the report is where we will consider strategies for fixing such errors).

Our target feature here will be price of the property sale, and in particular it will be of interest to see how the price fluctuates when compared with other features provided in the data set, such as Date of Sale, or County.

Question 1: Prepare a data quality report for your CSV file.

1(a) Initial review of data

The first step in preparing the data quality report for the CSV file is to import the various Python libraries that we are going to need in order to work with the provided data set.

The following libraries are expected to be used over the course of this assignment:

- **pandas**: to simplify working with, and navigation of, the dataset;
- **matplotlib**: to create various plots and charts so that the data can be visualised;
- **numpy**: to perform certain calculations; and
- **seaborn**: to perform some more advanced data visualisation that would have taken longer to do in matplotlib.

It is possible that some additional libraries may required later on, which will be imported as required.

In [1]:
```
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib
import numpy as np
import seaborn as sns

# The below allows plots to run directly in the notebook
%matplotlib inline
```

The next step is to actually read in the CSV file containing the data set into a dataframe variable, which will allow us to easily work with the data.

Please note that I have assumed that the CSV file is already UTF-8 encoded. I have based this assumption on the fact that there were no format issues with working with this dataset. If this was not the case, the file would need to be converted to UTF-8 and re-saved as a new file so that Pandas would be able to read it more clearly.

In [2]:
```
df = pd.read_csv("ppr-21200122.csv")
```

To get a sense of the size of the data set, we will check how many rows and columns it has.

In [3]:
```
print(f"The number of rows is: {df.shape[0]}")
print(f"The number of columns is: {df.shape[1]}")
```

```
The number of rows is: 10000
The number of columns is: 9
```

Now we will print the first and last 5 rows of the CSV file to get a feel for what the data looks like.

In [4]:
```
df.head(5)
```

Out[4]:

	Date of Sale (dd/mm/yyyy)	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property	Property Size Description
0	26/06/2020	1 GANDON PLACE, GANDON PARK, Lucan	NaN	Dublin	€370,044.05	No	Yes	New Dwelling house /Apartment	NaN
1	19/12/2014	72 ST ASSAMS PARK, RAHENY, DUBLIN 5	Dublin 5	Dublin	€480,000.00	No	No	Second-Hand Dwelling house /Apartment	NaN
2	11/02/2010	37 Ivy Hill, Gort Road, Ennis	NaN	Clare	€194,000.00	No	No	Second-Hand Dwelling house /Apartment	NaN
3	16/08/2018	14 TYRCONNELL PLACE, TYRCONNELL RD, INCHICORE ...	Dublin 8	Dublin	€275,000.00	No	No	Second-Hand Dwelling house /Apartment	NaN
4	27/02/2019	7 THE NEST, TUBBERCURRY, SLIGO	NaN	Sligo	€75,000.00	No	No	Second-Hand Dwelling house /Apartment	NaN

In [5]:
```
df.tail(5)
```

Out[5]:

	Date of Sale (dd/mm/yyyy)	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property	Prope S Descripti

	Date of Sale (dd/mm/yyyy)	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property	Prope S Descripti
9995	29/07/2019	SEFFIN, BIRR, CO OFFALY	NaN	Offaly	€100,000.00	No	No	Second-Hand Dwelling house /Apartment	N
9996	02/10/2020	74 CEDAR PLACE, RIDGEWOOD, SWORDS	NaN	Dublin	€316,091.00	No	No	Second-Hand Dwelling house /Apartment	N
9997	24/03/2017	APT 2 INISH HOUSE, IRISHTOWN, ATHLONE	NaN	Westmeath	€95,000.00	No	No	Second-Hand Dwelling house /Apartment	N
9998	24/04/2019	94 Castleguard Manor, Ardee	NaN	Louth	€230,000.00	No	Yes	New Dwelling house /Apartment	N
9999	10/08/2018	45 BUTTERFIELD DR, RATHFARNHAM, DUBLIN 14	Dublin 14	Dublin	€1,060,000.00	No	No	Second-Hand Dwelling house /Apartment	N

1(b) Analysis of datatypes

The features need to be converted to their appropriate datatypes - in order to do this, we must review the features and make a decision as to whether they would suited to "continuous" or "categorical" types.

We will begin by listing all the features and their current datatypes.

In [6]:
```
df.dtypes
```

Out[6]:
```
Date of Sale (dd/mm/yyyy)    object
Address                      object
Postal Code                  object
County                       object
Price (€)                    object
Not Full Market Price        object
VAT Exclusive                object
Description of Property      object
Property Size Description    object
dtype: object
```

We can see from the above output that all the features are currently set to the **object** datatype. To make a decision as to whether or not these features are "continuous" or "categorical", we will categorise them into the datatypes as described in Machine Learning for Predictive Data Analytics (2015, Kellegher et al) (p.79).

Please note that this analysis has been carried out from an initial review of the head an tail of the data set, and some of the features may need to be moved to other categories as we proceed with our analysis.

- *Date of Sale* is an **interval** datatype, as it is a value that represents a span of time.

- *Address* is a **textual** datatype, as addresses are short, free-form text data.
- *Postal Code* is a **categorical** datatype, as postal codes are finite values (there are a limited number of postal codes in Ireland) that cannot be ordered (there is no way to compare postal codes) and allow no arithmetic (you cannot subtract one postal code from another).
- *County* is a **categorical** datatype, for the same reasons listed for the *Postal Code* feature.
- *Price* is a **numeric** datatype, as it is a true numeric value that allows arithmetic operations (you can add and subtract prices, for example).
- *Not Full Market Price* is a **binary** datatype, as it only has two possible values - "yes" and "no".
- *VAT Exclusive* is a **binary** datatype, for the same reasons listed for the *Not Full Market Price* category.
- *Description of Property* is a **textual** datatype, as it is short, free-form text data.
- *Property Size Description* is a **textual** datatype, as it is short, free-form text data.

We can reduce these datatypes into *continuous* and *categorical* features as follows:

- *Continuous* features consist of numeric and interval types.
- *Categorical* features consist of categorical, ordinal, binary and textual types.

Therefore, our *continuous* features will consist of:

- Date of Sale; and
- Price.

Our *categorical* features will consist of:

- Address;
- Postal Code;
- County;
- Not Full Market Price;
- VAT Exclusive;
- Description of Property; and
- Property Size Description.

1(c) Converting categorical data to appropriate datatype

We will now mark the relevant features as "categorical". In order to do this, we will select all the features we marked as "categorical" in the section above and place them into a list. We will then loop through these features and mark each of them as being of type "category".

In [7]:
```
# Selecting categorical columns
categorical_columns = df[["Address", "Postal Code", "County", "Not Full Market Price", "VAT Exc

# Looping through columns and casting as category
for column in categorical_columns:
    df[column] = df[column].astype('category')
```

1(d) Converting continuous data to appropriate datatype

From an initial review of our chosen continuous features, we can see that there are some issues with the way the data is formatted that we will have solve before marking the columns as "continuous". These are as follows:

- *Price*: The Euro symbol and the decimal period need to be stripped out, and the resulting string needs to be converted to an integer.
- *Date of Sale*: The date needs to be converted to the standard pandas datetime format.

In [8]:
```
def change_currency_to_float(currency_value):
    """Converts currency string into a float"""
    return currency_value.replace("€", "").replace(",", "")

# Set all prices in dataframe to float value
df["Price (€)"] = df["Price (€)"].apply(change_currency_to_float).astype("float")
```

Next, we convert the Date of Sale feature to a pandas datetime datatype. We should also change the name of the feature to just say "Date of Sale" for brevity.

In [9]:
```
# Convert datetime string to datetime object
df["Date of Sale (dd/mm/yyyy)"] = pd.to_datetime(df["Date of Sale (dd/mm/yyyy)"], infer_datetim

# Rename datetime column
df.rename(columns={"Date of Sale (dd/mm/yyyy)": "Date of Sale"}, inplace=True)
```

Now that both features have been converted to their appropriate datatype, we will double check the head of the dataframe to ensure that both features are outputting as expected.

In [10]:
```
df.head()
```

Out[10]:

	Date of Sale	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property	Property Size Description
0	2020-06-26	1 GANDON PLACE, GANDON PARK, Lucan	NaN	Dublin	370044.05	No	Yes	New Dwelling house /Apartment	NaN
1	2014-12-19	72 ST ASSAMS PARK, RAHENY, DUBLIN 5	Dublin 5	Dublin	480000.00	No	No	Second-Hand Dwelling house /Apartment	NaN
2	2010-02-11	37 Ivy Hill, Gort Road, Ennis	NaN	Clare	194000.00	No	No	Second-Hand Dwelling house /Apartment	NaN
3	2018-08-16	14 TYRCONNELL PLACE, TYRCONNELL RD, INCHICORE ...	Dublin 8	Dublin	275000.00	No	No	Second-Hand Dwelling house /Apartment	NaN
4	2019-02-27	7 THE NEST, TUBBERCURRY, SLIGO	NaN	Sligo	75000.00	No	No	Second-Hand Dwelling house /Apartment	NaN

Now that we are satisfied that the conversion to appropriate datatype is successful, we can move these features into a "continuous_columns" variable for ease of access later on.

In [11]:
```
continuous_columns = df[["Date of Sale", "Price (€)"]].columns
```

Before moving on to the next step, we should print the datatypes of all features just to make sure that everything is working as intended.

In [12]:
```
df.dtypes
```

Out[12]:
```
Date of Sale                datetime64[ns]
Address                           category
Postal Code                       category
County                            category
Price (€)                          float64
Not Full Market Price             category
VAT Exclusive                     category
Description of Property           category
Property Size Description         category
dtype: object
```

1(e) Look for duplicate rows and columns

Initially, we will do a simple check to see if there are any duplicate rows or columns in the dataframe at all.

In [13]:
```
print(f"The total number of duplicate rows, excluding the original rows, is {df.duplicated().su
print(f"The total number of duplicate rows, including the original rows, is {df[df.duplicated(k
```

```
The total number of duplicate rows, excluding the original rows, is 1.
The total number of duplicate rows, including the original rows, is 2.
```

In [14]:
```
# Flip the dataframe so that the columns become rows
df_transposed = df.T
print(f"The total number of duplicate columns, excluding the original columns, is {df_transpose
print(f"The total number of duplicate columns, including the original columns, is {df_transpose
```

```
The total number of duplicate columns, excluding the original columns, is 0.
The total number of duplicate columns, including the original columns, is 0.
```

From the above output we can tell that there is one pair of duplicate rows, and no duplicate columns in the dataframe.

We should now review the duplicated rows. Our best way of checking for a duplicate in this dataframe would be to use the Date of Sale and Address features as a key, as per the below.

In [15]:
```
df[df.duplicated(subset=["Date of Sale", "Address"], keep = False)]
```

Out[15]:

	Date of Sale	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property	Property Size Description
339	2012-12-19	21 Crockanure Avenue, Kildare Town	NaN	Kildare	78000.0	No	No	Second-Hand Dwelling house /Apartment	NaN
4425	2012-12-19	21 Crockanure Avenue, Kildare Town	NaN	Kildare	78000.0	No	No	Second-Hand Dwelling house /Apartment	NaN

We can see from the above that there are two identical rows in our dataframe. It is very unlikely that the same house in the same address would be sold twice on the same date, and as such it is entirely possible that this row was entered in error.

For this reason, we can drop one of these entries from our dataframe.

In [16]:
```
# Drop the relevant index in place
df.drop(index = 4425, inplace = True)

# Reset the index
df = df.reset_index(drop = True)
```

We will now check for duplicates again to ensure that this worked correctly.

In [17]:
```
print(f"The total number of duplicate rows, including the original rows, is {df[df.duplicated(k
```

```
The total number of duplicate rows, including the original rows, is 0.
```

The above returned 0, therefore we can be certain that there are no further duplicate rows in our dataframe.

1(f) Check for constant columns

To check for constant columns, we will run the nunique() method to ensure that each feature has more than one unique value. If there is just one unique value, then the column is constant and is unlikely to provide us with any useful information.

In [18]:
```
df.nunique()
```

Out[18]:
```
Date of Sale                2763
Address                     9979
Postal Code                   22
County                        26
Price (€)                   2324
Not Full Market Price          2
VAT Exclusive                  2
Description of Property        3
Property Size Description      4
dtype: int64
```

The above tells us that each feature has more than one unique value, and as such, there are no constant columns.

1(g) Check the logical integrity of the data

We will now conduct some basic tests in connection with the logical integrity of the data. Please note that these are based on an initial basic analysis of the data, and as such more logic tests will be carried out once the descriptive statistics have been prepared.

Test 1: The Residential Property Price Register was introduced on 1 January 2010, and as such, the Date of Sale cannot be before 2010-01-01.

In [19]:
```
import datetime

test_1 = df[df["Date of Sale"] < np.datetime64("2010-01-01")]
print(f"The number of rows that fail test 1 is: {test_1.shape[0]}")
```

```
The number of rows that fail test 1 is: 0
```

Test 2: The Date of Sale cannot be in the future.

In [20]:
```
test_2 = df[df["Date of Sale"] > np.datetime64("2022-02-01")]
print(f"The number of rows that fail test 2 is: {test_1.shape[0]}")
```

```
The number of rows that fail test 2 is: 0
```

Test 3: The postal code must be a valid post code from differing counties in Ireland.

For this test, given the low number of unique values, we can simply try and print all the post codes to see if all of them match standard Irish postcodes.

In [21]:
```
for postal_code in df["Postal Code"].unique():
    print(postal_code)
```

```
nan
Dublin 5
Dublin 8
Dublin 14
Dublin 4
Dublin 6
Dublin 24
Dublin 13
Dublin 7
Dublin 15
Dublin 11
Dublin 16
Dublin 1
Dublin 17
Dublin 2
Dublin 12
Dublin 18
Dublin 3
Dublin 9
Dublin 22
Dublin 10
Dublin 6w
Dublin 20
```

From the above, we can see that all postal codes mentioned are Dublin-based, which would explain the large number of missing values in this column.

This could be considered a logical failure, as we had initially expected postal codes from all over the country to be included.

This is something that we should note in the data quality report. Aside from that, the postal codes listed all seem to be valid (although slightly out-of-date, given that they have been replaced with Eircodes).

While we are reviewing the postal codes, we should also probably check if there are any cases where the Dublin postal code is attached to an address that is not in Dublin.

In [22]:

```
# Loop through rows and print addresses that have a Dublin postcode but do not have Dublin writ
for index in range(df.shape[0]):
    if ("dublin" in df["County"][index].lower()) and ("dublin" not in df["Address"][index].lowe
        print(df["Address"][index])
```

```
1 GANDON PLACE, GANDON PARK, Lucan
2a Liffet Crescent, Lucan
AUBURN GROVE HOUSE, AUBURN, MALAHIDE
1 GLEANNTAN, LOUGHLINSTOWN
11 WARRINGTON MEWS, DONABATE
26 THE OAKS, RIDGEWOOD, SWORDS
No. 106, Upper Leeson Street
80 The Fairways, Seabrook Manor, Portmarnock
23 Clay Farm Avenue, Clay farm, Leopardstown
DARTRY, SANDYCOVE AVENUE EAST, DUN LAOGHAIRE
APT 2 THE PIERRE, VICTORIA TERRACE, DUN LAOGHAIRE
14 RICHMOND HALL, MOUNT ST.  ANNES, MILLTOWN
74 THE ELMS, MOUNT MERRION AVE, BLACKROCK
13 Broadfield Grove, Broadfield Manor, Rathcoole
46 GREEN VIEW, SEABROOK MANOR, PORTMARNOCK
168 Castle Park, Tallaght
41 THE GALLERY, TURVEY WALK, DONABATE
17 LINDENVALE, PROBY SQ, BLACKROCK
THE COTTAGE, DAMASTOWN, NAUL
34 Castle Street, Dalkey
226 Pace Road, Littlepace, Clonee
1  5  6 & 22 THE HAMPTON, APT 1 2 7 19 20 26 27 30 31 34-38, 41 42 47-49 BLOCK 1  THE MAIESTON
```

3 AIRPARK CLOSE, STOCKING LANE, RATHFARNHAM
66 Hampton Wood Square, Hampton Wood, St. Margaret's Road
12 Gandon Crescent, Gandon Park Lucan
189 Burnell Court, Northern Cross, Malahide Road
Apartment 280 Bushy Park House, Templeogue Road, Terenure
1 HOLYWELL LAWN, FELTRIM RD, SWORDS
2 THE COPSE, MILLERS GLEN, SWORDS
APT 8, TANDYS HALL, TANDYS LANE
53 Ballycullen Green, Old Court Road, Ballycullen
25 Dunsink Drive, Finglas
14 Blackberry Rise, Portmarnock
3A Bankside Cottages, Milltown
5 Bracken Close, Carpenterstown Road, Castleknock
Apartment 45, Millenium Tower, Charlotte Quay
23 Rivervalley View, Swords
40A Main Street, Belmayne, Balgriffin
11 GRACE PARK CLOSE, GRACE PARK WOOD, DRUMCONDRA
33 MALTON HOUSE, CUSTOM HOUSE SQ, MAYOR ST LOWER
21 Weavers Wood, Clonsilla
20 Millrace Court, Phoenix Park Racecourse, Castleknock
No 32 Thornwood, Beaumont
Block A3, The Quarter, Citywest
Apt 84, Block 5 Corr Castle, Howth Road
19 Goose Green, Grace Park Road, Drumcondra
1 THE AVENUE, MARNOCK'S BAY, PORTMARNOCK
57 Royston, Kimmage Road West
28 ST ANDREW'S WOOD, THE FAIRWAYS, LUCAN
5 APARTMENT BLOCK, OXMANSTOWN GREEN, BLACKHALL PLACE
10 THE COURT, CYPRESS DOWNS, TEMPLEOGUE
53 BELLISK, DONABATE, NORTH
17 Cedarhurst Road, Phoenix Park Racecourse, Castleknock
69 BRIGHTON ROAD, RATHGAR
30 The Green, Bellingsmore, Church Road
20 Raheen Drive, Ballyfermot
86 Ferrycarrig Road, Coolock
3 Beech Hall, Clay Farm, Leopardstown
54 Bracken Park Drive, Carpenterstown Rod, Castleknock
21 Barnwell Woods, Hansfield, Castleknock
64 BODEN HEATH, BALLYBODEN, RATHFARNHAM
26 Meadowbank Lane, Millers Glen, Swords
17 Somerton Way, Newcastle Road, Lucan
VERITAS HOUSE, DOMINICAN CONVENT, MUCKROSS PARK
9 FERNLEIGH VALE, Riverwood, Castleknock
221 Thomond Road, Ballyfermot
7 Fottrell Hall, Carpenterstown Road, Castleknock
13 COIS NA TRA, KILBUSH, RUSH
78 THE SQUARE, LARCH HILL, OSCAR TRAYNOR ROAD
36 THE FAIRWAYS, WOODBROOK GLEN, BRAY
20 CARMANHALL COURT, BURTON HALL RD, SANDYFORD
Apartment 176, Beechwood Court, Stillorgan
4 RALPH MEWS, NEWTOWNPARK AVE, BLACKROCK
135 Seafield Road East, Clontarf
17 Ardsolus, Old Naas Road, Brownsbarn
2 Somerton Place, Newcastle Road, Lucan
138 Malahide Marina Village, Malahide
4 Corbally Downs, Westbrook Glen, Saggart
16 Goose Green Court, Gracepark Road, Drumcondra
113 Bayside Boulevard North, Sutton
13 Ballygossan Park, Golf Links Road, Skerries
No. 12 Manor Hall, 1 1a Mount Brown, Kilmainham
4 Drumfinn Avenue, ballyfermot, dulin 10
44 CLONARD COURT, BALBRIGGAN
THE APARTMENT, 105 TREES ROAD, MOUNT MERRION
22 Maple Hall, Clay Farm, Leopardstown
10 Taylor Hill Boulevard, Balbriggan
92 Coultry Drive, Ballymun

Apartment 54 Third Floor, The Pavillion Roebuck Hill, Roebuck Road Mount Merrion
4 Proby Place, Proby Square, Carysfort Avenue
5 vauxhall street, balbriggan
4 Convent View Cottages, Ratoath Road, Cabra
36 Abbots Grove Park, Abbots Grove, Knocklyon
14 SWIFTBROOK CLOSE, FORTUNESTOWN, TALLAGHT
2 Dodderbrook Green, Dodderbrook, Oldcourt Road
4 Parkside Heath, Parkside, Malahide Road
31 Belmont Lawns, Stillorgan Road, Blackrock
13 Hampton Wood Way, Hampton Wood, Finglas
3 WATERSIDE GREEN, SWORDS RD, MALAHIDE
31 Delvin Banks, The Naul
6 Fottrell Hall, Carpenterstown Road, Castleknock
76 South Central, Rockbrook, Sandyford
Apartment 3 Block A, Mount Argus Apartments
Apartment 163, Grande Central, Rockbrook Sandyford
1 OAKTREE RD, STILLORGAN, BLACKROCK
61 THE FAIRWAYS, SEABROOK MANOR, PORTMARNOCK
The Whitehouse, Holmpatrick, Skerries
79 Bulfin Road, Inchicore
34 GLEN DR, THE PARK, CABINTEELY
52, The Drive, Semple Woods
9 Farmleigh Avenue, Farmleigh Woods, Castleknock
10 Taylor Hill Gardens, Naul Road, Balbriggan
8 Cuil Duin Walk, Cuil Duin, Citywest
12 Bellgree Heights, Tyrellstown
2 The Avenue, Aubrey Manor, Rathcoole
APARTMENT 83 FIRST FLOOR, NEW BANCROFT HALL, MAIN STREET TALLAGHT
13 Carnegie Avenue, North Main Street, Swords
48 The Grove, St. Marnock's Bay, Station Road
11 Stephen's Road, Inchicore
9 Castlegate Row, Adamstown Castle, Adamstown
25 Churchfields Close, Churchfields, Ashbourne
Apartment 28 Block 2, Highfield Court, Rathgar
41 The Lawn, Hansfield Wood
20 Grace Park Grove, Grace Park Wood, Grace Park Road
2 Barnwell Square West, Barnwell Square, Hansfield
APARTMENT 3, TURVEY GOLF CLUB, DONABATE
15 WILLSBOROUGH APTS, CLONSHAUGH ROAD, COOLOCK
10 Broadlands, Ballinclea Road, Killiney
35A METROPOLITAN APTS, INCHICORE ROAD, KILMAINHAM
4 SANDYFORD VIEW, SIMONS RIDGE, SANDYFORD
64 The Avenue, Scholarstown Wood, Scholarstown Road
35 Ashfield Avenue, Ridgewood, Swords
61 Ashley Hall, St. Edmunds, Palmerstown
3 Shackleton Grove, Shackleton Park, Lucan
25 Tynan Hall Park Tynan Hall, Kingswood, Ballymount Road Tallaght
7 Marrsfield Avenue, Belltree, Clongriffin
133 Wyckham Point, Wyckham Way, Dundrum
125 Fortunes Lawn, Citywest
21 LEALAND DRIVE, BAWNOGUE, CLONDALKIN
11 Somerton Grove, Newcastle Road, Lucan
3 PEMBROKE GROVE, IRISHTOWN
92 WHITEHALL ROAD WEST, PERRYSTOWN
14 Brookdale Avenue, Swords
Apt 15 Merton Hall, Mount Saint Anne's, MIlltown
12 Belltree Lane, Belltree, Clongriffin
103 Naas Road, Inchicore
13 Beechwood Heath, Beechwood, Hansfield
Apartment 72 Hanover Dock, Hanover Quay, Grand Canal Dock
Apartment 78 Pembroke Square, Upper Grand Canal Street
Unit 1 3 5 7 & 9, Rahillion Walk, Donabate
321 Argyle, Custom House Harbour, IFSC
124 Adamstown Avenue, Adamstown Square, Adamstown
4 Park Mews, Belltree, Clongriffin
APARTMENT 1, 4 MAIN STREET, HOWTH

139 Landen Road, Ballyfermot Road
3 THE FAIRWAYS, SEABROOK MANOR, PORTMARNOCK
49 Oaklands Avenue, Swords
11 Park House, Baldoyle
13a McKelvey Road, Finglas
MOLDOWNEY HOUSE, COAST RD, MALAHIDE
56 Weavers Hall, The Gallops, Sandyford
22A GATEWAY CRESCENT, GATEWAY, BALLYMUN
APT 80 RIDGE HALL, SHANGANNAH RD, BALLYBRACK VILLAGE
11 Somerton Drive, Somerton, Newcastle Road
13 AVOCA HALL, AVOCA PARK, BLACKROCK
3 Stillorgan Gate, Upper Kilmacud Road, Stillorgan
50 Barclay Court, Blackrock
12 HAMPTON PARK, ST HELEN'S WOOD, BOOTERSTOWN
37 Booterstown Wood, Booterstown Avenue, Booterstown
128 Glenageary Avenue, Glenageary
LARKFIELD, CORBALLIS, DONABATE
Apartment 7 Block A, Howth Junction Court, St.Donagh's Road
Apartment 18, 59 Pembroke Road, Ballsbridge
2 BROWNSBARN ORCHARD, OLD NAAS RD, KINGSWOOD
APt.49, New Row Square, New Row South Ward's Hill
8 Crosforge Close, Saggart
143 Charlestown Centre, St Margaret's Road, Finglas
6 ARD EADRAD APT, BLAKESTOWN ROAD, MULLHUDDART
43 BERESFORD AVENUE, DONABATE
6 Mount Prospect Drive, Clontarf
APARTMENT 90 CASTLE HALL, SWORDS CENTRAL, MAIN STREET SWORDS
24 Camden Lock, South Dock Road
68, Walkinstown Road, Walkinstown
68 RADCLIFF HALL, ST JOHNS ROAD, SANDYMOUNT
29 Idrone Drive, Knocklyon
1 HOLLYWELL, POPPINTREE PARK LANE WEST, POPPINTREE
45 Cuil Duin Way, Cuil Duin, Citywest
45 THE BELFRY, ENNISKERRY ROAD, STEPASIDE
6 oak lawn, royal oak, santry
apartment 4, Orwell Lodge, Rathgar
'Aurora', Mount Merrion Avenue, Blackrock
18 Graydon Crescent, Graydon, Newcastle
4 CARRICK COURT CLOSE, PORTMARNOCK
Apartment 13 Millrace View, Saggart
20 Abbey Street, Howth
2 BREMORE PASTURES WAY, HAMLET LANE, BALBRIGGAN
119 THE FORUM, BALLYMOSS RD, SANDYFORD
1b Ranelagh Mews, 1b Mount Pleasant Terrace, Ranelagh
12 BREMORE PASTURES DRIVE, HAMLETT LANE, BALBRIGGAN
13 The Avenue, Hansfield Wood
7 Bridge Avenue, Royal Canal Park, Ratoath Road Ashtown
LIOS NA GCLOCH, BEAVERSTOWN ROAD, DONABATE
APARTMENT 2, VERONA SEAFIELD ROAD, KILLINEY
No 11 Elder Heath Crescent, Elder Heath, Kiltipper Road
2 stanford, ardilea, clonskeagh
UNIT 35 FIFTH FLOOR, NORTHBANK, CASTLEFORBES RD
BALLYDOREEN, 201 ROCHESTOWN AVE, DUN LAOGHAIRE
9 THE STABLE YARD, BALBRIGGAN RD, SKERRIES
9 Goose Green, Gracepark Road, Drumcondra
26 Hampton Garden Drive, Balbriggan
8 Winton Avenue, Rathgar
28 Glen na Smól, Old Bawn, Tallaght.
APT 1, 10 CABRA ROAD, PHIBSBORO
4 Streamstown Cottages, Malahide
52 Offington Drive, Sutton
33 Belcamp Avenue, Priorswood
13 Castlechurch Park, Newcastle
54 Barnwell Woods, Barnwell Woods, Hansfield
157 BELGROVE PARK, MOUNT PROSPECT LAWNS, CLONTARF
35 ST JOHNS, PARK AVE, SANDYMOUNT

18 Brookfield, Malahide
LISANODE, 21 SANDYCOVE RD, SANDYCOVE
8 THE AVENUE, STATION MANOR, STATION ROAD
57 HAZELBURY GREEN, CLONEE, MEATH
14 BELLTREE AVENUE, CLONGRIFFIN
63 Belmayne Park Sout, Belmayne, Balgriffin
26 SEABROOK, BROOK LANE, RUSH
48 Fairhaven Walk, Castleknock Road, Castleknock
13 SLANE HOUSE, ARDILAUN COURT, PATRICK ST
29 Heywood Court, Northwood, Santry
41 Orwell Gate, Block A, Marianella
8 Kingsland, Donabate
14 The Way, The Paddocks, Hansfield
50 A Orwell Park, Rathgar
2 Claremont Avenue, Honeypark, Dun Laoghaire
10 Carrickmount Avenue, Rathfarnham
MARINO PARK HOUSE, MARINO PARK, BLACKROCK
Apartment 2, Phoenix Court, Infirmary road
343, Blackhorse Avenue
11a Grantham Place, South Circular Road
THE GABLES, KINSEALY LANE, KINSEALY
4 BERRYFIELD, FINNSTOWN PRIORY, LUCAN
29 ST UINCENTS PARK, TEMPLE HILL, BLACKROCK
3 ORCHARDTON MEWS, SWEETMOUNT AVE, DUNDRUM
16 Abbeywood Way, The Oaks, Lucan
11 The Northumberlands, Lower Mount Street
7 Heathfield Green, Cappagh Road, Finglas
79 CHERRYWOOD, LOUGHLINSTOWN, DUN LAOGHAIRE
99 ASHLEY HALL, ST EDMUND'S, PALMERSTOWN
1 DALTON MEWS, BACK RD, MALAHIDE
49 Waterside Drive, Swords Road, Malahide
15 old farm, carpenterstown
64C WALNUT CLOSE, KINGSWOOD, TALLAGHT
5 Barnwell Grove, Barnwell, Hansfield
6 Chapelhill, Chapelizod
13 Delhurst Mews, Ongar
8 Hamilton Walk, Ratoath Road, Royal Canal Park
7 Avoca Avenue, Blackrock
Stoney Lane, Rathcoole
Apartment 8, The Turnpike, Santry Cross
Apartment 23 Castlegate Hall, The Sentinel Building, Adamstown Lucan
4 The Lane, Millers Glen, Swords
APT 2 GARRYKNOCK BRACKEN PARK, CARPENTERSTOWN ROAD, CASTLEKNOCK
14 Inglenook Wood, Glenamuck Road, Carrickmines
21 PROSPECT AVENUE, PROSPECT MANOR, RATHFARNHAN
71 Palmerstown Green, Palmerstown
70 Kilkee House, Clare Village, Santry
1 Heathfield Close, Cappagh Road, Finglas
38 Brandon Road, Drimnagh
17, MOUNT AUBURN, KILLINE
2 MUILEANN COURT, KETTLES LANE, KINSEALY
34 WOODFORD, BREWERY RD, LEOPARDSTOWN
58 BOTANIC HALL, ADDISON PARK, GLASNEVIN
HIGHFIELD, DRYNAN, SWORDS
18 The Rise, Mount Merrion
80 Ivy Exchange, Parnell Street
11 Barnwell Gate, Barnwell Woods, Hansfield
2 Rowan House, Silverbanks, Baldoyle
34 Main Street, Belmayne, Balgriffin
3 St. Helen's Walk, St. Helen's, Adamstown
5 Hampton, Grace Park Road, Drumcondra
37 Hampton Wood Square, Hampton Wood, St. Margaret's Road
15 Grangeabbey, Abbey Road, Monkstown
75C Clontarf Park, Clontarf
47 St. Agnes Road, Crumlin
2 Cuil Duin Row, Citywest

31 Glenabbey Road, Mount Merrion
4 Parkside Way, Parkside, Malahide Road
13 A, Kenilworth Lane, Rathmines
70 Whitestown Avenue, Blanchardstown
2 YORK AVENUE, RATHMINES
Apartment 21 The Blake, Block 4 Level 6, Lansdowne Place
15 Hamilton Walk, Royal Canal Park, Ratoath Road
48dodsboro road, adamstown
apt 20 block d, doolin house, calre village
107 Fortunes Lawn, Citywest
6 Weavers Wood, Clonsilla
13 THe Walk, Bellingsmore, Kimartin
Apt 96 Abbot Court, Cualanor Upper Glenageary Road, Dun Laoghaire
32 Croftwell Square, School Road, Rathcoole
17 Shenick Road, Skerries
6 Shakelton Abbey, Lucan
35 Balally Avenue, Dundrum
4 PROBY PARK, BARNHILL RD, DALKEY
7 THE COURTYARD, HOLLYBANK AVE UPPER, RANELAGH
Cedar Lodge, Simmonsccourt Castle, Simmonscourt Road Ballsbridge
Apartment 5, Block 12 Station Way, Clongriffin
76 Heathfield Terrace, Finglas
49 Burnell Park Avenue, Castleknock
Apartment 5, 32 Mount Pleasant Square, Ranelagh
23 College Square, Wainsford Manor Drive, Terenure
7 The Dale, Citywest Village, Citywest
6 Hollywoodrath Green, Hollystown
5 Parklands Row, Parklands, CItywest
14 The Walk, Citywest Village, Citywest
4 Seabury Parade, Malahide
66 Burnell Green, Northern Cross, Malahide Road
19 BROADFIELD COURT, BROADFIELD MANOR, RATHCOOLE
164 TEMPLE COURT, NORTHWOOD, SANTRY
9 THE POPLARS, MONKSTOWN VALLEY, MONKSTOWN
6 Blackwood Park, Ongar Chase, Ongar
8 St. Patricks Cottages, Grange Road, Rathfarnham
32 THE HERON BUILDING, THORNWOOD, BOOTERSTOWN
129 Booterstown Avenue, Booterstown
86 THE ELMS, RIDGEWOOD, SWORDS
82 OTHERWISE KNOWN AS 81A, WOODLAWN PARK DRIVE, FIRHOUSE
22 The Way, The Paddocks, Hansfield
1A THE CAIRN, HAYSTOWN RD, RUSH
35 Seven Oaks, Drumcondra
33 Foxbourough Hill, Lucan
Apartment 21 Block D, Mount Argus Apartments
21 AMBER, THE GRANGE, STILLORGAN
154 RUSHEENY GROVE, RUSHEENY VILLAGE, CLONSILLA
6 SCHOOLHOUSE COURT, SCHOOLHOUSE LANE, SANTRY
12 LIFFEY CLOSE, LIFFEY VALLEY PARK, LUCAN
APARTMENT 119, BURNELL SQUARE, NORTHERN CROSS
18 CASTLE TERRACE COURT, ST MARGARETS AVE, MALAHIDE
7 Luttrellstown Oaks, Castleknock
Dunbo Lodge, Dunbo Hill, Howth
5 Muileann Place, Kettle's Lane, Kinsealy
12 Northwood Green, Santry
29 The Avenue, Carrickmines Green, Glenamuck Road
27 Northbrook Avenue, Ranelagh
23 SAINT MARGARETS AVENUE, RAHENY
54, Cedarview Northwood, Santry
206 Wyckham Point, Wyckham Way, Dundrum
38 THE WILLOWS, CLAREMONT RD, SANDYMOUNT
18 Temple Lawns, Northwood, Santry
10 The Paddocks Crescent, Adamstown
6 Brighton Place, Foxrock
18 The Rise, Millers Glen, Swords
19 ALDERWOOD PARK, SPRINGFIELD, TALLAGHT

40 Cardiffsbridge Avenue, Finglas
Apt 18 Pappan Grove, Northwood, Santry Demesne
25 The Gallops, Moorestown, Swords
'Woodside', 18 Shrewsbury Road
44 Beach Park, Portmarnock
403 THE EDGES, BEACON SOUTH QUARTER, SANDYFORD
3 Kennington Crescent, Templeogue
7 Hampton Crescent, Saint Helen's Wood, Booterstown Avenue Booterstown
35 Grande Central, Rockbrook, Sandyford
24 Daneswell Place, Botanic Road, Glasnevin
3 Barnwell Terrace, Hansfield, Dubin 15
44 THE GROVE, PHEASANT RUN, CLONEE
35 STILLORGAN HEATH, STILLORGAN
66 Annfield Drive, Annfield, Castleknock
Apt 93 Eustace Court, Cualanor, Upper Glenageary Road
2 Hampton Wood Lawn, Hampton Wood
3 Dodderbrook Way, Old Court Road, Ballycullen
84 ROSEHILL, CARYSFORT AVENUE, BLACKROCK
40 Hampton Park, St Helen's wood, Booterstown
3 2 BOYD HOUSE, HARBOUR RD, HOWTH
133 BROOKLAWN, STRANDVILLE AVE EAST, CLONTARF
150 PAIRC CHULAMBAIR, TEACH MEALOG, BAC 16
7 WALKER CLOSE, PARKSIDE, BALGRIFFIN
88 ROCKFIELD AVE, WHITEHALL RD
65 DUN SAITHNE CRESCENT, FLEMINGTON LANE, BALBRIGGAN
35 Adamstown Park, St Helens, Adamstown
22 Dargle Road, Drumcondra
9 upper prince edward terrace, blackrock
APT 9   49 ST. ANTHONY'S ROAD, HERBERTON, RIALTO
22 ORPHEN RISE, STILLORGAN GROVE, BLACKROCK
11 The Fulmar, Barnageeragh Cove, Skerries
2 Rogerstown Lane, Lusk
5 Cuil Duin Lawn, Cuil Duin, Citywest
25 Burford Drive, Honeypark, Dun Laoghaire
7 Woodside Demense, Aikens Village, Stepaside
18 Hallwell Drive, Hallwell, Adamstown
20 The Sycamore Elmfield, Ballyogan Road, Leopardstown
17 Baymeadows Avenue, Bay Meadows, Hollystown
21 Russell Avenue, East Wall
1 MOUNTTOWN VILLA, MOUNTTOWN RD LOWER, DUN LAOIGHRE
29 Lorcan Park, Santry
45 MELESIAN AVE, KINSEALY, SWORDS
20 Tullyhall Close, Lucan
Apartment 42 CastleparK Residence, Castlepark Road, Sandycove
Apt 9 Auburn Park, Auburn Avenue, Castleknock
57 College Sq, Wainsfort Manor Drive, Terenure
Apt. 401 Longboat Quay South, Hanover Quay
16 The Crescent, St. Marnock's Bay, Station Road
87 Thatch Road, Whitehall
39 St Martins Park, Kimmage
EDEN, 40 CASTLEPARK RD, DALKEY
23 Eustace Court, Cualanor, Upper Glenageary Road
28 Sion House, Tyrone Court, Thomas Davis Street West
2 Willmount Cottages, Glenalua Road, Killiney
BLOCK 5, 158 CUSTOM HALL, LOWER GARDINER ST
APT 2, 76 GROSVENOR SQ, LEINSTER RD
APT 409 STEPHENS HALL, EARLSFORT CENTRE, LR LEESON ST
25 Chapelwood Drive, Chapelwood, Hollystown
APT 21 LINDSAY HOUSE, BEAN COURT, PATRICK ST
3 PROBY PLACE, PROBY SQUARE, BLACKROCK
8 Roseland Ave, Cualanor, Dun Laoghaire
41 Longview Mews, Millers Glen, Swords
5 CHERRY LODGE, FARMLEIGH AVE, CASTLEKNOCK
347 Castlecurragh Heath, Blanchardstown
128 Cedar Place, Ridgewood, Swords
APT 60 THE BRAEBURN, THE OLD ORCHARD, LUCAN

34 Beechwood Heath, Beechwood, Hansfield
8 GLENALUA TERRACE, GLENALUA RD, KILLINEY
52 Ardilaun Court, Sybil Hill Road, Raheny
7 COLTHURST CRESCENT, HUNTINGTON GLEN, LUCAN
71 Ardilaun Court, Sybil Hill Road, Raheny
20 STANFORD, ARDILEA, CLONSKEAGH
24 Roseland Avenue, Cualanor, Dun Laoghaire
329 TRIMBLESTON, GOATSTOWN ROAD, GOATSTOWN
261 WYCKHAM POINT, WYCKHAM WAY, DUNDRUM
10 ESKER MEADOW GROVE, LUCAN, MEADOW GROVE
38 CEDAR PARK, RIDGEWOOD, SWORDS
14 TICKNOCK WAY, TICKNOCK HILL, SANDYFORD
49 Barnwell Place, Hansfield
DUNAENGUS, 1 SEAPOINT AVE, BLACKROCK
Apartment 17 Block B, Mount Argus Apartments
95 GRANITEFIELD MANOR, ROCHESTOWN AVE, DUN LAOGHAIRE
8 Goose Green Court, Gracepark Road, Drumcondra
9 Taney Manor, Dundrum
16 Old Court Avenue, Ferncourt, Ballycullen
7 burgage street, newcastle
APT 127 BLOCK C ARAS NA CLUAINE, WATERY LANE, CLONDALKIN
CARRIGFOYLE, MONKS MEADOW, COAST ROAD
3 John Dillon Street, Christchurch
4 Burleigh Mews, Burllington Road
48a Belmayne Avenue, Belmayne, Balgriffin
The Old Dispensary, Cabinteely
6 Boroimhe Aspen, Swords
63 The Bay, Elm Park, Merrion Road
35 FORT OSMOND, OLD COUNTY ROAD, CRUMLIN
7 ELDER HEATH COURT, KILTIPPER, TALLAGHT
LAUREL BANK, THORNHILL RD, BRAY
37 WALTON HALL, RIVERBANK, SWORDS
Apartment 13, 30 James's Walk, Rialto
16 Parkside Avenue, Parkside, Malahide Road
4 Heathfield Avenue, Cappagh Road, Finglas
apt 54, 20 christchurch place
6 The Green, Millers Glen, Swords
378 Pearse Street, Sallynoggin
35 Stewart Hall, Ryder Row, off Parnell Street
23 Phoenix Park Way, Phoenix Park Racecourse, Castleknock
Apartment Number 8, Cruise Park Hall, Tyrrelstown
29 THE GROVE, WOODBROOK GLEN, BRAY
16 Killart, Clonkeen Road, Cabinteely
26 Heywood Court, Northwood, Santry
79B Spiddal Road, Ballyfermot
16 wainsfort drive, terenure
4 Charleville Manor, Firhouse Road, Templeogue
45 The Elms, Ridgewood, Swords
FAIRWINDS, 53 QUAY ST, SKERRIES
15 Parkside Boulevard, Parkside, Malahide Road
18 Beresford Gardens, Beresford, Turvey Avenue
12 PARKLANDS PLACE, PARKLANDS, CITYWEST
11 Belclare Lawns, Poppintree
UNITS 1 - 28 BLOCK D, GANDON VIEW, LUCAN
26 Cabra Drive, Off Cabra Road
33 Muckross Green, Perrystown
85 Beechfield Avenue, Walkinstown
2 The Dale, Citywest Village
10 THE FORGE, LOWER MAIN ST, RUSH
12 Grace Park Close, Grace Park Wood, Grace Park Road
3 White Pines Dirve, White Pines Stocking Avenue, Rathfarnham
30 LAKELANDS CRESCENT, Stillorgan
Apt. 58 Block 3, Northwood Green, Santry
66 Parklands Place, Parklands, Citywest
4 Parc na Silla Lane, Loughlinstown
99 ABBEY PARK, BALDOYLE

25 GROVENOR HOUSE APTS, MONKSTOWN RD, MONKSTOWN
8 Rahillion Way, Donabate
78 The Steeples, Chapelizod
29 The Water Rill, Waterside, Malahide
16 THE GLEBE, ESKER, LUCAN
93 Slievenamon Road, Drimnagh
55 Griffith Avenue, Drumcondra
13 Carman's Court, Carman's Hall
6 THE DRIVE, STATION MANOR, PORTMARNOCK
3 THE WALK, ST MARNOCKS BAY, PORTMARNOCK
8 Deerpark Place, Kiltipper Way
29 CONNAWOOD GREEN, OLD CONNAUGHT AVE, BRAY
13 THE WALK, ROBSWALL, MALAHIDE
8 The Green, Kensington Lodge, Rochestown Avenue Dun Laoghaire
44 Ballynakelly Green, Crumlin
130 HOLYWELL, UPPER KILMACUD ROAD, DUNDRUM
33 BASKIN COTTAGES, BASKIN LANE, CLOGHRAD
15 Dun Emer Gardens, Lusk
26 Foxborough Road, Lucan
36 The Boulevard, Cruise Park, Tyrrelstown
6 HYDE HOUSE, CLONFADDA WOOD, MOUNT MERRION AVE
8 WARREN MANOR, THE WARREN, MALAHIDE
42 Clonsilla Road, Blanchardstown
36 WATERSIDE DRIVE, WATERSIDE SWORDS ROAD, MALAHIDE
APARTMENT 20, DEANSCOURT, CHRISTCHURCH
Apartment 30, Waterside Lawn, Waterside Swords Road
22 Parkview Green, Parkview Poppintree, Ballymun
8 Chancery Lane, 98 Bride Street
55 Lakelands Close, Kilmacud
CHIGWELL, CARRICKMINES
128 Botanic Road, Glasnevin
SPRINGFIELD HOUSE, SPRINGFIELD PARK, FOXROCK
APARTMENT NO 53, SMITHFIELD LOFTS, NORTH KING STREET
15 Grace Park View, Grace Park Wood, Grace Park Road
Apt 16 Lighthouse Apts, Church Road, East Wall
10 CRUISE PARK SQUARE, OAKLAND VILLAGE, TYRRELSTOWN
7 JAMES PLUNKETT HOUSE, GRATTON CRESCENT, INCHICORE
2 THE GREEN, SCHOLARSTOWN WOOD, RATHFARNHAM
38 Brennanstown Avenue, Brennanstown, Cabinteely
47 WOODSTOWN HEIGHTS, Knocklyon
39 THE OGHAM, GRANITEFIELD MANOR, ROCHESTOWN
10 HOLLYWOODRATH LAWNS, HOLLYSTOWN
9 Edenbrook Grove, Edenbrook, Citywest
15 AUBREY HOUSE APTS, QUINNS RD, SHANKILL
36 Richmond Square, Smithfield
11 Waterside Place, Waterside, Swords Road
17 Castleknock Dale, Castleknock
52 George's Avenue, Blackrock
Unit 9, Rathbourne Vale, Rathbourne
8 Cross Avenue, BOOTERSTOWN
13 Temple Hall, Mount St Annes, Milltown
38 Orwell Hall, Marianella, Rathgar
75 HARBOUR VIEW, CROFTON RD, DUN LAOGHAIRE
Apartment 13, 29 James's Walk, Herberton Rialto
Apt 133 Charlestown Centre, Charlestown Place, St Margaret's Road Finglas
57 Kennington Road, Templeogue
10 Cruise Park Court, Cruise Park
5 Glebe Crescent ors The Fair Green, Market Green, Balbriggan
15 Library Road, Shankill
149 Grange Lodge Avenue, Beaupark, Clongriffin
2 MOUNTAIN VILLA, CHURCH RD, KILLINEY
18 Athlumney Villas, Ranelagh
22 Elder Heath Drive, Kiltipper Road, Tallaght
7 LAUREL HOUSE, CARRICKMINES GREEN, GLENAMUCK RD CARRICKMINES
23 Somerton Way, Newcastle Road, Lucan
Talbot House, 137 Mount Merrion Avenue, Blackrock

Apartment 4 Charlotte Block, Honepark, Dun Laoghaire
16 The Walk, Skerries Rock, Skerries
15 Park Terrace South, Belltree, Clongriffin
127 NEW SESKIN COURT, WHITESTOWN WAY, TALLAGHT
41 Saddlers Court, Ardee Road, Rathmines
1 WHITECHAPEL ROAD, COOLMINE, CLONSILLA
58 Weavers Hall, The Gallops, Sandyford
17 Finnswood, Finnstown Cloisters, Lucan
39 Shrewsbury Square, Sandymount Ave, Sandymount
12 MONKSTOWN MANOR, MONKSTOWN FARM, MONKSTOWN
APT 27 THE AMBER, THE GRANGE, STILLORGAN
21 Whitechurch Avenue, Rathfarnham
15 Cuill Duin Lawn, Cuill Duin, Saggart
3 Green Park, Orwell Road, Rathagr
34 Bloomfield Park, Bloomfield Avenue, Donnybrook
20 New Row Place, New Row South
17 Waterside Drive, Swords Road, Malahide
182 THE GALLERY, TURVEY AVE, DONABATE
54 BLOCK D, ROEBUCK HILL, MOUNT ANVILLE ROAD
10 SHAWS LANE, BATH AVE, SANDYMOUNT
12 St Helens Grove, Adamstown, Lucan
20 Ballycullen Green, Oldcourt Road, Ballycullen
Plot E Drummans, Tobersool Lane, Balbriggan
1 THE CRESCENT, LARCH HILL OSCAR, TRAYNER ROAD
APT 8   BLOCK 4, BEAUPARK AVE, CLONGRIFFIN
1 Rossan Court, Waterville, Blanchardstown
4B Keadeen Avenue, Greenpark, Walkinstown
BRUNETTE, 40 ALBERT RD LOWER, SANDYCOVE
97 St. Jarlath's Road, Cabra
30 Meadowbank Road, Millers Glen, Swords
9 WATERSIDE WAY, SWORDS ROAD, MALAHIDE
3 Park Close, Gilford Road, Sandymount
26 CEDAR SQ, THE CEDARS, RIDGEWOOD
30 Beresford Crescent, Donabate
5 GLEBE CRESENT, MARKET GREEN, BALBRIGGAN
39 Taylor Hill Grange, Naul Road, Balbriggan
98 Adamstown Avenue, Adamstown Square, Adamstown
22 GLENDUN RD, LARKHILL, WHITEHALL
12 Iona Road, Glasnevin
42, Kilteragh Road, Foxrock
7 TANDYS HALL, TANDYS LANE, LUCAN
15 Robert St, Drumcondra
JASMINE LODGE, DOONSALLA DRIVE, DUN LAOGHAIRE
38 Halwell Road, Adamstown, Lucan
10 Vernon Avenue, Clontarf
5 Balgriffin Grove, Malahide Road, Balgriffin
4 DRAYTON CLOSE, MONKSTOWN RD, MONKSTOWN
3 pintail house, redcourt oaks, clontarf
21A Hole in the Wall Road, Donaghmede
6 Iona Villas, Glasnevin
62 THE BRAEBURN, THE ORCHARD, LEIXLIP RD
25 Bellevue Park, Booterstown
THE CHALET, 16 KILLINEY RD, KILLINEY
43 Park Court, Park Avenue, Sandymount
125 Grande Central, Rockbrook, Sandyford
58 CASTLELAND, BALBRIGGAN, PARK VIEW
9 Comyn lace, Drumcondra
Saol Nua, Ballyogan Road, Carrickmines
13 The Green, Skerries Rock, Skerries
25 LONGVIEW AVENUE, MILLERS GLEN, SWORDS
1 Castlegrange Lawn, Castaheany
1 Kitestown Road, Howth
2 peyton square, stoney lane, stoney road rathcoole
24 THE WATER RILL, WATERSIDE, SWORDS RD
14 FAIRLAWNS, SAVAL PARK RD, DALKEY
40 Meadowbank Road, Millers Glen, Swords

6 Barnwell Grove, Hansfield
65 Castleland Park Way, Balbriggan
52 Montrose Drive, Beaumont Park, Artane
21 Riverview, Oldbawn Bridge, Tallaght
21 THE SYCAMORES, GROVE RD, MALAHIDE
47 Cedar Grove, Ridgewood, Swords
56 The Paddocks, Adamstown, Lucan
25 Churchview Drive, Killiney
15 Knightsgate Crescent, Knightsgate, Rush
25 Obelisk Park, Blackrock
13 Seafield Crescent, Booterstown
11 GRIFFEEN WAY, ESKER SOUTH, LUCAN
10 Castleland Park Avenue, Balbriggan
6 ORCHARDSTOWN MEWS, SWEETMOUNT AVENUE, DUNDRUM
57 Belville Court, Johnstown Road, Cabinteely
1 Pembrin Wood, Beechpark Road, Foxrock
57 Merrion Grove, Stillorgan Road, Blackrock
9 The Avenue, Hazelbrook Square, Churchtown
247, Cruise Park Drive, Tyrrelstown
9 DRYNAM WALK, DRYNAM HALL, KINSEALY
12 properties at Baker's Yard, Portland Street North
11 Heathfield Way, Cappagh Road, Finglas
135 ARAS NA CLUAINE, YELLOW MEADOWS ROAD, WATERY LANE
APT 80, THE MOYLE, PROSPECT HILL
Forgney Oaks, Grove Avenue, Blackrock
80 CEDAR PLACE, RIDGEWOOD, SWORDS
13 Charnwood Park, Clonsilla
2 THE GATE, MILLERS GLEN, SWORDS
9 THE VIEW, ROBSWALL, MALAHIDE
47 Donnybrook Manor, Belmont Avenue, Donnybrook
21 PEMBROKE COURT, 75 PEMBROKE ROAD, BALLSBRIDGE
11 The Crescent, Scholarstown Wood, Scholarstown Road
2 Barnwell Road, Hansfield
36 Orwell Woods, Rathgar
11 Skylark Park Gate, St. Marnock's Bay, Station Road
34 Killeen Road, Ranelagh
15 17 18 and 19 ST COLUMCILLES DR, 1 1B 2 3 and 4 Seatown Road, Swords
16 THE WILLOWS, MONKSTOWN VALLEY, MONKSTOWN
124 BREGA, HAMLET LANE, BALBRIGGAN
APARTMENT NO 57, SMITHFIELD LOFTS, NORTH KING STREET
8 Barnwell Lawn, Barnwell, Hansfield
Apt 103, The Atrium Roebuck Hill, Mount Merrion
63 Parkview Road, Parkview Poppintree, Ballymun
COMMERCIAL PROPERTY AND, 423 RESI UNITS AT CLANCY QUAY, ISLANDBRIDGE
4 Magenta PLace, Glasthule
23 Beverton Heights, Beverton, Donabate
40 Woodview Park, Auburn Avenue, Castleknock
3 Shackleton Lodge, Shackleton Park, Lucan
28 MERRION PARK, HILL AVE SOUTH, BLACKROCK
78 St Peter's Terrace, Howth
APTS 1 2 & 4 GARVILLE HOUSE, GARVILLE AVENUE, RATHGAR
11 FAIRLAWNS, SAVAL PARK RD, DALKEY
10 Rathborne View, Rathborne Avenue, Ashtown
52 Oak Road, Donnycarney
15 SWAN HALL, BELGARD SQUARE, TALLAGHT
4 ELDER HEATH VALE, KILTIPPER, TALLAGHT
56 CHAPEL FARM AVE, CHAPEL FARM, LUSK
1 Orwell Court, Braemor Road, Churchtown
50 Parkside Crescent, Parkside, Malahide Road
68 BAILE NA LAOCHRA, POPPINTREE, BALLYMUN
16 Fairhaven Walk, Castleknock Road, Castleknock
Apartment 102 Leona Block, Honeypark, Dun Laoghaire
39 Lambay Road, Drumcondra
APT 25D, FERNLEIGH DR, CARPENTERSTOWN
Apt 26 Bridgefield, Northwood, Santry Demesne
410 Charlestown Centre, St Margaret's road, Finglas

Apartment 617 Cubes 7, Beacon South Quarter, Sandyford
4 CROFTON COURT, 7 CROFTON AVENUE, DUNLAOGHAIRE
31 Stapolin Avenue, The Coast, Baldoyle
ORMISTON, CHURCH RD, MALAHIDE
Apt 1 Adelaide House, 7/8 Haddington Terrace, Dun Laoghaire
11 The Tern, Barnageeragh Cove, Skerries
7 The Tern, Barnageeragh Cove, Skerries
39 Ashington Court, Navan Road
24 ROSELAND AVENUE, CUALANOR, DUN LAOGHAIRE
8 THE HALL, WENTWORTH COURT, TANDYS LANE
8 Seabrook, Brook Lane, Rush
51 The Avenue, Scholarstown Wood, Rathfarnham
71 Dunmore Lawns, Kingswood, Tallaght
30 LANA NA PAIRCE, BALLYOWEN LANE, LUCAN
22 Lohunda Park, Clonsilla
18 Beachfield Close, Walkinstown
15 Idrone Park, Knocklyon
Apartment 1 Castlegate Hall, The Sentinel Building, Adamstown Lucan
212 CLONKEEN RD, DEANSGRANGE, BLACKROCK
22 MONKSTOWN AVE, MONKSTOWN FARM, DUNLAOGHAIRE
22 Orwell Park, Rathgar
63 HOLYWELL VIEW, FELTRIM HALL, SWORDS
117 St Assams Avenue, Raheny
61, Cruise Park Drive, Tyrrelstown
51 Clonkeen Drive, Foxrock
5 Rossan Court, Waterville, Blanchardstown
164 THE GALLERY, TURVEY WALK, DONABATE
5 Kensignton Manor, Rochestown Avenue, Dun Laoghaire
NO.1A, OLD NAAS ROAD, BLUEBELL
APT 45 KILLWARDEN COURT, BOOT ROAD, CLONDALKIN
19 LES BUISSONNETS, SWEETMANS AVE, BLACKROCK
7 Odins Way, Taylors Lane, Rathfarnham
SITE NO. 31 ROYAL CANAL AVENUE, ROYAL CANAL PARK, RATOATH ROAD
4 The Hall, Scholarstown Wood, Scholarstown Road
124 EAST COURTYARD, TULLYVALE, CHERRYWOOD
CAPRI, ARD MHUIRE PARK, DALKEY
15 Park Drive, The Park, Cabinteely
33 Bishops Gate, Kiltiernan
Apartment 44, Grande Central, Rockbrook Sandyford
50 The Avenue, Semple Woods, Donabate
3 Red Arches Close, The Coast, Baldoyle
4 GOLDEN RIDGE WAY, SKERRIES RD, RUSH
2 Castlegrange Drive, Latchford, Castaheany
3 SYCAMORE GROVE, THE PARK, CABINTEELY
10 SEAGRAVE, THE LINKS, STATION RD
3 Vernon Mews, Vernon Avenue, Clontarf
14 ROWAN HOUSE, SILVERBANKS, THE COAST
74 CLONKEEN RD, DEANSGRANGE, BLACKROCK
30 Stanford Green, Walkinstown
15 Bannaville, Ranelagh
Apartment 88, St Margarets Road, Hampton Wood
11 Traverslea Wood, Lr Glenageary Road, Glenageary
54 Glenavon Park, Ballybrack
95 THE OAKS, RIDGEWOOD, SWORDS
12 Grace Park Grove, Grace Park Wood, Grace Park road
68 CASTLELAND PARK WAY, BALBRIGGAN
OBAN COTTAGE, BREMORE, BALBRIGGAN
23 Albany Road, Ranelagh
39 PAKENHAM, THE SLOPES, MONKSTOWN
24 Cruise Park Avenue, Cruise Park
49 Thornwood, Beaumont
8 DUNGRIFFIN VILLAS, HOWTH
52 BALLYCULLEN GREEN, OLD COURT ROAD, BALLYCULLEN
27 Block E3 Ennis House, Clare Village, Clare Hall
5 GLENSHANE DRIVE, BROOKFIELD, TALLAGHT
50 THE MILL, CROSSGUNS, PHIBSBORO

1 enderly, cunningham road, dalkey
Apt 30 Bridgefield, Northwood, Santry Demesne
6 Elder Heath Park, Elder Heath, Kiltipper Road
52 Cuil Duin Avenue, Edenbrook, Citywest
53 The Chandler Building, Rathborne Village, Ashtown
55 LAKELANDS CLOSE, UPPER KILMACUD ROAD, STILLORGAN
20 Trimbleston, Goatstown Road, Goatstown
32, DISWELLSTOWN WAY, CASTLEKNOCK
44 Greenville Terrace, South Circular Road
16 Foxborough Lawn, Lucan
30 BEECHFILED MEADOWS, CASTAHEANY, CLONEE
260 Galtymore Road, Drimnagh
36 BOYD HOUSE, MYRTLE AVENUE - THE COAST, BALDOYLE
43 LITTLEWOOD, BELARMINE, STEPASIDE
6 Barnhill Grove, Dalkey
3 GLEBE HALL, KILL AVE, DUN LAOGHAIRE
268 Beechwood Court, Stillorgan Road
83 ASHLEY HALL, ST EDMUND'S, PALMERSTOWN
138 Cremona Road, Ballyfermot
23 THE HEIGHTS, ROBSWALL, MALAHIDE
house 45, Drumnigh Manor, Portmarnock
49 Belarmine Grove, Belarmine Woods, Stepaside
39 Monastery Walk, Clondalkin
2 King Street, Clongriffin
6 Shackleton Glade, Shackleton Park, Lucan
Apartment 1B, Block 20, New Priory Hole in the wall road
1 THE OLD MILL, NAUL BRIDGE, NAUL
133 BREGA, HAMLET LANE, BALBRIGGAN
20 CLONFADDA WOOD, MOUNT MERRION AVE, BLACKROCK
Apartment 2, Castle Terrace Court Apartments, Malahide
4 BURFORD DRIVE, HONEYPARK, DUN LAOGHAIRE
HARWEN, ADELAIDE ROAD, GLENAGEARY
Apartment 96 Charlotte Block, Honeypark, Dun Laoghaire
43 Nutley Park, Donnybrook
39 THE KILNS, STATION RD, PORTMARNOCK
28 Parc Na Silla Rise, Loughlinstown
68 GREENVIEW APTS, SEABROOK MANOR, PORTMARNOCK
APT 13, HAZELBROOK, KILMACUD RD UPPER
42 Lower Kilmacud Road, Stillorgan
20 Vico Rock, Sorrento Road, Dalkey
16 SEFTON, ROSCHESTOWN AVE, DUN LAOGHAIRE
8 Elder Heath Green, Elder Heath, Kiltipper Road
58 OAK DOWNES, GREENPARK, CLONDALKIN
Cranwell, Rockville Crescent, Blackrock
Apt 807 The Gates, Beacon South Quarter, Sandyford
146 Rathborne Court, Ashtown
146 ADAMSTOWN AVENUE, ADAMSTOWN SQUARE, ADAMSTOWN
38 Foyle Road, Fairview
15 Taylor Hill View, Balbriggan
1 Ashwood Hall, Back Road, Malahide
Apartment 203 Bushy Park House, Templeogue Road, Terenure
38 MARTELLO COURT, PORTMARNOCK
26D FERNLEIGH DR, DISWELLSTOWN RD, CASTLEKNOCK
7 Willow Glen, Glenamuck Road, Carrickmines
96 Diswellstown Manon, Castleknock
5 Castlechurch Avenue, Newcastle
23 Corbally Avenue, Westbrook Lawns, Tallaght
1 Effra Terrace, Effra Road
33 ST MOCHTAS AVE, COOLMINE LODGE, CLONSILLA
29 Barnwell Grove, Hansfield, Barnwell
APT 11 CRUISE PARK SQUARE, TYRELLSTOWN
86 THE GALLERY, TURVEY WALK, DONABATE
45 ROCHDALE, HONEY PARK, KILL AVENUE
66 Airpark Avenue, Stocking Lane, Rathfarnham
79 THE GALLAN, GRANITEFIELD MANOR, ROCHESTOWN AVE
21 Coolnevaun, Stillorgan

22 Glendown Park, Templeogue
15 Clay Farm Grove, Clay Farm, Leopardstown
No.31 The Avenue, Robswall, Malahide
2 Cedarview, Northwood, Santry Demesne
53 FOREST BLVD, RIVERVALLEY, SWORDS
13 ROCHDALE, HONEY PARK, KILL AVENUE
9 MARINO PARK, MOUNT MERRION AVE, BLACKROCK
Kantara, Upr Albert Road, Glengeary
60 Crosforge, Saggart
MANDALA, DURHAM ROAD, SANDYMOUNT
Apartment 378 Block H, Castleforbes Square, Upper Mayor Street
11 THORMANBY COURT, THORMANBY RD, HOWTH
Interests in Grafton St, Sth William St Brittas Bay, Waterville Kerry
4 CARRIG COURT, FORTUNESTOWN LANE, SAGGART
Apt 6 Block A, Howth Junction Court, St. Donaghs Road
SCARABRAE, CORBAWN LANE, SHANKILL
453 Strand Road, Portmarnock
20 THE GROVE, HAYDENS PARK, LUCAN
Apt 439 Charlestown Centre, Charlestown Place, St Margaret's Road Finglas
APT 4 80 REUBEN STREET, RIALTO
61 MC CABE VILLAS, BOOTERSTOWN, BLACKROCK
24 Windmill Close, Rathcoole
24 SEA BROOK, BROOK LANE, RUSH
4 Brighton Square, Rathgar
51 Goatstown Road, Goatstown
NO 3 THE ORCHARD, CARPERSTOWN ROAD, CASTLEKNOCK
32 The Lawn, Hansfield Wood, Hansfield
58 Nutley Park, Donnybrook
69 Hampton Wood Square, Hampton Wood, St. Margaret's Road
67 Rosemount Estate, Dundrum
UNIT 1, 44 UPPER RATHMINES ROAD, RATHMINES
APT 12 THE OAKS, ROCKFIELD, DUNDRUM
5 Haydens Park Glade, Lucan
Site 39 Crinken Glen, Rathdown
33 Bay Meadows Avenue, Bay Meadows, Hollystown
28 Ailesbury Lawn, Dundrum
82 Diswellstown Manor, Castleknock
81 THE RILL, WATERSIDE, MALAHIDE
Slade View, Kilteel Road, Rathcoole
43 St. Davids court, Castle Avenue, Clontarf
28 Cuil Duin Green, Cuil Duin, Citywest
4 Shackleton Green, Shackleton Park, Lucan
Apt 218 Charlestown Centre, Charlestown Place, St. Margaret's Road Finglas
123 Mulvey Park, Dundrum
6 Newcourt Mews, Swords
Idrone Cottage, Newtown Avenue, Blackrock
APT 20 BLOCK B, THE GRANGE, STILLORGAN
64 Lombard Street West, South Circular Road
19 ABBOT DRIVE, CUALANOR, UPPER GLENAGEARY ROAD
5 Beechwood Gate, Beechwood, Hansfield
8 Allendale Lawn, Clonsilla
9 Bregia Road, Cabra
5 THE RISE, HUNTERS RUN, CLONEE
14 MOY GLAS RD, GRIFFEEN VALLEY, LUCAN
40 The Links, Elm Park, Merrion Road
5 Abbottstown House, Abbottstown Avenue, Finglas
39 The Grove, Hanstead, Lucan
26 MOUNT ROCHFORD RISE, FLEMINGTON LANE, BALBRIGGAN
52 The Terrace, Robswall, Malahide
17 Carriglea Gardens, Kill Avenue, Dun Laoghaire
26 Phoenix Park Avenue, Castleknock
10 Cherryfield Walk, Clonsilla
52 Forest Grove, Rivervalley, Swords
MAFRA, CHURCH RD, BALLYBRACK
7 ELEVEN ARCHES, GEORGES HILL, BALBRIGGAN
47 Whitethorn Road, Clonskeagh

47 Orwell Gardens, Rathgar
22 COMYN MANOR, SEATOWN ROAD, SWORDS
26 Graydon Crescent, Graydon, Newcastle
Apartment 21 Block E, Mount Argus Apartments
30 Milltown Avenue, Mount St. Annes, Milltown
25 The Quarry, Portmarnock
3 OLDFIELD HALL, TICKNOCK HILL, SANDYFORD
27 Hamilton Hill, Barnageeragh Cove
10 ASPEN RD, KINSEALY, SWORDS
34 Wainsfort Park, Terenure
7 Brook House, Corrig Avenue, Dun Laoghaire
4 THE OAKS, RIDGEWOOD, SWORDS
5 Wingifeld, Corke Abbey, Bray
819 Ratoath Road, Finglas
2 THORNLEIGH PARK, APPLEWOOD VILLAGE, SWORDS
33 Cuil Duin Lawn, Cuil Duin, Citywest
16 The Grove, Orlynn Park, Lusk
Apartment 437, Castleforbes Square, Mayor Street
23 Culmore Road, Palmerstown
72 DRYNAM HOUSE, DRYNAM DR, DRYNAM HALL
56 Eustace Court, Cualanor, Dun Laoghaire
77 Oak House, Silverbanks, The Coast
33 Myra Close, Emmet Road, Inchicore
57 HAZELWOOD, SANTRY AVE
26 Heathfield Walk, Cappagh Road, Finglas
APARTMENT 8, SHERBOURNE, AUNGIER STREET
69 The Oval, Tullyvale, Cabinteely
8 BOYD AVENUE, HONEYPARK, DUN LAOGHAIRE
14 The Copse, Millers Glen, Swords
10 The Lawn, Hansfield Wood, Hansfield
Apartment Number 7, Cruise Park Hall, Tyrrelstown
VALENCIA, 30 CARRICKBRACK RD, HOWTH
6 Kilakee Court, Firhouse
BEACH LANE, THE BURROW, PORTRANE
43 CARRIGMORE GREEN, CITYWEST, SAGGART
12 Ossory Road, North Strand
Apartment No.136 Abbot Court, Cualanor, Upper Glenageary Road
8 Stratton Way, Adamstown Square, Adamstown
37 BOOTERSTOWN WOOD, BOOTERSTOWN AVE, BOOTERSTOWN
APT 9, GANDON COURT, LOWER MAIN ST
3 HARBOUR COURT, GEORGES PLACE, DUN LAOGHAIRE
42 HEATHERVIEW PARK, AYLESBURY, TALLAGHT
39 The Terrace, Robswall, Malahide
21 Hazel Avenue, Stillorgan
75 DONOMORE AVENUE, TALLAGHT, .
33 Hamilton Hill, Barnageeragg Cove, Skerries
1 THE GREEN, WOODBROOK GLEN, BRAY
23 CORBALLY VALE, BLESSINGTON RD, SAGGART
7 The Brokerage, Townsend Street
92 BURNELL COURT, NORTHERN CROSS, MALAHIDE ROAD
GLANDORE MEWS, GLANDORE PARK, MOUNTTOWN LOWER
82 RIDGE HALL, SHANGANAGH RD, BALLYBRACK
APT 3, 4 GRANGE LODGE AVE, CLONGRIFFIN
21 Templemanor Grove, Walkinstown
21 Ravensdale park, Kimmage
9 Fairhaven Road, Castleknock Road, Castleknock
46 Shanglas Road, Beaumont
APARTMENT 4, OLD KILMAINHAM SQUARE, KILMAINHAM
Apt 14, Hollybrook Manor, Clontarf
17 GLENAAN ROAD, LARKHILL, WHITEHALL
14 Killary Grove, The Donahies, Donaghmede
2 St Marys Place, St Marys Road, Howth
2 Harolds Bridge Court, HArolds Cross
85 Olcovar, Shankill
13 ros mor view, scholarstown road, knocklyon
18 The Green, Millers Glen, Swords

8 Dun na Rí, Swords Road, Malahide
12 ESKER PINES, ESKER LANE, BALLYDOWD
75 La Rochelle, Christchurch
8 GRANGECRESCENT, POTTERY RD, DUNLAOGHRE
APT 2C, BELARMINE WAY, STEPASIDE
2 VERSCHOYLE CRESCENT, SAGGART ABBEY, BLESSINGTON RD
39 Glebe Hall, Kill Avenue, Dun Laoghaire
16 Silchester Park, Glenageary
Apartment 11 Wingfield, Enniskerry Road, Stepaside
2 Willow House, Silverbanks The Coast, Baldoyle
34 Ardilea Cresent, Heidelberg, Ardilea
13 Somerton Grove, Newcastle Road, Lucan
Apartment 20 Cubes 1, Beacon South Quarter, Sandyford
22 Parklands Boulevard, Parklands, Citywest
8 Shackleton Lawn, Shackleton Park, Lucan
9 Elder Heath Way, Elder Heath Kiltipper, Tallaght
13 DRYNAM RISE, DRYNAM HALL, KINSEALY
17 O'Neachtain Road, Drumcondra
APT 3, THE HERON BUILDING THORNWOOD, BOOTERSTOWN
20 Maple Hall, Clay Farm, Leopardstown
18 Ballyshannon Road, Coolock
APT 12, 4 HERBERTON STREET, HERBERTON
2 Cairnfort Green, Cairnfort, Stepaside
8 LIFFEY VALE, LIFFEY VALLEY PARK, LUCAN
Apartment 57, Grand Canal Wharf
20 Castleland Park Avenue, Balbriggan
BLOCK 9 APT 26 THE GRIFFITH, PROSPECT HILL, TOLKA VALLEY ROAD FINGLAS
12 Rosanule, Phoenix Park Racecourse, Castleknock
4 Dunville Court, Dunville Avenue, Ranelagh
62 Roseland Avenue, Cualanor
56 Stratton Walk, Adamstown Square, Adamstown
21 THE STARLINGS, SHANGANAGH ROAD, SHANKILL
41 Rossan Court, Waterville, Blanchardstown
14 Stocking Wood Hall, Stocking Avenue, Rathfarnham
5 THE LODGE, GOLDEN RIDGE, SKERRIES RD
ARD NA RI, 44 CARRICKBRACK RD, BAILY
17 Martello Road, Martello Tower, Balbriggan
9 GARLAND COURT, CHURCH ST, HOWTH
22 Cedarhurst Road, Phoenix Park Racecourse, Castleknock
2 Bolton Avenue, Bolton Park, Rathfarnham
26 HOWTH LODGE, HOWTH RD, HOWTH
6 THE GREEN, WOODBROOK GLEN, BRAY
5 The Place, Citywest Village, Citywest
Apartment 22 Lighthouse Apartments, Church Road, East Wall
3 BELVILLE COURT, JOHNSTOWN RD, DUN LAOGHAIRE
27 The Crescent, The Paddocks, Adamstown
3 PARK RD, GLENAGEARY HEIGHTS, GLENAGEARY
Apt 34 Abberley Square, High Street, Tallaght
APT 58, SMITHFIELD MARKET, BLOCK A SMITHFIELD
13 Somerton Copse, Newcastle Road, Lucan
Apartment 10 Park Lodge, 7-9 North Circular Road
Apartment 10 Block A, Mount Argus Apartments
5 Barnwell Square Sth, Barnwell Square, Hansfield
15 Hawthorn Walk, Bird Avenue, Clonskeagh
55 Castle Elms Court, Greencastle Road, Coolock
5 Ardsolus, Old Naas Road, Brownsbarn
Apartment 45, First Floor - Block A, Belvoir
11 Dodderbrook Terrace, Oldcourt Road, Ballycullen
APT 177, THE SWEEPSTAKES, BALLSBRIDGE
3 Laurence Court, Harolds Cross
80 THE ATRIUM, BEECHWOOD COURT, STILLORGAN
55 Beauvale Park, Beaumont
3 BROOKFIELD, MOUNT MERRION AVE, BLACKROCK
50 ANLEY COURT, ESKER LANE, LUCAN
28 Pintail, Redcourt Oaks, Clontarf
16 Knockcree, Glenamuck Road, Cafrrickmines

ST JUDES, COMMONS RD, SHANKILL
8 ELEVEN ARCHES, GEORGES HILL, BALBRIGGAN
THE PADDOCK, 39 SANDYFORD DOWNS, SANDYFORD
28 St. Edmunds Grove, Loman's Road
FAIRWAYS, 18 FOSTERS AVE, BLACKROCK
Unit 58 Main Street, Belmayne, Balgriffin
Apartment 6   Block 1, New Priory, The Hole in the Wall Road
55 Ahhfield Avenue, Ridgewaod, Forest Road
43 The Drive, Hunters Run, Clonee
126 Royal Canal Court, Royal Canal Way, Ashtown
Apt 15 New Bancroft Place, Main Street, Tallaght
APT 1 HOUSE 11, LINDEN SQUARE, GROVE AVENUE
33 The Park, Skerries Rock, Skerries
158 Castle avenue, Clontarf
Apartment 48, Second Floor - Block A, Belvoir
16 The Drive, St Marnock's Bay, Station Road
12 ROSEBANK VIEW, NANGOR ROAD, CLONDALKIN
6 Parklands Downs, Saggart
93 Culmore Road, Palmerstown
HARROW HOUSE, BALURE LANE, KILLINEY
64 Beech Park, Cabinteely
28 Grangeabbey, Monkstown
25 Willow Glen, Glenamuck Road, Carrickmines
17 Seagrave Close, Citygate, St. Margarets Rd
8 Citywest Village Avenue, Citywest Village, Citywest
APT 49 BRIGHTON, CASTLE COURT KILGOBBIN WOOD, SANDYFORD
2 Edenbrook Close, Edenbrook, Citywest
PEARSE MEWS, NEWBRIDGE AVENUE, SANDYMOUNT
Apartment 5 Hacienda, Wynnsward Drive, Clonskeagh
9 GRANITE HALL, SANDYCOVE, DUN LAOGHAIRE
5 THE HEIGHTS, ROBSWALL, MALAHIDE
176 THE GALLERY, TURVEY WALK, DONABATE
13 Parkside Way, Parkside, Malahide Road
4 Saddler's Avenue, Mulhuddart
6 Longview Square, Miller's Glen, Swords
11 Berwick Lawn, Swords
1 ALLEN PK RD, LOWER KILMACUD RD, STILLORGAN
21 BALRUDDERY WOOD, BALROTHERY, BALBRIGGAN
237 Clonliffe Road, Drumcondra
155 THE MEADOWS EAST, BELGARD HEIGHTS, TALLAGHT
60 THE FAIRWAYS, SEABROOK, PORTMARNOCK
13 Walker Green, Malahide Road, Balgriffin
82 NEWHAVEN BAY, FLEMINGTON LANE, BALBRIGGAN
22 THE VIEW, ROBSWALL, MALAHIDE
8 Broadfield Grange, Broadfield Manor, Rathcoole
St. Mary's, Victoria Road, Dalkey
11 Coill Dubh Grove, Malahide
Apartment 31 Temple Hall, Mount St. Annes, Milltown
APT 70, MILLBANK THE LINKS, STATION RD
35 Castle Riada Grove, Lucan
16 TRAMWAY COURT, OLD BLESSINGTON ROAD, TALLAGHT
Apt 14, 30 St James Walk, Rialto
27 Grace Park Grove, Grace Park Wood, Grace Park Road
KILMORE, KILLINEY ROAD, KILLINEY
Lucca House, 2 Ailesbury Way, Ailesbury Road
THE GABLES, BARNHILL RD, DALKEY
1 Broadstone Avenue, Phibsboro
21 Belarmine Way, Stepaside
1 Foxhill Green, Donaghamede
13 Shackleton Gate, Shackleton Park, Lucan
Apt 7 Pappan Grove, Northwood, Santry Demesne
59 Belville Court, Johnstown Road, Cabinteely
175 ROSELAWN ROAD, CASTLEKNOCK
1 Muileann Mews, Kettle's Lane, Kinsealy
22 Grace Park Way, Grace Park Wood, Grace Park Road
IVY HOUSE, MAIN STREET, RATHCOOLE

15 The Laurels, Clontarf
24 The Green, Bellingsmore, Church Road
2 Liffey Walk, Lifey Valley Park, Lucan
Apartment 7 Block C, Mount Argus Apartments
14 Farmhill Road, Goatstown
33 ORWELL HALL, MARINELLA, RATHGAR
24 Haydens Park Grove, Lucan
11A, HOME VILLAS, DONNYBROOK
135 Templeville Drive, Templeogue
41 Belltree Avenue, Belltree, Clongriffin
42 STRATTAN WALK, ADAMSTOWN SQUARE, ADAMSTOWN
Apartment 7 Block C, Smithfield Market, Smithfield
52 Templeville Drive, Templeogue
2 CASTLELANDS, HYDE RD, DALKEY
22 Ring Terrace, Inchicore
38 Kilmore Avenue, Coolock
24 BROOKDALE WAY, RIVERVALLEY, SWORDS
19 FOXFORD, BALLYOWEN LANE, LUCAN
11 CHURCH SQ, CHURCH RD, EAST WALL
11 Ashe Lane, Knocksedan Demesne, Swords
11 MAPLE HALL, MOUNT SAINT ANNES, MILLTOWN
APARTMENT 63 LEONA BLOCK, HONEYPARK, DUN LAOGHAIRE
67 THE TALLOW, RATHBORNE VILLAGE, ASHTOWN
12 CLONUSKE GREEN, HAMLET LANE, BALBRIGGAN
40 HOLYWELL WOOD, GORSE HILL, SWORDS
Apartment 178 Neptune Block, Honeypark, Dun Laoghaire
45 Ardilaun Court, Sybil Hill Road, Raheny
10 Cruise Park Drive, Cruise Park
APT 307, BLOCK 6, PARKLANDS
9 The Hall, Scholarstown Wood, Scholarstown Road
19 Elder Heath walk, elder Heath, Kiltipper Road
16 Belmont Drive, Aiken's Village, Stepaside
26 THE CRESCENT, ROBSWALL, MALAHIDE
Apartment 11 Watermark Building, Block C Royal Canal Park, Ashtown
ST EDWARD, 20B ADELAIDE ST, DUNLAOGHAIRE
7 Arundel, Monkstown Valley, Monkstown
Apt 28 Laurel House, Carrickmines Green, Glenamuck Road
The Old Lodge, Thormanby Road, Howth
11 Blackwood Crescent, Ongar Chase, Ongar
11 Parkside Boulevard, parkside, malahide road
FERNDALE, AVOCA AVE, BLACKROCK
32 THE OAKS, ELMFIELD, LEOPARDSTOWN
APT 2 SUTTON CASTLE, SHIELMARTIN ROAD, SUTTON
Apartment No.24, Abelard Square, Phoenix Park Racecourse
21 Whitebeam Hall, Clay Farm, Leopardstown
13 Linnett, Barnageeragh Cove, Barnageeragh Skerries
12 ROCHDALE, HONEY PARK, KILL AVENUE
22 Meadowbank Green, Millers Glen, Swords
54, The Drive, Semple Woods
6 Redwood Drive, Kilnamanagh, Tallaght
4 LIFFEY RD, LIFFEY VALLEY PARK, LUCAN
62 Ardlea Road, Artane
24 Parklands Boulevard, Parklands, Citywest
28 Weavers Hall, The Gallops
7 FERRYCARRIG DRIVE, FAIRFIELD, COOLOCK
APPARTMENT 2, ROEBUCK PARK HOUSE, LARCHFIELD PARK GOATSTOWN
16 Park Avenue, Millers Glen, Swords
14 Knightsgate Walk, Rush
Apartment No.161 Abbot Court, Cualanor, Upper Glenageary Road
The Cottage, Walshestown
74 HANNAH SQR, ST EDMUNDS, PALMERSTOWN
APT 50, AVENDALE HOUSE, SEVEN OAKS
13 Taney Green, Taney Road, Dundrum
16 Bishops Gate, Enniskerry Road
4 ARUNDEL, MONKSTOWN VALLEY, MONKSTOWN
Apt 227 Charlestown Centre, Charlestown Place, St. Margaret's Road Finglas

4 St Esra Close, Killester Avenue, Killester
1 Shackleton Lawn, Shackleton Park, Lucan
No 7 An Riasc, Farnham Drive, Finglas
Annaville Residence, Dundrum Road
8 Heathfield Cresent, Cappagh Road, Finglas
5 BARNWELL SQUARE WEST, HANSFIELD SQUARE
231 Palmerstown Woods, Clondalkin
68 Cherryfield Road, Walkinstown
Finlieve, Hoar Rock, Skerries
APT 61 BRIGHTON WOOD, BRIGHTON ROAD, FOXROCK
10 BALLYCULLEN GREEN, OLDCOURT ROAD, BALLYCULLEN
46 Fairhaven Avenue, Castleknock Road, Castleknock
2 LARAGH, HAMPTON ST, BALBRIGGAN
49 Kildonan Drive, Finglas
NO 7, THE TAP APARTMENTS, KING STREET
APT 77, STEWART HALL, PARNELL ST
8 HAYDENS PARK VIEW, LUCAN, PARK VIEW
24 St Pancras Avenue, Mount Tallant Avenue, Terenure
174 Oakcourt Drive, Palmerstown
14 South Central, Rockbrook, Sandyford
62B Strand Street, Skerries
11 LIFFEY VALLEY AVE, LIFFEY VALLEY PARK, LUCAN
FINISTERE, BALLYGIHEN AVE, SANDYCOVE
NIRVANA, SUNDRIVE RD, RUSH
82 Holmpatrick, Skerries
3 TOWER HALL, SWORDS CENTRAL, MAIN ST
38 Lambay Road, Drumcondra
apt 7 the village, stepaside
8 Rathmill Glen, Rathmill Manor, Rathcoole
Apartment 3, The Brickworks, Station Road
5 Maple Lodge, Farmleigh Woods, Castleknock
12 The Fulmar, Barnageeragh Cove, Skerries
12 GANDON LANE, KINSEALY
36 HEATHFIELD WAY, CAPPAGH ROAD, FINGLAS
APT 75 KNOCKMAREE, ST. LAURENCES ROAD, CHAPELIZOD
3 Condor, Church Road, Killiney
Apartment 34, The Village, Stepaside
Apt 229 Charlestown Centre, Charlestown Place, St. Margaret's Road Finglas
23 The Rise, Skerries
2 Shelton Grove, Kimmage
24 CLYDE HOUSE, SERPENTINE AVENUE, BALLSBRIDGE
6 Yeats Hall, Carrickmines Wood, Brennanstown Rd
Apt18 Abbot Court Cualanor, Upper Glenageary Road, Dun Laoghaire
14 The Dale, Citywest Village, Citywest
112 Carrs Mill, Portrane Road, Donabate
82B Aras Na Cluaine, Watery Lane, Clondalkin
Apt 8, Three Rock Grove, Harolds Grange Road Rathfarnham
29 Farmleigh Avenue, Stillorgan
8 BOROIMHE ASPEN, FOREST RD, SWORDS
5 SYNGE COURT, NEWTOWN VILLAS, CHURCHTOWN
THE COACHOUSE, 288 BANNOW RD, CABRA
Apartment 509, Block B1 Beacon South Quarter, Sandyford
25 hollyville lawn, palmerstown
12 OWENDOHER LODGE, BALLYBODEN RD, RATHFARNHAM
34 Achill Square, Waterville, Blanchardstown
Apartment 87 C remorne, Greenmount Road, Terenure
APT 1, 82 NORTHUMBERLAND RD, BALLSBRIDGE
64 Clanree Road, Donnycarney
14 Graydon Glen, Graydon, Newcastle
1 Lambourne Village, Clontarf
47 Marrsfield Avenue, Belltree, Clongriffin
15 THE WILLOW, ROCKFIELD, DUNDRUM
Apartment No.26, Abelard Square, Phoenix Park Racecourse
COISLAN, BALLINASCORNEY, BRITTAS
9 Fforster Terrace, willsbrook road, lucan
76 TREES ROAD, MOUNT MERRION

Property at Townsend Street, Bracken's Lane and Moss Street
Apt. 16 Dunluce, 21 Anglesea Road, Ballsbridge
Apt 8 Clonbern, Phoenix Park Racecourse, Castleknock
81 HANNAH SQUARE, ST EDMOUNDS, PALMERSTOWN
32 CASTLE COURT, BOOTERSTOWN, BLACKROCK
16 Villa Park Road, Navan Road
BROADMEADOW, LOWER COMMONS, GARRISTOWN
112 Shrewsbury Lawn, Cabinteely
21 Edenbrook Green, Edenbrook, Citywest
14 LIFFEY AVE, LIFFEY VALLEY PARK, LUCAN
76 ROSSE COURT HEIGHTS, BALGADDY RD, LUCAN
120, Upper Leeson Street
27 Belarmine Walk, Belarmine Woods, Stepaside
9 St. Helen's Mews, St. Helen's, Adamstown
CHAGFORD, 61 ST MARGARETS RD, MALAHIDE
4 Belmayne Avenue, Belmayne, Balgriffin
APARTMENT 2 BLOCK 32, MARRSFIELD AVENUE, CLONGRIFFIN
6 Rathborne Park, Rathborne, Ashtown
3 Adelphi Manor, Upper Georges Street, Dun Laoghaire
21 Villa Park Gardens, Navan Road
19 Hollywoodrath Close, Hollywoodrath, Hollystown
62 Bracken Park Drive, CarpenterstownRoad, Castleknock
92 Park Avenue, Swords
6 Slade Castle Avenue, Saggart
APT 20 TURNSTONE BLOCK B, THORNWOOD BOOTERSTOWN AVE, BOOTERSTOWN
57 ELMFIELD COURT, NINTH LOCK RD, CLONDALKIN
APARTMENT 10 PACKENHAM HOUSE, DEAN COURT PATRICK STREET
45 Mulvey Park, Dundrum
6 Loftus Square, Rathfarnham Gate, Rathfarnham
71 RADCLIFF HALL, ST JOHNS ROAD, SANDYMOUNT
60 KIMMAGE GROVE, KIMMAGE
17 The Walk, Citywest Village, Citywest
30 The Grove, Melrose Park, Kinsealy
Apartment 605 Cubes 6, Beacon South Quarter, Sandyford
36 Rialto Street, Rialto
1 Hamilton Hill, Barnageeragh Cove, Skerries
Apartment No. 66 St. Steevens Gate, 126/133 James Street
87 Ashfield Avenue, Ridgewood, Swords
7 Barnwell Walk, Hansfield, Barnwell
2 Ashfield Grove, Ridgewood, Swords
7 Gandon Green, Gandon Park, Lucan
Apt. 1 Eustace Court, Cualanor, Upper Glenageary Road
45 Castlegate Park, Adamstown Castle, Adamstown
ST GERRARDS, CORBALLIS, DONABATE
17 REDBERRY, FINNSTOWN PRIORY, LUCAN
1Glenbrook Park, Rathfarnham
3 Claremont Road, Sandymount
Apartment No. 54 & Car Space 54, The Forum, Sandyford
27 Eaton Brae, Orwell Road, Rathgar
46 TRIMLESTON, HAMLET LANE, BALBRIGGAN
6 THE ACADEMY BUILDING, PARKWEST POINT, PARKWEST
26 Parkside Boulevard, Parkside, Malahide Road
35 Shanglas Road, Beaumont
48 IONA, PROSPECT HILL, FINGLAS ROAD
Apartment 114, MILLTOWN HALL, MOUNT SAINT ANNES
3 Shackleton Wood, Shackleton Park, Lucan
11 Dodderbrook Rise, Old Court Road, Ballycullen
17 Killakee Lawns, Firhouse
53 Strand Road, Baldoyle
1 Somerton Way, Newcastle Road, Lucan
47 BREMORE PASTURES CRESCENT, BREMORE PASTURES, BALBRIGGAN
UNIT NO. 29 2ND FLOOR, NORTHBANK, CASTLEFORBES ROAD
46 THE HAMPTON, ST. PAPPINS STREET, BALLYMUN
13 Somerton Avenue, Newcastle Road, Lucan
39 Longview Mews, Millers Glen, Swords
Market Square, Bailieborough

26 Merton Avenue, South Circular Road
32 LINNETT, BARNAGEERAGH COVE, SKERRIES
10 LIFFEY DALE, LIFFEY VALLEY PARK, LUCAN
65 Brookdale Lawns, Rivervalley, Swords
7 Drumanagh Court, Kilbush Lane, Rush
2 Barnwell Square East, Barnwell Square, Hansfield
Alliance Building, The Gasworks
Apt 5 New Bancroft Centre, Main Street, Tallaght
Unit 100 Royal Canal Court, Royal Canal Way, Ashtown
3 Darley Lane, Belmont, Aikens Village
33 Seafield Court, Killiney
11 Rokeby Park, Lucan
Apartment No.169 Abbot Court, Cualanor, Upper Glenageary Road
161 Hampton Wood Road, Finglas
1 Waterfall Court, 145/147 Richmond Road
4 Rathborne Walk, Royal Canal Park, Ratoath Rd Ashtown
Apt 1 Auburn Park, Auburn Avenue, Castleknock
Apartment 3 Rexton Court, Sandymount Avenue, Ballsbridge
ARDMORE, LOWER GLENAGEARY RD, DUN LAOGHAIRE
33 HOLYWELL SQ, FELTRIM HALL, SWORDS
40 MELVILLE VIEW, CITYSIDE, FINGLAS
18 Rathgar Park, Rathgar
14 Glenfarne Road, Raheny
Apartment 46 The Anchorage, Crofton Road, Dun Laoghaire
15 EATON GREEN, MAIN ST, RATHCOOLE
3 MOUNT SANDFORD, MILLTOWN ROAD, MILLTOWN
Apartment 14, 43 St. Anthony's Road, Rialto
Apt . 37 The garden House, Waterfall Avenue, Drumcondra
15 THE HAMPTON, SANTRY CROSS, BALLYMUN
14 COLLEGE VIEW, SILLOGE RD, BALLYMUN
'Avoca' 55 Hainault Rd, Foxrock
116 Grattan Lodge, Hole-in-the Wall Road, Donaghmede
34 Rockford Manor, Stradbrook Road, Blackrock
9 Belarmine Walk, Belarmine Woods, Stepaside
70 CIANLEA, RATHBEALE RD, SWORDS
6 ROSSE COURT GROVE, BALGADDY RD, LUCAN
16 Slade Castle Walk, Saggart
14 DUN EMER PLACE, DUN EMER, LUSK
9 ST HELENS NORTH, MARINE PARADE, SANDYCOVE
Apt 19, 64 Mountjoy Sq.
1 Werburgh's Court, Drynam Road, Swords
68 Offington Park, Sutton
14 SWANWARD COURT, PARNELL RD, HAROLDS CROSS
1 Rathmill Park, Rathmill Manor, Rathcoole
8 DE VESCI HOUSE, LONGFORD PLACE, MONKSTOWN
15 Hollywoodrath Rise, Hollystown
3 SUMMERHILL PLACE, OFF RUTLAND STREET LOWER
66 LEVMOSS AVENUE, LEVMOSS PARK, THE GALLOPS
1 Seapark Hill, Malahide
APT. 306 STEPHENS HALL, EARLSFORT COURT, LOWER LEESON STREET
4 SEA VIEW TERRACE, CLARENCE ST, DUN LAOGHAIRE
APT 3 BALLYGIHEN HOUSE, SANDYCOVE RD, SANDYCOVE
Stradbrook Lodge, Stradbrook Road, Blackrock
51 Hampton Wood Park, Hampton Wood, Finglas
Apt. 70 Old Kilmainham Village, Bowe Lane, Kilmainham
2 FOREST VIEW, RATHINGLE, SWORDS
21 Adamstown Boulevard, Adamstown Square, Adamstown
9 Vale View Lawn, Cabinteely
14 south dock place, ringsend
14 ROCHDALE, HONEY PARK, KILL AVENUE
55 Woodbrook Hall, Castleknock
1 DALE DR, KILMACUD, STILLORGAN
Apartment 59 Leona Block, Honeypark, Dun Laoghaire
Santa Maria B, Seamount Road, Malahide
28 LOUVAIN, ARDILEA, ROEBUCK ROAD
78, Bangor Road, Crumlin

```
14 Strandmill Avenue, Portmarnock
30 FAIRVIEW CLOSE, RICHMOND AVE, FAIRVIEW
Apt 119  Alliance Building, The Gasworks Barrow Street, South Lotts Road
Rye Cottage, Clondalkin
115 Adamstown Way, Adamstown Square, Adamstown
WHITEWATER, HOWTH SUMMIT, BAILEY HOWTH
43 Killester Park, KIllester
22 ACHILL HOUSE, CUSTOM HOUSE SQ, MAYOR ST LOWER
11 Westminister Wood, Foxrock
8 NEWTOWN HALL, HAZELBROOK SQUARE, CHURCHTOWN
11 Valentia Road, Drumcondra
1 Belltree Avenue, Belltree, Clongriffin
Apt 314 Charlestown Centre, Charlestown Place, St Margaret's Road  Finglas
91 Landscape Park, Churchtown
12 St. Pancras House, Mount Tallant Avenue, Terenure
74 CEDAR PLACE, RIDGEWOOD, SWORDS
```

While the resulting output is quite large, all the addresses seem to refer to areas or suburbs of Dublin. As such, we do not have any postal codes incorrectly referencing a non-Dublin address.

Test 4 The county names must be valid and there must be no duplicates.

For this test, given the low number of unique values, we can simply try and print all the post codes to see if all of them match standard Irish postcodes.

In [23]:
```
for county_name in df["County"].unique():
    print(county_name)
```

```
Dublin
Clare
Sligo
Waterford
Cork
Meath
Carlow
Limerick
Westmeath
Leitrim
Donegal
Laois
Tipperary
Galway
Offaly
Mayo
Wexford
Kildare
Wicklow
Kilkenny
Kerry
Louth
Monaghan
Roscommon
Cavan
Longford
```

From the above, we can see that these are all valid county names.

Test 5: The price cannot be negative.

In [24]:
```
test_5 = df[df["Price (€)"] < 0]
print(f"The number of rows that fail test 5 is: {test_5.shape[0]}")
```

```
The number of rows that fail test 5 is: 0
```

1(h) Save clean dataframe to new CSV file

Now that we have done an initial analysis of the dataset, we can save the cleaned dataframe into a new CSV file.

In [25]:
```
df.to_csv("q1_cleaned_dataframe.csv", index=False)
```

1(i) Descriptive Statistics for Categorical Data

Before preparing descriptive statistics for the categorical and continuous data, we should read in our new clean data into a variable and work with that instead.

Note that when reading in the .csv file, pandas will set the categories back to "object" - as such we should reconvert these to ensure that they all match what we had originally envisaged.

In [26]:
```
df = pd.read_csv("q1_cleaned_dataframe.csv")
```

In [27]:
```
# Selecting categorical columns
categorical_columns = df[["Address", "Postal Code", "County", "Not Full Market Price", "VAT Exc

# Looping through columns and casting as category
for column in categorical_columns:
    df[column] = df[column].astype('category')

# Setting Date of Sale to datetime format
df["Date of Sale"] = pd.to_datetime(df["Date of Sale"], infer_datetime_format=True)
```

To prepare the categorical features, we will need to obtain the following in relation to each feature:

- First Mode (frequency and proportion)
- Second Mode (frequency and proportion)
- Percentage of Instances Missing
- Cardinality (aka unique instances)

We begin by calling the "describe" function in respect of the categorical features. This will provide us with the cardinality, first mode, and first mode frequency.

In [28]:
```
df_table_categorical = df.select_dtypes(["category"]).describe().T

df_table_categorical
```

Out[28]:

	count	unique	top	freq
Address	9999	9979	1 ROSE TERRACE, FRANCIS ST, WEXFORD	2
Postal Code	1856	22	Dublin 15	241
County	9999	26	Dublin	3195
Not Full Market Price	9999	2	No	9524
VAT Exclusive	9999	2	No	8364
Description of Property	9999	3	Second-Hand Dwelling house /Apartment	8338
Property Size Description	1039	4	greater than or equal to 38 sq metres and less...	731

To ensure that the column names are consistent with the data we are trying to find, we will rename the attributes in accordance with the naming convention in this notebook.

In [29]:
```
# Renaming columns to match textbook
df_table_categorical.rename(columns={'count': 'Number of Instances', 'unique': 'Cardinality', '

df_table_categorical
```

Out[29]:

	Number of Instances	Cardinality	1st Mode	1st Mode (Frequency)
Address	9999	9979	1 ROSE TERRACE, FRANCIS ST, WEXFORD	2
Postal Code	1856	22	Dublin 15	241
County	9999	26	Dublin	3195
Not Full Market Price	9999	2	No	9524
VAT Exclusive	9999	2	No	8364
Description of Property	9999	3	Second-Hand Dwelling house /Apartment	8338
Property Size Description	1039	4	greater than or equal to 38 sq metres and less...	731

Next, we will calculate the "Percentage of Instances Missing" and assign these values to a variable for ease of access.

In [30]:
```
# Add up all null values, divide by number of rows, multiply by 100
category_columns_missing = 100 * (df[categorical_columns].isnull().sum()/df.shape[0])

# Add columns to dataframe
df_category_missing = pd.DataFrame(category_columns_missing, columns=["Missing Values (%)"])

df_category_missing
```

Out[30]:

	Missing Values (%)
Address	0.000000
Postal Code	81.438144
County	0.000000
Not Full Market Price	0.000000
VAT Exclusive	0.000000
Description of Property	0.000000
Property Size Description	89.608961

To calculate the proportion of the mode, we will need to find the percentage of the frequency of the attribute compared with the overall count.

In [31]:
```
# Find proportion of 1st mode in total number of instances
first_mode_proportion = (df_table_categorical["1st Mode (Frequency)"]/df_table_categorical["Num

# Add first mode proportion to dataframe
df_category_first_mode_proportion = pd.DataFrame(first_mode_proportion, columns=["1st Mode (Pro

df_category_first_mode_proportion
```

Out[31]:

	1st Mode (Proportion (%))
Address	0.020002
Postal Code	12.984914
County	31.953195
Not Full Market Price	95.249525
VAT Exclusive	83.648365
Description of Property	83.388339
Property Size Description	70.356112

We will now merge what we have so far together before making a start on calculating the second mode.

In [32]:

```
# Merge current values
df_data_quality_category = pd.concat([df_table_categorical, df_category_first_mode_proportion,

df_data_quality_category
```

Out[32]:

	Number of Instances	Cardinality	1st Mode	1st Mode (Frequency)	1st Mode (Proportion (%))	Missing Values (%)
Address	9999	9979	1 ROSE TERRACE, FRANCIS ST, WEXFORD	2	0.020002	0.000000
Postal Code	1856	22	Dublin 15	241	12.984914	81.438144
County	9999	26	Dublin	3195	31.953195	0.000000
Not Full Market Price	9999	2	No	9524	95.249525	0.000000
VAT Exclusive	9999	2	No	8364	83.648365	0.000000
Description of Property	9999	3	Second-Hand Dwelling house /Apartment	8338	83.388339	0.000000
Property Size Description	1039	4	greater than or equal to 38 sq metres and less...	731	70.356112	89.608961

To find the second mode, as well as its frequency and proportion, we will need to use the value_counts() method of pandas, which will return an object containing the number of occurrences of each feature. We will convert the keys to a list and select the second most occurring value for the second mode. The frequency will be the value associated with that key. To calculate the proportion, the method is largely identical to the one used to calculate the proportion of the first mode.

In [33]:

```
# Initialise list to hold second mode keys
second_mode_keys_list = []

# Loop through columns and add to keys list
for i in range(len(categorical_columns)):
    second_mode_keys_list.append(df[categorical_columns[i]].value_counts().index.tolist()[1])

# Create new column
df_data_quality_category["2nd Mode"] = second_mode_keys_list
```

In [34]:

```
# Create list to hold second mode values
second_mode_values_list = []
```

```
# Get second mode values
for i in range(len(categorical_columns)):
    second_mode_values_list.append(df[categorical_columns[i]].value_counts()[1])

# Set second mode values to list
df_data_quality_category["2nd Mode (Frequency)"] = second_mode_values_list
```

In [35]:

```
# Get proportion of second mode values and add to dataframe
df_data_quality_category["2nd Mode (Proportion (%))"] = (df_data_quality_category["2nd Mode (Fr

df_data_quality_category
```

Out[35]:

	Number of Instances	Cardinality	1st Mode	1st Mode (Frequency)	1st Mode (Proportion (%))	Missing Values (%)	2nd Mode	2nd Mode (Frequency)
Address	9999	9979	1 ROSE TERRACE, FRANCIS ST, WEXFORD	2	0.020002	0.000000	HORETOWN, FOULKSMILLS, WEXFORD	2
Postal Code	1856	22	Dublin 15	241	12.984914	81.438144	Dublin 24	145
County	9999	26	Dublin	3195	31.953195	0.000000	Cork	1113
Not Full Market Price	9999	2	No	9524	95.249525	0.000000	Yes	475
VAT Exclusive	9999	2	No	8364	83.648365	0.000000	Yes	1635
Description of Property	9999	3	Second-Hand Dwelling house /Apartment	8338	83.388339	0.000000	New Dwelling house /Apartment	1660
Property Size Description	1039	4	greater than or equal to 38 sq metres and less...	731	70.356112	89.608961	greater than 125 sq metres	132

Initial Thoughts

The main two points of interest here are the low cardinalities of the "Description of Property" and the "Property Size Description" features. We will try and examine them below to see if we can determine the reasoning for this.

In [36]:

```
df["Description of Property"].unique()
```

Out[36]:

```
['New Dwelling house /Apartment', 'Second-Hand Dwelling house /Apartment', 'Teach/Árasán Cónait
he Atháimhe']
Categories (3, object): ['New Dwelling house /Apartment', 'Second-Hand Dwelling house /Apartmen
t', 'Teach/Árasán Cónaithe Atháimhe']
```

From the above, we can see that one of the values that is repeated is in Gaeilge, and translates roughly to House / Apartment. It looks to be that we are actually looking at a binary datatype here - this column seems

to suggest whether or not the house or apartment is new or second hand. This is something to be discussed in more detail in the data quality report.

In [37]:
```
for description in df["Property Size Description"].unique():
    print(description)
```

```
nan
greater than or equal to 125 sq metres
greater than or equal to 38 sq metres and less than 125 sq metres
less than 38 sq metres
greater than 125 sq metres
```

From the above, it seems that that this feature represents categories of property size, and can probably be simplified further. This is something to be discussed in more detail in the data quality report.

It is also interesting to see that both first and second modes are houses located in Wexford - however, given the fact that the frequency of both modes is merely "2", this could merely be a coincidence, as it is not unusual for two homes to be resold in the same decade.

1(j) Descriptive Statistics for Continuous Data

To prepare the continuous features, we will need to obtain the following in relation to each feature:

- Minimum
- 1st Quartile
- Mean
- Median
- 3rd Quartile
- Maximum
- Standard Deviation
- Number of Instances
- Percentage of Instances Missing
- Cardinality (Unique Instances)

Please note that the operations in this section are very similar to the ones carried out for the categorical data. As such, no comments have been added to the code in this section to avoid redundancy.

We begin by calling the "describe" function in respect of the continuous features. This will provide us with all the information we need except for percentage of instances missing and cardinality.

We will ensure that the datetime_is_numeric parameter is switched to True, so that the the Date of Sale is factored into the calculations.

In [38]:
```
df_table_continuous = df[continuous_columns].describe(datetime_is_numeric=True).T

df_table_continuous
```

Out[38]:

	count	mean	min	25%	50%	75%	max	std
Date of Sale	9999	2017-01-28 08:40:54.005400320	2010-01-04 00:00:00	2014-11-14 00:00:00	2017-06-19 00:00:00	2019-08-20 00:00:00	2022-01-14 00:00:00	NaN
Price (€)	9999.0	274402.924038	5953.0	120000.0	200000.0	310000.0	139165000.0	1472944.686009

From the above we can see that the standard deviation of the date of sale was not calculated properly. This

is because we need to explicitly tell Pandas to include datetime objects when calculating the standard deviation and to convert it into number of days. We will make this calculation and reinsert into the dataframe.

Because there is just one NaN value in the column, we can use the "fillna" function to replace this with the new calculation.

In [39]:
```
date_of_sale_std = df["Date of Sale"].std(numeric_only=False).days

df_table_continuous.fillna(date_of_sale_std, inplace=True)

df_table_continuous
```

Out[39]:

	count	mean	min	25%	50%	75%	max	std
Date of Sale	9999.0	2017-01-28 08:40:54.005400320	2010-01-04 00:00:00	2014-11-14 00:00:00	2017-06-19 00:00:00	2019-08-20 00:00:00	2022-01-14 00:00:00	1.139000e+03
Price (€)	9999.0	274402.924038	5953.0	120000.0	200000.0	310000.0	139165000.0	1.472945e+06

To ensure that the column names are consistent with the data we are trying to find, we will rename the attributes with the naming convention in this notebook.

In [40]:
```
df_table_continuous.rename(columns={'count': 'Number of Instances', 'mean': 'Mean', 'min': 'Mir

df_table_continuous
```

Out[40]:

	Number of Instances	Mean	Minimum	1st Quartile	Median	3rd Quartile	Maximum	Standard Deviation (Days)
Date of Sale	9999.0	2017-01-28 08:40:54.005400320	2010-01-04 00:00:00	2014-11-14 00:00:00	2017-06-19 00:00:00	2019-08-20 00:00:00	2022-01-14 00:00:00	1.139000e+03
Price (€)	9999.0	274402.924038	5953.0	120000.0	200000.0	310000.0	139165000.0	1.472945e+06

Next, we will find the cardinality of the continuous features by using the unique function.

In [41]:
```
continuous_features_card = df[continuous_columns].nunique()

df_continuous_card_table = pd.DataFrame(continuous_features_card, columns=["Cardinality"])

df_continuous_card_table
```

Out[41]:

	Cardinality
Date of Sale	2763
Price (€)	2324

Finally, we will calculate the "Percentage of Instances Missing" using the same method that we used for the categorical features.

In [42]:
```
continuous_columns_missing = 100 * (df[continuous_columns].isnull().sum()/df.shape[0])
```

```
df_continuous_missing = pd.DataFrame(continuous_columns_missing, columns=["Missing Values (%)"]

df_continuous_missing
```

Out[42]:

	Missing Values (%)
Date of Sale	0.0
Price (€)	0.0

Now that we have all the values that we need, we will merge the calculations into one table.

In [43]:
```
df_data_quality_continuous = pd.concat([df_table_continuous, df_continuous_card_table, df_conti

df_data_quality_continuous
```

Out[43]:

	Number of Instances	Mean	Minimum	1st Quartile	Median	3rd Quartile	Maximum	Standard Deviation (Days)	Cardina
Date of Sale	9999.0	2017-01-28 08:40:54.005400320	2010-01-04 00:00:00	2014-11-14 00:00:00	2017-06-19 00:00:00	2019-08-20 00:00:00	2022-01-14 00:00:00	1.139000e+03	2
Price (€)	9999.0	274402.924038	5953.0	120000.0	200000.0	310000.0	139165000.0	1.472945e+06	2

Initial Thoughts

The main points of interest here are located in the "Price" feature, and specifically the presence of extreme outliers and large standard deviation. To see what could be the cause of this, we will try and print some rows that have a price below 10,000, and some rows that have a price of over 3,000,000.

In [44]:
```
df[df["Price (€)"] < 10000]
```

Out[44]:

	Date of Sale	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property	Property Size Description
273	2013-12-20	20 Millrace Court, Phoenix Park Racecourse, Ca...	Dublin 15	Dublin	8000.00	No	Yes	New Dwelling house /Apartment	greater than or equal to 38 sq metres and less...
345	2013-12-20	17 Cedarhurst Road, Phoenix Park Racecourse, C...	Dublin 15	Dublin	8000.00	No	Yes	New Dwelling house /Apartment	greater than or equal to 38 sq metres and less...
913	2018-12-20	CARRICKBOYLE, BRINLACK, CO. DONEGAL	NaN	Donegal	6000.00	No	No	Second-Hand Dwelling house /Apartment	NaN
1045	2012-10-26	COMMONS ROAD, NAVAN	NaN	Meath	8750.00	Yes	No	Second-Hand Dwelling house /Apartment	NaN

	Date of Sale	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property	Property Size Description
3134	2019-11-08	27 DAVIS ST, TIPPERARY TOWN, TIPPERARY	NaN	Tipperary	5953.00	No	No	Second-Hand Dwelling house /Apartment	NaN
3368	2013-12-20	23 Phoenix Park Way, Phoenix Park Racecourse, ...	Dublin 15	Dublin	8000.00	No	Yes	New Dwelling house /Apartment	greater than or equal to 38 sq metres and less...
5987	2015-01-20	APARTMENT 7, 63 GREAT WATER STREET, CO LONGFORD	NaN	Longford	6847.31	No	No	Second-Hand Dwelling house /Apartment	NaN
7041	2013-12-20	12 Rosanule, Phoenix Park Racecourse, Castleknock	Dublin 15	Dublin	8000.00	No	Yes	New Dwelling house /Apartment	greater than or equal to 38 sq metres and less...
7153	2013-12-20	22 Cedarhurst Road, Phoenix Park Racecourse, C...	Dublin 15	Dublin	8000.00	No	Yes	New Dwelling house /Apartment	greater than or equal to 38 sq metres and less...
7271	2017-12-11	77 SWORDS RD, DRUMCONDRA, DUBLIN 9	Dublin 9	Dublin	7500.00	No	No	Second-Hand Dwelling house /Apartment	NaN
7888	2018-02-14	TOWNAHOUSEY, KILLARGUE, DROMAHAIRE	NaN	Leitrim	9000.00	Yes	No	Second-Hand Dwelling house /Apartment	NaN
8551	2014-04-29	DONERAILE ROAD, CASTLETOWNROCHE, MALLOW	NaN	Cork	8000.00	No	No	Second-Hand Dwelling house /Apartment	NaN
9490	2014-09-15	ARDMORE, LOWER GLENAGEARY RD, DUN LAOGHAIRE	NaN	Dublin	6000.00	No	No	Second-Hand Dwelling house /Apartment	NaN
9973	2016-06-28	Apartment 45, Broad Leaf, Broad Street	NaN	Limerick	7222.00	No	Yes	New Dwelling house /Apartment	less than 38 sq metres

Given the average prices of property in Ireland it is extremely unlikely that some of these entries have been entered correctly. For example, it would be difficult to believe that a house in Dun Laoghaire would be sold for €6,000. This is something that should be considered in the data quality report.

In [45]:

```
df[df["Price (€)"] > 3000000]
```

Out[45]:

	Date of Sale	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property	Property Size Description
121	2019-12-18	1 5 6 & 22 THE HAMPTON, APT 1 2 7 19 20 26 2...	Dublin 11	Dublin	3.919000e+06	No	No	Second-Hand Dwelling house /Apartment	NaN
286	2021-05-25	Block A3, The Quarter, Citywest	Dublin 24	Dublin	1.076268e+07	No	Yes	New Dwelling house /Apartment	NaN
421	2021-09-03	VERITAS HOUSE, DOMINICAN CONVENT, MUCKROSS PARK	NaN	Dublin	4.900000e+06	No	No	Second-Hand Dwelling house /Apartment	NaN
2164	2021-07-13	Apartment 21 The Blake, Block 4 Level 6, Lans...	Dublin 7	Dublin	3.259912e+06	No	Yes	New Dwelling house /Apartment	NaN
3493	2020-06-24	UNITS 1 - 28 BLOCK D, GANDON VIEW, LUCAN	NaN	Dublin	8.629956e+06	No	Yes	New Dwelling house /Apartment	NaN
3705	2019-03-20	9 An Móinéar, Murrough, Renmore	NaN	Galway	3.237000e+06	No	Yes	New Dwelling house /Apartment	NaN
4775	2018-03-07	15 17 18 and 19 ST COLUMCILLES DR, 1 1B 2 3 an...	NaN	Dublin	3.400000e+06	No	No	Second-Hand Dwelling house /Apartment	NaN
4823	2018-06-22	COMMERCIAL PROPERTY AND, 423 RESI UNITS AT CLA...	Dublin 8	Dublin	1.391650e+08	Yes	No	Second-Hand Dwelling house /Apartment	NaN
4875	2020-12-21	2 GREEG COURT, PARNELL STREET, DUBLIN 1	Dublin 1	Dublin	9.600000e+06	No	No	Second-Hand Dwelling house /Apartment	NaN
5803	2019-04-17	HOUSE 1, MIDDLETOWN HOUSE RETIREMENT VILLAGE, ...	NaN	Wexford	3.200000e+06	No	No	Second-Hand Dwelling house /Apartment	NaN
7564	2014-10-03	DODDER GROUND, MILLTOWN BRIDGE RD, MILLTOWN DU...	Dublin 14	Dublin	4.000000e+06	No	No	Second-Hand Dwelling house /Apartment	NaN

	Date of Sale	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property	Property Size Description
8271	2013-08-13	Annaville Residence, Dundrum Road	NaN	Dublin	7.675771e+06	No	No	Second-Hand Dwelling house /Apartment	NaN
9406	2013-10-31	Alliance Building, The Gasworks	Dublin 4	Dublin	3.870000e+07	Yes	No	Second-Hand Dwelling house /Apartment	NaN

From the above we can see that some of the entries seem to assign one entry to the sale of several properties, which seems to be the cause of at least some of the outliers. This makes sense, given the tendency in Ireland for investors to purchase a portfolio of properties for sale, and it was also something that was noted in the information note of the RPPR. It will, however, definitely have an effect on the logical integrity of the data. This is something that should be noted in the data quality report. We can also treat this as a failure of a logical integrity test.

Test 6: There must be one property sale recorded per entry in the dataset.

This test has failed for reasons set out above.

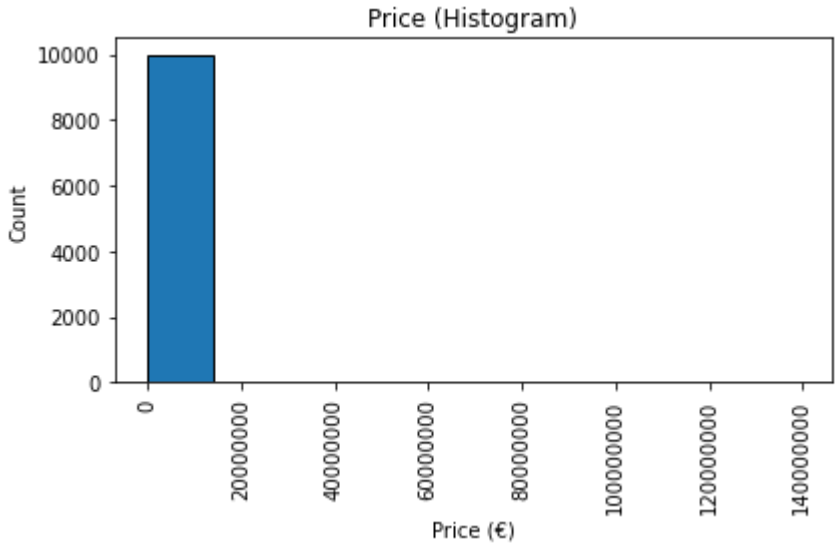
1(k) Plot histograms for continuous features

A histogram will need to be plotted for each continuous feature. There will be further discussion on each graph in the Data Quality Report.

Please note that the way the graphs are displayed using matplotlib / seaborn is largely the same - as such, comments on the code for displaying graphs will only be provided once to avoid redundancies. The same comments will apply to the rest of the code that display graphs.

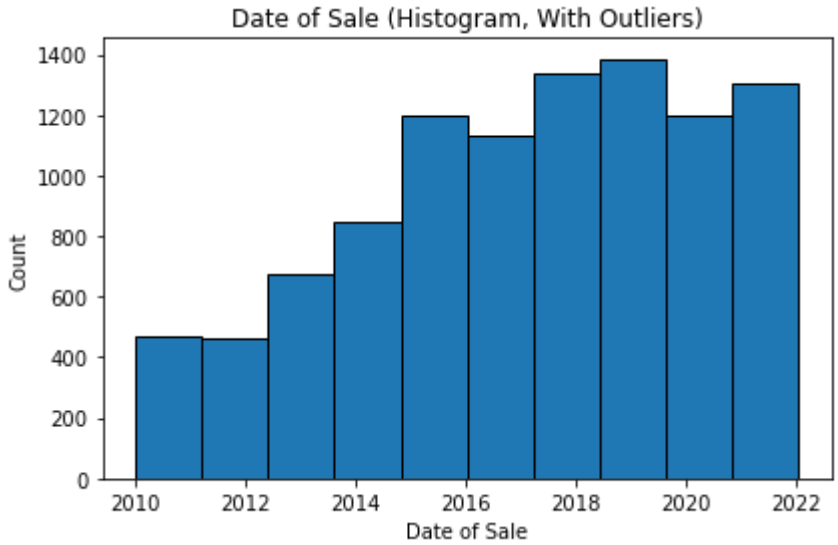
In [46]:

```
# Plotting histogram
plt.hist(df["Price (€)"], edgecolor="black")

# Setting title, xlabel, ylabel.
plt.title("Price (Histogram)")
plt.xlabel("Price (€)")
plt.ylabel("Count")

# Rotating x labels by 90 degrees for readability
plt.xticks(rotation = 90)

# Suppressing scientific notation
plt.ticklabel_format(style="plain")

# Ensuring that the entire graph fits into a single png and displaying
plt.tight_layout()
plt.savefig('q1_price_hist_outliers.png', dpi=300, bbox_inches="tight")
plt.show()
```



In [47]:

```
plt.hist(df["Date of Sale"], edgecolor="black")
plt.title("Date of Sale (Histogram, With Outliers)")
plt.xlabel("Date of Sale")
plt.ylabel("Count")
plt.tight_layout()
plt.savefig('q1_date_of_sale_hist.png', dpi=300, bbox_inches="tight")
plt.show()
```



The Date of Sale histogram looks largely in order. We can see that property sales seem to have been relatively low at the start of the decade, and have grown steadily since.

With regard to the Price histogram, however, we can see that the presence of outliers has greatly skewed the data, making it very difficult to glean any useful information from it.

It will be useful to, in the first instance, clamp the outliers, and re-run the histogram to see what the distribution looks like then.

We will use the method discussed in the textbook in order to clamp our values, which consists of clamping the data using the interquartile range.

In [48]:

```
# Find the first and third quartiles
first_quartile = df.quantile(0.25)
third_quartile = df.quantile(0.75)

# Find the interquartile range
inter_quartile_range = third_quartile - first_quartile
```

```
# Find the minimum and maximum clamps using the formula explained in the textbook
min_clamp = abs(first_quartile - (inter_quartile_range * 1.5))
max_clamp = third_quartile + (inter_quartile_range * 1.5)

print(f"The minimum clamp is {min_clamp}")
print(f"The maximum clamp is {max_clamp}")
```

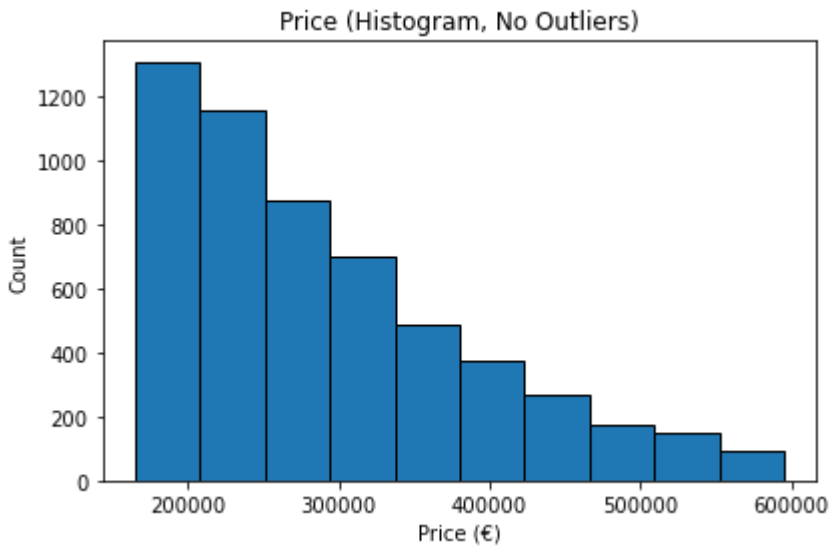
```
The minimum clamp is Price (€)    165000.0
dtype: float64
The maximum clamp is Price (€)    595000.0
dtype: float64
```

In [49]:

```
# Create a new dataframe to ensure that we are not making any modifications to the existing dat
df_without_outliers = df.copy()

# Get values outside the range of the clamp
minimum_rows = df_without_outliers[df_without_outliers["Price (€)"]  < int(min_clamp)].index
maximum_rows = df_without_outliers[df_without_outliers["Price (€)"]  > int(max_clamp)].index

# Drop values that are outside of the clamp
df_without_outliers.drop(minimum_rows, inplace = True)
df_without_outliers.drop(maximum_rows, inplace = True)
```

In [50]:

```
plt.hist(df_without_outliers["Price (€)"], edgecolor="black")
plt.title("Price (Histogram, No Outliers)")
plt.xlabel("Price (€)")
plt.ylabel("Count")
plt.ticklabel_format(style="plain")
plt.tight_layout()
plt.savefig('q1_price_hist_no_outliers.png', dpi=300, bbox_inches="tight")
plt.show()
```



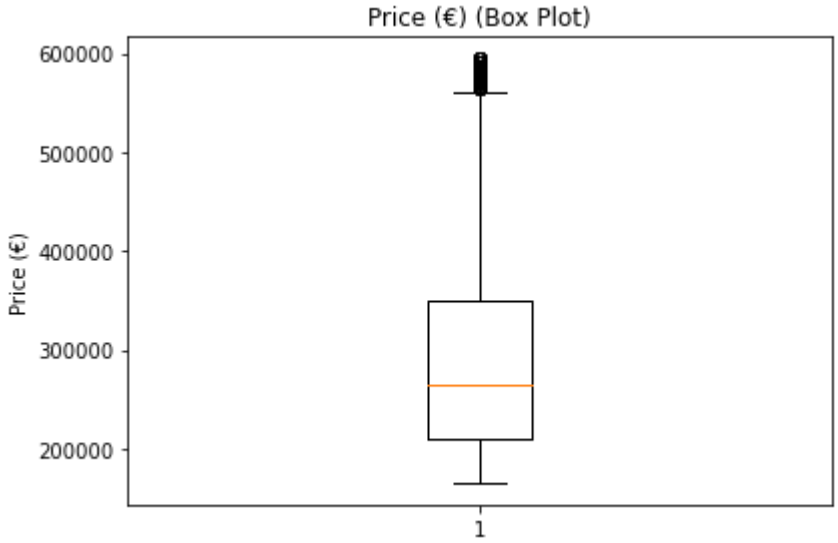
We can see from the above that the histogram is largely what we would expect, in the sense that it is skewed to the left. This means that the majority of the properties were purchased for a certain sum, and as we go up the price range, there are increasingly fewer number of sales consisting of property sales for larger sums (representing the "tail" of the histogram).

We should also note, however, that our clamp is quite severe using the interquartile range method, and a large portion of the dataset is lost when we visualise our data in this way. Therefore, in our data quality report we should consider an alternative way of dealing with outliers that preserves the data.

1(l) Plot box plots for continuous features

Please note that for the purposes of creating the box plot, we will once again be using the copy dataframe with no outliers in order to get a better idea of what the distribution looks like.

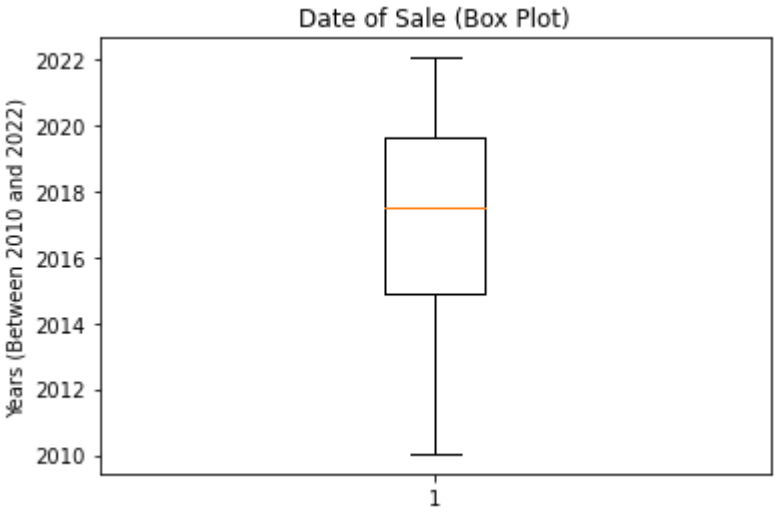
In [51]:
```
plt.boxplot(df_without_outliers["Price (€)"])
plt.title("Price (€) (Box Plot)")
plt.ylabel("Price (€)")
plt.tight_layout()
plt.savefig('q1_price_box_no_outliers.png', dpi=300, bbox_inches="tight")
plt.show()
```



The resulting box plot, without outliers, looks largely in order and in line with the results shown in the price histogram. The spread between second and fourth quartiles seems to be between 20k-40k, with some large outliers over 56k. However, as discussed above, this particular feature should be revisited once we have figured out a better way of treating outliers.

Please note that it is not possible to create a box plot of a datetime feature. What we can do instead, however, is to extract the number of years from our existing feature and plot a box plot based on those values. In order to obtain these values, we divide by 8.64e+13, the number of nanoseconds in a day, then by 365, to get the number of years, and then subtract 40 (as the timestamp gives us the number of seconds since 1970). Finally, we add 2010, to get an accurate representation of the year.

In [52]:
```
df_datetime_numeric = df.copy()
df_datetime_numeric = pd.to_numeric(df_datetime_numeric["Date of Sale"]) / 8.64e+13 / 365 - 40
plt.boxplot(df_datetime_numeric)
plt.title("Date of Sale (Box Plot)")
plt.ylabel("Years (Between 2010 and 2022)")
plt.savefig('q1_date_of_sale_box.png', dpi=300, bbox_inches="tight")
plt.show()
```



Again, this is inline with the information presented in the Date of Sale histogram, and the spread of values between second and fourth quartiles was between 2015 and 2020.

1(m) Plot bar charts for categorical features

A bar plot will need to be plotted for each categorical feature. Please note that due to the enormous cardinality of the Address feature it was decided that a bar chart would be (a) inappropriate and (b) impossible to glean any information from. As a compromise, a bar chart of the top 15 most frequent values was plotted, although the utility of this bar chart is questionable. Please also note that stacked bar plots were considered for boolean values but ultimately decided against as the readability of these was not as good as just the regular ones.

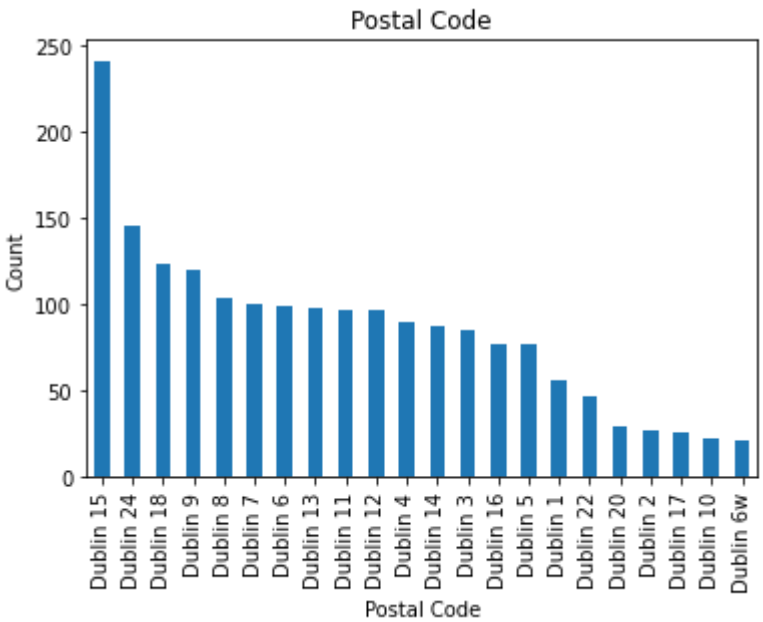
An attempt was made to visualise this feature using the Folium library but this was ultimately unsuccessful due to the quality of the address strings - more information on this will be provided later in the assignment.

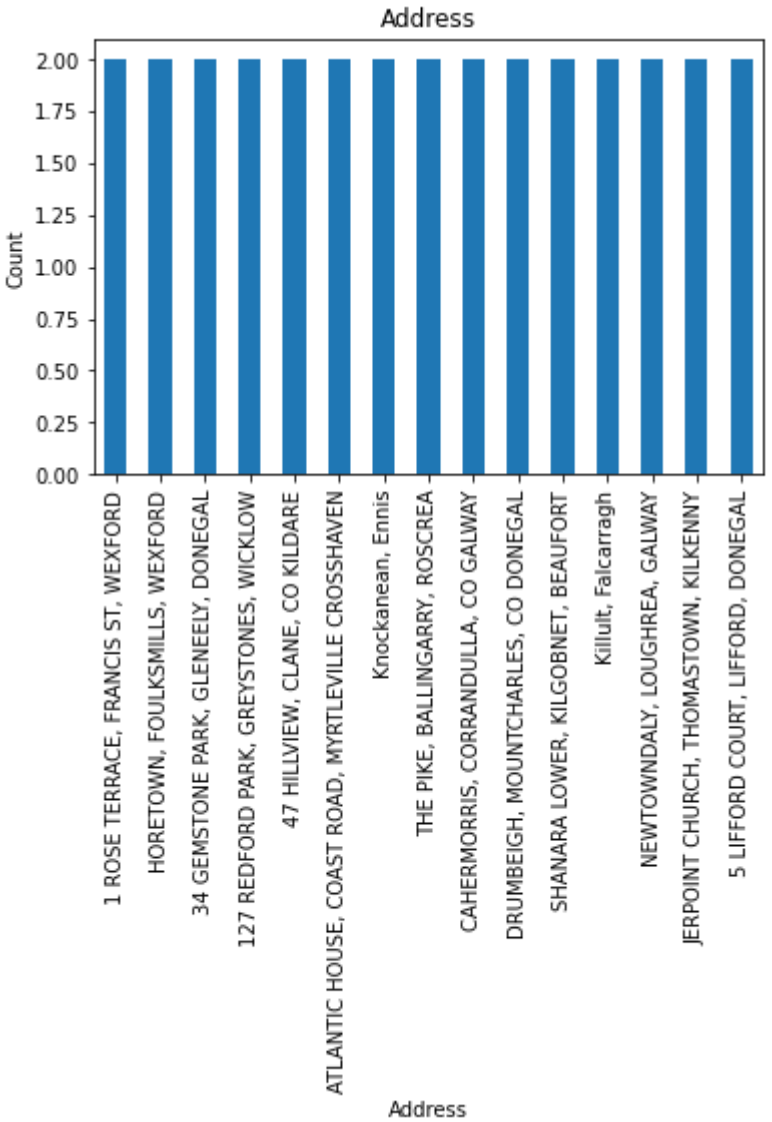
In [53]:
```
categorical_columns
```

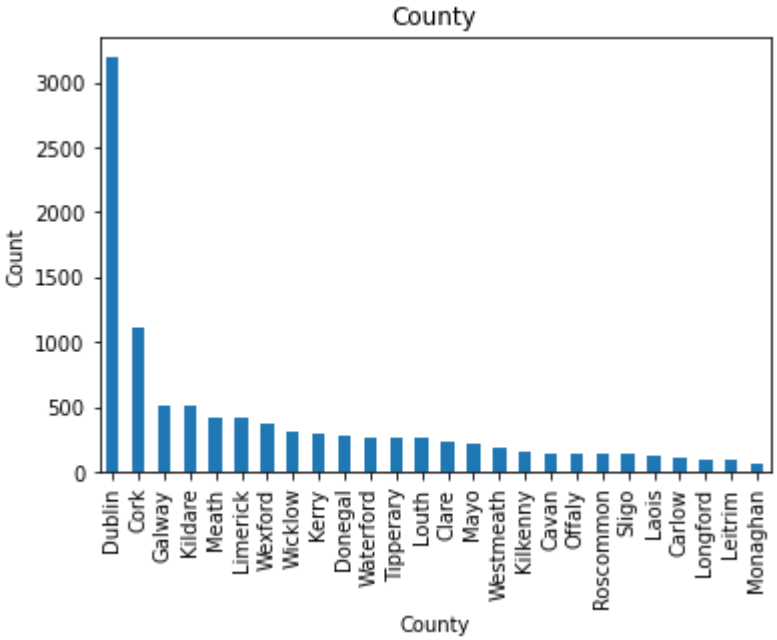
Out[53]:
```
Index(['Address', 'Postal Code', 'County', 'Not Full Market Price',
       'VAT Exclusive', 'Description of Property',
       'Property Size Description'],
      dtype='object')
```

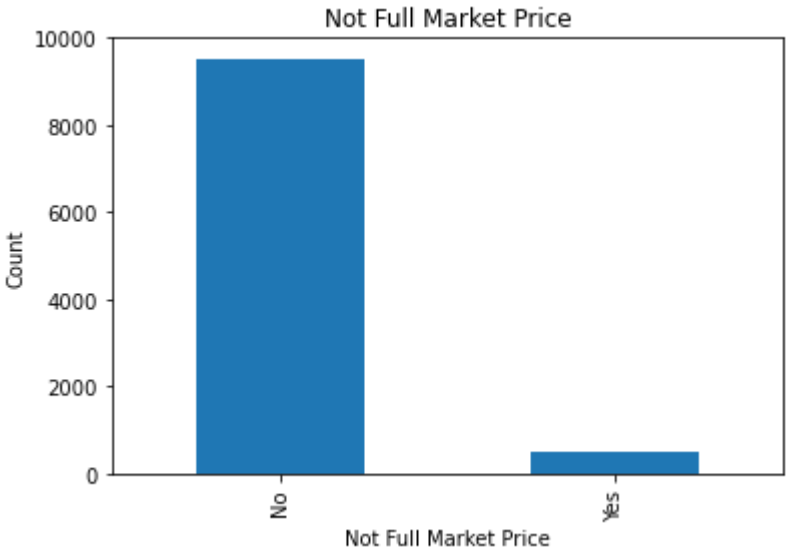
In [54]:
```
df["Postal Code"].value_counts().plot(kind = "bar")
plt.xlabel("Postal Code")
plt.ylabel("Count")
plt.title("Postal Code")
plt.savefig('q1_postal_code_bar.png', dpi=300, bbox_inches="tight")
plt.show()

df["Address"].value_counts()[:15].plot(kind = "bar")
plt.xlabel("Address")
plt.ylabel("Count")
plt.title("Address")
plt.savefig('q1_address_bar.png', dpi=300, bbox_inches="tight")
plt.show()

df["County"].value_counts().plot(kind = "bar")
plt.xlabel("County")
plt.ylabel("Count")
plt.title("County")
plt.savefig('q1_county_bar.png', dpi=300, bbox_inches="tight")
plt.show()

df["Not Full Market Price"].value_counts().plot(kind = "bar")
plt.xlabel("Not Full Market Price")
```

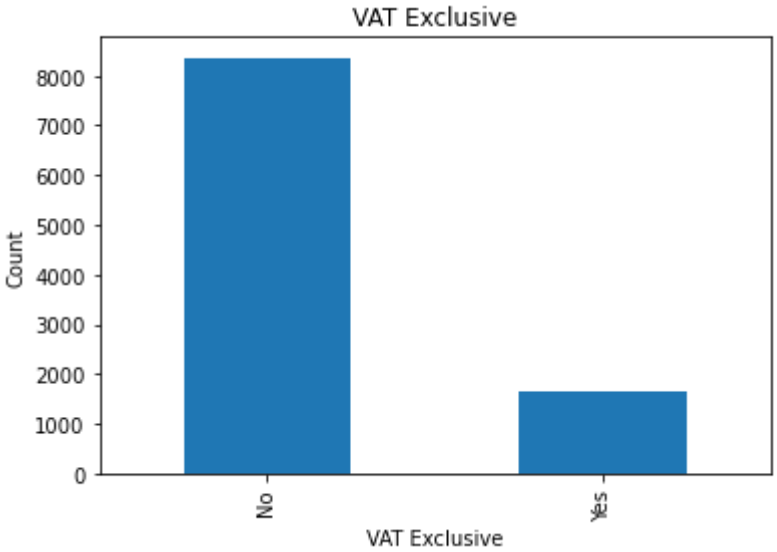
```
plt.ylabel("Count")
plt.title("Not Full Market Price")
plt.savefig('q1_market_price_bar.png', dpi=300, bbox_inches="tight")
plt.show()

df["VAT Exclusive"].value_counts().plot(kind = "bar")
plt.xlabel("VAT Exclusive")
plt.ylabel("Count")
plt.title("VAT Exclusive")
plt.savefig('q1_vat_bar.png', dpi=300, bbox_inches="tight")
plt.show()

df["Description of Property"].value_counts().plot(kind = "bar")
plt.xlabel("Description of Property")
plt.ylabel("Count")
plt.title("Description of Property")
plt.savefig('q1_description_of_property_bar.png', dpi=300, bbox_inches="tight")
plt.show()

df["Property Size Description"].value_counts().plot(kind = "bar")
plt.xlabel("Property Size Description")
plt.ylabel("Count")
plt.title("Property Size Description")
plt.savefig('q1_property_size_description_bar.png', dpi=300, bbox_inches="tight")
plt.show()
```

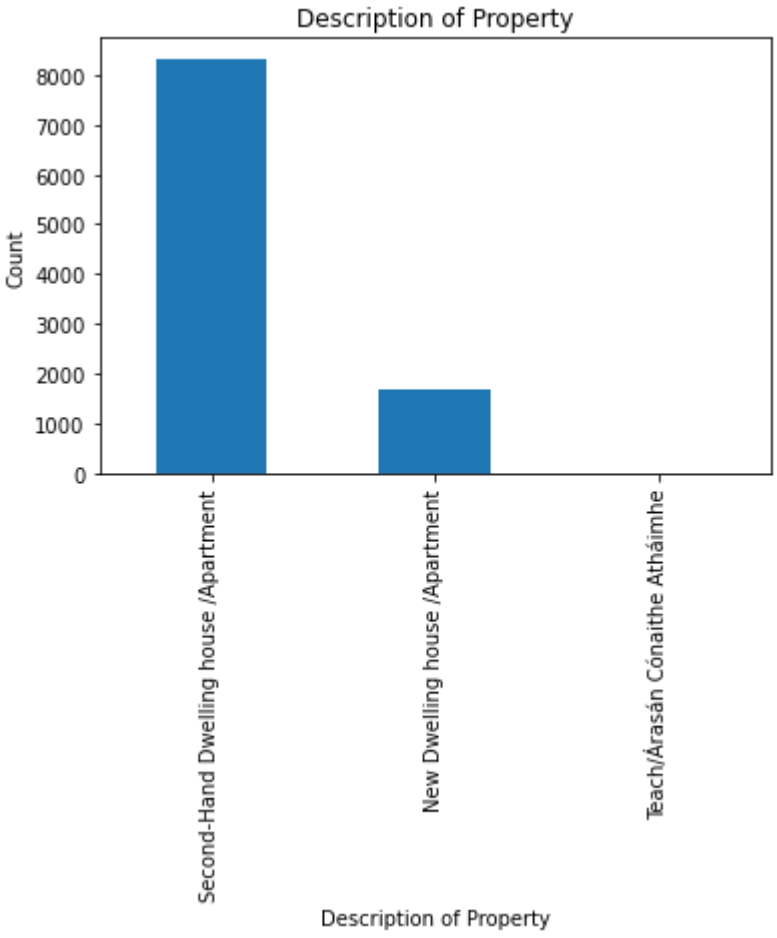


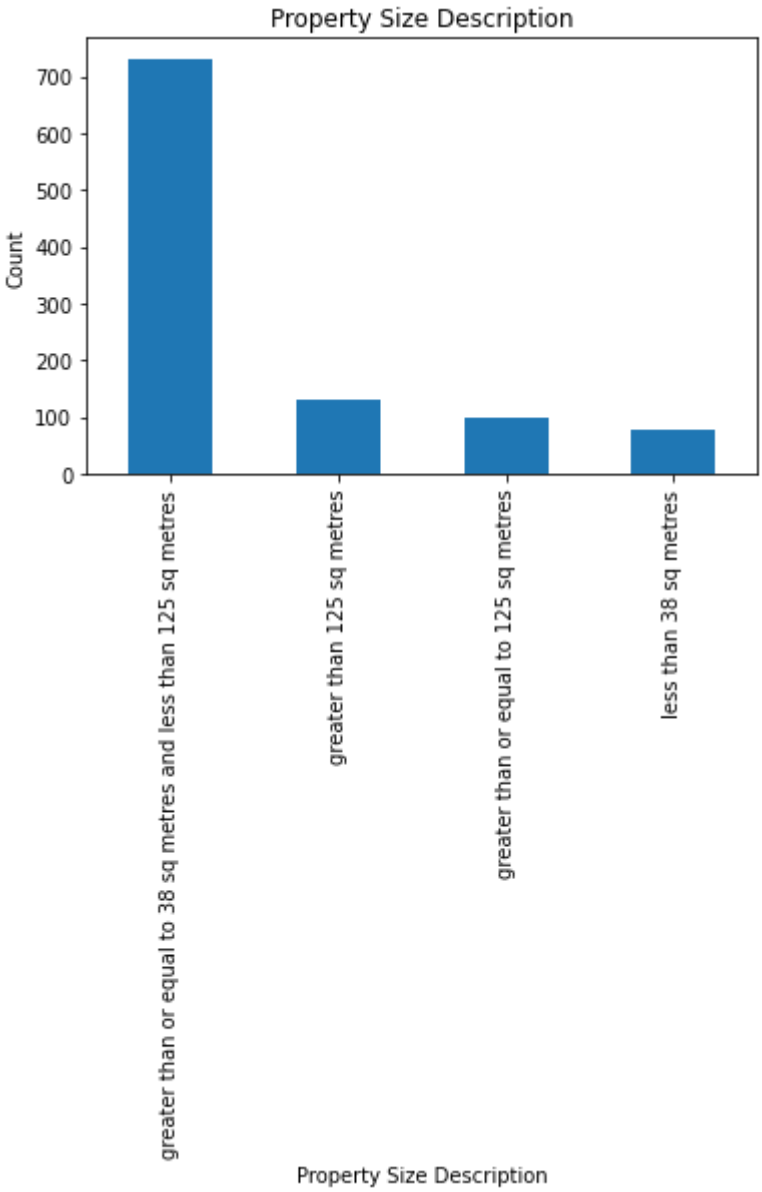

Address
Count
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
1 ROSE TERRACE, FRANCIS ST, WEXFORD
HORETOWN, FOULKSMILLS, WEXFORD
34 GEMSTONE PARK, GLENEELY, DONEGAL
127 REDFORD PARK, GREYSTONES, WICKLOW
47 HILLVIEW, CLANE, CO KILDARE
ATLANTIC HOUSE, COAST ROAD, MYRTLEVILLE CROSSHAVEN
Knockanean, Ennis
THE PIKE, BALLINGARRY, ROSCREA
CAHERMORRIS, CORRANDULLA, CO GALWAY
DRUMBEIGH, MOUNTCHARLES, CO DONEGAL
SHANARA LOWER, KILGOBNET, BEAUFORT
Killult, Falcarragh
NEWTOWNDALY, LOUGHREA, GALWAY
JERPOINT CHURCH, THOMASTOWN, KILKENNY
5 LIFFORD COURT, LIFFORD, DONEGAL
Address


County
Count
3000
2500
2000
1500
1000
500
0
Dublin
Cork
Galway
Kildare
Meath
Limerick
Wexford
Wicklow
Kerry
Donegal
Waterford
Tipperary
Louth
Clare
Mayo
Westmeath
Kilkenny
Cavan
Offaly
Roscommon
Sligo
Laois
Carlow
Longford
Leitrim
Monaghan
County


Not Full Market Price
10000
8000
6000
4000
2000
0
Count
No
Yes
Not Full Market Price


VAT Exclusive
8000
7000
6000
5000
4000
3000
2000
1000
0
Count
No
Yes
VAT Exclusive


Description of Property
8000
7000
6000
5000
4000
3000
2000
1000
0
Count
Second-Hand Dwelling house /Apartment
New Dwelling house /Apartment
Teach/Árasán Cónaithe Athláimhe
Description of Property



Question 2: Prepare a data quality plan for the cleaned CSV file.

Note: Please refer to the Data Quality Report contained in this directory for an explanation and consideration of various issues relating to this dataset. The report should be read in conjunction with this section of the notebook.

The data quality report contained in this folder contains a detailed explanation of the various different solutions to data quality issues discovered with the report. Please see below a summary of which handling strategy was chosen for each issue.

We will then proceed by working through implementing each handling strategy, together with an explanation as to why it was chosen.

Just to flag, there seems to be some overlap between the Data Quality Plan and the Data Quality Report as set out in the assignment. I was initially under the impression that the Data Quality Report should contain a discussion of initial feature plots and potential strategies of dealing with issues, as well as their advantages and disadvantages, and the notebook should contain the Data Quality Plan, consisting of a summary plan table, an explanation as to why certain handling strategies were chosen, and the implementation in code.

Just to be absolutely sure that I was satisfying the requirements set out in the assignment, I have included a data quality plan PDF in this folder, which consists of the summary table below, and slightly reworked

versions of the markdown cells in this section of the notebook explaining why a particular handling strategy was chosen. The PDF and the notebook contain largely the same information but I have provided both versions just to make sure (although I think the notebook version is a slightly better representation of my findings, as it easier to cross-refer to code in this section).

Summary Data Quality Plan

Feature Name	Issue	Handling Strategy
Price	Presence of Entries that Refer to Multiple Properties	Split entries out into multiple rows
Price	Presence of high outliers	Replace values below and above threshold to certain quantile
Address	Data has a high cardinality and non-standardised input	Do nothing
Postal Code	Large number of missing values	Change feature name to Postal Code (Dublin)
Not Full Market Price	Feature framed as a negative	Replace "yes" with "no" and vice-versa
Not Full Market Price	Non-boolean values used for boolean category	Change values to boolean values
VAT Exclusive	Non-boolean values used for boolean category	Change values to boolean values
VAT Exclusive / Description of Property	Potential reference to same facet of data	Do nothing
Description of Property	Values can simplified to a simple boolean in relation to second-hand property	Replace category with boolean category
Property Size	Feature has large amount of missing values	Drop feature

Price

This is the feature that requires the most amount of cleaning. We will begin with the first entry in our data quality plan, which refers to splitting entries that relate to multiple values into multiple rows.

Let us review the table that contains prices of over 3,000,000 again.

In [55]:
```
df[df["Price (€)"] > 3000000]
```

Out[55]:

	Date of Sale	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property	Property Size Description
121	2019-12-18	1 5 6 & 22 THE HAMPTON, APT 1 2 7 19 20 26 2...	Dublin 11	Dublin	3.919000e+06	No	No	Second-Hand Dwelling house /Apartment	NaN
286	2021-05-25	Block A3, The Quarter, Citywest	Dublin 24	Dublin	1.076268e+07	No	Yes	New Dwelling house /Apartment	NaN
421	2021-09-03	VERITAS HOUSE, DOMINICAN CONVENT, MUCKROSS PARK	NaN	Dublin	4.900000e+06	No	No	Second-Hand Dwelling house /Apartment	NaN

	Date of Sale	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property	Property Size Description
2164	2021-07-13	Apartment 21 The Blake, Block 4 Level 6, Lans...	Dublin 7	Dublin	3.259912e+06	No	Yes	New Dwelling house /Apartment	NaN
3493	2020-06-24	UNITS 1 - 28 BLOCK D, GANDON VIEW, LUCAN	NaN	Dublin	8.629956e+06	No	Yes	New Dwelling house /Apartment	NaN
3705	2019-03-20	9 An Móinéar, Murrough, Renmore	NaN	Galway	3.237000e+06	No	Yes	New Dwelling house /Apartment	NaN
4775	2018-03-07	15 17 18 and 19 ST COLUMCILLES DR, 1 1B 2 3 an...	NaN	Dublin	3.400000e+06	No	No	Second-Hand Dwelling house /Apartment	NaN
4823	2018-06-22	COMMERCIAL PROPERTY AND, 423 RESI UNITS AT CLA...	Dublin 8	Dublin	1.391650e+08	Yes	No	Second-Hand Dwelling house /Apartment	NaN
4875	2020-12-21	2 GREEG COURT, PARNELL STREET, DUBLIN 1	Dublin 1	Dublin	9.600000e+06	No	No	Second-Hand Dwelling house /Apartment	NaN
5803	2019-04-17	HOUSE 1, MIDDLETOWN HOUSE RETIREMENT VILLAGE, ...	NaN	Wexford	3.200000e+06	No	No	Second-Hand Dwelling house /Apartment	NaN
7564	2014-10-03	DODDER GROUND, MILLTOWN BRIDGE RD, MILLTOWN DU...	Dublin 14	Dublin	4.000000e+06	No	No	Second-Hand Dwelling house /Apartment	NaN
8271	2013-08-13	Annaville Residence, Dundrum Road	NaN	Dublin	7.675771e+06	No	No	Second-Hand Dwelling house /Apartment	NaN
9406	2013-10-31	Alliance Building, The Gasworks	Dublin 4	Dublin	3.870000e+07	Yes	No	Second-Hand Dwelling house /Apartment	NaN

Unfortunately, we can see from the above that there is no uniform way of checking which row represents multiple values. There are different ways in which multiple sales have been noted, ranging from printing out the numbers of the property addresses one by one, to separating them with a hyphen.

However, one thing seems to be common among these properties, and that is that the first field of the address contains more than one word / character separated by a space. Let's test this approach by splitting the addresses by the comma delimiter, and printing out those rows where the first field contains more than 3 words / characters.

In [56]:

```
# Loop through every row
for index in range(df.shape[0]):
    current_address = df["Address"][index]

    # Break up the address by string
    current_address_separated_by_comma = current_address.strip().split(", ")
    first_field_in_address = current_address_separated_by_comma[0].split()

    # Print address where the first field is great than four items
    if len(first_field_in_address) > 4:
        print(index, current_address)
```

```
31 72 A - J.K.L. STREET, EDENDERRY, CO OFFALY
86 6 7 8 9 ST ANNES TERRACE, SHORTCOURSE, WATERFORD
121 1  5  6 & 22 THE HAMPTON, APT 1 2 7 19 20 26 27 30 31 34-38, 41 42 47-49 BLOCK 1  THE MAIES
TON
141 5 DUN NA RIOGH GREEN, MONREAD RD, NAAS
149 Apartment 280 Bushy Park House, Templeogue Road, Terenure
193 29 Rockview (Site No. 50), Mountrath Road, Portlaoise
306 Apt 29  Block C  An Radharc, Maryborough Ridge, Maryborough Hill  Douglas
440 Apt 41  Block D  An Radharc, Maryborough Ridge, Maryborough Hill  Douglas
459 4 PARC NA SILLA LANE, LOUGHLINSTOWN, DUBLIN
466 95 GREAT WILLIAM OBRIEN ST, CORK, CORK
513 19 Block A  Silverbridge Court, Knock Road, Claremorris
608 Apt 34 Block B Clares Court, Main Street/ Church Street
773 169 Block 9  The Schooner, Harty's Quay, Rochestown
850 9 THE LINKS HOLIDAY VILLAGE, LAHINCH, HOLIDAY VILLAGE
875 APT 1 MARYSTONE CENTRE GLADSTONE ST, CLONMEL, TIPPERARY
914 NO 1 OLD DOWNS ROAD, WILLOW GROVE, DELGAMY
935 25 Tynan Hall Park  Tynan Hall, Kingswood, Ballymount Road  Tallaght
1026 NO 36 CUIRT NA HABHAINN, CLAREGALWAY, CO GALWAY
1059 Unit 1  3  5  7 & 9, Rahillion Walk, Donabate
1127 18 CAISEAL NA RI GOLDEN RD, CASHEL, TIPPERARY
1237 THE OLD RECTORY - MANSE, HOLLYMOUNT, CLAREMORRIS
1392 UNIT 6 6 ST.ANTHONY'S ROAD, RIALTO, DUBLIN 8
1444 14 ST JOHNS WOOD PARK, ASHBOURNE, MEATH
1486 No 11 Elder Heath Crescent, Elder Heath, Kiltipper Road
1534 APT 40 GLEANN NA RIOGH GROVE, NAAS, COUNTY KILDARE
1550 APT 4 IVY LEAF HOUSE, BRIAN BORU SQUARE, FERMOY
1692 105  104 and 99 The Fairways, Tubbercurry
1872 APT 2 GARRYKNOCK  BRACKEN PARK, CARPENTERSTOWN ROAD, CASTLEKNOCK
2023 4 UPPER ST JOSEPH'S TERRACE, ATHY, CO. KILDARE
2191 1 PRIMROSE CRESCENT THE PASTURES, LOVE LANE, CHARLEVILL
2287 168 WEST AVE PARK GATE FRANKFIELD, DOUGLAS, CORK
2296 ST DECLANS 11 WILTON LAWN, GLASHEEN RD, CORK
2351 11 ABBOTT COURT THE HERMIYAGE, HOLYCROSS, THURLES
2461 82 OTHERWISE KNOWN AS 81A, WOODLAWN PARK DRIVE, FIRHOUSE
2510 APT 53 BRU NA SIONNA, SHANNON
2528 Apt 82 Leas na Mara, Ballymoneen Road
2622 115 BELFRY HALL BLOCK C, BELFRY HALL, CITYWEST RD DUBLIN 24
2687 No. 5 Gort na dTulach, Tullagh Hill, Loughrea
2739 69 CILL FOIRENNN JOHNSTOWN NAVAW, MEATH
2762 APT 23 THE WATER FRONT, HANOVER QUAY, DUBLIN 2
2808 9 upper prince edward terrace, blackrock
2809 APT 9   49 ST. ANTHONY'S ROAD, HERBERTON, RIALTO
2867 203 CRATLOE WOOD STUDENT VILLAGE, OLD CRATLOE RD, LIMERICK
2896 22 & 33 Loughcrew Hills, Oldcastle
2938 7 CNOC NA GREINE PARK, KILCULLEN, KILDARE
2943 Apt. 401 Longboat Quay South, Hanover Quay
```

2996 LES BUISSONMETS 12 PINE VALLEY AVE, RATHFARNHAM, DUBLIN 16
3179 NO. 16 THE WELLNESS CENTRE, GORTUMULLEN, KENMARE
3226 APT 127 BLOCK C ARAS NA CLUAINE, WATERY LANE, CLONDALKIN
3272 68 BANNOW RD CABRA WEST, DUBLIN
3296 42 FATHER PAUL MURPHY ST, EDENDERRY, OFFALY
3493 UNITS 1 - 28 BLOCK D, GANDON VIEW, LUCAN
3544 FLAT 1 TO FLAT 10, 90 LOWER BAGGOT STREET, DUBLIN 2
3623 4 Parc na Silla Lane, Loughlinstown
4024 5 Glebe Crescent ors The Fair Green, Market Green, Balbriggan
4143 APT 13 LAR NA CATHRACH, ABBEY TRINITY, TUAM
4170 29 PHILIP ONEILL PALCE COBH, CORK
4209 No 4 College View Place, Westport Road, Castlebar
4319 Apt 61 Leas na Mara, Ballymoneen Road, Galway
4327 7 & 7A RINN LIA, OLD BALLYBRIT RD, BALLYBRIT
4347 21 ARBURN WAY CLUAIN RI, ATHLONE ROAD, BALLYMAHON
4436 21A Hole in the Wall Road, Donaghmede
4460 82 CLEARSTREAM COURT MCKEE AVE, FINGLAS DUBLIN 1, DUBLIN
4463 Nos. 18 20 24 25 33 & 34, An Bonnan Bui Way, Blacklion
4512 APT 3 ST KEVINS HOUSE, BLOOMFIELD AVENUE, DUBLIN 8
4612 No 4 Church View Gardens, Watergrasshill, Cork
4615 197 CRATLOE WOOD STUDENT VILLAGE, OLD CRATLOE RD, LIMERICK
4698 12 properties at Baker's Yard, Portland Street North
4747 Apartment No 67 Leitrim Marina, Leitrim Village
4775 15 17 18 and 19 ST COLUMCILLES DR, 1 1B 2 3 and 4 Seatown Road, Swords
4851 3 7 DE BRUIN COURT, POLEBERRY, WATERFORD
4887 APTS 1 2 & 4 GARVILLE HOUSE, GARVILLE AVENUE, RATHGAR
4978 16 & 17 ANNESLEY BRIDGE ROAD, NORTH STRAND ROAD, DUBLIN 3
5021 27 JOHN MC CORMACK AVE, WALKINSTOWN, DUBLIN 12
5098 18 CARRIGWEIR WEIR RD TUAM, GALWAY
5169 13 J.K.L STREET APT B, EDENDERRY, CO OFFALY
5175 144 ST JOHNS WOOD WEST, CLONDALKIN, DUBLIN 22
5180 APT 48 OLD KILMAINHAM VILLAGE, BOWE LANE, DUBLIN CITY
5219 36 The Beeches The Beeches, Archerstown Demesne, Ashbourne County Meath
5227 155 Block 9 The Schooner, Harty's Quay, Rochestown
5246 31 AND 35 ROMAN ST, CORK CITY, CORK
5251 SITE NO. 31 ROYAL CANAL AVENUE, ROYAL CANAL PARK, RATOATH ROAD
5289 APT 21 MERCHANTS SQUARE APARTMENTS, ENNIS, COUNTY CLARE
5503 27 Block E3 Ennis House, Clare Village, Clare Hall
5526 5 THE LAWN BLOCK A, ABBEYLANDS, CLANE
5730 28 Parc Na Silla Rise, Loughlinstown
5799 21 ACHILL SOUND HOLIDAY VILLAGE, (BARRA NA FARRAIGE), ACHILL SOUND
5826 Apartment 203 Bushy Park House, Templeogue Road, Terenure
5854 166 CRATLOE WOOD STUDENT VILLAGE, OLD CRATLOE RD, LIMERICK
5896 APT2 BLOCK 1 ARD AOIBHINN, CRADDOCKSTOWN ROAD, NAAS
5899 APT 11 CRUISE PARK SQUARE, TYRELLSTOWN
5905 FLAT 5   1 HOLLYBANK ROAD, DRUMCONDRA, DUBLIN 9
5948 166 LOUISA PARK STATION RD, LEIXLIP, KILDARE
6043 Kylemore Estate Ard Na Gaoithe, Cashel Road, Clonmel
6057 APARTMENTs 2 3 and 4, MORRISTOWNBILLAR, NEWBRIDGE
6098 APT 4 80 REUBEN STREET, RIALTO
6110 House No. 83 Dundoogan (Sector 1), Haynestown, Dundalk
6121 83 Cois Na Mara Golf Links Rd, Bettystown, Co Meath
6210 3 An Cuanin Ard Glas, Ferrybank
6353 32 Crest of The Wave, Magheracar, Bundoran
6438 House No. 76 Dundoogan (Sector 1), Haynestown, Dundalk
6459 22 THE OLD MILL RACE, ATHGARVAN, KILDARE
6514 Apt 34 Block Q An Clarin, Athenry
6519 83 HILL OF DOWN HOUSE, SPENCER DOCK, DUBLIN 1
6541 10 ARD MOR GREEN FORTUNESTOWN LANE, TALLAGHT, DUBLIN
6599 NO 28 CUIRT NA HABHAINN, CLAREGALWAY, GALWAY
6696 4 SCRIBBLESTOWN RD FINGLAS SOUTH, DUBLIN 11, DUBLIN
6705 No. 19 Tullan Strand Rd, Bundoran
6707 APT 13 PARKVIEW 113 HAROLDS X RD, DUBLIN 6, DUBLIN
6742 68 THE VIEW ST WOLSTANS ABBEY, CELBRIDGE, KILDARE
6827 APT ABOVE THE MAPLE BAR, SAINT BRENDANS STREET, PORTUMNA
6888 No. 32 Lana and Bhaile, Craughwell

```
6897 No. 5 Garrai an Droichead, Killora, Craughwell
6994 31 THE PARK BLUE CEDARS, BALLYBOFEY, DONEGAL
7037 BLOCK 9 APT 26 THE GRIFFITH, PROSPECT HILL, TOLKA VALLEY ROAD  FINGLAS
7066 138 THE COPPICE WOODFARM ACRES, PALMERSTOWN DUBLIN 20, DUBLIN
7080 NO. 2 THE SCHOOL YARD, PETER STREET   -   KNIGHTSTOWN, VALENTIA ISLAND
7092 9 APT BLOCK 2 COIS ABHAINN CLANE, KILDARE
7157 APT 6 ST MARYS COURT, ST MARYS AVE, THURLES
7164 APT 107 OLD COUNTY ROAD, CRUMLIN, DUBLIN 12
7208 APT22 GRANGE HOUSE TAYLORS HILL, COLLEGE RD, RATHFARNHAM DUBLIN 16
7231 33 SHANNONVALE OLD CRATLOE RD, LIMERICK
7303 4 6 & 8 Nephin Court, Ballina, County Mayo
7351 56  57  59  60 Ivy Grove, Knockbeg West, Collooney
7409 Apt 15 New Bancroft Place, Main Street, Tallaght
7474 9 WILLOW LAWN PRIM RD CELBRDIGE, KILDARE
7603 35 2 4 GEORGES QUAY, DUBLIN 2, DUBLIN
7642 48 SLI GHEAL BALLYMONEEN RD, KNOCKNACARRA, GALWAY
7745 106 BELGROVE PK MT PROPSECT LAWNS, MOUNT PROSPECT AVE CLONTARF, DUBLIN CITY
7805 PT 76 BLOCK G THE BRACKEN, MARINA VILLAGE, GREYSTONES
7923 APT. 1 WATERPARK RUGBY CLUB, LOWER MAYPARK LANE, BALLINAKILL
8176 84 SAND DUNE HOLIDAY COTTAGES, BANNA, KERRY
8192 79 BARLEY FIELD  WICKLOW HILLS, NEWTOWNMOUNTKENNEDY, WICKLOW
8268 47 JOHN MC CORMACK AVENUE, WALKINSTOWN, DUBLIN 12
8299 Apt 2  75A THOMAS MOORE ROAD, WALKINSTOWN, DUBLIN 12
8342 74 ABBEY COURT KILLALA RD, BALLINA, CO MAYO
8402 40  40A  40B & 40C Carrowmore Drive, Knock
8405 8 PRINCE EDWARD TCE LR, BLACKROCK, CO. DUBLIN
8413 80 CARRS MILL PORTRANE RD, DONABATE, DUBLIN
8432 No 1 Manor Hill Terrace, Waterford
8496 apt 38 dun na ri, Gort
8718 UNIT 7 GLASAN STUDENT VILLAGE, BALLYBANE RD, MERVUE
8766 23 St Columba's Road Lower, Drumcondra, Dublin
8785 Apt 5 Ard Na Dara, Ardmore Road, Mullingar
8828 8 LOCK MILLS GROVE ISLAND, CORBALLY, LIMERICK
8934 JAMESWELL COURT APT 1 BLOCK 2, NEWBRIDGE, KILDARE
8986 APT 1 6 CATHERINE STREET, WATERFORD, CO WATERFORD
8995 E 105 AN TSEAN MHARGADH, GREENLANES, DROGHEDA
9058 APT 20 TURNSTONE BLOCK B, THORNWOOD BOOTERSTOWN AVE, BOOTERSTOWN
9076 60 CUSTOM HOUSE HARBOUR APTS, CUSTOM HOUSE HARBOUR, DUBLIN 1
9115 Apartment No. 66 St. Steevens Gate, 126/133 James Street
9200 Apartment No. 54 & Car Space 54, The Forum, Sandyford
9249 Apt 34  Block C  An Radharc, Maryborough Ridge, Maryborough Hill  Douglas
9293 UNIT NO. 29 2ND FLOOR, NORTHBANK, CASTLEFORBES ROAD
9311 121 THE GLEBE CARLANSTOWN RD KELLS, MEATH
9337 19 An Bunnan Bui Way, Blacklion
9409 Apt 5 New Bancroft Centre, Main Street, Tallaght
9415 Unit 100 Royal Canal Court, Royal Canal Way, Ashtown
9436 2 DROM NA NANE PK, BEAUMONT, DUBLIN 9
9484 NO 11 ARD NA BHAILE, BUTTEVANT, CO CORK
9561 Apt . 37 The garden House, Waterfall Avenue, Drumcondra
9597 No. 8 Gort na dTulach, Tullagh Hill, Loughrea
9666 APT. 3 - AN CUIRT, MONARD, CO TIPPERARY
9667 3 THE AVENUE RATHDALE ENFIELD, MEATH
9773 Apt. 70  Old Kilmainham Village, Bowe Lane, Kilmainham
9775 BLOCK H 165 CRANN NUA, EDENDERRY RD, PORTARLINGTON
9811 APT 2 23 HIGH STREET, LIMERICK
9852 Number 18 The Old Mill Race, Monasterevin
9857 No. 6 The Avenue  Inse Bay, Laytown  County Meath
9911 22 the Belfry (site only), Trim
9929 2 MICHAEL O REGAN PLACE, RATHASS, TRALEE
```

The above gives us a significantly shorter list to review. While it is not full guaranteed that it will capture every single invalid row, it will certainly allow us to cut down on the amount of outliers. In any event, it is not feasible to work through every single row in the dataframe to check, so this seems to be the best option given the size of the dataset.

Now, we will print the entries that represent the sale of multiple properties. We can see that some valid rows have been captured by our review - for example, where the first line in an address is particularly long, such as "22 the Belfry (site only)" - so we will ignore these rows.

In [57]:

```
multiple_sale_indices = [86, 121, 1059, 1692,
                         2896, 3493, 3544, 4327,
                         4463, 4698, 4887, 4978,
                         4775, 5246, 6057, 7303,
                         7351, 7603, 8402]

for multiple_sale_index in multiple_sale_indices:
    print(df.loc[multiple_sale_index])
    print("********************************************************************************")

print(df.loc[86])
```

```
Date of Sale                                          2020-03-06 00:00:00
Address                     6 7 8 9 ST ANNES TERRACE, SHORTCOURSE, WATERFORD
Postal Code                                                            NaN
County                                                           Waterford
Price (€)                                                          40000.0
Not Full Market Price                                                   No
VAT Exclusive                                                           No
Description of Property                 Second-Hand Dwelling house /Apartment
Property Size Description                                              NaN
Name: 86, dtype: object
********************************************************************************
Date of Sale                                          2019-12-18 00:00:00
Address                     1  5  6 & 22 THE HAMPTON, APT 1 2 7 19 20 26 2...
Postal Code                                                       Dublin 11
County                                                               Dublin
Price (€)                                                         3919000.0
Not Full Market Price                                                   No
VAT Exclusive                                                           No
Description of Property                 Second-Hand Dwelling house /Apartment
Property Size Description                                              NaN
Name: 121, dtype: object
********************************************************************************
Date of Sale                                          2019-04-18 00:00:00
Address                     Unit 1  3  5  7 & 9, Rahillion Walk, Donabate
Postal Code                                                            NaN
County                                                              Dublin
Price (€)                                                         1231602.74
Not Full Market Price                                                  Yes
VAT Exclusive                                                          Yes
Description of Property                     New Dwelling house /Apartment
Property Size Description                                              NaN
Name: 1059, dtype: object
********************************************************************************
Date of Sale                                       2021-01-12 00:00:00
Address                     105  104 and 99 The Fairways, Tubbercurry
Postal Code                                                       NaN
County                                                          Sligo
Price (€)                                                     105726.0
Not Full Market Price                                              No
VAT Exclusive                                                     Yes
Description of Property                 New Dwelling house /Apartment
Property Size Description                                         NaN
Name: 1692, dtype: object
********************************************************************************
Date of Sale                                  2014-12-30 00:00:00
Address                     22 & 33 Loughcrew Hills, Oldcastle
Postal Code                                                 NaN
County                                                    Meath
```

```
Price (€)                                          52863.44
Not Full Market Price                                 No
VAT Exclusive                                        Yes
Description of Property          New Dwelling house /Apartment
Property Size Description           greater than 125 sq metres
Name: 2896, dtype: object
**************************************************************************
Date of Sale                                2020-06-24 00:00:00
Address                      UNITS 1 - 28  BLOCK D, GANDON VIEW, LUCAN
Postal Code                                                   NaN
County                                                     Dublin
Price (€)                                               8629956.0
Not Full Market Price                                        No
VAT Exclusive                                               Yes
Description of Property               New Dwelling house /Apartment
Property Size Description                                     NaN
Name: 3493, dtype: object
**************************************************************************
Date of Sale                                          2021-05-06 00:00:00
Address                      FLAT 1 TO  FLAT 10, 90 LOWER BAGGOT STREET, DU...
Postal Code                                                           NaN
County                                                             Dublin
Price (€)                                                       1525000.0
Not Full Market Price                                                 No
VAT Exclusive                                                         No
Description of Property               Second-Hand Dwelling house /Apartment
Property Size Description                                             NaN
Name: 3544, dtype: object
**************************************************************************
Date of Sale                                2019-10-22 00:00:00
Address                      7 & 7A RINN LIA, OLD BALLYBRIT RD, BALLYBRIT
Postal Code                                                   NaN
County                                                     Galway
Price (€)                                               195000.0
Not Full Market Price                                        No
VAT Exclusive                                                No
Description of Property            Second-Hand Dwelling house /Apartment
Property Size Description                                     NaN
Name: 4327, dtype: object
**************************************************************************
Date of Sale                                          2015-10-08 00:00:00
Address                      Nos. 18 20 24 25 33 & 34, An Bonnan Bui Way, B...
Postal Code                                                           NaN
County                                                              Cavan
Price (€)                                                       202643.0
Not Full Market Price                                                 No
VAT Exclusive                                                        Yes
Description of Property                          New Dwelling house /Apartment
Property Size Description    greater than or equal to 38 sq metres and less...
Name: 4463, dtype: object
**************************************************************************
Date of Sale                                          2014-12-09 00:00:00
Address                      12 properties at Baker's Yard, Portland Street...
Postal Code                                                        Dublin 1
County                                                               Dublin
Price (€)                                                       1229331.9
Not Full Market Price                                                 No
VAT Exclusive                                                        Yes
Description of Property                          New Dwelling house /Apartment
Property Size Description    greater than or equal to 38 sq metres and less...
Name: 4698, dtype: object
**************************************************************************
Date of Sale                                          2021-05-20 00:00:00
Address                      APTS 1  2 & 4 GARVILLE HOUSE, GARVILLE AVENUE,...
Postal Code                                                        Dublin 6
County                                                               Dublin
```

```
Price (€)                                                          500000.0
Not Full Market Price                                                    No
VAT Exclusive                                                            No
Description of Property                  Second-Hand Dwelling house /Apartment
Property Size Description                                               NaN
Name: 4887, dtype: object
**************************************************************************
Date of Sale                                              2017-08-24 00:00:00
Address                       16 & 17 ANNESLEY BRIDGE ROAD, NORTH STRAND ROA...
Postal Code                                                             NaN
County                                                               Dublin
Price (€)                                                          842000.0
Not Full Market Price                                                    No
VAT Exclusive                                                            No
Description of Property                  Second-Hand Dwelling house /Apartment
Property Size Description                                               NaN
Name: 4978, dtype: object
**************************************************************************
Date of Sale                                              2018-03-07 00:00:00
Address                       15 17 18 and 19 ST COLUMCILLES DR, 1 1B 2 3 an...
Postal Code                                                             NaN
County                                                               Dublin
Price (€)                                                         3400000.0
Not Full Market Price                                                    No
VAT Exclusive                                                            No
Description of Property                  Second-Hand Dwelling house /Apartment
Property Size Description                                               NaN
Name: 4775, dtype: object
**************************************************************************
Date of Sale                                    2015-12-04 00:00:00
Address                        31 AND 35 ROMAN ST, CORK CITY, CORK
Postal Code                                                    NaN
County                                                        Cork
Price (€)                                                 68750.0
Not Full Market Price                                          No
VAT Exclusive                                                  No
Description of Property      Second-Hand Dwelling house /Apartment
Property Size Description                                      NaN
Name: 5246, dtype: object
**************************************************************************
Date of Sale                                              2017-07-20 00:00:00
Address                       APARTMENTs 2  3 and 4, MORRISTOWNBILLAR, NEWBR...
Postal Code                                                             NaN
County                                                           Kildare
Price (€)                                                          270000.0
Not Full Market Price                                                    No
VAT Exclusive                                                            No
Description of Property                  Second-Hand Dwelling house /Apartment
Property Size Description                                               NaN
Name: 6057, dtype: object
**************************************************************************
Date of Sale                                              2015-07-31 00:00:00
Address                             4 6 & 8 Nephin Court, Ballina, County Mayo
Postal Code                                                             NaN
County                                                                Mayo
Price (€)                                                         222026.43
Not Full Market Price                                                    No
VAT Exclusive                                                           Yes
Description of Property                        New Dwelling house /Apartment
Property Size Description    greater than or equal to 38 sq metres and less...
Name: 7303, dtype: object
**************************************************************************
Date of Sale                                              2015-10-16 00:00:00
Address                       56  57  59  60 Ivy Grove, Knockbeg West, Collo...
Postal Code                                                             NaN
County                                                               Sligo
```

```
Price (€)                                                          595060.3
Not Full Market Price                                                    No
VAT Exclusive                                                           Yes
Description of Property                          New Dwelling house /Apartment
Property Size Description    greater than or equal to 38 sq metres and less...
Name: 7351, dtype: object
**********************************************************************
Date of Sale                                    2017-05-17 00:00:00
Address                      35 2 4 GEORGES QUAY, DUBLIN 2, DUBLIN
Postal Code                                                Dublin 2
County                                                       Dublin
Price (€)                                                 233000.0
Not Full Market Price                                           No
VAT Exclusive                                                   No
Description of Property      Second-Hand Dwelling house /Apartment
Property Size Description                                      NaN
Name: 7603, dtype: object
**********************************************************************
Date of Sale                                           2015-06-23 00:00:00
Address                           40  40A  40B & 40C Carrowmore Drive, Knock
Postal Code                                                              NaN
County                                                                 Mayo
Price (€)                                                          127000.0
Not Full Market Price                                                    No
VAT Exclusive                                                           Yes
Description of Property                          New Dwelling house /Apartment
Property Size Description    greater than or equal to 38 sq metres and less...
Name: 8402, dtype: object
**********************************************************************
Date of Sale                                          2020-03-06 00:00:00
Address                      6 7 8 9 ST ANNES TERRACE, SHORTCOURSE, WATERFORD
Postal Code                                                              NaN
County                                                             Waterford
Price (€)                                                            40000.0
Not Full Market Price                                                    No
VAT Exclusive                                                            No
Description of Property                Second-Hand Dwelling house /Apartment
Property Size Description                                                NaN
Name: 86, dtype: object
```

We can add more detail to these entries by placing the number of properties sold in that particular entry as another element in the array.

In [58]:

```
multiple_sale_indices_with_number_of_properties = [(86,4), (121,23), (1059,5), (1692,3),
                                                   (2896,2), (3493,28), (3544,11), (4327,2),
                                                   (4463,6), (4698,12), (4887,3), (4978,2),
                                                   (4775,9), (5246,2), (6057,3), (7303,3),
                                                   (7351,4), (7603,3), (8402,4)]


total_new_rows = 0
for number_of_properties in multiple_sale_indices_with_number_of_properties:
    total_new_rows += number_of_properties[1]
print(f"The total number of new rows is {total_new_rows}")
```

```
The total number of new rows is 129
```

In order to efficiently split these rows out, we will have to write a short function that will:

- Take in the number of properties represented by the row.
- Calculate the average price of each property (we are making an assumption here that each property had the same purchase price. We are making this assumption because we do not have access to the domain expert at this time and taking the average seems to be the most rational estimate that can be made)

- Add new rows containing the new price to a dataframe. We will not make any changes to the description of the property so that we can tell whether or not this property is part of a group purchase. Perhaps at a later stage we can create a feature on this basis.

In [59]:
```
def get_new_row(df, index, average_price):
    """Function that creates a new row with a specific price in our dataframe"""
    new_row = {"Date of Sale": df.loc[index]["Date of Sale"],
               "Address": df.loc[index]["Address"],
               "Postal Code": df.loc[index]["Postal Code"],
               "County": df.loc[index]["County"],
               "Price (€)": average_price,
               "Not Full Market Price": df.loc[index]["Not Full Market Price"],
               "VAT Exclusive": df.loc[index]["VAT Exclusive"],
               "Description of Property": df.loc[index]["Description of Property"]
               }
    return new_row
```

In [60]:
```
def get_split_out_properties(df, indices_and_number_of_properties):
    """Function that splits out multiple properties"""
    split_out_properties = []

    # Loop through indices
    for index_and_number_of_properties in indices_and_number_of_properties:

        # Get index and number of properties
        index = index_and_number_of_properties[0]
        number_of_properties = index_and_number_of_properties[1]

        # Get average price by dividing price by number of properties, create new row
        average_price = df["Price (€)"][index] / number_of_properties
        new_row = get_new_row(df, index, average_price)

        # Create necessary number of rows depending on how many properties are in the entry
        current_property = 1
        while current_property <= number_of_properties:
            split_out_properties.append(new_row)
            current_property+=1

    df_split_out_properties = pd.DataFrame(split_out_properties, columns=["Date of Sale","Addre
                                                                          "Postal Code","Co
                                                                          "Price (€)","Not
                                                                          "VAT Exclusive","
    return df_split_out_properties
```

In [61]:
```
# Split out the properties
df_split_out_properties = get_split_out_properties(df, multiple_sale_indices_with_number_of_pro
df_split_out_properties
```

Out[61]:

	Date of Sale	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property
0	2020-03-06	6 7 8 9 ST ANNES TERRACE, SHORTCOURSE, WATERFORD	NaN	Waterford	10000.000000	No	No	Second-Hand Dwelling house /Apartment

	Date of Sale	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property
1	2020-03-06	6 7 8 9 ST ANNES TERRACE, SHORTCOURSE, WATERFORD	NaN	Waterford	10000.000000	No	No	Second-Hand Dwelling house /Apartment
2	2020-03-06	6 7 8 9 ST ANNES TERRACE, SHORTCOURSE, WATERFORD	NaN	Waterford	10000.000000	No	No	Second-Hand Dwelling house /Apartment
3	2020-03-06	6 7 8 9 ST ANNES TERRACE, SHORTCOURSE, WATERFORD	NaN	Waterford	10000.000000	No	No	Second-Hand Dwelling house /Apartment
4	2019-12-18	1 5 6 & 22 THE HAMPTON, APT 1 2 7 19 20 26 2...	Dublin 11	Dublin	170391.304348	No	No	Second-Hand Dwelling house /Apartment
...	...	...	...	...	...	...	...	...
124	2017-05-17	35 2 4 GEORGES QUAY, DUBLIN 2, DUBLIN	Dublin 2	Dublin	77666.666667	No	No	Second-Hand Dwelling house /Apartment
125	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	No	Yes	New Dwelling house /Apartment
126	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	No	Yes	New Dwelling house /Apartment
127	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	No	Yes	New Dwelling house /Apartment
128	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	No	Yes	New Dwelling house /Apartment

129 rows × 8 columns

We can see from the above that the new function seems to have worked. We can see some odd outlier values here still, but as a whole the logical integrity of the address section is now intact.

Now we just need to drop the original outlier rows from our dataframe, reset the index, and append the new dataframe onto it.

In [62]:

```
# Remove original rows and rest index
for multiple_sale_index in multiple_sale_indices:
```

```
    df.drop(index = multiple_sale_index, inplace = True)
    df = df.reset_index(drop = True)

df = pd.concat([df, df_split_out_properties], ignore_index = True)

df.head()
```

Out[62]:

	Date of Sale	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property	Property Size Description
0	2020-06-26	1 GANDON PLACE, GANDON PARK, Lucan	NaN	Dublin	370044.05	No	Yes	New Dwelling house /Apartment	NaN
1	2014-12-19	72 ST ASSAMS PARK, RAHENY, DUBLIN 5	Dublin 5	Dublin	480000.00	No	No	Second-Hand Dwelling house /Apartment	NaN
2	2010-02-11	37 Ivy Hill, Gort Road, Ennis	NaN	Clare	194000.00	No	No	Second-Hand Dwelling house /Apartment	NaN
3	2018-08-16	14 TYRCONNELL PLACE, TYRCONNELL RD, INCHICORE ...	Dublin 8	Dublin	275000.00	No	No	Second-Hand Dwelling house /Apartment	NaN
4	2019-02-27	7 THE NEST, TUBBERCURRY, SLIGO	NaN	Sligo	75000.00	No	No	Second-Hand Dwelling house /Apartment	NaN

In [63]:

```
df.tail()
```

Out[63]:

	Date of Sale	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property	Property Size Description
10104	2017-05-17	35 2 4 GEORGES QUAY, DUBLIN 2, DUBLIN	Dublin 2	Dublin	77666.666667	No	No	Second-Hand Dwelling house /Apartment	NaN
10105	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	No	Yes	New Dwelling house /Apartment	NaN
10106	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	No	Yes	New Dwelling house /Apartment	NaN
10107	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	No	Yes	New Dwelling house /Apartment	NaN

	Date of Sale	Address	Postal Code	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property	Property Size Description
10108	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	No	Yes	New Dwelling house /Apartment	NaN

We can see now by viewing the head and the tail of the dataframe that the rows seem to have been split out successfully.

Please note that it is entirely possible that some entries showing multiple property purchases in our dataset may have slipped through the method referred to above, for example in situations where the first line of the address had less than four fields in it, or, as the *Information Note* states, there was a group purchase of both residential and commercial properties, which case only the residential would have been recorded in the dataset.

However, the above method should have filtered out a large amount of the multiple properties, so in the interest of efficiency we will proceed on this basis.

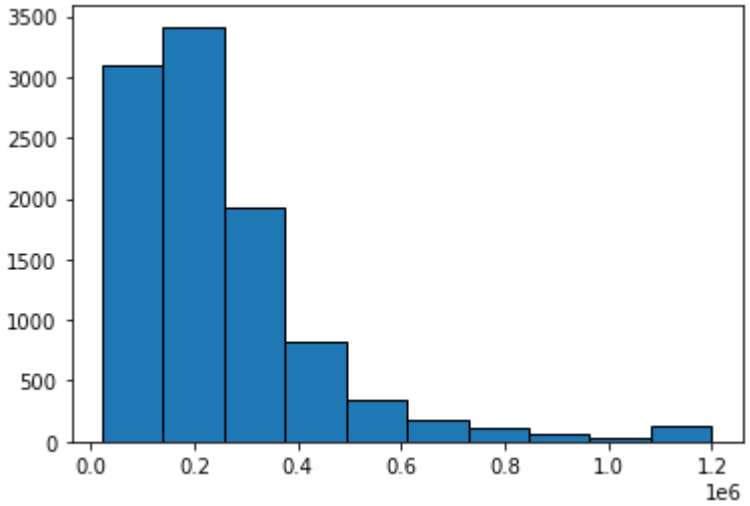
Next, we will need to decide what to do about the main outliers in the dataset, of which there are many, both low and high.

It would make sense to take the 1st percentile and the 99th percentile, and set all values below or higher that value to the value of the relevant percentile.

In [64]:
```
low_clamp = df["Price (€)"].quantile(0.01)
high_clamp  = df["Price (€)"].quantile(0.99)

print(f"The proposed lower threshold is €{low_clamp}")
print(f"The number of rows affected by low clamp is", df[df["Price (€)"]<low_clamp].shape[0])
print(f"The proposed higher threshold is €{high_clamp}")
print(f"The number of rows affected by high clamp is", df[df["Price (€)"]>high_clamp].shape[0])
```

```
The proposed lower threshold is €21513.28
The number of rows affected by low clamp is 102
The proposed higher threshold is €1199200.0000000007
The number of rows affected by high clamp is 102
```

In [65]:
```
df["Price (€)"].clip(low_clamp, high_clamp, inplace = True)

plt.hist(df["Price (€)"], edgecolor="black")
plt.show()
```



This distribution now looks more in line with what we originally saw with the more aggressive clamping in the report section. We can see a few outliers still towards the right end of the histogram, which is a result of us setting certain values to the 99th percentile, but overall, this now looks like a dataset that we can work with.

Address

An attempt was made to use the Folium library in order to transform the address field - however, the library seemed to have had quite a lot of issues with parsing some of the address strings. At the end of the process, only 15%-20% of strings were successfully parsed - this was decided to be too low a value to be considered satisfactory as a replacement for the "Address" feature. Any attempts to manipulate the address data to see if this would allow Geopy to parse successfully - such as changing the case of the address, or adding "Ireland" to the end of the string, proved to be unsuccessful.

As such, the Address feature was left the way it was. The reason that it was not dropped altogether is that perhaps at a later stage a new library will be released, or even a new update for Folium, that will be able to successfully parse the data in this dataset, in which case we will make use of this feature.

The resulting addresses were saved to a new csv file entitled "Property_Addresses.csv" - it is possible that we may still make use of this when considering new features for this dataset.

Attempts to parse the address feature can be seen in the cells below. It should be noted that due to the rules of API usage, a wait of one second is required between each address processing. Because of this, the entire process to work through the whole dataset takes approximately three hours, so these cells are commented out, so that the individual running this notebook isn't interrupted by this process.

In [66]:

```
"""
import pandas as pd
import geopy
import time


print(df.shape[0])

df_addresses_list = []
df_addresses = pd.DataFrame()
df_addresses["Address"] = 0
print(df_addresses)

for index in range(df.shape[0]):
    try:
        current_address = df["Address"][index]
        geolocator = geopy.Nominatim(user_agent="vrakhmanin1@gmail.com")
```

```
        location = geolocator.geocode(current_address, exactly_one=True)
        df_addresses = df_addresses.append({"Address": location}, ignore_index = True)
        df_addresses_list.append(location)
    except:
        pass
    time.sleep(1)

df_addresses.to_csv('Property_Addresses.csv', index=False)
print("Done!")
"""
```

Out[66]:
```
'\nimport pandas as pd\nimport geopy\nimport time\n\n\nprint(df.shape[0])\n\ndf_addresses_list
= []\ndf_addresses = pd.DataFrame()\ndf_addresses["Address"] = 0\nprint(df_addresses)\n\nfor in
dex in range(df.shape[0]):\n    try:\n        current_address = df["Address"][index]\n        g
eolocator = geopy.Nominatim(user_agent="vrakhmanin1@gmail.com")\n        location = geolocator.
geocode(current_address, exactly_one=True)\n        df_addresses = df_addresses.append({"Addres
s": location}, ignore_index = True)\n        df_addresses_list.append(location)\n    except:\n
pass\n    time.sleep(1)\n\ndf_addresses.to_csv(\'Property_Addresses.csv\', index=False)\nprint
("Done!")\n'
```

In [67]:
```
"""
import folium
m = folium.Map(location = [53.1424, 7.6921])
"""
```

Out[67]:
```
'\nimport folium\nm = folium.Map(location = [53.1424, 7.6921])\n'
```

In [68]:
```
"""
for i in range(len(df_addresses_list)):
    try:
       folium.Marker(
          location=[df_addresses_list[i].latitude, df_addresses_list[i].longitude],
          popup=df_addresses_list[i].address,
       ).add_to(m)
    except:
        pass

m
"""
```

Out[68]:
```
'\nfor i in range(len(df_addresses_list)):\n    try:\n       folium.Marker(\n          location
=[df_addresses_list[i].latitude, df_addresses_list[i].longitude],\n          popup=df_addresses
_list[i].address,\n       ).add_to(m)\n    except:\n        pass\n        \nm\n'
```

It should be noted that the addresses that *were* parsed successfully were able to be successfully placed on a Folium map. A screenshot of this can be seen below (in case this screenshot doesn't work, it is saved down as folium.png in this folder) - it is hard to tell precisely, but it seems that the distribution is relatively even, with a significantly higher density in Dublin. It is also interesting to see that there is an outlier with a property in Belfast - it looks like Geopy and/or Folium parsed this address incorrectly. We know that this is an issue with the library as opposed to the dataset because a search was run for "Belfast" within the .csv file and this search did not return any results.

Postal Code

It does not make sense to drop the entire Postal Code feature, as the missing values just represent the fact that the property is not located in Dublin. For this reason, we will rename the feature to ensure that it more accurately describes the data contained therein. We will keep the non-Dublin entries as "NaN", which will represent non-Dublin properties.

In [69]:
```
df.rename(columns={'Postal Code': 'Postal Code (Dublin)'}, inplace=True)
df.head()
```

Out[69]:

	Date of Sale	Address	Postal Code (Dublin)	County	Price (€)	Not Full Market Price	VAT Exclusive	Description of Property	Property Size Description
0	2020-06-26	1 GANDON PLACE, GANDON PARK, Lucan	NaN	Dublin	370044.05	No	Yes	New Dwelling house /Apartment	NaN
1	2014-12-19	72 ST ASSAMS PARK, RAHENY, DUBLIN 5	Dublin 5	Dublin	480000.00	No	No	Second-Hand Dwelling house /Apartment	NaN
2	2010-02-11	37 Ivy Hill, Gort Road, Ennis	NaN	Clare	194000.00	No	No	Second-Hand Dwelling house /Apartment	NaN
3	2018-08-16	14 TYRCONNELL PLACE, TYRCONNELL RD, INCHICORE ...	Dublin 8	Dublin	275000.00	No	No	Second-Hand Dwelling house /Apartment	NaN
4	2019-02-27	7 THE NEST, TUBBERCURRY, SLIGO	NaN	Sligo	75000.00	No	No	Second-Hand Dwelling house /Apartment	NaN

Description of Property

As mentioned above, the "Description of Property" feature currently contains only three values, one of which is in Irish. These values ultimately state if a property is either new or second-hand. We should flag that the value that is in Irish does not seem to specify whether or not the property is second-hand or not, so we will make the assumption that it is referring to a new property. In any event this assumption relates to one entry only, so even if it turns out to be incorrect, it should not cause any significant distortion of the data.

For these reasons, we can clean the "Description of Property" feature by transforming it into a simple boolean "Second-Hand" column which will indicate whether or not the purchase is of a second-hand property. The reason we are using "Second-Hand" as opposed to "New" as our feature name is because there are more second-hand properties being sold in our dataframe than new ones, and as such it makes sense to frame the feature using the majority value.

We will be replacing the current "Description of Property Feature" when we transform it. This is because it is very difficult to glean any information from the feature as it currently stands, and it requires the user to review the actual text values of the entries, which is unintuitive. For these reasons, there is no point in keeping the original feature - we can replace it with our new one.

In [70]:
```
df["Description of Property"].unique()
```

Out[70]:
```
array(['New Dwelling house /Apartment',
       'Second-Hand Dwelling house /Apartment',
       'Teach/Árasán Cónaithe Atháimhe'], dtype=object)
```

In [71]:
```
df.rename(columns={'Description of Property': 'Second-Hand'}, inplace=True)

df["Second-Hand"] = df["Second-Hand"].map({"Second-Hand Dwelling house /Apartment": True, "New
df.head()
```

Out[71]:

	Date of Sale	Address	Postal Code (Dublin)	County	Price (€)	Not Full Market Price	VAT Exclusive	Second-Hand	Property Size Description
0	2020-06-26	1 GANDON PLACE, GANDON PARK, Lucan	NaN	Dublin	370044.05	No	Yes	False	NaN
1	2014-12-19	72 ST ASSAMS PARK, RAHENY, DUBLIN 5	Dublin 5	Dublin	480000.00	No	No	True	NaN
2	2010-02-11	37 Ivy Hill, Gort Road, Ennis	NaN	Clare	194000.00	No	No	True	NaN
3	2018-08-16	14 TYRCONNELL PLACE, TYRCONNELL RD, INCHICORE ...	Dublin 8	Dublin	275000.00	No	No	True	NaN
4	2019-02-27	7 THE NEST, TUBBERCURRY, SLIGO	NaN	Sligo	75000.00	No	No	True	NaN

VAT Exclusive / Description of Property

As mentioned in the Data Quality Report, there is a possibility that both VAT Exclusive and Description of Property features are merely referring to the same facet of the data, namely because VAT is payable mostly on new homes, not second-hand ones. If the Description of Property merely refers to whether or not the property is second hand, then we can potentially drop this feature for redundancy.

However, upon further review, it was discovered that there are, in fact certain circumstances in which VAT could be payable on second-hand homes. For example, guidance from the Irish Revenue Commissioners linked here states that a vendor and purchaser may, for instance, decide to include VAT in a second-hand property purchase to avoid clawback.

To make an accurate assessment, we should first check how many property sale entries we have where VAT was payable on a second hand home, or vice-versa.

In [72]:
```
counter = 0

# Loop through rows
for index in range(df.shape[0]):

    # Check the value of VAT and set to variable
    if df.iloc[index]["VAT Exclusive"] == "Yes":
        vat = True
    else:
        vat = False

    # If second-hand matches VAT, the two categories are at odds and counter should be updated
    if df.iloc[index]["Second-Hand"] == vat:
        counter += 1

# Print counter to show number of rows were there is a difference between VAT and second hand
print(f"There are {counter} instances of mismatch between second-hand and VAT exclusive")
```

```
There are 26 instances of mismatch between second-hand and VAT exclusive
```

As we can see from the above, there are 26 rows where our assumption that both features point to the same facet of the data proves incorrect. We should note that:

- Irish tax law is not our area of expertise;
- it is not possible to contact the domain expert to clarify whether or not the mismatched entries are errors; and
- we have discovered reliable guidance from Revenue to suggest at least one reason as to why there might be a mismatch between Second-Hand and VAT Exclusive features.

For these reasons, it was decided to keep both features in, as it would seem that there is ambiguity as to whether or not these two features are one and the same, and we wish to minimise loss of data where possible.

Property Size Description

After careful consideration, I decided that keeping this data would not be beneficial to glean any new information for the data set. While there are a few entries from which I could potentially learn about specific property sizes, there are so few of these that it would be impossible to make any general statement about the relationship of the property size to any other feature in the data set.

For these reasons, I will be dropping this feature.

In [73]:
```
df.drop(columns=["Property Size Description"], inplace = True)

df.head()
```

Out[73]:

	Date of Sale	Address	Postal Code (Dublin)	County	Price (€)	Not Full Market Price	VAT Exclusive	Second-Hand
0	2020-06-26	1 GANDON PLACE, GANDON PARK, Lucan	NaN	Dublin	370044.05	No	Yes	False
1	2014-12-19	72 ST ASSAMS PARK, RAHENY, DUBLIN 5	Dublin 5	Dublin	480000.00	No	No	True
2	2010-02-11	37 Ivy Hill, Gort Road, Ennis	NaN	Clare	194000.00	No	No	True
3	2018-08-16	14 TYRCONNELL PLACE, TYRCONNELL RD, INCHICORE ...	Dublin 8	Dublin	275000.00	No	No	True
4	2019-02-27	7 THE NEST, TUBBERCURRY, SLIGO	NaN	Sligo	75000.00	No	No	True

Not Full Market Price

In order to make the information easier to work with and to understand, I will frame the feature as a positive as opposed to a negative (as saying "Not Full Market Price" is less intuitive than just saying "Full Market Price". For the same reasons I will map the values to boolean values).

In [74]:
```
df.rename(columns={'Not Full Market Price': 'Full Market Price'}, inplace=True)
df["Full Market Price"] = df["Full Market Price"].map({"No": True, "Yes": False})
df.head()
```

Out[74]:

	Date of Sale	Address	Postal Code (Dublin)	County	Price (€)	Full Market Price	VAT Exclusive	Second-Hand
0	2020-06-26	1 GANDON PLACE, GANDON PARK, Lucan	NaN	Dublin	370044.05	True	Yes	False

	Date of Sale	Address	Postal Code (Dublin)	County	Price (€)	Full Market Price	VAT Exclusive	Second-Hand
1	2014-12-19	72 ST ASSAMS PARK, RAHENY, DUBLIN 5	Dublin 5	Dublin	480000.00	True	No	True
2	2010-02-11	37 Ivy Hill, Gort Road, Ennis	NaN	Clare	194000.00	True	No	True
3	2018-08-16	14 TYRCONNELL PLACE, TYRCONNELL RD, INCHICORE ...	Dublin 8	Dublin	275000.00	True	No	True
4	2019-02-27	7 THE NEST, TUBBERCURRY, SLIGO	NaN	Sligo	75000.00	True	No	True

VAT Exclusive

For reasons similar to the ones described in the "Not Full Market Price" section, we will transform the values here to boolean values.

In [75]:
```
df["VAT Exclusive"] = df["VAT Exclusive"].map({"Yes":True, "No":False})
df.head()
```

Out[75]:

	Date of Sale	Address	Postal Code (Dublin)	County	Price (€)	Full Market Price	VAT Exclusive	Second-Hand
0	2020-06-26	1 GANDON PLACE, GANDON PARK, Lucan	NaN	Dublin	370044.05	True	True	False
1	2014-12-19	72 ST ASSAMS PARK, RAHENY, DUBLIN 5	Dublin 5	Dublin	480000.00	True	False	True
2	2010-02-11	37 Ivy Hill, Gort Road, Ennis	NaN	Clare	194000.00	True	False	True
3	2018-08-16	14 TYRCONNELL PLACE, TYRCONNELL RD, INCHICORE ...	Dublin 8	Dublin	275000.00	True	False	True
4	2019-02-27	7 THE NEST, TUBBERCURRY, SLIGO	NaN	Sligo	75000.00	True	False	True

Save to New Dataframe

We will now save down the resulting dataframe into a new .csv file, as requested in the question.

In [76]:
```
df.to_csv("q2_dataframe_after_data_quality_report.csv", index=False)
```

Question 3: Preparing relationships between feature pairs

Before starting this question, we must first read in our new cleaned dataframe. It should be noted that we will need to convert our features to categorical again.

In [77]:
```
df = pd.read_csv("q2_dataframe_after_data_quality_report.csv")

# Selecting categorical columns
categorical_columns = df[["Address", "Postal Code (Dublin)", "County"]].columns
boolean_columns = df[["Full Market Price", "VAT Exclusive", "Second-Hand"]].columns

# Looping through columns and casting as category
for column in categorical_columns:
```

```
    df[column] = df[column].astype('category')

# Looping through columns and casting as bool
for column in boolean_columns:
    df[column] = df[column].astype(bool)

# Setting Date of Sale to datetime format
df["Date of Sale"] = pd.to_datetime(df["Date of Sale"], infer_datetime_format=True)
```

In order to discuss relationships between the various features, we must first consider which of our categories will yield interesting correlations (or lack thereof). For reference, we will print out our categories once more below.

In [78]:
```
df.dtypes
```

Out[78]:
```
Date of Sale            datetime64[ns]
Address                       category
Postal Code (Dublin)          category
County                        category
Price (€)                      float64
Full Market Price                 bool
VAT Exclusive                     bool
Second-Hand                       bool
dtype: object
```

Both matplotlib and seaborn were considered for visualising the data. Ultimately, it was decided that seaborn added extra functionality to matplotlib that would be extremely beneficial for these purposes (i.e. catplot), so seaborn was used to plot the feature pairs referred to below.

Plots Considered (General)

Our target feature is Price, which is a continuous feature. The vast majority of our features are categorical, and as such most of our relationship plots will be categorical vs continuous. Different types of plots were considered to visualise this type of relationship, and specifically the following:

- Scatter Plot
- Box Plot
- Bar Plot
- Boxen Plot
- Violin Plot

Ultimately, only box and bar plots were chosen for the purposes of visualising these relationships. This is because bar plots will show us a general fluctuation between price when compared to a different feature, and the box plot will also show us the general spread of such values. By combining the two, we should get a relatively rich visual representation of the data set. It should be noted that for the bar plot, the estimator being shown is the mean. While it would have been possible to put in the median as the parameter, so that we have both values in bar plot form, it was decided not to do so as the median would have been shown clearly on the box plot anyway.

Plots Considered (Boolean)

One final point to flag is that for comparison of categorical features with boolean features, there were three different types of plots prepared, taking into account that all three boolean features have an overwhelmingly large first mode:

- a bar plot showing the proportion of both boolean features per category;

- a bar plot showing the minority boolean feature on its own per category (i.e. second mode feature, as it would not have been visible in the plot referred to in bullet point one); and
- a bar plot showing the percentage of the first mode per feature.

As in the sections shown above, stacked bar plots were considered, but given that the boolean features were overwhelmingly dominated by a particular value, such stacked bar plots were deemed not as readable as regular ones.

For the second bullet point, this type of graph was only included if the minority values became unreadable on the overall graph at bullet point one.

All three types of plots in combination should give us a solid understanding of the correlation, or lack thereof, between such features.

General Comments

Below are the feature pairs that were chosen for comparison, together with the reason for why they were chosen. The main feature pairs that we wish to take notice of are ones that interact with our target feature, which is price. These are listed in the first table. However, in order to more accurately understand our target domain, it is important to also select some feature pairs that we think may yield interesting results. It will not directly assist us with predicting the target feature, but it may allow us to understand the Irish property market more deeply - these are listed in the second table below.

Just to flag, not every possible feature pair was marked for plotting - only the ones that promised to yield interesting results. For example, VAT Exclusive we expect to have very similar results to Second-Hand, given how similar the two columns are, and as such, VAT Exclusive was only plotted against Second-Hand in the additional section.

As requested in the assignment, there is a short paragraph after each set of graphs briefly describing the findings so far. There is a summary table at the end of this question with an overall summary also.

Note: The assignment sheet did not specify that we should save the resulting plots as separate png files or a separate word document. For these reasons, I have included both the plots and the analysis in this notebook. I should flag that I have tested the code several times and have found that it reproduced the same results each time.

Target Feature Pairs

Feature Pair	Reason for Inclusion
Date of Sale and Price	Fluctuation of market prices over time
Postal Code and Price	Relationship between location of property and price
County and Price	Relationship between location of property and price
Second Hand and Price	Check if there is a relationship between price and whether or not the property is second hand
VAT Exclusive and Price	Check if there is a relationship between price and whether or not the property is VAT exclusive
Full Market Price and Price	Check if houses sold for full market price are more expensive than ones that are not

Additional Feature Pairs

Feature Pair	Reason for Inclusion
Date of Sale and Full Market Price	Check if the market has changed over time in favour of second hand or new

Feature Pair	Reason for Inclusion
Postal Code and Full Market Price	Check if more second hand or new homes are being sold in particular location
County and Full Market Price	Check if more second hand or new homes are being sold in particular location
Date of Sale and Second Hand	Check if purchase of second hand homes grew over time
Postal Code and Second Hand	Check if purchase of second hand homes was more prevalent in a particular postal code
County and Second Hand	Check if purchase of second hand homes was more prevalent in a particular county
Full Market Price and Second Hand	Relationship between full market price and second hand
VAT Exclusive and Second Hand	Check similarity between the two features

It is very important to note that all of our statements made below with relation to correlation between features are based on this dataset alone and from initial basic research into the Irish property market.

Before confirming any of the statements made herein, it will be useful to: (a) test our approach on a much larger dataset, to ensure that the correlations remain; and (b) consult with a domain expert to ensure that the statements we are making make sense within the context of our domain.

Target Feature Pairs

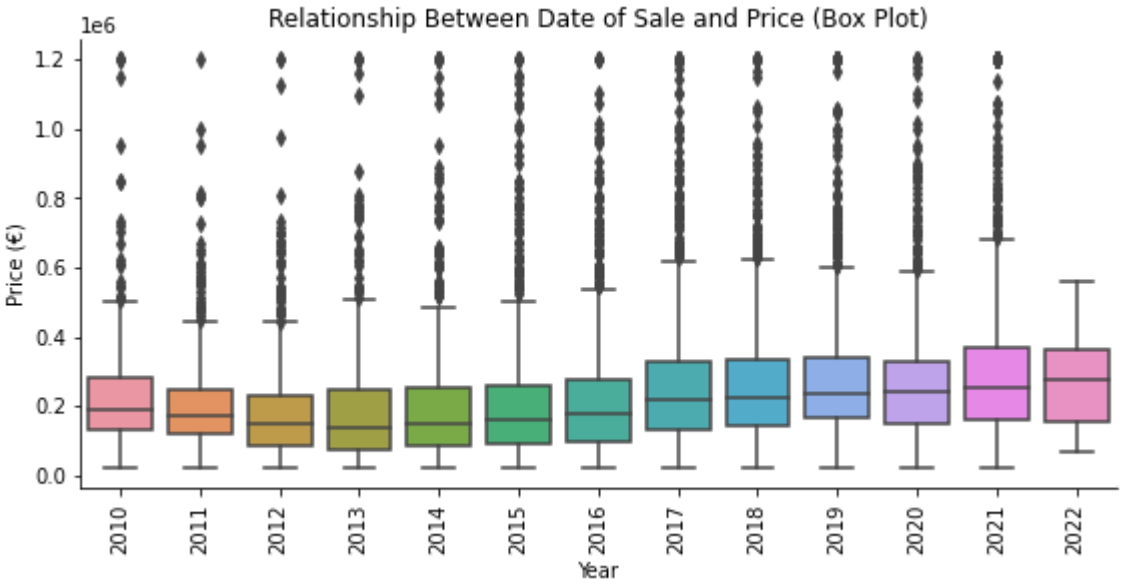
Date of Sale and Price

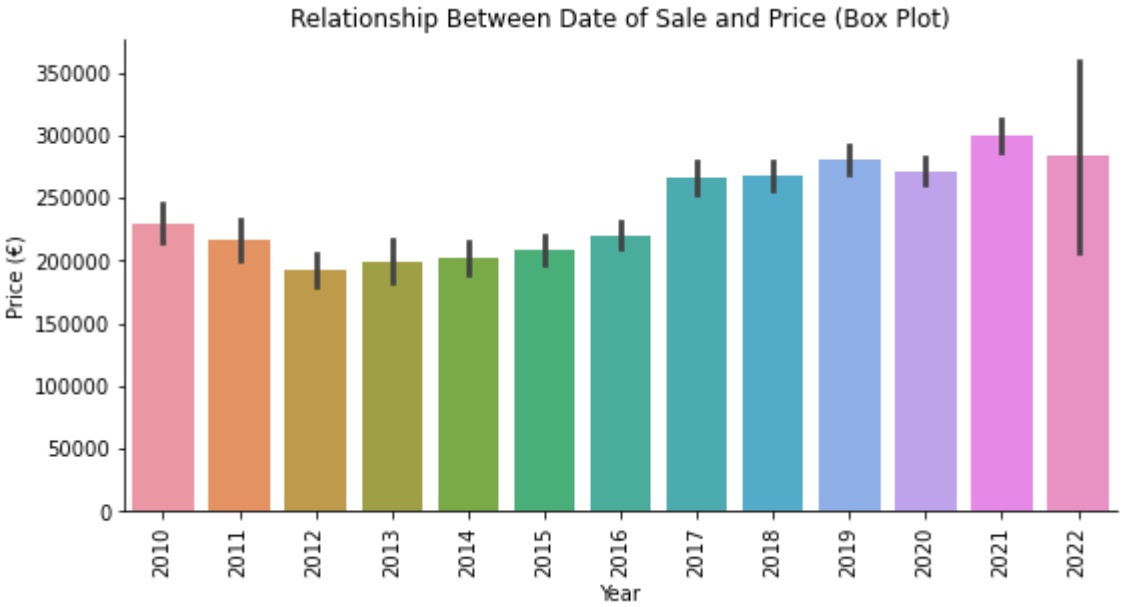
Initially, a continuous vs continuous scatter plot was considered for visualising this relationship. However, the resulting scatter plot was quite difficult to read. Following a discussion with a demonstrator during the labs, it was decided that a better way of visualising this relationship would be to extract the years from the Date of Sale feature, and to plot against the years instead. Converting Date of Sale to year causes the feature to become categorical, so we can use our standard methodology described above. We can also view this as us "binning" our data, i.e. making our continuous data discrete by placing each entry into a bin consisting of a single year.

While of course there is some loss of data here, as we do not get the level of detail here that we would have gotten had we used every specific date that we had, it was decided that this granularity was ultimately unnecessary for visualisation purposes, and viewing the price growth per year was more than sufficient for the purposes of this analysis.

Note: Whenever reference is made to certain prices (or any category) being higher in a category, this is just shorthand for saying "on average", or in accordance to the mean, as there may of course be prices in a category lower than in a different category even if the mean there is higher.

In [79]:
```
df['Year'] = pd.DatetimeIndex(df['Date of Sale']).year
sns.catplot(x="Year", y="Price (€)", kind="box", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between Date of Sale and Price (Box Plot)")
plt.show()

sns.catplot(x="Year", y="Price (€)", kind="bar", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between Date of Sale and Price (Box Plot)")
plt.show()
```





Discussion of Outcomes

Despite the rather large number of outliers that yet remain in this graph, there is a clear tendency that can be seen throughout, suggesting a correlation between year and price. It seems that on average, property prices grew steadily over the past decade, with a minor dip in 2012/2013. While the increase in price is not particularly dramatic, it is very clear that property prices are growing nonetheless, which makes sense as the country continues its recovery from the economic crisis.

We can also see from the bar chart that there are significantly fewer prices recorded for the year 2022, which makes sense, given the fact that as of the timing of writing this assignment we are only in the month of March. However, the mean so far does show that the prices, while a little lower than last year, are still in line with the growth shown in previous years.

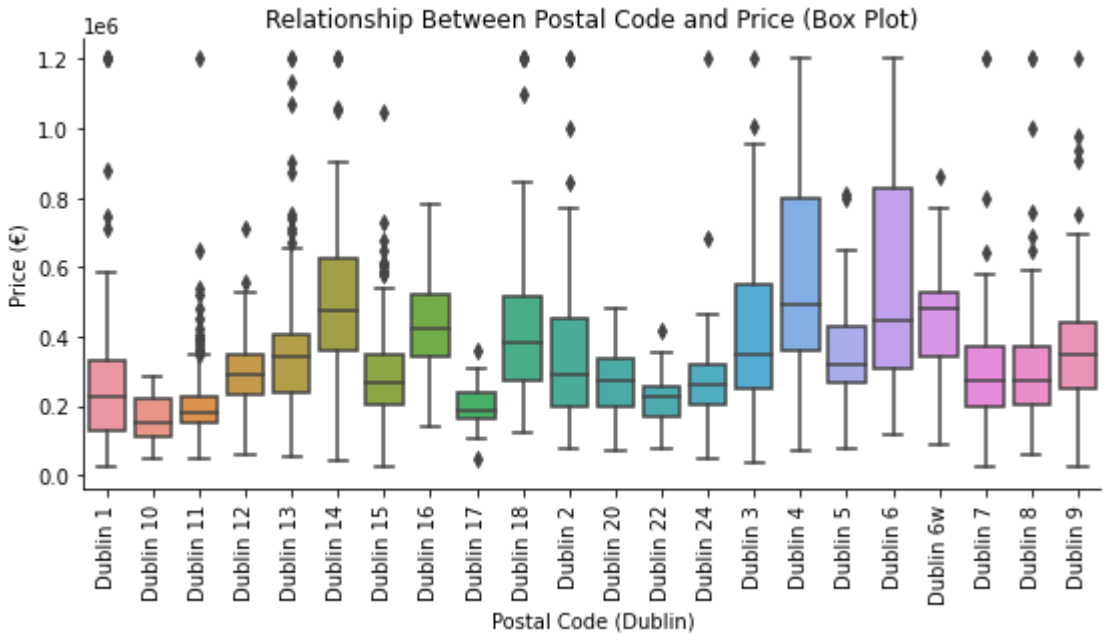
The spread seems to be consistent between all years.

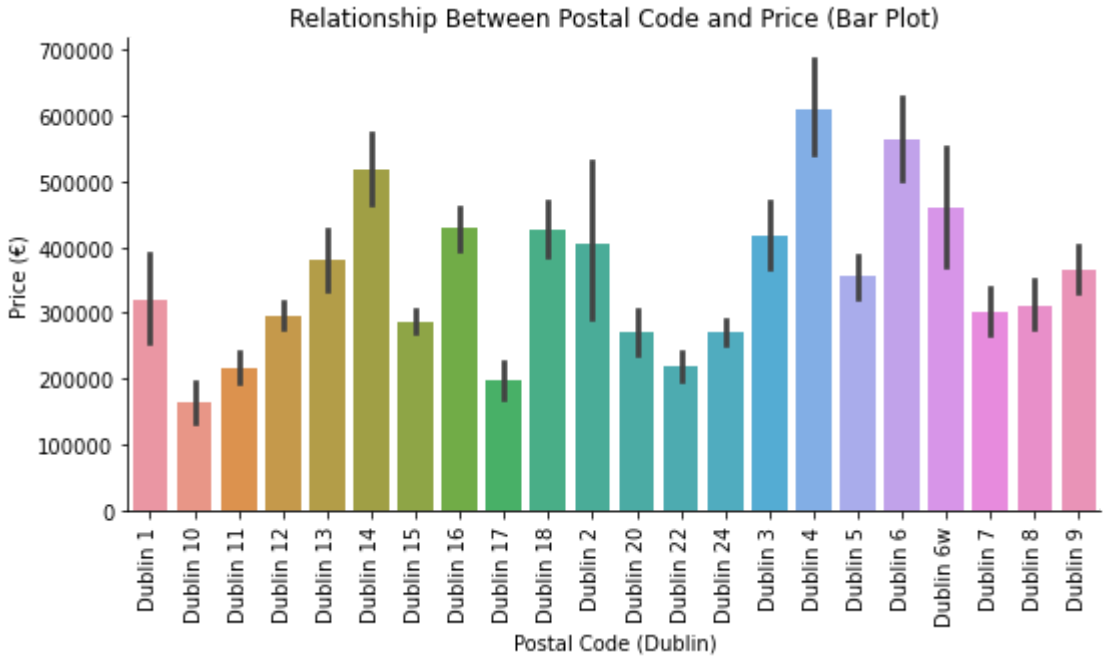
Postal Code and Price

In [80]:

```
sns.catplot(x="Postal Code (Dublin)", y="Price (€)", kind="box", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between Postal Code and Price (Box Plot)")
plt.show()

sns.catplot(x="Postal Code (Dublin)", y="Price (€)", kind="bar", data=df, height=4, aspect=2)
```

```
plt.xticks(rotation=90)
plt.title("Relationship Between Postal Code and Price (Bar Plot)")
plt.show()
```





Discussion of Outcomes

The above graphs suggest a correlation between geographic location and price.

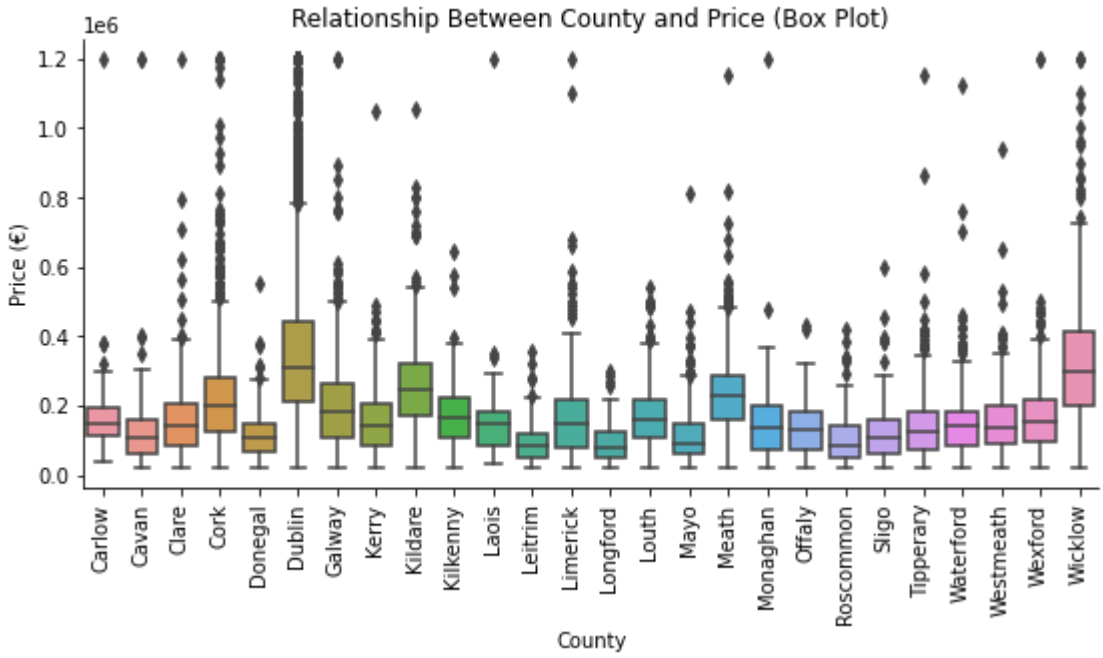
The two graphs show us that overall, the most expensive properties tend to be located in Dublin 4, Dublin 6, Dublin 6W and Dublin 14. The cheapest properties were sold in Dublin 11, Dublin 17, and Dublin 10. This is inline with expectation, as the cheapest properties are located in postal codes that are further out from the city center, and the most expensive properties are ones that are located in postal codes that are closer to city center but not in the inner city.

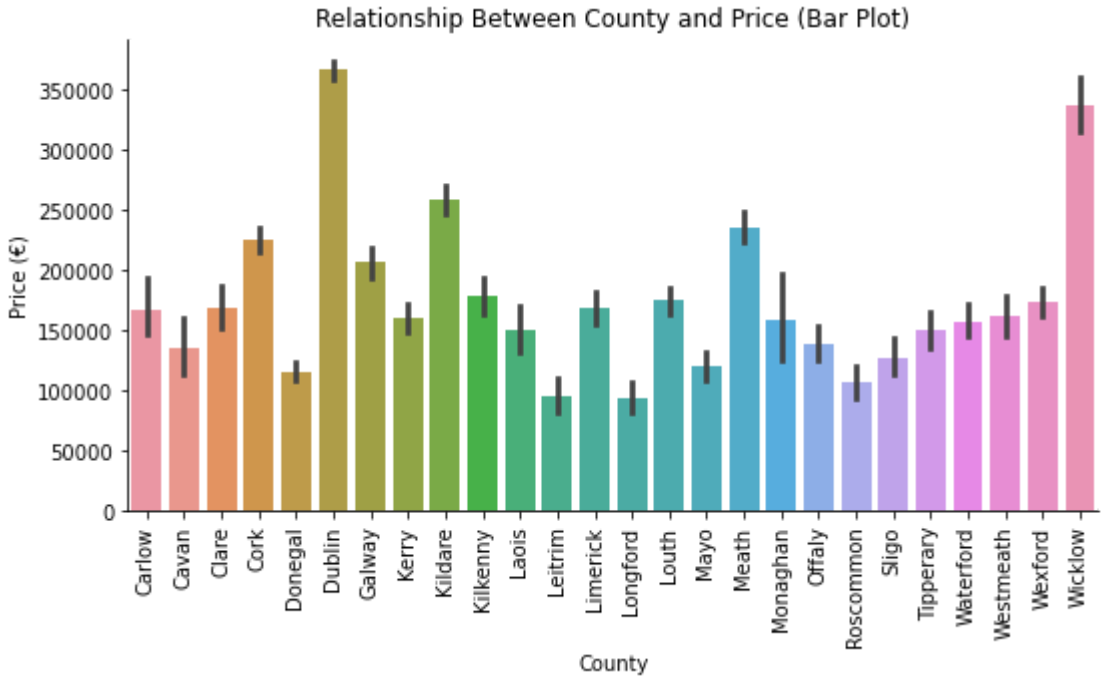
There also seems to be a larger spread of property prices overall in Dublin 6 and Dublin 4, which would seem somewhat unusual given that these locations are, on average, where the most expensive properties are located.

County and Price

In [81]:

```
sns.catplot(x="County", y="Price (€)", kind="box", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between County and Price (Box Plot)")
plt.show()

sns.catplot(x="County", y="Price (€)", kind="bar", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between County and Price (Bar Plot)")
plt.show()
```





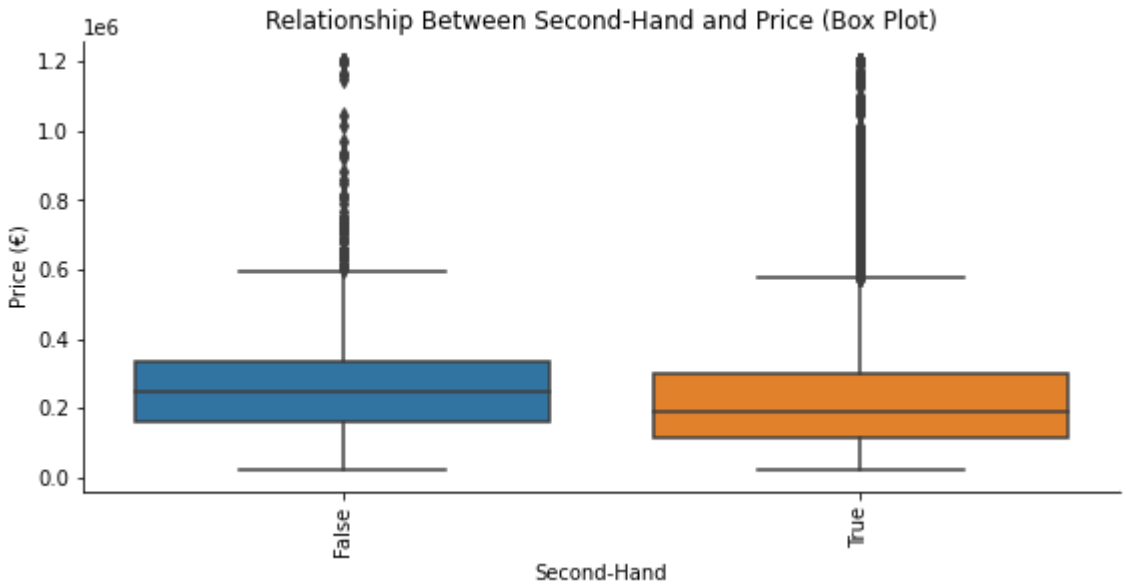
Discussion of Outcomes

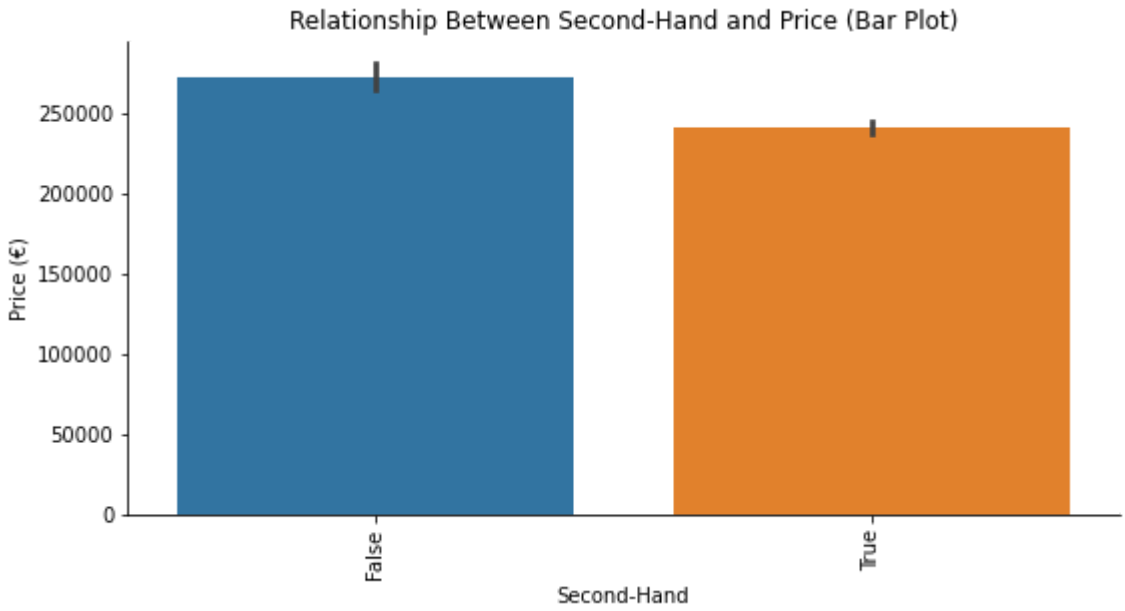
The above graphs suggest a correlation between geographic location and price.

From the highest priced properties, we can see that Dublin and Wicklow have the highest prices, followed by Kildare and Meath, with the lowest prices being located in Leitrim, Longford and Donegal. It is curious that counties containing large cities such as Cork and Galway seem to have lower prices that counties like Meath or Kildare - perhaps this is something that should be discussed with a domain expert to see if there is a reason for this.

Second Hand and Price

In [82]:
```
sns.catplot(x="Second-Hand", y="Price (€)", kind="box", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between Second-Hand and Price (Box Plot)")
plt.show()

sns.catplot(x="Second-Hand", y="Price (€)", kind="bar", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between Second-Hand and Price (Bar Plot)")
plt.show()
```





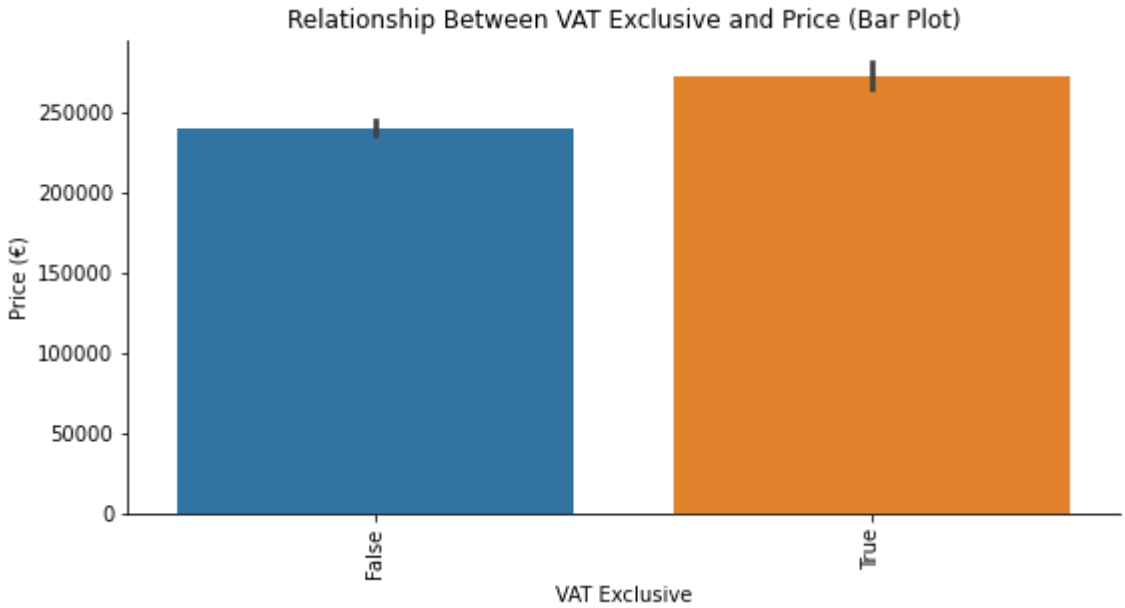
Discussion of Outcomes

From the above, we can see that while the difference between second hand and new homes is not large, on average new homes do seem to be a little more expensive than second-hand ones. This is in line with expectations, as one would expect that a second-hand home would be sold at a slight discount as opposed to a newly-built one. Therefore, there seems to be a weak correlation between second-hand and price. The spread between the two features seems largely the same.

VAT Exclusive and Price

In [83]:
```
sns.catplot(x="VAT Exclusive", y="Price (€)", kind="box", data=df, height=4, aspect=2)
```

```
plt.xticks(rotation=90)
plt.title("Relationship Between VAT Exclusive and Price (Box Plot)")
plt.show()

sns.catplot(x="VAT Exclusive", y="Price (€)", kind="bar", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between VAT Exclusive and Price (Bar Plot)")
plt.show()
```



Discussion of Outcomes

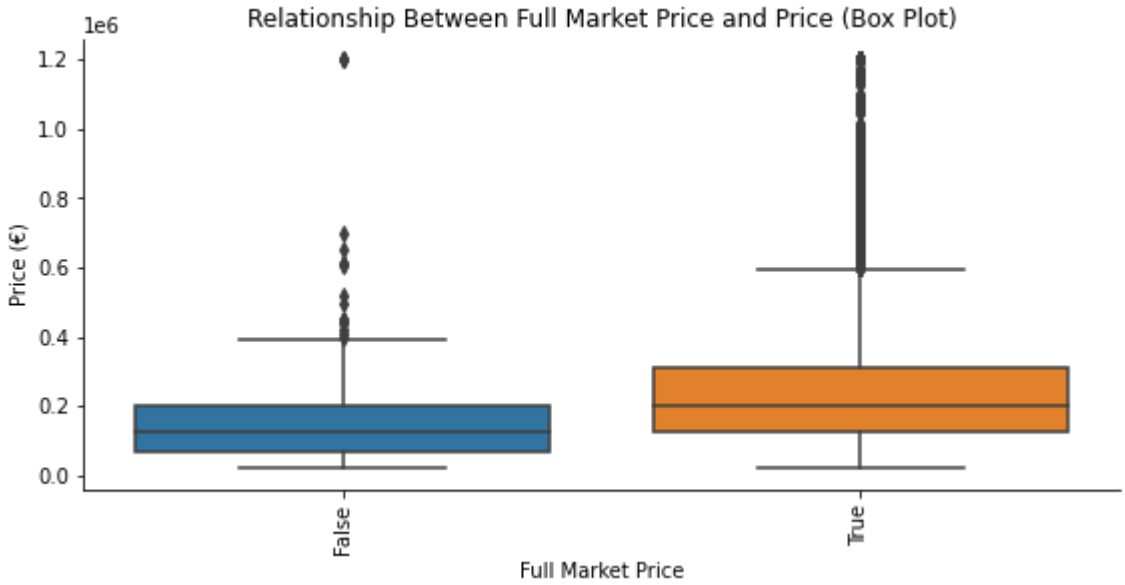
As expected, and it seems that homes were VAT is payable are slightly more expensive than those that are not, suggesting that there is a slight correlation between the two features. This makes sense, as homes where VAT is payable are likely to be new, which would drive the price up. The spread between the two features seems largely the same.

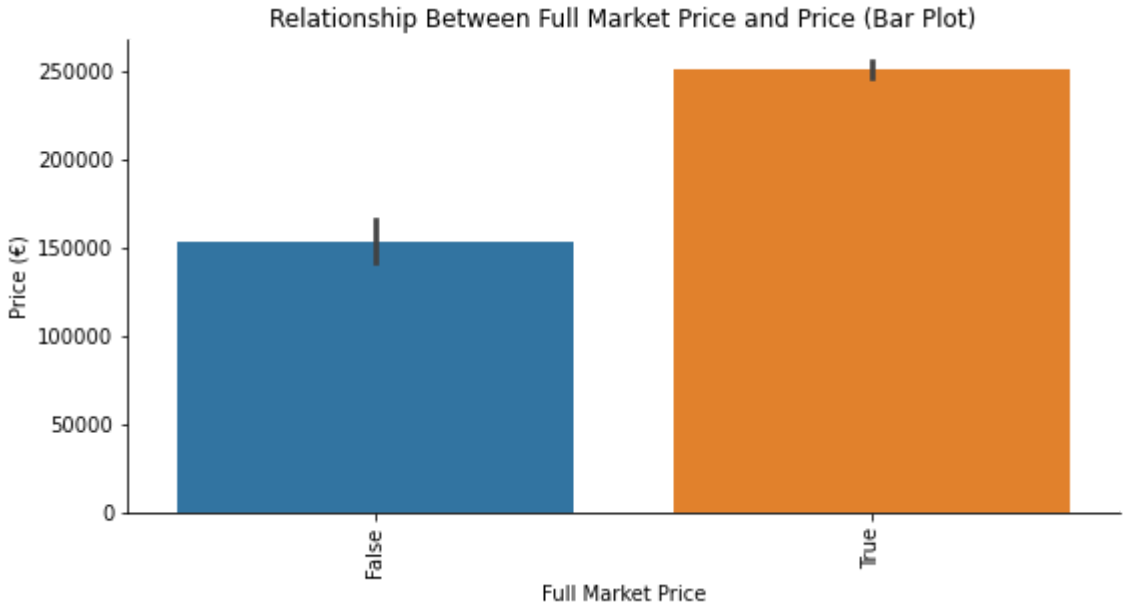
Full Market Price and Price

In [84]:

```
sns.catplot(x="Full Market Price", y="Price (€)", kind="box", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between Full Market Price and Price (Box Plot)")
plt.show()

sns.catplot(x="Full Market Price", y="Price (€)", kind="bar", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
```

```
plt.title("Relationship Between Full Market Price and Price (Bar Plot)")
plt.show()
```





Discussion of Outcomes

The outcome here is very unsurprising - properties marked as being sold at "Full Market Price" were sold for a higher price than those that are not. In fact, we would be surprised if the opposite was true.

However, it was useful to check this point just to make sure that this feature is functioning as expected.

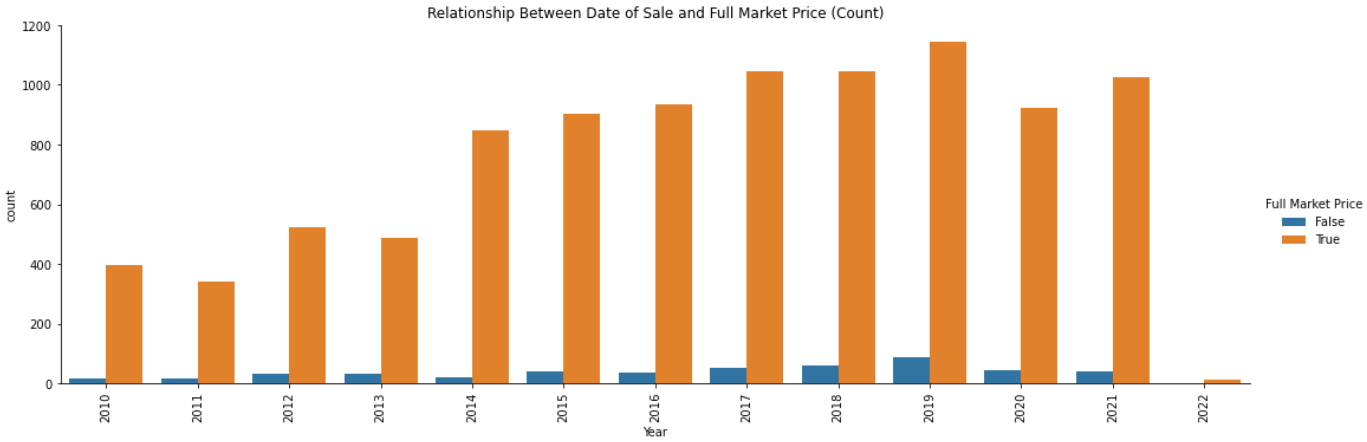
Additional Feature Pairs

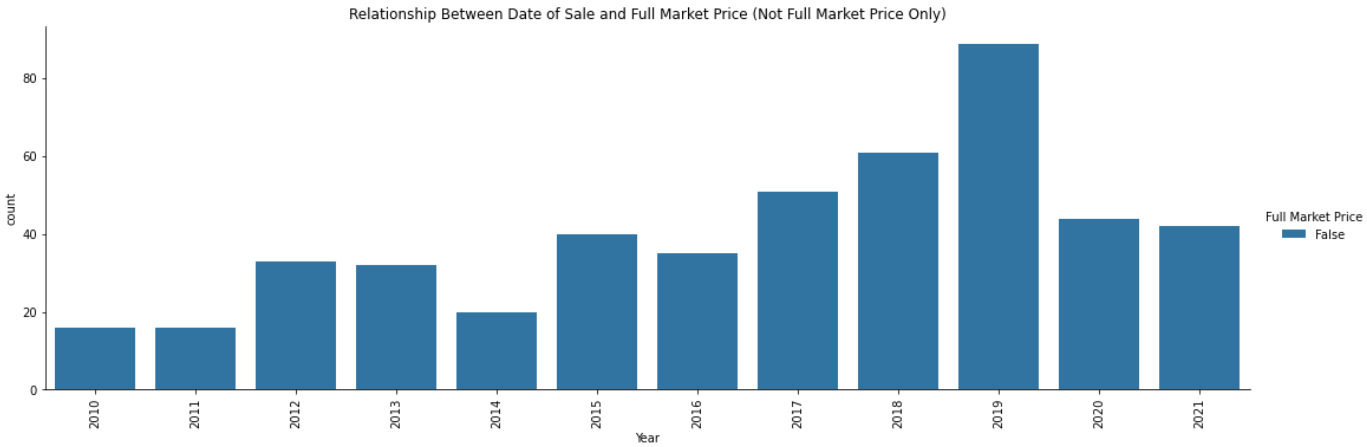
Date of Sale and Full Market Price

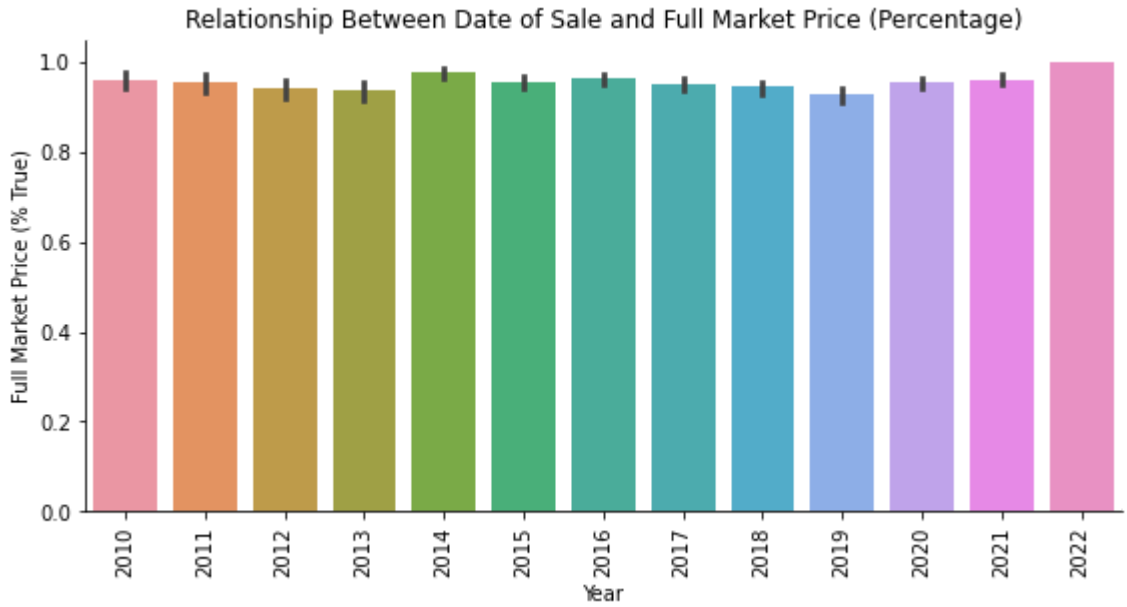
In [85]:

```
df_not_full_market_price = df.loc[df['Full Market Price'] == False]

sns.catplot(x="Year", kind="count", hue="Full Market Price", data=df, height=5, aspect=3)
plt.xticks(rotation = 90)
plt.title("Relationship Between Date of Sale and Full Market Price (Count)")
plt.show()

sns.catplot(x="Year", kind="count", hue="Full Market Price", data=df_not_full_market_price, hei
plt.xticks(rotation = 90)
plt.title("Relationship Between Date of Sale and Full Market Price (Not Full Market Price Only)
```

```
plt.show()

sns.catplot(x="Year", y="Full Market Price", kind="bar", data=df, height=4, aspect=2)
plt.title("Relationship Between Date of Sale and Full Market Price (Percentage)")
plt.xticks(rotation=90)
plt.ylabel("Full Market Price (% True)")
plt.show()
```







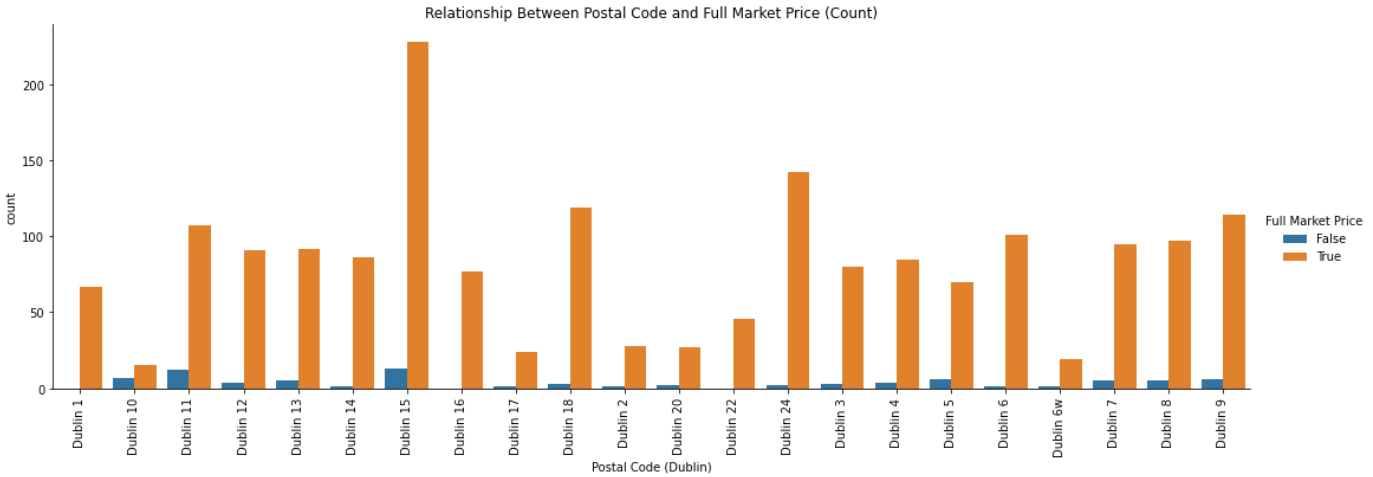
Discussion of Outcomes

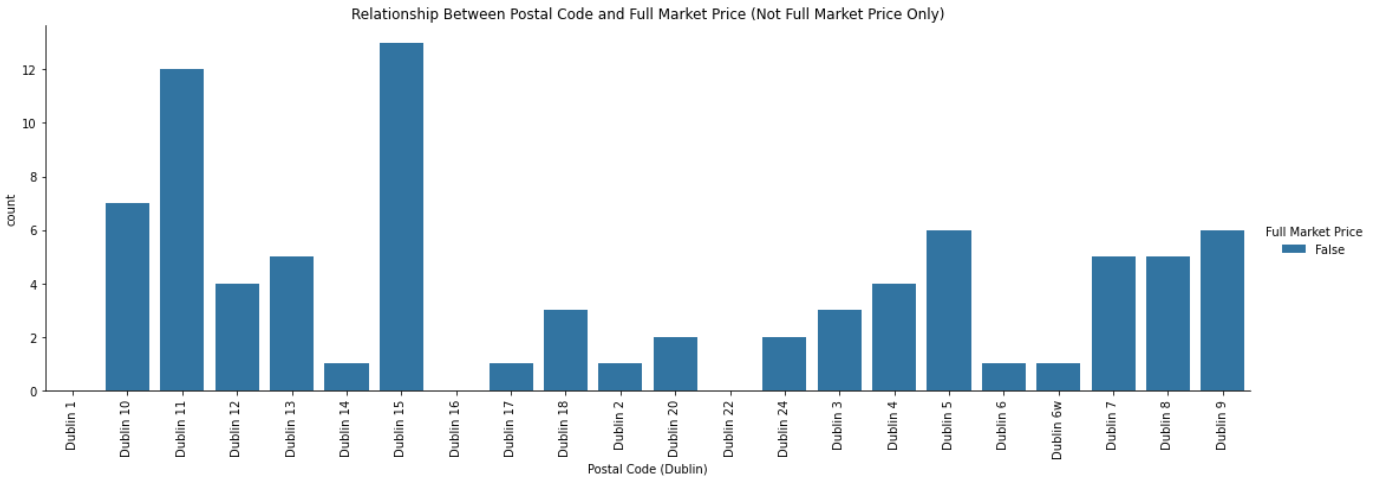
From the above we can see that there is not any particularly strong correlation between Date of Sale and Full Market Price. When we look at the third chart, which shows the percentage of Full Market Price properties being sold, we can see that the actual percentages remain largely the same over the past decade, so this feature is not particularly affected by the date of sale.

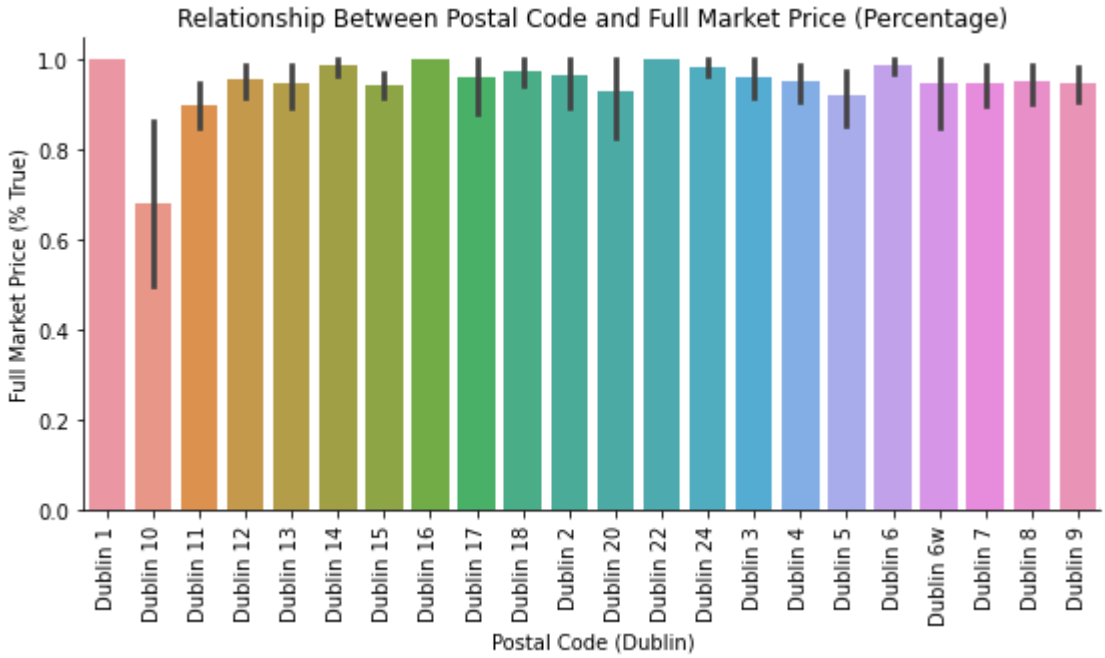
Postal Code and Full Market Price

In [86]:

```
sns.catplot(x="Postal Code (Dublin)", kind="count", hue="Full Market Price", data=df, height=5,
plt.xticks(rotation = 90)
plt.title("Relationship Between Postal Code and Full Market Price (Count)")
plt.show()

sns.catplot(x="Postal Code (Dublin)", kind="count", hue="Full Market Price", data=df_not_full_m
plt.xticks(rotation = 90)
plt.title("Relationship Between Postal Code and Full Market Price (Not Full Market Price Only)"
plt.show()

sns.catplot(x="Postal Code (Dublin)", y="Full Market Price", kind="bar", data=df, height=4, asp
plt.xticks(rotation=90)
plt.title("Relationship Between Postal Code and Full Market Price (Percentage)")
plt.ylabel("Full Market Price (% True)")
plt.show()
```







Discussion of Outcomes

From the above, we can see that there isn't a particularly strong correlation between properties being sold at full market price and a particular postal code. Generally speaking, most postal codes have a similar percentage of full and non-full market price property sales. There are some outliers - for example, it is clear that Dublin 1, 16 and 22 have the most Full Market Price sales, whereas Dublin 10 has by far the least.

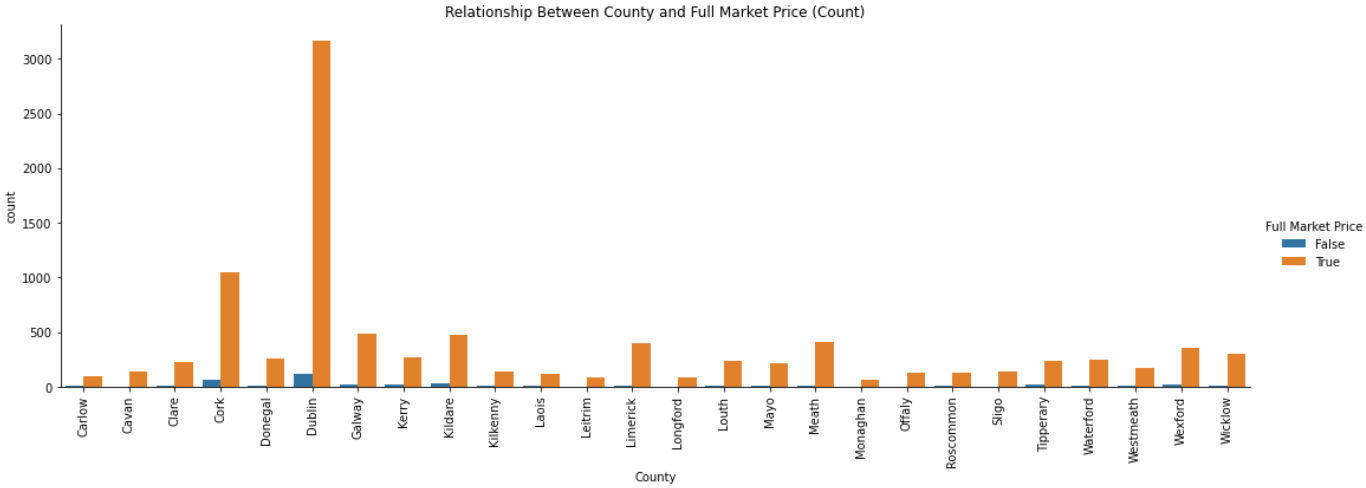
It is not entirely clear why Dublin 16 and 22 have such a high number of Full Market Price sales, but with regard to Dublin 1, this makes sense as it is in the city center, and as such the properties being sold are most likely new developments, and as such will likely be sold for full market price (i.e. unlikely that a newly built property will be sold at a discount).

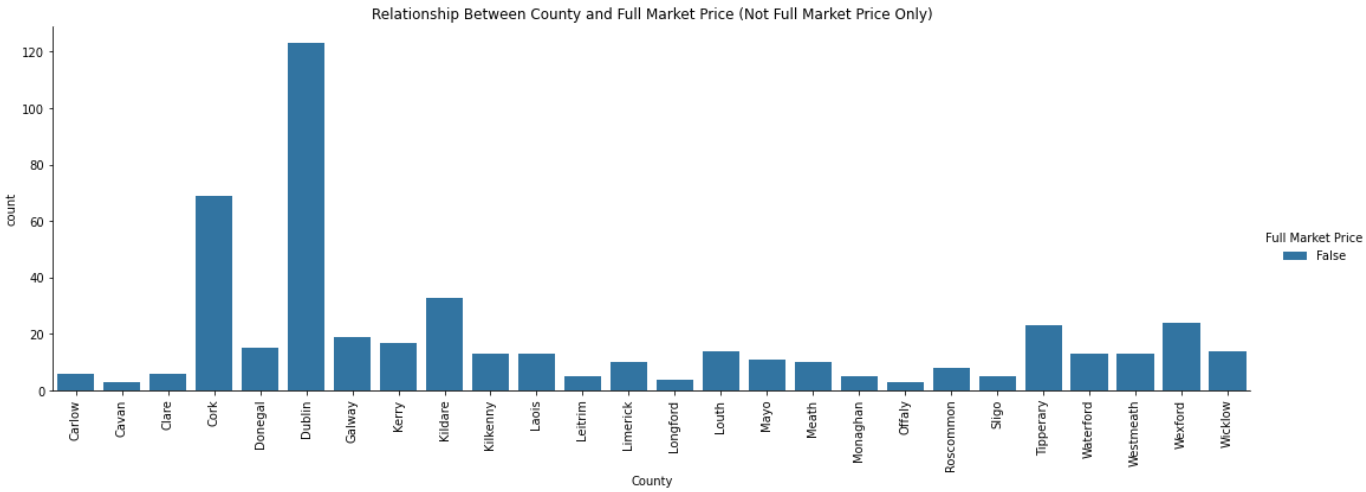
Dublin 10 having the least full market price sales initially seemed a little unusual - however on second glance it was discovered that this area simply just had the least sales out of any other postal code, so it is possible that this is merely an outlier.

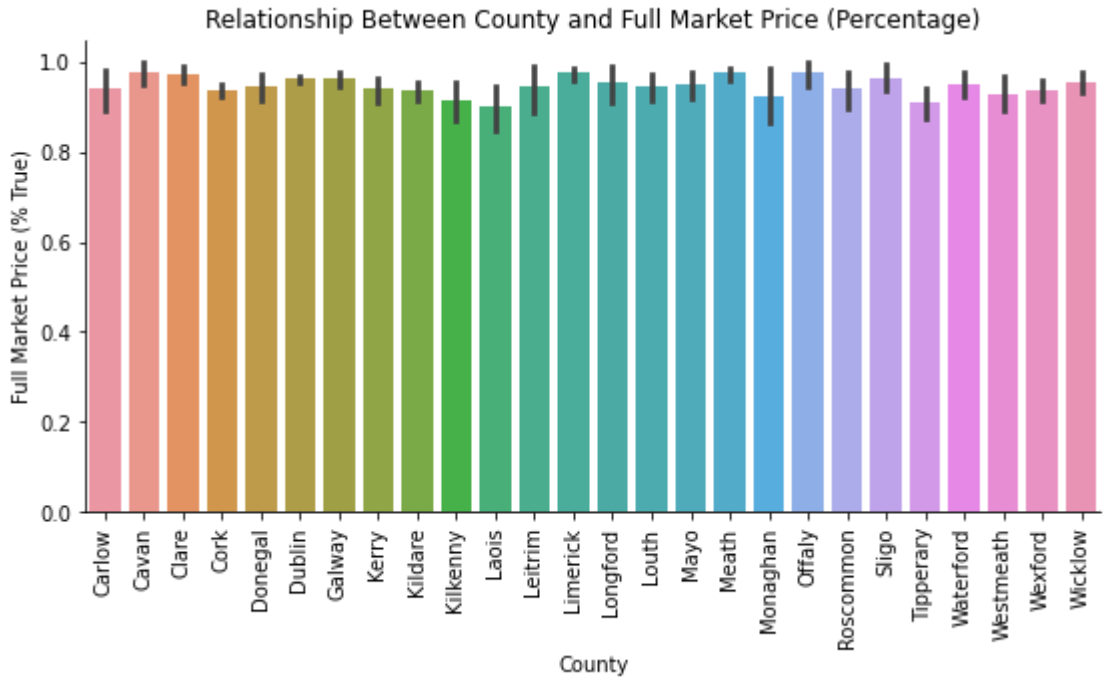
County and Full Market Price

In [87]:

```
sns.catplot(x="County", kind="count", hue="Full Market Price", data=df, height=5, aspect=3)
plt.xticks(rotation = 90)
plt.title("Relationship Between County and Full Market Price (Count)")
plt.show()

sns.catplot(x="County", kind="count", hue="Full Market Price", data=df_not_full_market_price, h
plt.xticks(rotation = 90)
plt.title("Relationship Between County and Full Market Price (Not Full Market Price Only)")
plt.show()

sns.catplot(x="County", y="Full Market Price", kind="bar", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between County and Full Market Price (Percentage)")
plt.ylabel("Full Market Price (% True)")
plt.show()
```







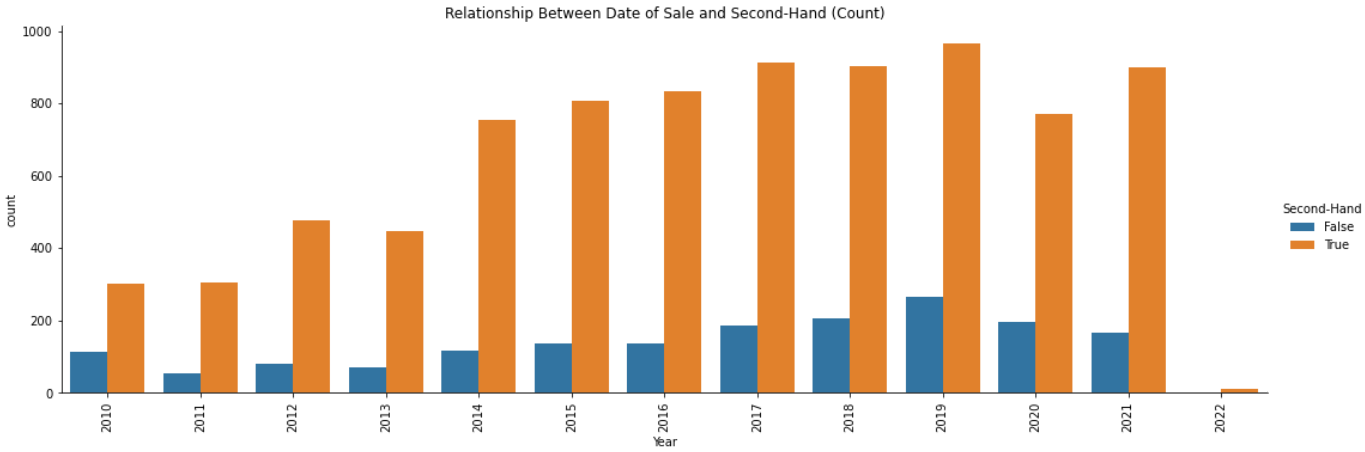
Discussion of Outcomes

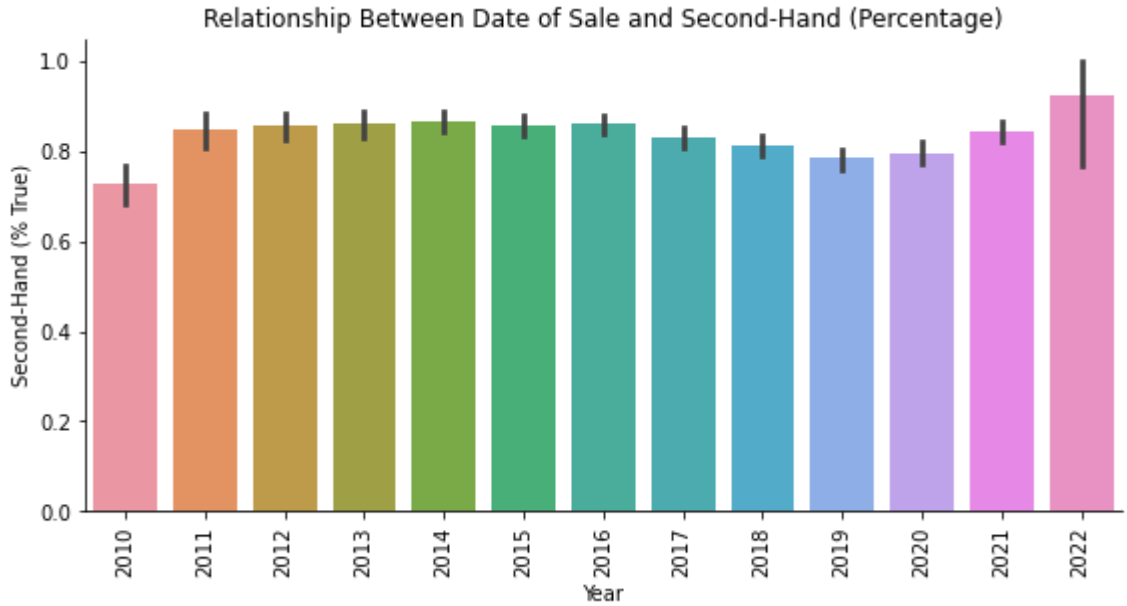
As with the section above, it does not seem that the relationship between County and Full Market Price shows any significant correlation. The third bar chart here is particularly useful - it shows us that the vast majority of properties sold were sold at full market price, and the difference from county to county seems largely negligible. It does seem that Laois has the least number of properties sold at full market price, but again, the difference between these counties is not significant.

Date of Sale and Second Hand

In [88]:

```
sns.catplot(x="Year", kind="count", hue="Second-Hand", data=df, height=5, aspect=3)
plt.xticks(rotation = 90)
plt.title("Relationship Between Date of Sale and Second-Hand (Count)")
plt.show()

sns.catplot(x="Year", y="Second-Hand", kind="bar", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between Date of Sale and Second-Hand (Percentage)")
plt.ylabel("Second-Hand (% True)")
plt.show()
```





Discussion of Outcomes

From the above, it looks like there is a minor correlation between the two features. We can see that the proportion of second hand properties remained largely the same throughout the first half of the decade, following which there was a small dip between 2018 and 2021. It looks like sales of second hand properties are on the rise again in 2022, although it is difficult to say at the start of the year.

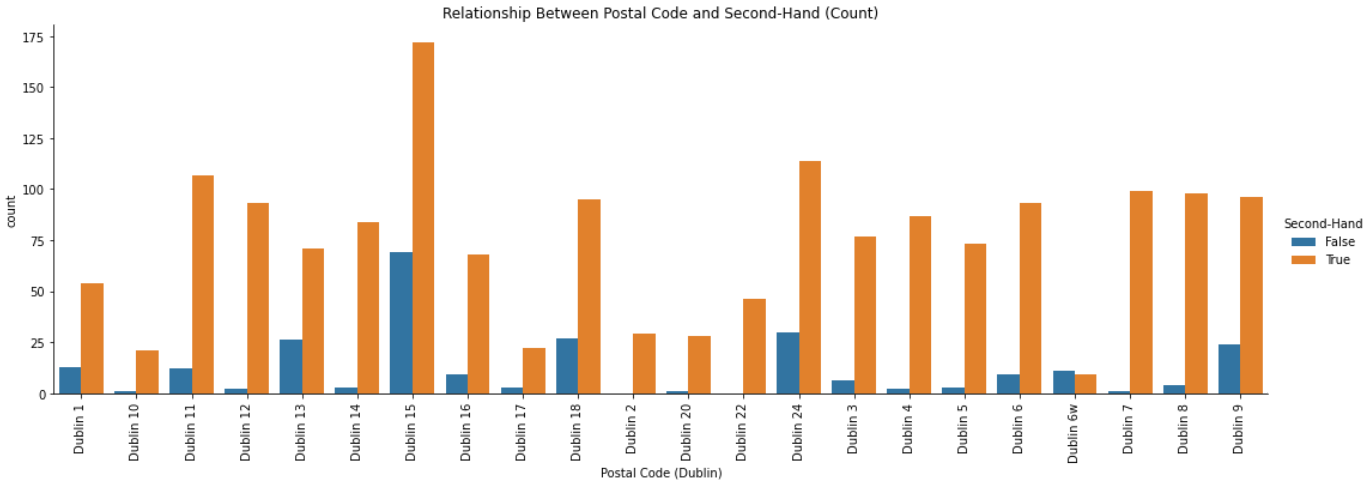
It is possible that this dip can be seen due to an increase in economic recovery from the financial crisis and an increase in development of properties, leading to more sales of new properties. It is entirely possible that such development started taking place a few years before this dip, as it of course takes time for properties to be built.

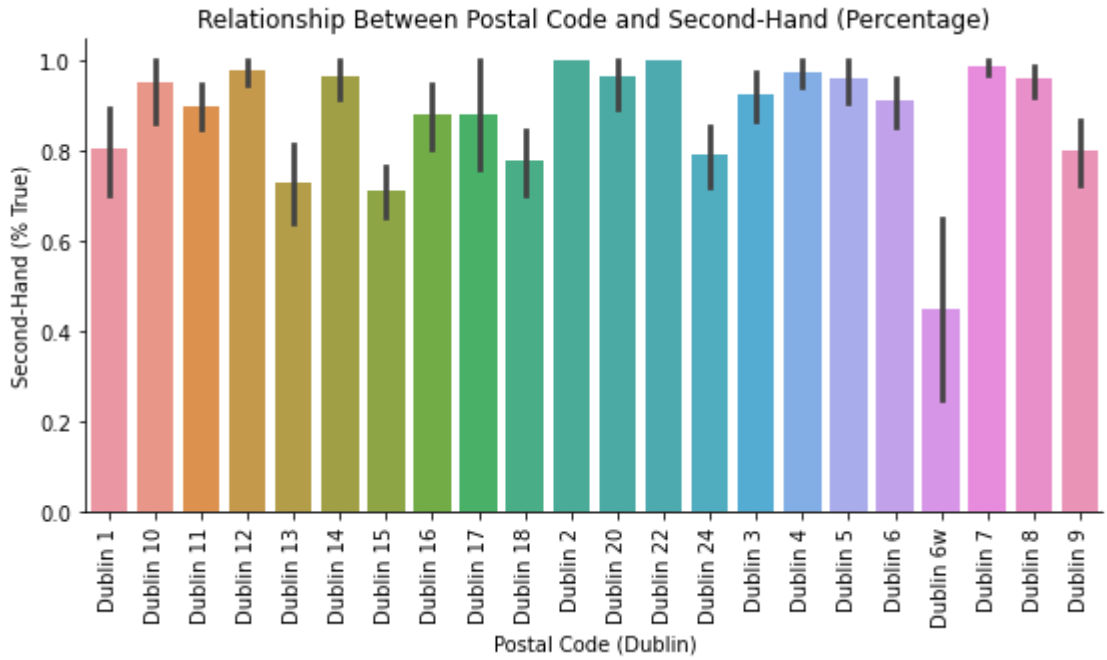
Postal Code and Second Hand

In [89]:

```
sns.catplot(x="Postal Code (Dublin)", kind="count", hue="Second-Hand", data=df, height=5, aspec
plt.xticks(rotation = 90)
plt.title("Relationship Between Postal Code and Second-Hand (Count)")
```

```
plt.show()

sns.catplot(x="Postal Code (Dublin)", y="Second-Hand", kind="bar", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between Postal Code and Second-Hand (Percentage)")
plt.ylabel("Second-Hand (% True)")
plt.show()
```





Discussion of Outcomes

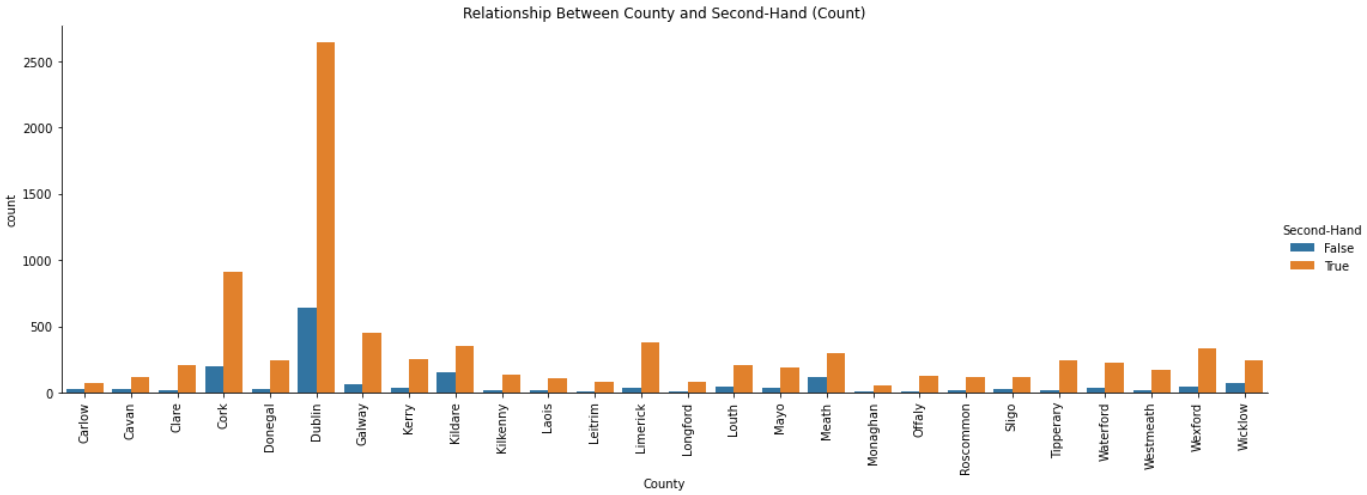
We can see a minor correlation between these two features. Specifically, there are certain postal codes that have noticeably less second hand property sales than others, specifically Dublin 6W, 13 and 15. Of particular interest is the fact that there seem to have been more second hand property sales in Dublin 6W than new ones. However, as discussed above, this postal code has had the least property sales overall, and as such this may just be an outlier.

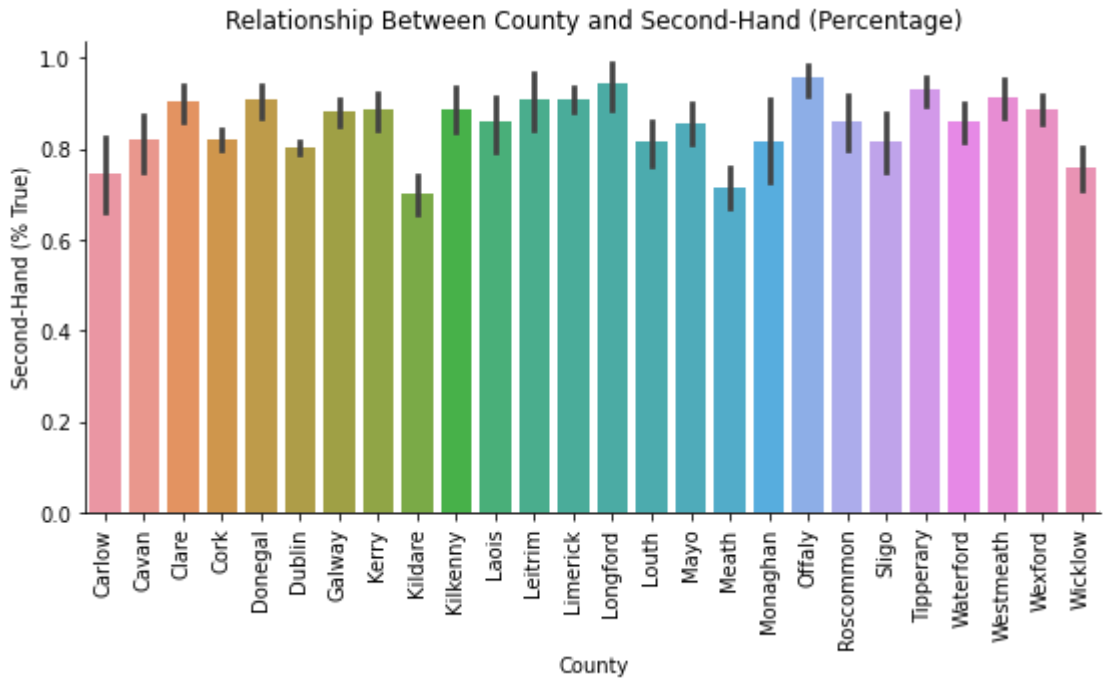
County and Second Hand

In [90]:

```
sns.catplot(x="County", kind="count", hue="Second-Hand", data=df, height=5, aspect=3)
plt.xticks(rotation = 90)
plt.title("Relationship Between County and Second-Hand (Count)")
plt.show()

sns.catplot(x="County", y="Second-Hand", kind="bar", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between County and Second-Hand (Percentage)")
```

```
plt.ylabel("Second-Hand (% True)")
plt.show()
```





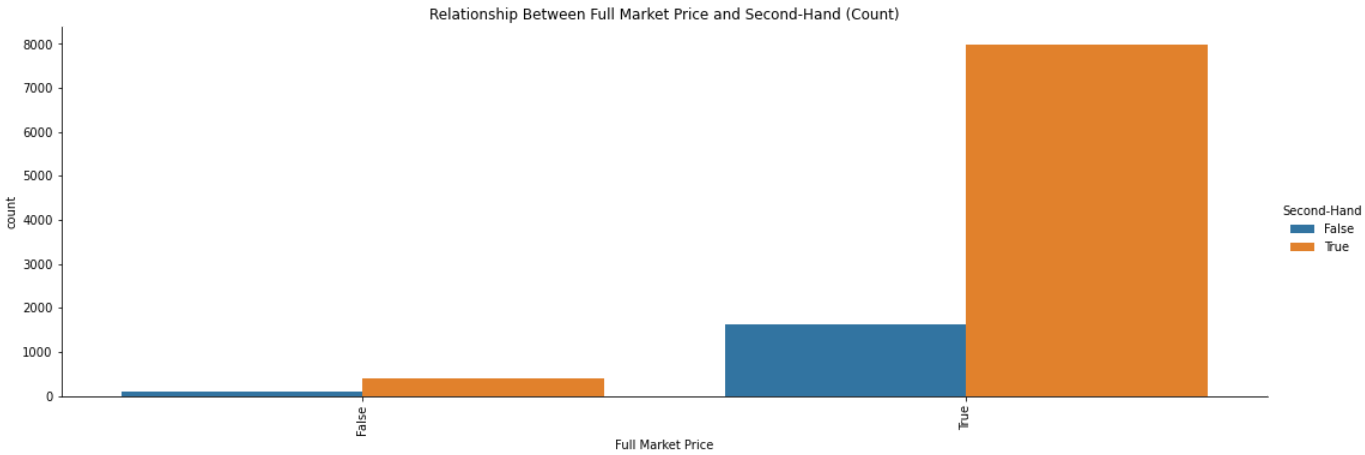
Discussion of Outcomes

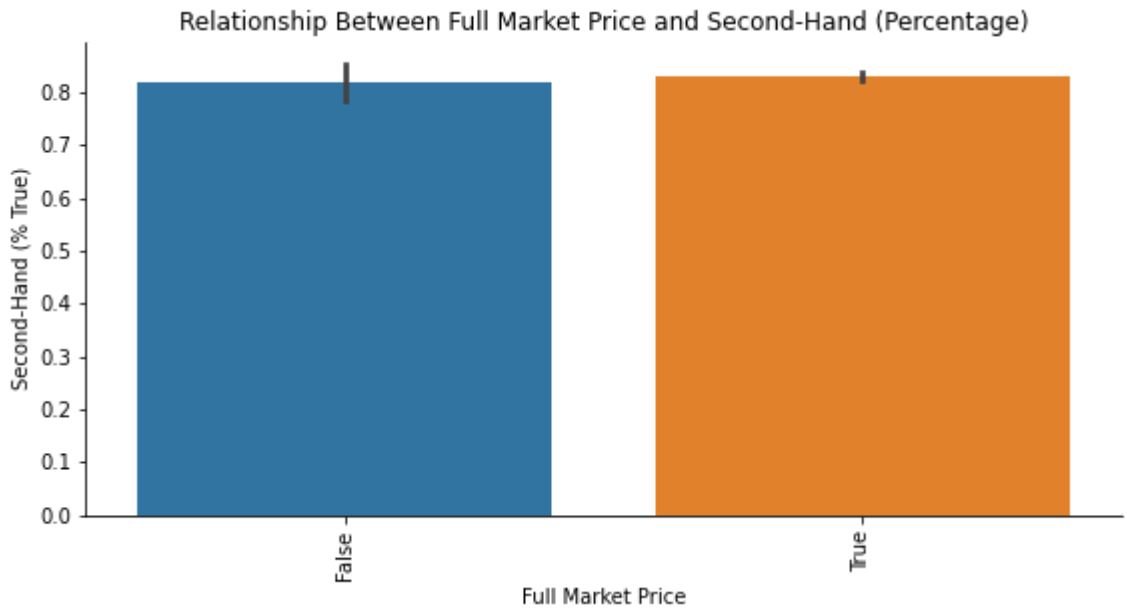
There seems to be a weak correlation between these two features - mostly they seem to be around the same level, with the exception of certain dips in Meath, Carlow and Kildare. It is hard to say why this is the case without consulting with a domain expert.

Full Market Price and Second Hand

In [91]:

```
sns.catplot(x="Full Market Price", kind="count", hue="Second-Hand", data=df, height=5, aspect=3
plt.xticks(rotation = 90)
plt.title("Relationship Between Full Market Price and Second-Hand (Count)")
plt.show()

sns.catplot(x="Full Market Price", y="Second-Hand", kind="bar", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between Full Market Price and Second-Hand (Percentage)")
plt.ylabel("Second-Hand (% True)")
plt.show()
```





Discussion of Outcomes

This graph suggests that there is almost no correlation between these two features, suggesting that whether or not the property was sold for full market price has no bearing on whether or not the property was second-hand. This is quite an interesting discovery, as it would have been fair to make an assumption that a second-hand home would likely have been sold at an undervalued price - however, this does not seem to be the case.

The fact that property prices are not affected by whether or not said property is new is quite interesting, as most other types of property (such as cars, for example) tend to lose value if they are second-hand. This is likely one of the factors that leads to real property being a solid investment.

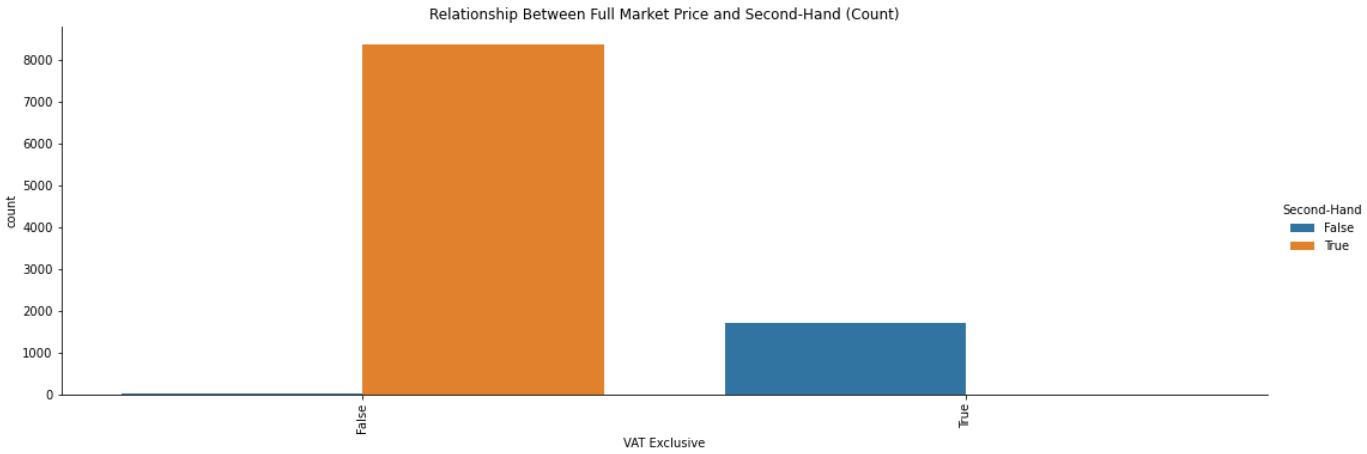
It should be noted here again, however, that the RPPR note on this feature was not clear, and as such there are many different reasons as to why a particular property may be marked as not being sold at full market price - it is possible that there are a variety of factors at play here, and for these reasons we should be careful when making assumptions relating to this feature.

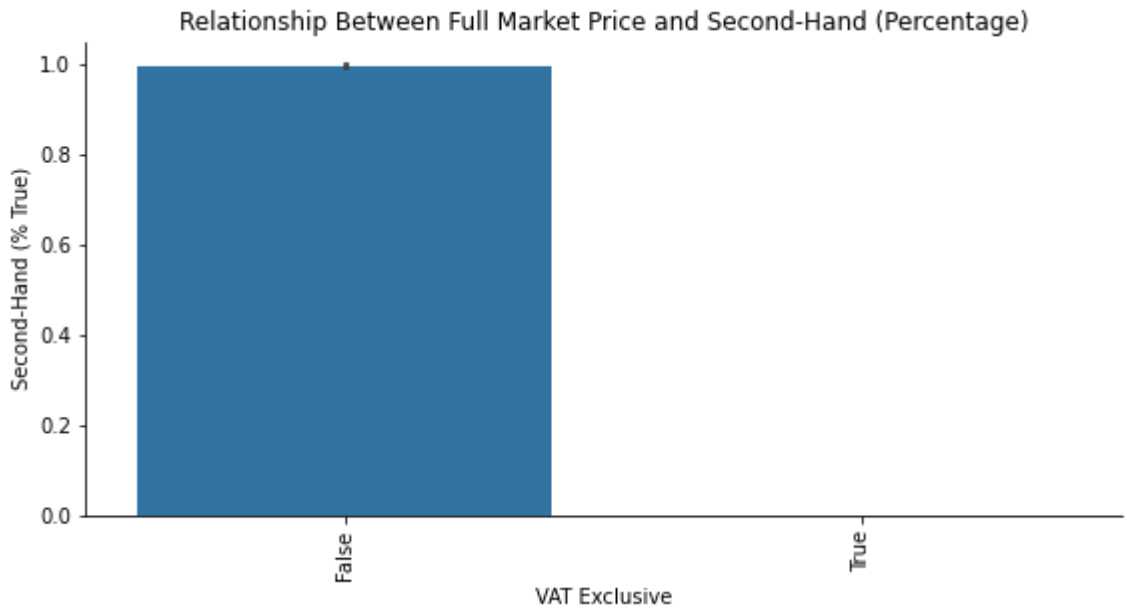
VAT Exclusive and Second Hand

In [92]:

```
sns.catplot(x="VAT Exclusive", kind="count", hue="Second-Hand", data=df, height=5, aspect=3)
plt.xticks(rotation = 90)
plt.title("Relationship Between Full Market Price and Second-Hand (Count)")
plt.show()

sns.catplot(x="VAT Exclusive", y="Second-Hand", kind="bar", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
```

```
plt.title("Relationship Between Full Market Price and Second-Hand (Percentage)")
plt.ylabel("Second-Hand (% True)")
plt.show()
```





Discussion of Outcomes

As discussed above, there was initially a concern as to whether or not Full Market Price and Second-Hand features were, in fact, representing the same facet of data. As we can see from these graphs, the values here match incredibly closely, suggesting that that out assumption was largely true and that there is a very tight correlation between the two features. As expected, there is only a small proportion of values that are different between the two features. However, as we discussed previously, it makes sense to keep both of these features in our dataset as there are some potential differences in exceptions where VAT is being paid on second-hand properties. We do not wish to lose any data where this occurs.

Summary

The above graphs describe in detail our findings and whether or not there are correlations between the various features. In order to summarise our position, it is useful to display a short table explaining whether or not we discovered a correlation between the various features.

Feature Pair	Correlation
Date of Sale and Price	Yes, correlation exists
Postal Code and Price	Yes, correlation exists
County and Price	Yes, correlation exists
Second Hand and Price	Yes, but correlation is weak

Feature Pair	Correlation
VAT Exclusive and Price	Yes, but correlation is weak
Full Market Price and Price	Yes, correlation exists
Date of Sale and Full Market Price	Correlation doesn't exist / is very weak
Postal Code and Full Market Price	Correlation doesn't exist / is very weak
County and Full Market Price	Correlation doesn't exist / is very weak
Date of Sale and Second Hand	Yes, but correlation is weak
Postal Code and Second Hand	Yes, correlation exists
County and Second Hand	Yes, correlation exists
Full Market Price and Second Hand	Correlation doesn't exist / is very weak
VAT Exclusive and Second Hand	Yes, correlation exists

The main takeaway from this exercise is that there seems to be a correlation, whether it be weak or strong, between Price, our target feature, and most of the other features in our dataset. This is reassuring, as it means that all the features that we have in our dataset at the moment will assist us with predicting the target feature. Our additional feature pairs also led to some interesting discoveries that helped us with understanding the Irish residential property market in more detail.

Question 4: Create a few new features (at least 3) with the aim to better capture the problem domain and the target outcome.

In order to create a new feature, we must first:

- Consider how such potential feature will help us better understand the dataset;
- Consider which existing feature(s) will allow us to create a new feature; and
- Consider whether or not we should add the feature as a new feature or to transform an existing one.

Below are set out the proposed new features with the corresponding analysis referred to above. Also provided is an "Initial Thoughts" section which describes features that were considered but ultimately discarded.

Please note that the graphs provided for the new features are (a) bar plots and (b) graphs showing the feature's correlation with Price. The reason why further graphs showing correlations with other features were not shown was because the questions states that the features need ot capture "the target outcome" - as such, while of interest, for the purposes of this question the feature's correlation with other features within the dataset is not particularly relevant.

It should be noted that one feature has already been transformed as part of the data cleaning - from Description of Property to Second-Hand. However, it was ultimately decided that this transformation was more of a data cleaning exercise as opposed to a fully new feature creation, and as such, it was decided to not to included it in this section of the assignment.

Considered Features

Initially, a feature that would feature the price including the VAT was considered. Based on the VAT Exclusive and the Description of Property features, it is possible to determine what the appropraite level of VAT is for a property depending on whether it is new or second-hand. Then, if VAT has not been included in a particular entry, we would include this VAT. This would have given a more complete picture of the amount that was

actually paid for a property. However, because this transformation constitutes an edit to the target feature, being "Price", this was deemed an inappropriate feature for inclusion, and was not included in the edited dataset.

In any event, it would have been difficult to tell how much VAT precisely would have been payable on each property sale, for reasons already discussed in previous sections.

New Feature 1: Quarter

This feature is a conversion of the date of sale into financial quarters. This will be useful as we can check price fluctuations based on a particular time period within the year easily - for example, if there is a tendency for properties to be purchased in Summer as opposed to Winter.

We would still like to keep the "Date of Sale" continuous feature in our dataframe in order to ensure that we have granular detail if we need it, so we will create this as a new feature using the built-in pandas "PeriodIndex" converter. We will also ensure that this feature is set to "category", as this feature is categorical.

In [93]:
```
df["Quarter"] = df["Date of Sale"].dt.quarter
df["Quarter"] = df["Quarter"].astype('category')
```

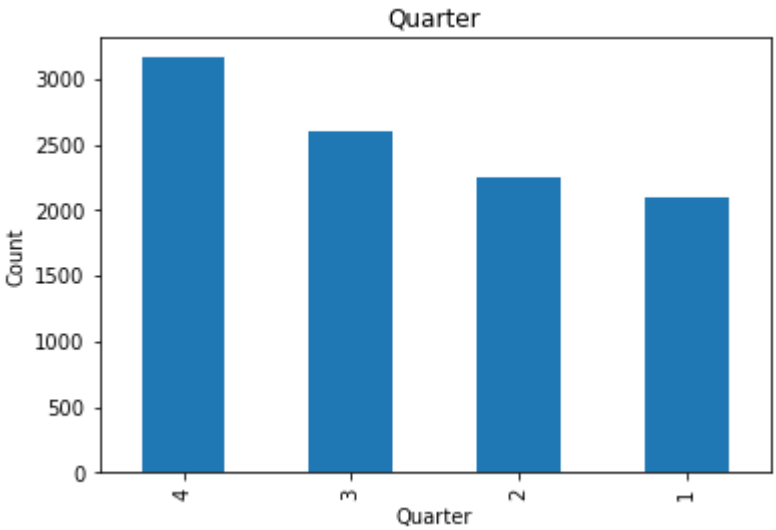
In [94]:
```
df.head()
```

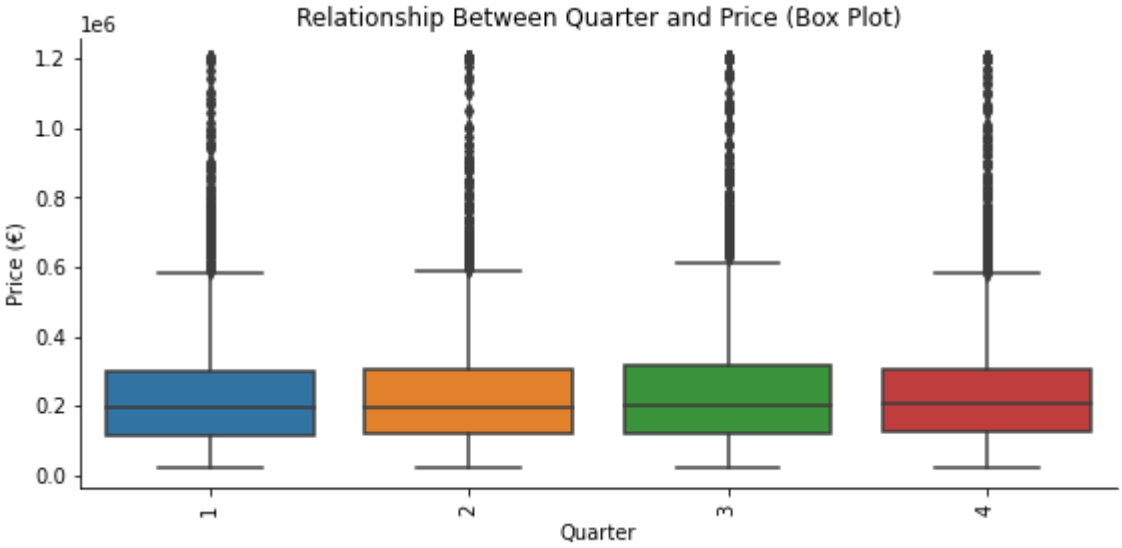
Out[94]:

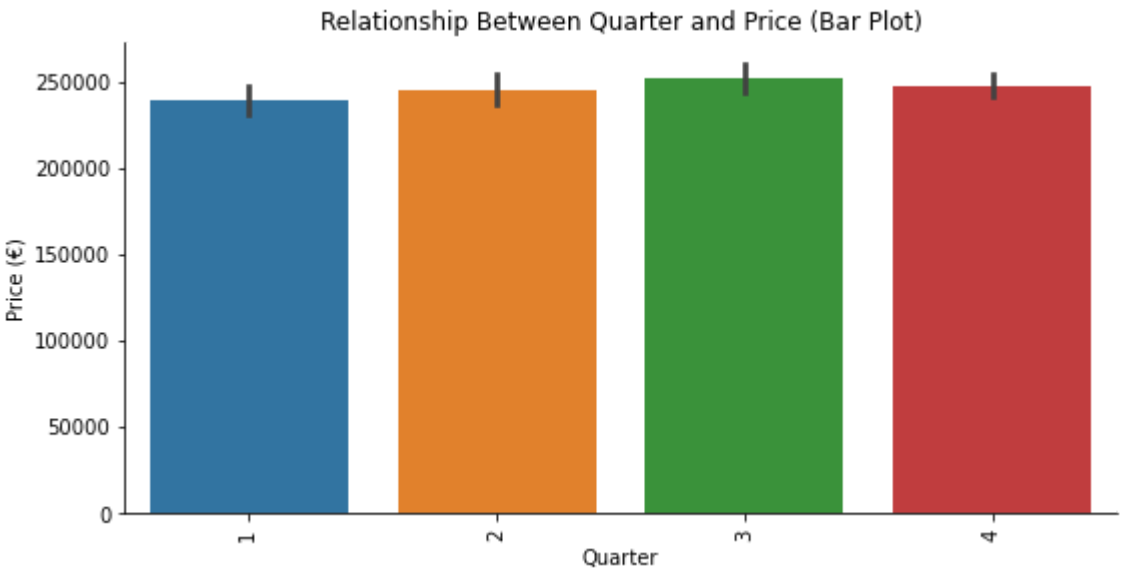
	Date of Sale	Address	Postal Code (Dublin)	County	Price (€)	Full Market Price	VAT Exclusive	Second-Hand	Year	Quarter
0	2020-06-26	1 GANDON PLACE, GANDON PARK, Lucan	NaN	Dublin	370044.05	True	True	False	2020	2
1	2014-12-19	72 ST ASSAMS PARK, RAHENY, DUBLIN 5	Dublin 5	Dublin	480000.00	True	False	True	2014	4
2	2010-02-11	37 Ivy Hill, Gort Road, Ennis	NaN	Clare	194000.00	True	False	True	2010	1
3	2018-08-16	14 TYRCONNELL PLACE, TYRCONNELL RD, INCHICORE ...	Dublin 8	Dublin	275000.00	True	False	True	2018	3
4	2019-02-27	7 THE NEST, TUBBERCURRY, SLIGO	NaN	Sligo	75000.00	True	False	True	2019	1

Plotting New Graphs With Feature

In [95]:
```
df["Quarter"].value_counts().plot(kind = "bar")
plt.xlabel("Quarter")
plt.ylabel("Count")
plt.title("Quarter")
plt.show()

sns.catplot(x="Quarter", y="Price (€)", kind="box", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between Quarter and Price (Box Plot)")
plt.show()

sns.catplot(x="Quarter", y="Price (€)", kind="bar", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between Quarter and Price (Bar Plot)")
plt.show()
```







Discussion of Outcomes

The above suggests that most properties tend to be bought in Q4, and the least number of properties tend to be bought in Q1. The spread for each quarter seems to be roughly the same.

The above also suggests that there is a very weak correlation between price and quarter. We do, however, see that the most expensive properties tend to be bought in Q3, whereas the least expensive properties tend to be bought in Q1.

This is a useful piece of information, so we will keep this new feature in a dataset.

New Feature 2: Multiple Property Purchase

Another interesting feature to consider is whether or not a property was part of a group sale. It is useful to know the purchasing dynamics in our domain, and particularly if single properties are preferred over bulk purchases, and how these interact with price.

As we can see from our answers above, the total number of split out rows - which represent the sales that contain multiple properties - is 129. These were all appended the end of the dataframe. First, let us just double check the tail 129 rows in our dataframe just to be sure that these are, in fact, the entries that represent multiple properties.

In [96]:
```
df.tail(129)
```

Out[96]:

	Date of Sale	Address	Postal Code (Dublin)	County	Price (€)	Full Market Price	VAT Exclusive	Second-Hand	Year	Quarter
9980	2020-03-06	6 7 8 9 ST ANNES TERRACE, SHORTCOURSE, WATERFORD	NaN	Waterford	21513.280000	True	False	True	2020	1
9981	2020-03-06	6 7 8 9 ST ANNES TERRACE, SHORTCOURSE, WATERFORD	NaN	Waterford	21513.280000	True	False	True	2020	1
9982	2020-03-06	6 7 8 9 ST ANNES TERRACE, SHORTCOURSE, WATERFORD	NaN	Waterford	21513.280000	True	False	True	2020	1
9983	2020-03-06	6 7 8 9 ST ANNES TERRACE, SHORTCOURSE, WATERFORD	NaN	Waterford	21513.280000	True	False	True	2020	1
9984	2019-12-18	1 5 6 & 22 THE HAMPTON, APT 1 2 7 19 20 26 2...	Dublin 11	Dublin	170391.304348	True	False	True	2019	4
...	...	...	...	...	...	...	...	...	...	...
10104	2017-05-17	35 2 4 GEORGES QUAY, DUBLIN 2, DUBLIN	Dublin 2	Dublin	77666.666667	True	False	True	2017	2
10105	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	True	True	False	2015	2
10106	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	True	True	False	2015	2

	Date of Sale	Address	Postal Code (Dublin)	County	Price (€)	Full Market Price	VAT Exclusive	Second-Hand	Year	Quarter
10107	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	True	True	False	2015	2
10108	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	True	True	False	2015	2

129 rows × 10 columns

The above looks to be in order. As such, we will now need to create the new column. We will initialise it to "False" by default, and then set the values of the bottom rows to "True".

In [97]:
```
# Create column and set to false
df["Multiple Property Purchase"] = False

# Loop through last 129 rows and set values to true
for index in range(df.shape[0]-129, df.shape[0], 1):
    df.at[index, "Multiple Property Purchase"] = True
```

In [98]:
```
df.head()
```

Out[98]:

	Date of Sale	Address	Postal Code (Dublin)	County	Price (€)	Full Market Price	VAT Exclusive	Second-Hand	Year	Quarter	Multiple Property Purchase
0	2020-06-26	1 GANDON PLACE, GANDON PARK, Lucan	NaN	Dublin	370044.05	True	True	False	2020	2	False
1	2014-12-19	72 ST ASSAMS PARK, RAHENY, DUBLIN 5	Dublin 5	Dublin	480000.00	True	False	True	2014	4	False
2	2010-02-11	37 Ivy Hill, Gort Road, Ennis	NaN	Clare	194000.00	True	False	True	2010	1	False
3	2018-08-16	14 TYRCONNELL PLACE, TYRCONNELL RD, INCHICORE ...	Dublin 8	Dublin	275000.00	True	False	True	2018	3	False
4	2019-02-27	7 THE NEST, TUBBERCURRY, SLIGO	NaN	Sligo	75000.00	True	False	True	2019	1	False

In [99]:
```
df.tail()
```

Out[99]:

	Date of Sale	Address	Postal Code (Dublin)	County	Price (€)	Full Market Price	VAT Exclusive	Second-Hand	Year	Quarter	Multiple Property Purchase

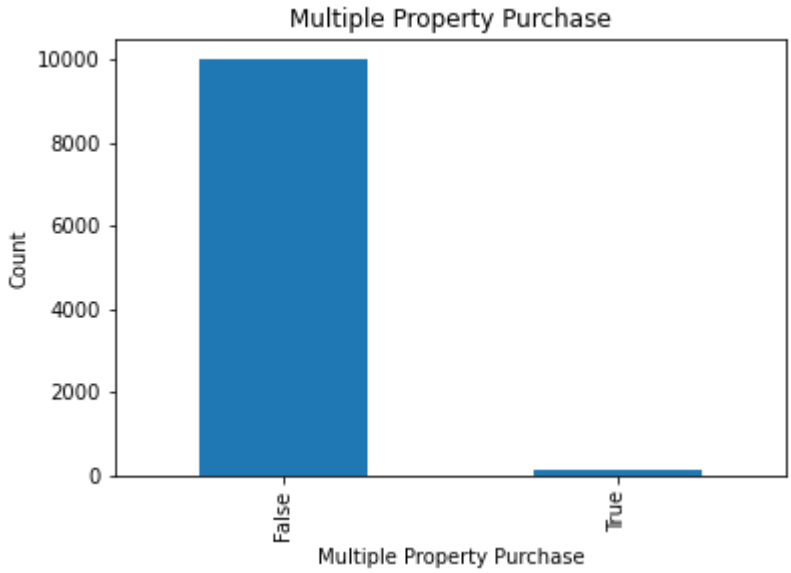
	Date of Sale	Address	Postal Code (Dublin)	County	Price (€)	Full Market Price	VAT Exclusive	Second-Hand	Year	Quarter	Multiple Property Purchase
10104	2017-05-17	35 2 4 GEORGES QUAY, DUBLIN 2, DUBLIN	Dublin 2	Dublin	77666.666667	True	False	True	2017	2	True
10105	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	True	True	False	2015	2	True
10106	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	True	True	False	2015	2	True
10107	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	True	True	False	2015	2	True
10108	2015-06-23	40 40A 40B & 40C Carrowmore Drive, Knock	NaN	Mayo	31750.000000	True	True	False	2015	2	True

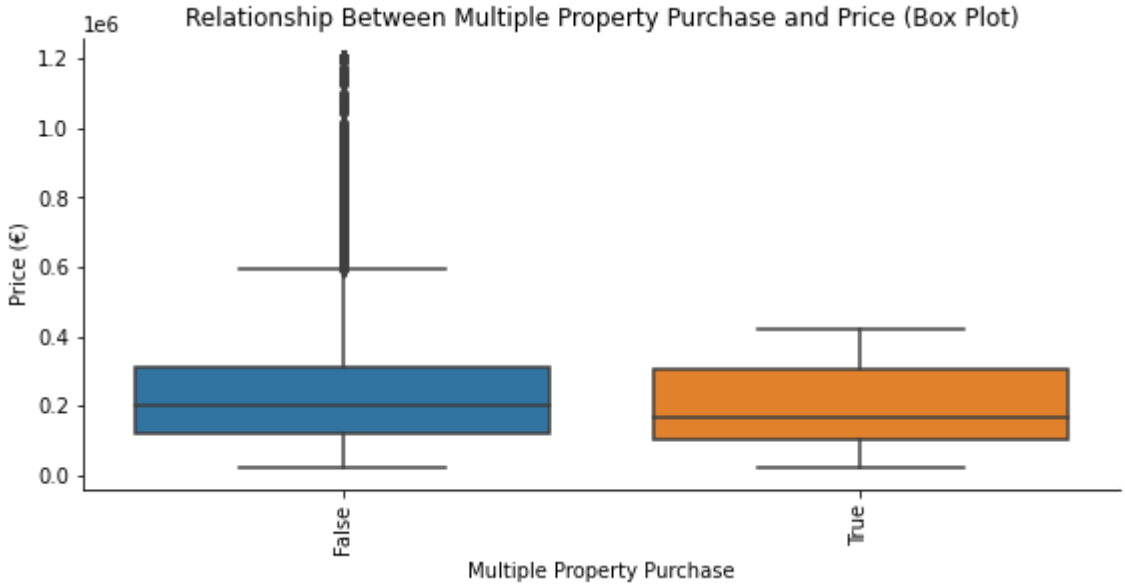
As we can see from the above, the new feature has been successfully implemented.

Plotting New Graphs With Feature

In [100…

```
df["Multiple Property Purchase"].value_counts().plot(kind = "bar")
plt.xlabel("Multiple Property Purchase")
plt.ylabel("Count")
plt.title("Multiple Property Purchase")
plt.show()

sns.catplot(x="Multiple Property Purchase", y="Price (€)", kind="box", data=df, height=4, aspec
plt.xticks(rotation=90)
plt.title("Relationship Between Multiple Property Purchase and Price (Box Plot)")
plt.show()

sns.catplot(x="Multiple Property Purchase", y="Price (€)", kind="bar", data=df, height=4, aspec
plt.xticks(rotation=90)
plt.title("Relationship Between Multiple Property Purchase and Price (Bar Plot)")
plt.show()
```





Discussion of Outcomes

The above suggests that the vast majority of entries in our dataset are not purchases of multiple properties. In terms of the correlation between this feature and price, we can see that, as a whole, non-multiple property purchases tend to be slightly more expensive. This is likely due to the fact that group property purchases receive an overall discount per property (given that the overall price is likely to be very large anyway). The

spread is largely the same between values, although the non-multiple property purchases seem to have a much larger number of outliers.

This is a useful piece of information, so we will keep this new feature in a dataset.

New Feature 3: Inflation Index

In order to link our dataframe with the wider economic context in Ireland, it is useful to consider the inflation that the country has undergone over the past decade. The assumption is that housing prices will rise and full in direct correlation with the inflation index.

In order to obtain the inflation data, the World Bank API was accessed link here and data was obtained in respect of the inflation index in Ireland between 2010 and 2020.

Unfortunately, there is no way of selecting individual years or countries, and as such we should read in the downloaded csv file into a new pandas dataframe and extract the required information.

In [101…
```
inflation_df = pd.read_csv("inflation.csv")
```

In [102…
```
inflation_df.head()
```

Out[102…

	Country Name	Country Code	Indicator Name	Indicator Code	1960	1961	1962	1963	1964	1965	…	2012	201
0	Aruba	ABW	Inflation, consumer prices (annual %)	FP.CPI.TOTL.ZG	NaN	NaN	NaN	NaN	NaN	NaN	…	0.627472	-2.37206
1	Africa Eastern and Southern	AFE	Inflation, consumer prices (annual %)	FP.CPI.TOTL.ZG	NaN	NaN	NaN	NaN	NaN	NaN	…	9.158707	5.74694
2	Afghanistan	AFG	Inflation, consumer prices (annual %)	FP.CPI.TOTL.ZG	NaN	NaN	NaN	NaN	NaN	NaN	…	6.441213	7.38577
3	Africa Western and Central	AFW	Inflation, consumer prices (annual %)	FP.CPI.TOTL.ZG	NaN	NaN	NaN	NaN	NaN	NaN	…	4.578375	2.43920
4	Angola	AGO	Inflation, consumer prices (annual %)	FP.CPI.TOTL.ZG	NaN	NaN	NaN	NaN	NaN	NaN	…	10.277905	8.77781

5 rows × 66 columns

In [103…
```
irish_inflation_df = inflation_df.loc[inflation_df["Country Name"] == "Ireland"]
irish_inflation_df
```

Out[103...

	Country Name	Country Code	Indicator Name	Indicator Code	1960	1961	1962	1963	1964	1965	...
111	Ireland	IRL	Inflation, consumer prices (annual %)	FP.CPI.TOTL.ZG	0.450837	2.756999	4.263727	2.453621	6.71729	4.988124	...

1 rows × 66 columns

In [104...

```
irish_columns = irish_inflation_df[["2010", "2011", "2012", "2013", "2014", "2015", "2016", "20
```

In [105...

```
irish_inflation_df = irish_inflation_df[irish_columns]
irish_inflation_df
```

Out[105...

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	202
111	-0.922096	2.557189	1.696209	0.508715	0.182542	-0.289879	0.008306	0.340532	0.48837	0.939044	-0.33458

In [106...

```
# Create new inflation column and set to 0
df["Inflation Index"] = 0

# Loop through rows
for index, row in df.iterrows():

    # Get current year, and do nothing if the year is 2021 or 2022 (as we do not have the value
    current_year = str(row["Year"])
    if current_year == "2021" or current_year == "2022":
        value_to_insert = None

    # For all other years insert the value of the inflation index
    else:
        value_to_insert = float(irish_inflation_df[current_year])
    df.at[index, "Inflation Index"] = value_to_insert
```

In [107...

```
df.head()
```

Out[107...

	Date of Sale	Address	Postal Code (Dublin)	County	Price (€)	Full Market Price	VAT Exclusive	Second-Hand	Year	Quarter	Multiple Property Purchase	In
0	2020-06-26	1 GANDON PLACE, GANDON PARK, Lucan	NaN	Dublin	370044.05	True	True	False	2020	2	False	-0.3
1	2014-12-19	72 ST ASSAMS PARK, RAHENY, DUBLIN 5	Dublin 5	Dublin	480000.00	True	False	True	2014	4	False	0.1
2	2010-02-11	37 Ivy Hill, Gort Road, Ennis	NaN	Clare	194000.00	True	False	True	2010	1	False	-0.9

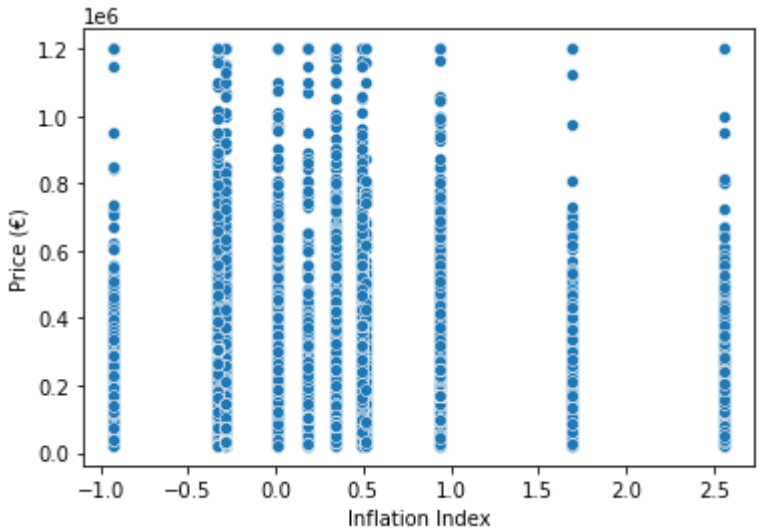
	Date of Sale	Address	Postal Code (Dublin)	County	Price (€)	Full Market Price	VAT Exclusive	Second-Hand	Year	Quarter	Multiple Property Purchase	In
3	2018-08-16	14 TYRCONNELL PLACE, TYRCONNELL RD, INCHICORE ...	Dublin 8	Dublin	275000.00	True	False	True	2018	3	False	0.4
4	2019-02-27	7 THE NEST, TUBBERCURRY, SLIGO	NaN	Sligo	75000.00	True	False	True	2019	1	False	0.9

Plotting New Graphs With Feature

One would think that with a dataset like inflation, the datatype would be continuous. As we can see from the graph below, however, the data behaves like a categorical feature. This is likely due to the fact that we have a single value for each year, which is significantly less than the total number of Price entries that we have.

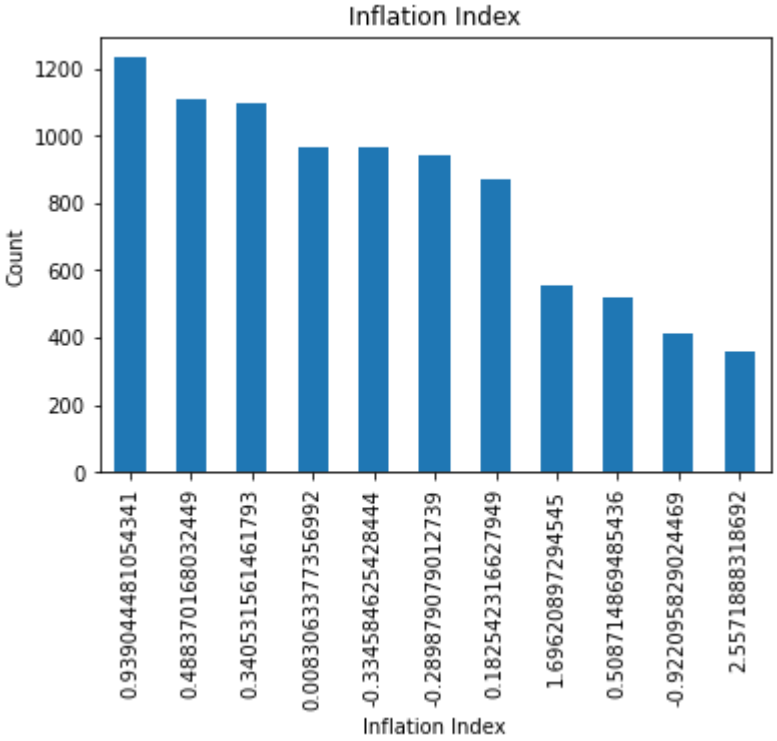
In [108…
```
sns.scatterplot(data=df, x="Inflation Index", y="Price (€)")
```

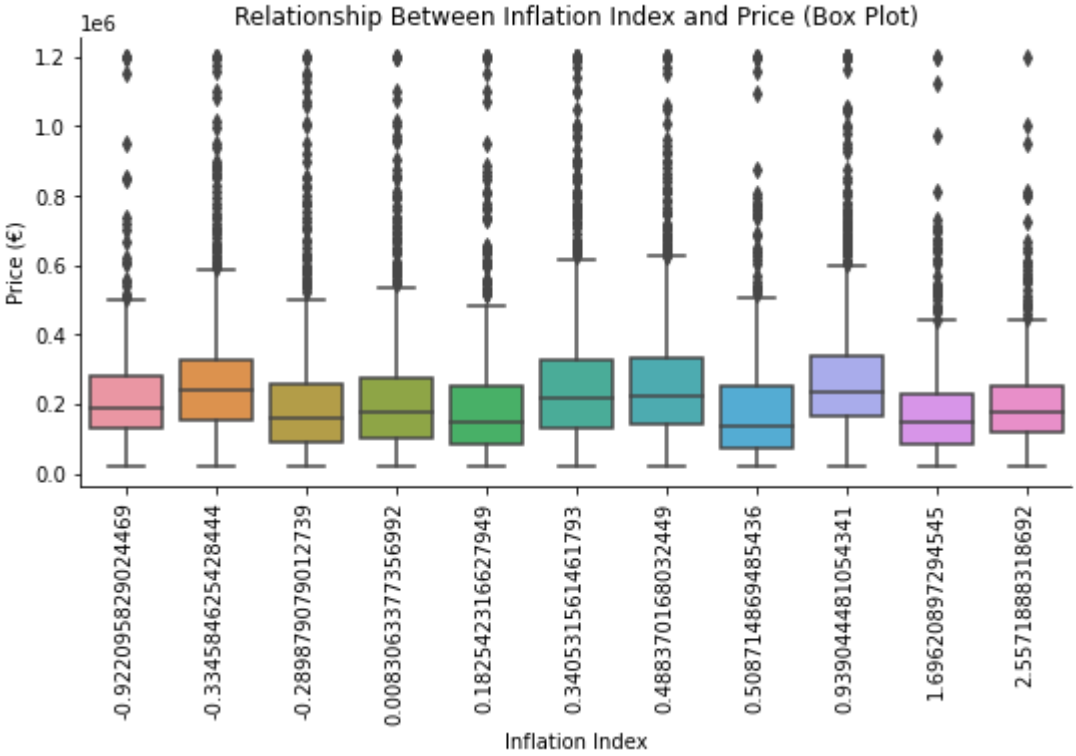
Out[108…
```
<AxesSubplot:xlabel='Inflation Index', ylabel='Price (€)'>
```

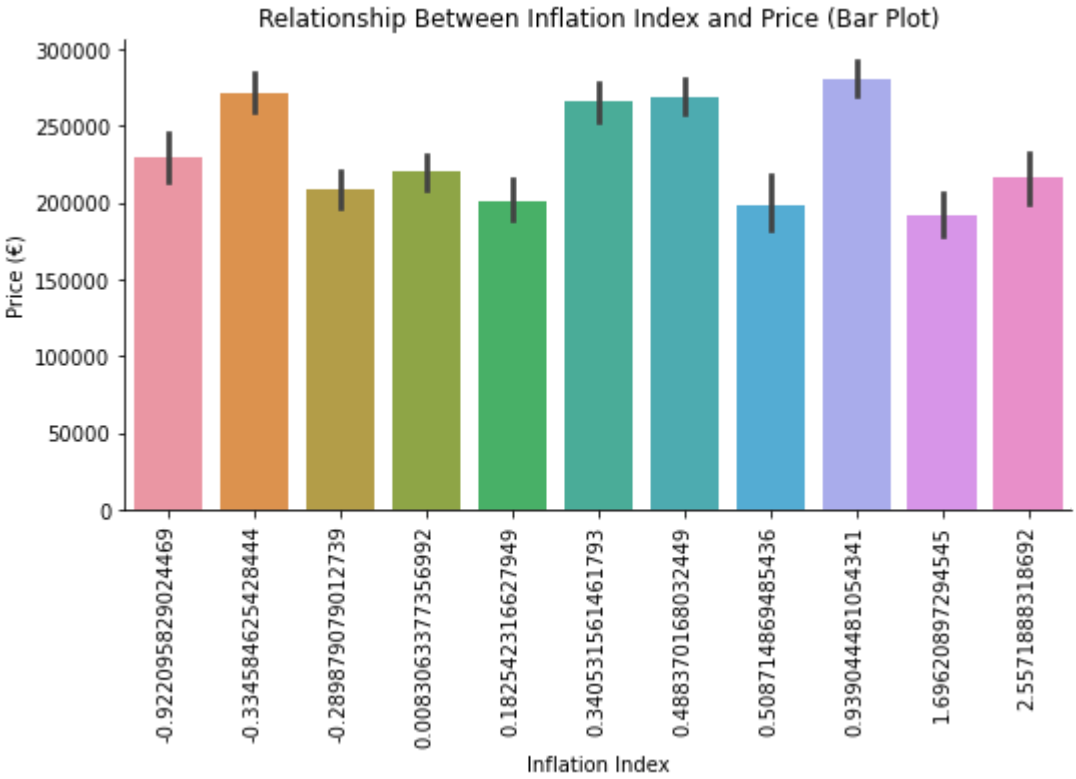


For these reasons, in order to visualise our data, it makes sense to treat inflation as more of a categorical feature, and simply run box and bar plots instead (i.e. the data is already "binned").

In [109…
```
df["Inflation Index"].value_counts().plot(kind = "bar")
plt.xlabel("Inflation Index")
plt.ylabel("Count")
plt.title("Inflation Index")
plt.show()

sns.catplot(x="Inflation Index", y="Price (€)", kind="box", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between Inflation Index and Price (Box Plot)")
plt.show()

sns.catplot(x="Inflation Index", y="Price (€)", kind="bar", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between Inflation Index and Price (Bar Plot)")
plt.show()
```


Inflation Index
Count
0
200
400
600
800
1000
1200
0.9390444810543 41
0.4883701680324 49
0.3405315614617 93
0.0083063377356 992
-0.3345846254284 44
-0.2898790790127 39
0.1825423166279 49
1.6962089729454 5
0.5087148694854 36
-0.9220958290244 69
2.5571888318692
Inflation Index


Relationship Between Inflation Index and Price (Box Plot)
1e6
Price (€)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
-0.9220958290244 69
-0.3345846254284 44
-0.2898790790127 39
0.0083063377356 992
0.1825423166279 49
0.3405315614617 93
0.4883701680324 49
0.5087148694854 36
0.9390444810543 41
1.6962089729454 5
2.5571888318692
Inflation Index



Discussion of Outcomes

Due to the low sample of inflation data, it is difficult to say whether or not there is a correlation between the data.We do note, however, that the largest prices seem to be bunched up between the 0.3 and 0.9 values, with some seemingly random spikes in the -0.33 area. We also see a dip in the 1.69 and 2.55 values, although looking at our feature count, we can see that these two values occur significantly less frequently than any of the others - it is possible, therefore, that we are merely looking at outliers here. The spread was largely the same between values.

It is difficult to say for certain for the reasons outlined above, but it would seem that the price of properties rises as inflation does, which is what we would expect. It should be noted that this statement should not necessarily be relied on, and instead more granular data should be tracked down and inserted into the dataframe for better results. We will still use this new feature in our dataset, as it is useful to show the correlation between property prices and the wider Irish economic context.

New Feature 4: Province

Further to our exercise with Geopy and Folium, we still have a generated CSV file containing some normalised addresses. In order to make use of it, it might be helpful to extract the province from those few addresses that we found and incorporate them into our dataframe.

This is useful as it allows us to view the geographical location of our addresses on a large scale. We already mentioned that addresses have too high a cardinality to visualise property, and as such, plotting the province seems like a useful compromise.

We begin by reading in our CSV file.

In [110…

```
df_addresses = pd.read_csv("Property_Addresses.csv")
```

Next, we create a short function that will return the name of the province that an address belongs to, if it finds this. From our earlier exercise, we note that Geopy appends the name of the province to every address string that it normalises, and as such we expect to get results for every string in our csv file.

In [111…

```
def get_province_of_address(current_address):
    """Searches the normalised address string and returns the name of the province, if found"""
    if "Leinster" in current_address:
        return "Leinster"
    if "Connacht" in current_address:
        return "Connacht"
    if "Munster" in current_address:
        return "Munster"
    if "Ulster" in current_address:
        return "Ulster"
    return
```

Next, we loop through our dataframe and replace the existing addresses with the province name. Note that we will still have our original csv file in case the normalised addresses are required again.

In [112…

```
# Looping through rows
for index in range(df_addresses.shape[0]):
    current_address = df_addresses["Address"][index]

    # If address is not null, call the get_province_of_address function and set it to the curre
    if type(current_address) != float:
        current_address = df_addresses["Address"][index]
        try:
            df_addresses["Address"][index] = get_province_of_address(current_address)
        except:
            pass
```

We will also change the column name from "Address" to "City", to make appending to our dataframe easier.

In [113…

```
df_addresses.rename(columns={"Address": "Province"}, inplace=True)
```

Finally, we will concatenate the new feature on to our existing one.

In [114…

```
df = pd.concat([df, df_addresses], axis=1)
df.head()
```

Out[114…

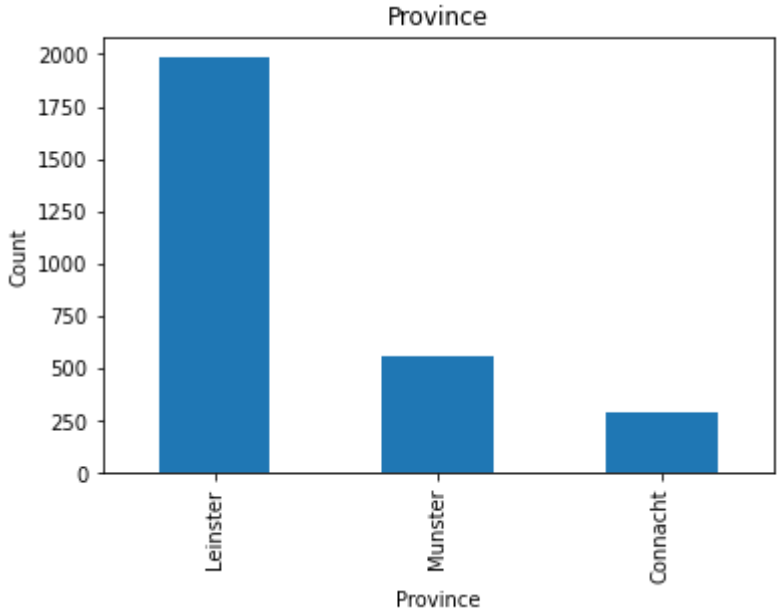
	Date of Sale	Address	Postal Code (Dublin)	County	Price (€)	Full Market Price	VAT Exclusive	Second-Hand	Year	Quarter	Multiple Property Purchase	In
0	2020-06-26	1 GANDON PLACE, GANDON PARK, Lucan	NaN	Dublin	370044.05	True	True	False	2020	2	False	-0.3
1	2014-12-19	72 ST ASSAMS PARK, RAHENY, DUBLIN 5	Dublin 5	Dublin	480000.00	True	False	True	2014	4	False	0.1
2	2010-02-11	37 Ivy Hill, Gort Road, Ennis	NaN	Clare	194000.00	True	False	True	2010	1	False	-0.9
3	2018-08-16	14 TYRCONNELL PLACE, TYRCONNELL RD, INCHICORE ...	Dublin 8	Dublin	275000.00	True	False	True	2018	3	False	0.4
4	2019-02-27	7 THE NEST, TUBBERCURRY, SLIGO	NaN	Sligo	75000.00	True	False	True	2019	1	False	0.9

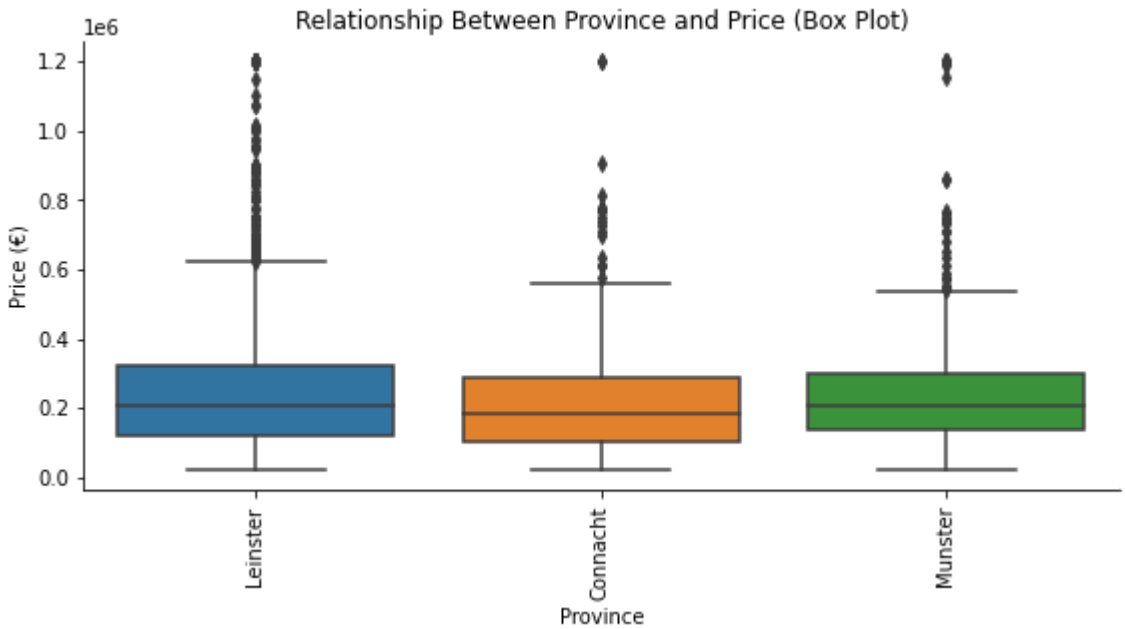
Plotting New Graphs With Feature

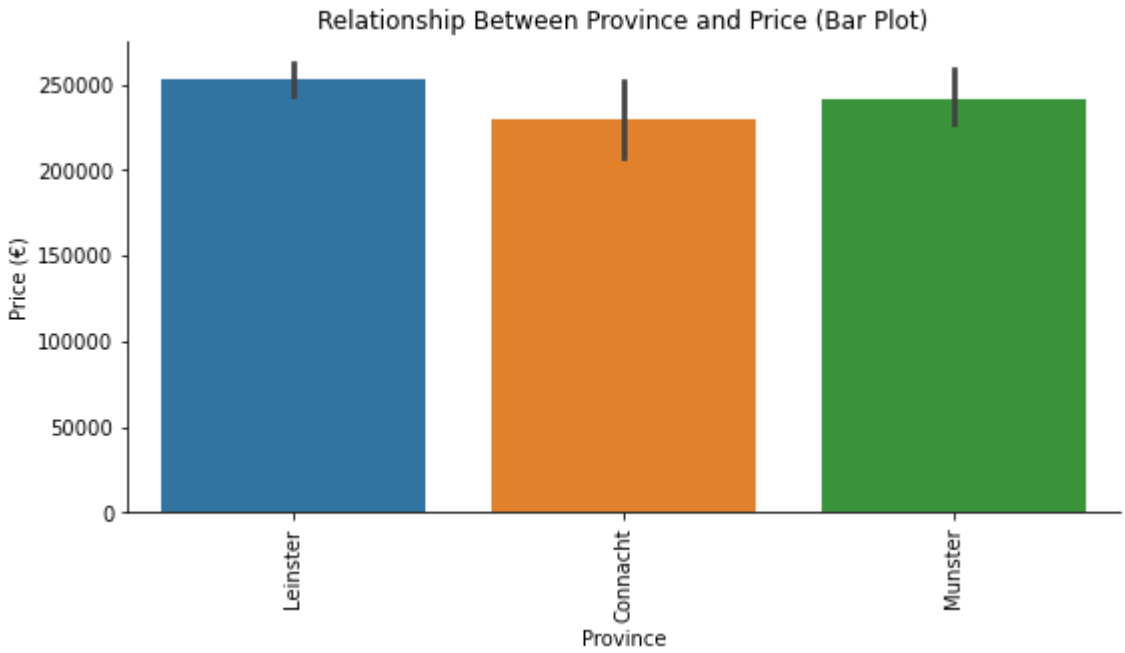
We will first double check to see the size of unique values within the dataframe.

In [115…

```
df["Province"].value_counts().plot(kind = "bar")
plt.xlabel("Province")
plt.ylabel("Count")
plt.title("Province")
plt.show()

sns.catplot(x="Province", y="Price (€)", kind="box", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between Province and Price (Box Plot)")
plt.show()

sns.catplot(x="Province", y="Price (€)", kind="bar", data=df, height=4, aspect=2)
plt.xticks(rotation=90)
plt.title("Relationship Between Province and Price (Bar Plot)")
plt.show()
```







Discussion of Outcomes

We can see from the above graphs that there were significantly more properties sold in Leinster than in any other province. In terms of correlation with price, it would seem that on average Leinster is the more expensive province, followed by Munster, and then finishing with Connacht. This adds up to our analysis earlier where we found that the most expensive counties were Dublin and Cork (located in Leinster and Munster respectively). The spread generally seems largely the same in all provinces.

It is also interesting to note that there are no property sales in our dataset sold in Ulster counties that are part of the Republic of Ireland.

As noted earlier, however, due to the fact that Geopy was not able to process a large number of our address strings, it is possible that a larger sample is required to confirm our deductions from these graphs.

We will keep this feature as this information is useful at showing how price is affected by larger geographical categories.

Save to CSV

Our final step for this assignment is to save down the resulting dataframe, together with all the new features that were created in this section.

In [116…

```
df.to_csv("final_dataframe_with_new_features.csv", index=False)
```

Conclusion

Over the course of this assignment, the randomised data provided to use was reviewed, cleaned, and modified to include new features that are of interest. While there was some issues relating to massive outliers with price, these were dealt with and correlations with other features successfully plotted. While outside the scope of this question, it is now possible to implement some sort of predictive model in connection with our target feature, price, such as linear regression.

Overall, the final .csv file is now ready to be used as an input for a machine learning algorithm, which is the aim that was set out in the beginning of this assignment

END